# $\Pi^0_4$ conservation of a Carlson-Simpson lemma for 1-variable words

Quentin Le Houérou Ludovic Patey

July 30, 2026

**Abstract**

Carlson and Simpson proved that for every finite coloring of the 1-variable words over a finite alphabet $A$, there is an infinite $\omega$-variable word on which all the 1-variable words are monochromatic. This statement for $\ell$-colorings, written $\mathsf{CSL}^1_\ell$, is known to be strictly weaker than $\mathsf{ACA}_0$. We prove that $\mathsf{RCA}_0 + \mathsf{CSL}^1_2$ is a $\forall\Pi^0_4$-conservative extension of $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$. Among its consequences, it implies that neither the indivisibility of the universal triangle-free Henson graph for 2-colorings, nor the tree theorem for pairs and two colors, imply $\Sigma^0_2$-induction. This answers a question of Chong, Li, Wang and Yang [10].

## 1 Introduction

A *1-variable word* over a finite alphabet $A$ is a finite word over the alphabet $A \sqcup \{\star\}$ where the variable $\star$ appears at least once. An *$\omega$-variable word* over $A$ is an infinite sequence over the alphabet $A \sqcup \{x_0, x_1, \dots\}$ where each variable kind $x_i$ appears at least once, and the first occurrence of $x_i$ is before the first occurrence of $x_{i+1}$. Given an $\omega$-variable word $W$ and a 1-variable word $v$ of length $|v|$, we denote by $W[v]$ the 1-variable word obtained by replacing every occurrence of $x_i$ in $W$ by the $i$th element of $v$, and cutting at the first occurrence of the variable $x_{|v|}$. The Carlson-Simpson Lemma [7] for 1-variable words ($\mathsf{CSL}^1_\ell$) is the following statement. See Section 2 for a formal definition.

> For every $\ell$-coloring of the 1-variable words over a finite alphabet $A$, there is an $\omega$-variable word $W$ over $A$ and some color $c < \ell$ such that, for every 1-variable word $v$ over $A$, $W[v]$ has color $c$.

The Carlson-Simpson Lemma is of interest for its connections with the Dual Ramsey theorem [7], the existence of big Ramsey degrees for universal structures [33], and its relation to Hindman's theorem [28]. In particular, $\mathsf{CSL}^1_\ell$ restricted to unary alphabets can be formulated as follows, where $[\mathbb{N}]^{<\omega}$ denotes the collection of all finite sets of integers.

> For every coloring $f : [\mathbb{N}]^{<\omega} \to \ell$, there is an infinite sequence of non-empty (finite or infinite) sets $X_0, X_1, \dots$ with $\min X_i < \min X_{i+1}$

and some color $c < \ell$ such that for every finite set $F \subseteq \mathbb{N}$ of size at least 2, $f(\{0, \dots, \min X_{\max F}\} \cap \bigcup_{i \in F} X_i) = c$.

Note that if we furthermore require that $\max X_i < \min X_{i+1}$, then each $X_i$ is finite, and the resulting statement is equivalent to Hindman's theorem (see Proposition 2.8).

We are interested in the meta-mathematics of the Carlson-Simpson Lemma for 1-variable words from a reverse-mathematical viewpoint. Reverse mathematics is a foundational program whose goal is to find optimal axioms to prove theorems from ordinary mathematics. It uses the framework of subsystems of second-order arithmetic, with a base theory, $\mathsf{RCA}_0$, capturing *computable mathematics*. The theory $\mathsf{RCA}_0$ consists of the axioms of Robinson arithmetic, together with the $\Delta^0_1$-comprehension scheme and $\Sigma^0_1$-induction scheme. See Simpson [51] or Dzhafarov and Mummert [22] for an introduction to reverse mathematics.

Among the systems of interest, $\mathsf{WKL}_0$ is the extension of $\mathsf{RCA}_0$ with König's lemma for binary trees, $\mathsf{ACA}_0$ is the extension of $\mathsf{RCA}_0$ with the arithmetic comprehension scheme, or equivalently with the statement $\forall X \exists Y (Y = X')$, where $X' = \{e : \Phi^X_e(e) \downarrow\}$ is the Turing jump of $X$. Last, $\mathsf{ACA}^+_0$ is the extension of $\mathsf{RCA}_0$ with the statement $\forall X \exists Y (Y = X^{(\omega)})$, where $X^{(\omega)}$ is the $\omega$-jump of $X$ defined inductively as follows: $X^{(0)} = X$, $X^{(n+1)} = (X^{(n)})'$, and $X^{(\omega)} = \bigoplus_n X^{(n)}$. These systems are linearly ordered by logical strength:

$$\mathsf{RCA}_0 < \mathsf{WKL}_0 < \mathsf{ACA}_0 < \mathsf{ACA}^+_0$$

The system $\mathsf{ACA}^+_0$ is famous for being the best known upper bound to Hindman's theorem [54] and, until recently, to the Carlson-Simpson Lemma for 1-variable words [1]. On the other hand, Hindman's theorem implies $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$ [4].

In this article, we prove the following theorem, where a $\forall \Pi^0_n$-formula is of the form $\forall X \varphi(X)$ for some $\Pi^0_n$-formula $\varphi$, and $\mathsf{B}\Sigma^0_n$ denotes the collection (or bounding) scheme for $\Sigma^0_n$-formulas.

**Main Theorem 1.1.** $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ *is* $\forall \Pi^0_4$*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

As $\mathsf{RCA}_0 \vdash \mathsf{CSL}^1_2 \to \mathsf{CSL}^1_\ell$ for every standard $\ell \in \omega$, it follows that $\mathsf{WKL}_0 + \mathsf{CSL}^1_\ell$ is also $\forall \Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$. On the other hand, the statement $\forall \ell \mathsf{CSL}^1_\ell$ implies $\mathsf{B}\Sigma^0_3$ over $\mathsf{RCA}_0$, and therefore is not even $\Pi_1$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ (see Corollary 2.7). By Main Theorem 1.1, $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ does not imply $\mathsf{ACA}_0$ or even $\Sigma^0_2$-induction. The former fact was proven by Liu and Patey [43], and shows a fundamental difference with Hindman's theorem. In particular, since $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ proves the indivisibility of the universal triangle-free Henson graph for 2-colorings, it follows that the latter does not imply $\Sigma^0_2$-induction (see Section 6.1).

## 1.1 Organization of the paper

The proof of Main Theorem 1.1 uses a refinement of the indicator method of Kirby and Paris [36] using a parameterized variant of Ketonen and Solovay's

$\alpha$-largeness [35]. In Section 2, we review the Ramsey theory of variable words, and their analysis in reverse mathematics. Then, in Section 3, we introduce the framework of parameterized largeness. In Section 4 and Section 5, we prove the combinatorial core of our theorem by showing that this notion of largeness is closed under applications of the Graham-Rothschild theorem and the Carlson-Simpson theorem for 1-variable words. In Section 6, we prove Main Theorem 1.1 using these combinatorial results, and study its relations with the indivisibility of the universal triangle-free graph and the tree theorem for pairs. Last, in Section 7, we conclude with open questions.

# 2 Ramsey theory of variable words

In this section, we formally introduce the concepts and theorems from Ramsey theory that will be used throughout the paper[1]. See Graham, Rothschild and Spencer [26] or Todorcevic [52] for general introduction to Ramsey theory, and Dodos and Kanellopoulos [19] for a particular focus to Ramsey's theory for product spaces.

## 2.1 Variable words

We identify a non-negative integer $k \in \mathbb{N}$ with the set $\{0, \dots, k-1\}$. A *word* over a finite alphabet $A$ is a finite ordered sequence $w = \langle a_0, \dots, a_{n-1} \rangle \in A^n$ for some $n \in \mathbb{N}$. An *infinite word* over $A$ is a function $W : \mathbb{N} \to A$. We denote by $A^{<\omega}$ and $A^\omega$ the sets of finite and infinite words over $A$, respectively. For $w = \langle a_0, \dots, a_{n-1} \rangle$, we write $|w|$ for the *length* $n$ of the word $w$ and given $i < n$, we let $w(i) = a_i$.

An *$n$-variable word* over $A$ is a word $w$ over the alphabet $A \sqcup \{x_j : j < n\}$ where each $x_j$ appears at least once and the first occurrence of $x_j$ appears before the first occurrence of $x_{j+1}$. If the last occurrence of $x_j$ appears before the first occurrence of $x_{j+1}$, we say that $w$ is an *ordered $n$-variable word*. We call $n$ the *dimension* of the $n$-variable word $w$. Denote by $A^{<\omega,n}$ and $A^{<\omega,n}_<$ the sets of unordered and ordered $n$-variable words over $A$, respectively. Note that $A^{<\omega} = A^{<\omega,0}$, and that $A^{<\omega,n} \subseteq (A \sqcup \{x_0, \dots, x_{n-1}\})^{<\omega}$.

An *$\omega$-variable word* over $A$ is an infinite word $W$ over the alphabet $A \sqcup \{x_j : j \in \mathbb{N}\}$ where each variable kind $x_j$ appears at least once and the first occurrence of $x_j$ appears before the first occurrence of $x_{j+1}$. The notion of *ordered $\omega$-variable word* is defined accordingly. Note that each variable kind in an $\omega$-variable word can appear infinitely often, while they occur necessarily finitely often in an ordered $\omega$-variable word. We write $A^{\omega,\omega}$ and $A^{\omega,\omega}_<$ for the set of all unordered and ordered $\omega$-variable words over $A$, respectively.

[1] There is a difference of terminology between the literature from combinatorics and from reverse mathematics. In combinatorics, an $n$-variable word is ordered by default, while an unordered $n$-variable word is called an $n$-parameter word. The Ordered Variable Word theorem studied in reverse mathematics is called the Carlson-Simpson theorem in combinatorics, which should not be confused with the Carlson-Simpson Lemma studied in reverse mathematics. The latter corresponds to the unordered version of the Ordered Variable Word theorem.

Given an (ordered) $\omega$-variable word $W$ over $A$ and an (ordered) $n$-variable word $u$ over $A$, we write $W[u]$ for the finite (ordered) $n$-variable word over $A$ where each occurrence of $x_j$ is replaced by $u(j)$, and cut before the first occurrence of $x_{|u|}$. In particular, letting $\epsilon$ be the empty word, $W[\epsilon]$ is the initial segment of $W$ before the first occurrence of $x_0$. The substitution notation $W[a]$ must not be confused with $W(i)$: the former notation denotes the finite word obtained by substitution of all the occurrences of $x_0$ by the letter $a$ and cutting before the first occurrence of $x_1$, while the latter notation denotes the letter in $W$ at position $i$. The variable kind universe $V = \{x_0 < x_1 < \dots\}$ might differ between variable words to avoid name clash or for simplicity. It might be with other symbols $\{y_0 < y_1 < \dots\}$ or range over a subset $\{x_{i_0} < x_{i_1} < \dots\}$. It will always be disjoint from the alphabet $A$.

Given $n \in \mathbb{N}$ and an $\omega$-variable word $W$, we let

$$\mathsf{USub}_A^{n,\star}(W) = \{W[u] : u \in A^{<\omega,n}\}$$

The ordered counterpart of $\mathsf{USub}_A^{n,\star}(W)$ is written $\mathsf{OSub}_A^{n,\star}(W)$.

## 2.2 Infinite variable word theorems

Carlson and Simpson [7] studied a dual version of Ramsey's theorem, and proved for this a purely combinatorial statement about colorings of variable words, known in reverse mathematics as the Carlson-Simpson Lemma:

**Theorem 2.1** (Higher-order Carlson-Simpson Lemma)**.** *Fix $n \geqslant 0$ and $\ell \geqslant 1$. For every finite alphabet $A$ and every finite partition $C_0 \sqcup \dots \sqcup C_{\ell-1} = A^{<\omega,n}$, there is some color $i < \ell$ and an infinite $\omega$-variable word $W$ such that $\mathsf{USub}_A^{n,\star}(W) \subseteq C_i$.*

We write $\mathsf{CSL}_\ell^n(k)$ for the statement of theorem 2.1 restricted to $\ell$-colorings of $n$-variable words over finite alphabets of size $k$. If a parameter is omitted, then it will be assumed to be universally quantified. Thus, $\mathsf{CSL}_2^1$ is the statement $\forall k \mathsf{CSL}_2^1(k)$. The Higher-order Carlson-Simpson Lemma admits a simple inductive proof on $n$, starting from its zero-dimensional version [7].

The Carlson-Simpson lemma belongs to a long line of tree partition theorems, such as the tree theorem [15, 16, 46, 10, 14] and Milliken's tree theorem [45, 2], among others. These partition theorems serve as a backbone of many other combinatorial theorems, such as statements of the existence of finite big Ramsey degrees. In particular, $\mathsf{CSL}$ was used to reprove a theorem by Dobrinen [18] stating that the universal triangle-free Henson graph admits finite big Ramsey degree [33, 1].

The higher-order Carlson-Simpson lemma admits an ordered variable word counterpart, whose case $n = 0$ was proven by Carlson and Simpson [7], and which follows from Carlson's theorem [6] in the general case:

**Theorem 2.2** (Higher-order Ordered Variable Word theorem)**.** *Fix $n \geqslant 0$ and $\ell \geqslant 1$. For every finite alphabet $A$ and every finite partition $C_0 \sqcup \dots \sqcup C_{\ell-1} = A_<^{<\omega,n}$, there is some color $i < \ell$ and an infinite ordered $\omega$-variable word $W$ such that $\mathsf{OSub}_A^{n,\star}(W) \subseteq C_i$.*

We write $\mathsf{OVW}^n_\ell(k)$ for the statement of theorem 2.2 restricted to $\ell$-colorings of ordered $n$-variable words over finite alphabets of size $k$.

The statement $\mathsf{OVW}^0_\ell(k)$ was originally proven using a Baumgartner-style argument [7], and admits multiple quantitative refinements, such as a density version [20] and a piecewise syndetic version. More recently, Liu and Patey [43] gave a proof of $\mathsf{OVW}^0_\ell(k)$ following the style of Towsner [54].

The statement $\mathsf{OVW}^1_\ell(1)$ is a slight variation of the Finite Union theorem, or equivalently Hindman's theorem [28], which is also known to admit a purely combinatorial version due to Towsner [54]. On the other hand, the general statement is a consequence of Carlson's theorem for extracted variable words [6] whose only known proofs use topological dynamics or ultrafilters, except for its zero-dimensional version [5]. Note that Carlson's theorem admits an inductive proof using its one-dimensional version as base case, so the existence of a combinatorial proof of Carlson's theorem and of $\mathsf{OVW}^n_\ell(k)$ is reduced to the existence of a combinatorial proof of $\mathsf{CT}^1_\ell(k)$. Thankfully, the known applications of the higher-order variable word theorems are already consequences of their unordered versions.

## 2.3 Finite variable word theorems

All these variable word statements admit finitary counterparts thanks to a compactness argument. There exist also direct proofs in the spirit of the celebrated Hales-Jewett theorem. Following the new proof of the Hales-Jewett theorem by Shelah [50], all the considered finitary theorems admit primitive recursive bounds [19].

Given an (ordered) $n$-variable word $w$ over $A$ and an (ordered) $m$-variable word $u$ over $A$, with $|u| \leqslant n$, we write $w[u]$ for the finite (ordered) $m$-variable word over $A$ where each occurrence of $x_j$ is replaced by $u(j)$, and cut before the first occurrence of $x_{|u|}$ if $|u| < n$. If $|u| = n$, then $w[u]$ is not cut and $|w[u]| = |w|$. Given $m \leqslant n \in \mathbb{N}$ and a $n$-variable word $w$, we let

$$\mathsf{USub}^{m,\star}_A(w) = \{w[u] : u \in A^{<n,m}\} \quad \text{and} \quad \mathsf{USub}^{m,=}_A(w) = \{w[u] : u \in A^{n,m}\}$$

We also write $\mathsf{USub}^{=}_A(w)$ for $\bigcup_{m \leqslant n} \mathsf{USub}^{m,=}_A(w)$. The ordered counterpart of $\mathsf{USub}^{\square}_A(w)$ is written $\mathsf{OSub}^{\square}_A(w)$, were $\square$ is a placeholder for the various notations.

We shall be interested in the two following finitary theorems: The first theorem is the Graham-Rothschild theorem about unordered variable words [26], whose primitive recursive bounds were proven by Shelah [50]. There exists an ordered variable word counterpart, proven by Dodos and Kanellopoulos [19, Theorem 2.15].

**Theorem 2.3** (Graham-Rothschild, $\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $GR(k, d, m, \ell)$ such that for every $n \geqslant GR(k, d, m, \ell)$, every $n$-variable word $w$ over an alphabet $A$ of size $k$ and every coloring $f : \mathsf{USub}^{m,=}_A(w) \to \ell$, there exists some $v \in \mathsf{USub}^{d,=}_A(w)$ such that $\mathsf{USub}^{m,=}_A(v)$ is $f$-monochromatic.*

The second theorem is the finitary version of the Ordered Variable Word theorem, proven by Dodos, Kanellopoulos and Tyros [20].

**Theorem 2.4** (Finite OVW, $\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $OVW(k,d,m,\ell)$ such that for every $n \geqslant OVW(k,d,m,\ell)$, every ordered $n$-variable word $w$ over an alphabet $A$ of size $k$ and every coloring $f : \mathsf{OSub}_A^{m,\star}(w) \to \ell$, there exists some $v \in \mathsf{OSub}_A^{d,=}(w)$ such that $\mathsf{OSub}_A^{m,\star}(v)$ is $f$-monochromatic.*

We are actually interested in the unordered counterpart of Theorem 2.4 for $m = 0$ and $m = 1$. The unordered version of Theorem 2.4 does not seem to appear in the literature in its full generality. Thankfully, the cases $m = 0$ and $m = 1$ are immediate consequences of Theorem 2.4 since 1-variable words and ordered 1-variable words coincide (see the paragraph after Lemma 3.14).

## 2.4 Variable word theorems in reverse mathematics

Theorems about variable words have been extensively studied in reverse mathematics, due to their connection with many other combinatorial theorems. We now review the reverse-mathematical literature around these statements and their relations to variable word theorems.

*Dual Ramsey theorem and variable words.* In order to prove a dual version of Ramsey's theorem, Carlson and Simpson [7] proved the zero-dimensional version of the Ordered Variable Word theorem [7, Theorem 6.3] and used it as a base principle to inductively prove the now-called Carlson-Simpson Lemma [7, Lemma 2.4]. Miller and Solomon [44] studied both the Dual Ramsey theorem and $\mathsf{OVW}^0$ from the viewpoint of reverse mathematics. They showed that $\mathsf{OVW}^0_2(2)$ is not provable over $\mathsf{WKL}_0$. Dzhafarov, Flood, Solomon and Westrick [21] studied thoroughly the Dual Ramsey theorem and proved that $\mathsf{CSL}$ holds over $\Pi^1_1{-}\mathsf{CA}_0$. Liu, Monin and Patey [42] proved that $\mathsf{OVW}^0(2)$ holds over $\mathsf{ACA}$, a strengthening of $\mathsf{ACA}_0$ with full induction. They also showed that $\mathsf{OVW}^0(2)$ follows from Hindman's theorem over $\mathsf{RCA}_0$. Later, Anglès d'Auriac, Liu, Mignoty and Patey [1] proved that $\mathsf{OVW}^0$ holds over $\mathsf{ACA}_0$ and deduced that for every $n \in \omega$, the statement $\mathsf{CSL}^n$ holds over $\mathsf{ACA}^+_0$, the strengthening of $\mathsf{ACA}_0$ with the axiom stating the existence of the $\omega$-jump of any set. Le Houérou, Patey and Yokoyama [31] proved that $\mathsf{RCA}_0 + \mathsf{OVW}^0$ is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$, from which it follows that $\mathsf{OVW}^0$ does not imply $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$. More recently, Liu and Patey [43] gave a Towsner-style proof of $\mathsf{OVW}^0$ and showed that it implies neither $\mathsf{ACA}_0$, nor Ramsey's theorem for pairs over $\mathsf{RCA}_0$. They also proved that for every $n \in \omega$, the statement $\mathsf{CSL}^n$ holds over $\mathsf{ACA}_0$.

*Hindman's theorem and variable words.* Hindman's theorem ($\mathsf{HT}$) is notoriously equivalent to the Finite Union theorem ($\mathsf{FUT}$), which states, for every $\ell \geqslant 1$ and for every coloring $f : [\mathbb{N}]^{<\omega} \to \ell$, the existence of an infinite sequence of finite sets $\vec{F} = F_0 < F_1 < \dots$ such that $\mathsf{FU}(\vec{F})$ is monochromatic. Here, $\mathsf{FU}(\vec{F})$ denotes the set of all non-empty finite unions over $\vec{F}$. The Finite Union theorem is closely related to the 1-dimensional version of the Ordered Variable Word theorem for unary alphabets. Indeed, $\mathsf{OVW}^1_\ell(1)$ can be reformulated as follows: "For every coloring $f : [\mathbb{N}]^{<\omega} \times \mathbb{N} \to \ell$, there is an infinite sequence of finite sets $\vec{F} = F_0 < F_1 < \dots$ and a color $i < \ell$ such that for

every $E \in \mathsf{FU}(\vec{F})$, and every $s$ such that $\max E < \min F_s$, $f(E, \min F_s) = i$." It is not difficult to see that $\mathsf{FUT}_\ell$ and $\mathsf{OVW}^1_\ell(1)$ are equivalent over $\mathsf{RCA}_0$ (see Proposition 2.8). Hindman's theorem admits multiple proofs, which were systematically analyzed in reverse mathematics: First, the original proof by Hindman [28], which was later simplified by Baumgartner [3]. Both proofs were analyzed by Blass, Hirst, and Simpson [4], who proved that the former holds over $\mathsf{ACA}^+_0$, while the latter holds over the much stronger system $\Pi^1_2$-$\mathsf{TI}_0$. In terms of lower bounds, they also proved that $\mathsf{HT}_2$ implies $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$. It follows from the equivalence between $\mathsf{HT}_2$ and $\mathsf{OVW}^1_2(1)$ that $\mathsf{OVW}^1_2(1)$ implies $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$. Galvin and Glazer [29] proved Hindman's theorem using idempotent ultrafilters. Their proof, analyzed by Towsner [53] holds in $\Sigma^1_1$-$\mathsf{TI}_0$. Last, Towsner [54] gave a simple combinatorial proof of Hindman's theorem, which holds over $\mathsf{ACA}^+_0$.

*Big Ramsey degrees and tree partition theorems.* Structural Ramsey theory studies extensions of Ramsey's theorem to infinite mathematical structures. An infinite structure $\mathfrak{M}$ admits *finite big Ramsey degree* if for every finite substructure $\mathfrak{F}$, there is some number $k_\mathfrak{F}$ such that for every finite coloring of $[\mathfrak{M}]^\mathfrak{F}$ (the isomorphic copies of $\mathfrak{F}$ in $\mathfrak{M}$), there is a sub-copy $\mathfrak{N}$ of $\mathfrak{M}$ such that $[\mathfrak{N}]^\mathfrak{M}$ uses at most $k_\mathfrak{F}$ colors. For instance, Devlin's theorem [17] states that $(\mathbb{Q}, <)$ admits finite big Ramsey degree, while Sauer [48] proved that the random graph admits finite big Ramsey degree. The existence of finite big Ramsey degrees is usually proven using tree partition theorems. For instance, the two previous theorems are reduced to Milliken's tree theorem [2], which holds in $\mathsf{ACA}_0$. The existence of big Ramsey degrees for the universal triangle-free Henson graph was proven by Dobrinen [18] using forcing with strong coded trees. Hubička [33] gave another proof using the Carlson-Simpson Lemma. This was used by Anglès d'Auriac, Liu, Mignoty and Patey [1] to show that the existence of finite big Ramsey degrees for the universal triangle-free Henson graph holds in $\mathsf{ACA}^+_0$. Recently, Cholak, Dobrinen and Towsner [9] proved that it holds over $\mathsf{ACA}_0$ and Cholak, Dobrinen and McCoy [8] showed that it implies $\mathsf{ACA}_0$ as soon as we consider colorings of finite sub-structures of size at least 2. Liu and Patey [43] proved that the indivisibility of the universal triangle-free Henson graph admits cone avoidance, and therefore does not imply $\mathsf{ACA}_0$. In this paper, we show that this indivisibility statement for 2-colorings is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

## 2.5 Preliminary reductions

We conclude this section with some simple reductions, both over $\mathsf{RCA}_0$ and over strong Weihrauch reducibility. The goal is to show to which extent Main Theorem 1.1 is optimal:

- It does not hold when allowing more colors, even when restricted to unary alphabets: We shall see that $\forall\ell\mathsf{CSL}^1_\ell(1)$ implies $\mathsf{B}\Sigma^0_3$ over $\mathsf{RCA}_0$, and therefore is not even $\Pi_1$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.
- It does not hold when asking for the $\omega$-variable word to be ordered, even for unary alphabets. Indeed, $\mathsf{OVW}^1_2(1)$ implies $\mathsf{ACA}_0$ over $\mathsf{ACA}_0$.

**Definition 2.5.** *A problem* $\mathsf{P}$ *is* strongly Weihrauch reducible *to another problem* $\mathsf{Q}$ *(written* $\mathsf{P} \leqslant_{sW} \mathsf{Q}$*) if there are two Turing functionals* $\Phi, \Psi$ *such that for every* $\mathsf{P}$*-instance* $X$*,* $\Phi^X$ *is a* $\mathsf{Q}$*-instance* $\hat{X}$ *such that for every* $\mathsf{Q}$*-solution* $\hat{Y}$ *to* $\hat{X}$*,* $\Phi^{\hat{Y}}$ *is a* $\mathsf{P}$*-solution to* $X$*.*

Given a set $X$ and some $n \in \mathbb{N}$, we write $[X]^n$ for the set of all $n$-subsets of $X$. Ramsey's theorem for $n$-tuples and $\ell$-colorings ($\mathsf{RT}^n_\ell$) is the statement "For every coloring $f : [\mathbb{N}]^n \to \ell$, there is an infinite set $H \subseteq \mathbb{N}$ such that $[H]^n$ is $f$-monochromatic." The following proposition was proven by Liu and Patey [43, Lemma 8.5] using multiple reductions involving intermediary principles. We give a more direct proof:

**Proposition 2.6.** *For every* $n, \ell \geqslant 0$, $\mathsf{RT}^{n+1}_\ell \leqslant_{sW} \mathsf{CSL}^n_\ell(1)$*. Moreover,* $\mathsf{RCA}_0 \vdash \forall n, \ell(\mathsf{CSL}^n_\ell(1) \to \mathsf{RT}^{n+1}_\ell)$*. The same holds when replacing* $\mathsf{CSL}^n_\ell(1)$ *with* $\mathsf{OVW}^n_\ell(1)$*.*

*Proof.* Fix $n, \ell \geqslant 0$. Let $f : [\mathbb{N}]^{n+1} \to \ell$ be an instance of $\mathsf{RT}^{n+1}_\ell$. Let $A = \{a\}$ and define $g : A^{<\omega, n} \to \ell$ by letting $g(w) = f(i_0, \dots, i_{n-1}, |w|)$, where $i_s$ is position of the first occurrence of the variable $x_s$ in $w$. Let $W$ be an $\omega$-variable word over $A$ such that $\mathsf{USub}^{n,\star}_A(W)$ is $g$-monochromatic for some color $c < \ell$. For clarity, say that the variable kinds of $W$ range over $\{y_0, y_1, \dots\}$. Let $H = \{h_0 < h_1 < \dots\}$ be such that $h_s$ is the position of the first occurrence of $y_s$ in $W$.

We claim that $[H]^{n+1}$ is $f$-monochromatic for color $c$. Let $\{h_{i_0} < \dots < h_{i_n}\} \in [H]^{n+1}$. Let $v$ be the variable word of length $i_n$ defined by $v(s) = x_j$ if $s = i_j$, and $v(s) = a$ otherwise. Let $w = W[v]$. Note that $|w| = h_{i_n}$, and that the first occurrence of the variable $x_j$ in $w$ is at position $h_{i_j}$. By definition of $g$, $g(w) = f(h_{i_0}, \dots, h_{i_n}) = c$.

The proof holds over $\mathsf{RCA}_0$, and is exactly the same for ordered variable words. This completes our proof of Proposition 2.6. □

**Corollary 2.7.** $\mathsf{RCA}_0 \vdash \forall \ell \mathsf{CSL}^1_\ell(1) \to \mathsf{B}\Sigma^0_3$*. In particular,* $\forall \ell \mathsf{CSL}^1_\ell(1)$ *is not* $\Pi_1$*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$*.*

*Proof.* Hirst [30, Theorem 6.11] proved that $\mathsf{RCA}_0 \vdash \forall \ell \mathsf{RT}^2_\ell \to \mathsf{B}\Sigma^0_3$. It follows by Proposition 2.6 that $\mathsf{RCA}_0 \vdash \forall \ell \mathsf{CSL}^1_\ell(1) \to \mathsf{B}\Sigma^0_3$. Finally, $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_3$ is not $\Pi_1$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$, as $\mathsf{RCA}_0 \vdash \mathsf{B}\Sigma^0_3 \to \mathsf{Con}(\mathsf{I}\Sigma^0_1)$ but $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \nvdash \mathsf{Con}(\mathsf{I}\Sigma^0_1)$, and $\mathsf{Con}(\mathsf{I}\Sigma^0_1)$ is a $\Pi_1$-sentence. □

The next proposition clarifies the relationship between the ordered variable word theorem for unary alphabets, and the finite union theorem.

**Proposition 2.8.** $\mathsf{RCA}_0 \vdash \forall \ell(\mathsf{OVW}^1_\ell(1) \leftrightarrow \mathsf{FUT}_\ell)$.

*Proof.* Fix $\ell \geqslant 1$. Given a (finite or infinite) variable word $w$ and some variable kind $x_i$, we let $\mathsf{occ}(w, x_i)$ be the set of positions of all the occurrences of $x_i$ in $w$. Whenever $w$ is a 1-variable word, we simply write $\mathsf{occ}(w)$.

- $\mathsf{OVW}^1_\ell(1) \to \mathsf{FUT}_\ell$. Let $f : [\mathbb{N}]^{<\mathbb{N}} \to \ell$ be a coloring. Define the coloring $g : \{0\}^{<\omega, 1} \to \ell$ as follows: $g(w) = f(\mathsf{occ}(w))$. Let $W$ be an ordered

$\omega$-variable word over $\{0\}$ such that $\mathsf{OSub}^{1,\star}_{\{0\}}(W)$ is $g$-monochromatic for some color $c < \ell$. For every $s \in \mathbb{N}$, let $F_s = \mathsf{occ}(W, x_s)$. Note that for every $s \in \mathbb{N}$, $\max F_s < \min F_{s+1}$ as $W$ is ordered. We claim that for every $H \in \mathsf{FU}(F_0, F_1, \dots)$, $f(H) = c$. Indeed, for every such $H$, there is a 1-variable word $w \in \mathsf{OSub}^{1,\star}_{\{0\}}(W)$ such that $H = \mathsf{occ}(w)$. Then $f(H) = g(w) = c$.

- $\mathsf{FUT}_\ell \to \mathsf{OVW}^1_\ell(1)$. By Blass, Hirst, and Simpson [4], $\mathsf{RCA}_0 \vdash \mathsf{FUT}_\ell \to \mathsf{ACA}_0$. Let $A = \{a\}$ be a unary alphabet, and $f : A^{<\omega,1} \to \ell$ be a coloring. For every $v \in A^{<\omega,1}$ and $c < \ell$, let $R_{v,c}$ be the set of all $n \in \mathbb{N}$ such that $f(v \cdot a^{n-|v|}) = c$. Here, $a^x$ denotes the word of length $x$ composed only of occurrences of the letter $a$. By $\mathsf{COH}$, which holds over $\mathsf{ACA}_0$, there is an infinite $\vec{R}$-cohesive set $C = \{p_0 < p_1 < \dots\} \subseteq \mathbb{N}$, that is, an infinite set $C \subseteq \mathbb{N}$ such that for every $v \in A^{<\omega,1}$ and $c < \ell$, either $C \subseteq^* R_{v,c}$, or $C \subseteq^* \overline{R}_{v,c}$. Here, $\subseteq^*$ means "included up to finite changes". Note that for every $u_0, u_1 \in A^{<\omega,1}$ such that $\mathsf{occ}(u_0) = \mathsf{occ}(u_1)$, $u_0$ and $u_1$ are comparable, and for every $c < \ell$, $R_{u_0,c} =^* R_{u_1,c}$, where $=^*$ means "equal up to finite changes". Moreover, for every $u \in A^{<\omega,1}$, there is a unique $c < \ell$ such that $C \subseteq^* R_{u,c}$. It follows that for every finite $H \subseteq C$, there is a unique color $c < \ell$ such that for every 1-variable word $u$ satisfying $\mathsf{occ}(u) = H$, $C \subseteq^* R_{u,c}$. Let $g : \mathcal{P}_{fin}(C) \to \ell$ be defined by $g(H) = c$ for this unique color $c$. By $\mathsf{FUT}_\ell$, there is an infinite sequence $F_0 < F_1 < \dots$ of finite subsets of $C$ and some color $c < \ell$ such that for every $H \in \mathsf{FU}(F_0, F_1, \dots)$, $g(H) = c$. Let $W$ be the $\omega$-variable word defined for every $s \in \mathbb{N}$ by $W(s) = x_i$ if $s \in F_i$, and $W(s) = a$ otherwise. Then, for every $u \in \mathsf{OSub}^{1,\star}_A(W)$, $\mathsf{occ}(u) \in \mathsf{FU}(F_0, F_1, \dots)$, hence $g(\mathsf{occ}(u)) = c$. By definition of $g$, $C \subseteq^* R_{u,c}$, so $\lim_{n \in C} f(u \cdot a^{n-|u|}) = c$. By a greedy algorithm, one can $W \oplus f$-compute an $\omega$-variable word $V \in \mathsf{OSub}_A(W)$ such that for every $u \in \mathsf{OSub}^{1,\star}_A(V)$, $f(u) = c$.

□

**Corollary 2.9.** $\mathsf{RCA}_0 \vdash \mathsf{OVW}^1_2(1) \to \mathsf{ACA}_0$. *In particular,* $\mathsf{OVW}^1_2(1)$ *is not* $\Pi_1$*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

*Proof.* By Blass, Hirst, and Simpson [4], $\mathsf{RCA}_0 \vdash \mathsf{FUT}_2 \to \mathsf{ACA}_0$, so we conclude by Proposition 2.8. Moreover, $\mathsf{ACA}_0 \vdash \mathsf{Con}(\mathsf{I}\Sigma^0_1)$. □

Note the difference between $\mathsf{OVW}^1_2$ and $\mathsf{CSL}^1_2$, as the former implies $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$ by Corollary 2.9, while the latter is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ by Main Theorem 1.1.

# 3 Largeness below $\omega^\omega$

Ketonen and Solovay [35] defined a notion of $\alpha$-largeness for $\alpha < \varepsilon_0$ quantifying the size of finite sets, to generalize the celebrated Paris-Harrington theorem. This notion of largeness was later used as a refinement of the Kirby and

Paris [36] indicator method to produce $\forall\Pi^0_3$-conservation theorems over various subsystems of arithmetic [47, 38, 37, 55]. More recently, the authors and Yokoyama defined a parameterized version of largeness to prove partial conservation theorems over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ for $\forall\Pi^0_4$-sentences [39, 41].

We now present this notion of largeness, parameterized by a fixed $\Delta^0_0$-formula $\theta(x, y, z)$.

**Definition 3.1.** *Two finite sets $X < Y$ are* $\theta$-apart *if*

$$\forall x < \max X \exists y < \min Y \forall z < \max Y \theta(x, y, z)$$

Note that $\theta$-apartness is a transitive relation. Moreover, if $X < Y$ are $\theta$-apart and $X_0 \subseteq X$ and $Y_0 \subseteq Y$, then $X_0, Y_0$ are $\theta$-apart. Contrary to Ketonen and Solovay's notion of largeness, we only consider ordinals $\alpha$ of the form $\omega^n \cdot k$, with $n, k \in \mathbb{N}$.

**Definition 3.2** ([39])**.** *A set $X \subseteq_{\mathtt{fin}} \mathbb{N}$ is*

- $\omega^0$-large$(\theta)$ *if $X \neq \varnothing$.*
- $\omega^{(n+1)}$-large$(\theta)$ *if $X \setminus \min X$ is $(\omega^n \cdot \min X)$-large$(\theta)$*
- $\omega^n \cdot k$-large$(\theta)$ *if there are $k$ pairwise $\theta$-apart $\omega^n$-large$(\theta)$ subsets of $X$*

$$X_0 < X_1 < \cdots < X_{k-1}$$

**Proposition 3.3** ([39])**.** *For every $n \in \omega$, $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \forall x \exists y \forall z \theta(x, y, z)$ proves that for every $b \geqslant 1$, $\omega^n \cdot b$-largeness$(\theta)$ is a largeness notion.*

A set $X$ is $\mathsf{RT}^n_k$-$\omega^n$-*large$(\theta)$* if for every coloring $f : [X]^n \to k$, there is an $\omega^n$-large$(\theta)$ subset $H \subseteq X$ such that $[H]^n$ is $f$-monochromatic. Le Houérou, Patey and Yokoyama [39, 40] proved the following closure property:

**Theorem 3.4** ([39, 40], $\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{RT}^2_2} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\omega^{f_{\mathsf{RT}^2_2}(n)}$-large$(\theta)$ set is $\mathsf{RT}^2_2$-$\omega^n$-large$(\theta)$.*

In a follow-up article [32], the authors proved that such a function $f_{\mathsf{RT}^2_2}$ can be chosen to be polynomial.

## 3.1 Basic combinatorics

**Lemma 3.5** ([31], $\mathsf{RCA}_0$)**.** *For every $a, b$ and for every $\omega^{a+2b+1}$-large$(\theta)$ set $X$, then there are some $k \in \mathbb{N}$ and some $\omega^a$-large$(\theta)$ pairwise $\theta$-apart subsets $X_0 < \cdots < X_{k-1}$ of $X$ such that every $H \in \prod_{i<k} X_i$ is $\omega^b$-large$(\theta)$ (here, we abuse the product notation and see $H$ as a subset of $\mathbb{N}$ intersecting every $X_i$ exactly once rather than a tuple in the product $X_0 \times \cdots \times X_{k-1}$).*

Given an increasing function $g : \mathbb{N} \to \mathbb{N}$, a set $X \subseteq \mathbb{N}$ is $g$-*sparse* if for every $x, y \in X$ with $x < y$, then $g(x) < y$. We shall only consider $g$-sparsity for primitive recursive functions $g$.

**Lemma 3.6** (Folklore, see Ketonen and Solovay [35]). *For every primitive recursive function $g : \mathbb{N} \to \mathbb{N}$, there exists some $a \in \omega$ such that every $\omega^a$-large (and a fortiori $\omega^a$-large($\theta$)) set $X$ satisfies $|X| > g(\min X - 1)$.*

Combining Lemma 3.6 with Lemma 3.5, one can always assume that an $\omega^n$-large($\theta$) set is $g$-large for some fixed primitive recursive function $g$, at the cost of a polynomial overhead in the exponent:

**Lemma 3.7** ($\mathsf{RCA}_0$). *For every primitive recursive function $g : \mathbb{N} \to \mathbb{N}$, there is some $a \in \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\omega^{2n+a}$-large($\theta$) set contains a $g$-sparse $\omega^n$-large($\theta$) subset.*

*Proof.* Fix $g$. By Lemma 3.6, there is some $a \in \mathbb{N}$ such that every $\omega^a$-large($\theta$) set $X$ satisfies $|X| > g(\min X - 1)$. Let $X$ be an $\omega^{2n+a}$-large($\theta$) set. By Lemma 3.5, there are some $\omega^a$-large($\theta$) pairwise $\theta$-apart subsets $X_0 < \cdots < X_{k-1}$ of $X$ such that every $H \in \Pi_{i<k} X_i$ is $\omega^n$-large($\theta$). In particular, the set $H = \{\max X_i : i < k\}$ is $\omega^n$-large($\theta$). Moreover, for every $x, y \in H$ such that $x < y$, the interval $]x, y]$ contains an $\omega^a$-large($\theta$) set, so $g(x) < y$. It follows that $H$ is $g$-sparse. □

## 3.2 Largeness and variable words

Largeness is defined in terms of sets of integers, while the considered theorems are statements about words and variable words. We bridge the two notions by defining the notions of ordered and unordered $X$-variable word over a finite set $X \subseteq \mathbb{N}$.

**Definition 3.8.** *Let $w$ be an (ordered) $n$-variable word over a finite alphabet $A$. We write $\mathsf{set}(w)$ for the set of positions of first occurrences of the variable kinds. We say that $w$ is $\alpha$-large($\theta$) if $\mathsf{set}(w)$ is $\alpha$-large($\theta$).*

Given a finite set $X$, an (ordered) $|X|$-variable word $w$ over $A$ such that $\mathsf{set}(w) = X$ is called an *(ordered) $X$-variable word*. We are now ready to define largeness versions of the Graham-Rothschild theorem, of the Ordered Variable Word theorem and the Carlson-Simpson lemma.

**Definition 3.9.** *A set $X \subseteq_{fin} \mathbb{N}$ is said to be $\mathsf{GR}^n_\ell(k)$-$\alpha$-large($\theta$) if for every $X$-variable word $w$ over an alphabet $A$ of size $k$ and every coloring $f : \mathsf{USub}^{n,=}_A(w) \to \ell$, there exists some $\alpha$-large($\theta$) $v \in \mathsf{USub}^{=}_A(w)$ such that $\mathsf{USub}^{n,=}_A(v)$ is $f$-monochromatic. The set $X$ is $\mathsf{GR}^n$-$\alpha$-large($\theta$)* if it is $\mathsf{GR}^n_\ell(k)$-$\alpha$-large($\theta$) for every $k, \ell < \min X$.

Given two sets $X_0 < X_1$ and an $X_0 \cup X_1$-variable word $w$, it will often be useful to decompose $w$ into $w_0 \cdot w_1$ with $|w_0| < \min X_1$, where $w_0$ is an $X_0$-variable word and $w_1$ is a *located* $X_1$-variable word. By *located word*, we mean a word $u$ whose domain is an interval in $\mathbb{N}$ rather than an initial segment of $\mathbb{N}$. In this case, we write $|u|$ for the least upper bound of its domain, and consider the concatenation of two located words only if the disjoint union of their domain forms an interval.

**Definition 3.10.** *A set $X \subseteq_{fin} \mathbb{N}$ is said to be* $\mathsf{OVW}^n_\ell(k)$-$\alpha$-large($\theta$) *if for every ordered $X$-variable word $w$ over an alphabet $A$ of size $k$ and every coloring $f : \mathsf{OSub}^{n,\star}_A(w) \to \ell$, there exists some $\alpha$-large($\theta$) $v \in \mathsf{OSub}^{=}_A(w)$ such that $\mathsf{OSub}^{n,\star}_A(v)$ is $f$-monochromatic. The set $X$ is* $\mathsf{OVW}^n$-$\alpha$-large($\theta$) if it is $\mathsf{OVW}^n_\ell(k)$-$\alpha$-large($\theta$) for every $k, \ell < \min X$.

**Definition 3.11.** *A set $X \subseteq_{fin} \mathbb{N}$ is said to be* $\mathsf{CSL}^n_\ell(k)$-$\alpha$-large($\theta$) *if for every $X$-variable word $w$ over an alphabet $A$ of size $k$ and every coloring $f : \mathsf{USub}^{n,\star}_A(w) \to \ell$, there exists some $\alpha$-large($\theta$) $v \in \mathsf{USub}^{=}_A(w)$ such that $\mathsf{USub}^{n,\star}_A(v)$ is $f$-monochromatic. The set $X$ is* $\mathsf{CSL}^n$-$\alpha$-large($\theta$) if it is $\mathsf{CSL}^n_\ell(k)$-$\alpha$-large($\theta$) for every $k, \ell < \min X$.

For every $n$-variable word $w$ over an alphabet $A$ of size $k$, there exists a canonical isomorphism

$$\iota_w : \begin{cases} [k]^{<n,d} \to \mathsf{USub}^{d,\star}_A(w) & \text{for every } d < n \\ [k]^{n,d} \to \mathsf{USub}^{d,=}_A(w) \end{cases}$$

which maps the unique element of $[k]^{|w|,|w|}$ to $w$. Accordingly, if $w$ is an ordered $n$-variable word, there is an isomorphism

$$\iota_{w,<} : \begin{cases} [k]^{<n,d}_< \to \mathsf{OSub}^{d,\star}_A(w) & \text{for every } d < n \\ [k]^{n,d}_< \to \mathsf{OSub}^{d,=}_A(w) \end{cases}$$

Thanks to these isomorphisms, one can simplify the definition of the largeness versions of our theorems:

**Lemma 3.12.** *A set $X = \{x_0 < \cdots < x_{a-1}\}$ is $\mathsf{GR}^n_\ell(k)$-$\alpha$-large($\theta$) iff for every coloring $f : [k]^{a,n} \to \ell$, there is some $v \in \mathsf{USub}^{=}_{[k]}([k]^{a,a})$ such that $\{x_i : i \in \mathsf{set}(v)\}$ is $\alpha$-large($\theta$) and $\mathsf{USub}^{n,=}_{[k]}(v)$ is $f$-monochromatic.*

**Lemma 3.13.** *A set $X = \{x_0 < \cdots < x_{a-1}\}$ is $\mathsf{OVW}^n_\ell(k)$-$\alpha$-large($\theta$) iff for every coloring $f : [k]^{<a,n}_< \to \ell$, there is some $v \in \mathsf{OSub}^{=}_{[k]}([k]^{a,a})$ such that $\{x_i : i \in \mathsf{set}(v)\}$ is $\alpha$-large($\theta$) and $\mathsf{OSub}^{n,\star}_{[k]}(v)$ is $f$-monochromatic.*

**Lemma 3.14.** *A set $X = \{x_0 < \cdots < x_{a-1}\}$ is $\mathsf{CSL}^n_\ell(k)$-$\alpha$-large($\theta$) iff for every coloring $f : [k]^{<a,n} \to \ell$, there is some $v \in \mathsf{USub}^{=}_{[k]}([k]^{a,a})$ such that $\{x_i : i \in \mathsf{set}(v)\}$ is $\alpha$-large($\theta$) and $\mathsf{USub}^{n,\star}_{[k]}(v)$ is $f$-monochromatic.*

Since the notions of ordered and unordered $n$-variable words coincide for $n = 0$ and $n = 1$, it follows from the previous lemmas that for $n \in \{0, 1\}$, any $\mathsf{OVW}^n$-$\alpha$-large($\theta$) set is $\mathsf{CSL}^n$-$\alpha$-large($\theta$). The authors [31] proved explicit bounds for $\mathsf{GR}^0$-$\alpha$-large($\theta$)[2] and $\mathsf{OVW}^0$-$\alpha$-large($\theta$) for $\alpha$ for the form $\omega^n \cdot r$:

**Theorem 3.15** ([31], $\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{GR}^0} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\omega^{f_{\mathsf{GR}^0}(n)}$-large($\theta$) set is $\mathsf{GR}^0$-$\omega^n$-large($\theta$).*

[2] Actually, the authors used an ordered version of the Graham-Rotschild theorem, say $\mathsf{OGR}^0$-$\alpha$-largeness($\theta$). Simply note that any $\mathsf{OGR}^0$-$\alpha$-large($\theta$) set is $\mathsf{GR}^0$-$\alpha$-large($\theta$).

**Theorem 3.16** ([31], $\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{OVW}^0} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\mathfrak{w}^{f_{\mathsf{OVW}^0}(n)}$-large($\theta$) set is $\mathsf{OVW}^0$-$\mathfrak{w}^n$-large($\theta$).*

**Corollary 3.17.** *There exists a primitive recursive function $f_{\mathsf{CSL}^0} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\mathfrak{w}^{f_{\mathsf{CSL}^0}(n)}$-large($\theta$) set is $\mathsf{CSL}^0$-$\mathfrak{w}^n$-large($\theta$).*

*Proof.* Let $f_{\mathsf{CSL}^0} = f_{\mathsf{OVW}^0}$ since any $\mathsf{OVW}^0$-$\mathfrak{w}^n$-large($\theta$) set is $\mathsf{CSL}^0$-$\mathfrak{w}^n$-large($\theta$). □

Let $X_0 < X_1$ be two sets, and $X = X_0 \cup X_1$. Let $w$ be an $X$-variable word over an alphabet $A$. For clarity, we assume that the variables space ranges over $\{x_n : n \in X\}$. We can write $w = w_0 \cdot w_1$, where $|w_0| < \min X$, so that $w_0$ is an $X_0$-variable word over $A$ and $w_1$ is a located $X_1$-variable word over the extended alphabet $B = A \sqcup \{x_p : p \in X_0\}$. For every $v_0 \in \mathsf{USub}_A^{0,=}(w_0)$, $v_0 \cdot w_1[x_p \mapsto v_0(p)]$ is an $X_1$-variable word over $A$, where $w_1[x_p \mapsto v_0(p)]$ means that every occurrence of the variable $x_p$ with $p \in X_0$ is replaced with the value of $v_0$ at position $p$.

# 4 Graham-Rothschild theorem

The goal of this section is to prove the following theorem:

**Theorem 4.1** ($\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{GR}^1} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\mathfrak{w}^{f_{\mathsf{GR}^1}(n)}$-large($\theta$) set is $\mathsf{GR}^1$-$\mathfrak{w}^n$-large($\theta$).*

The core of the argument lies in the following technical proposition, which state sufficient largeness conditions for "grouping" blocks of solutions into a larger solution.

**Proposition 4.2.** *Let $n \in \mathbb{N}$, let $r : \mathbb{N} \to \mathbb{N}$ be some non-decreasing function, and suppose that the following induction hypothesis holds:*

> *There exists some $a \in \mathbb{N}$ such that every $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^a$-large($\theta$) set $X$ is $\mathsf{GR}^1$-$\mathfrak{w}^n \cdot r(\min X)$-large($\theta$).*

*Let $k, \ell \in \mathbb{N}$, and $X_0 < \cdots < X_{k-1}$ be $\theta$-apart and $\mathfrak{w}^{f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(a))}$-large($\theta$) sets such that $\ell^k < \min X_0$. Let $X = X_0 \cup \cdots \cup X_{k-1}$ and let $w = w_0 \cdots w_{k-1}$ be an $X$-variable word over an alphabet $A$ with $|A| < \min X_0$.*

*Then, for every coloring $f : \mathsf{USub}_A^{1,=}(w) \to \ell$, there exist $\mathfrak{w}^n \cdot r(\min X_i)$-large($\theta$) subsets $W_i \subseteq X_i$ and some $W_0 \cup \cdots \cup W_{k-1}$-variable word $v \in \mathsf{USub}_A^{=}(w)$ such that the color of any 1-variable word $u \in \mathsf{USub}_A^{1,=}(v)$ only depends on which $X_i$ the first occurrence of the variable belongs.*

*Proof.* We say that some 1-variable word $u \in \mathsf{USub}_A^{1,=}(w)$ is of *type* $i$ (for some $i < k$) if the first occurrence of the variable belongs to $X_i$. Given a set $Y \subseteq X$, the 1-variable words of type $i$ are $Y$-*invariant* if their color do not depend on the valuations of the variables $\{x_n : n \in Y\}$.

**Proof outline:** The construction will follow $3k$ steps, in the following order:

Step $(k-1)$.a, Step $(k-1)$.b, Step $(k-1)$.c, ..., Step 0.a, Step 0.b, Step 0.c.

Let $i < k$, and assume $W_j \subseteq X_j$ have been defined for $j \in \{i+1, \dots, k-1\}$.

- In Step i.a, we are going to find some subset $Y_i \subseteq X_i$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Y_i \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,a} \in \mathsf{USub}_A^=(w^{i+1,c})$ (with the convention that $w^{k,c} = w$) such that the 1-variable words of type $< i$ are $Y_i$-invariant.
- In Step i.b, we will find some subset $Z_i \subseteq Y_i$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Z_i \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,b} \in \mathsf{USub}_A^=(w^{i,a})$ such that the 1-variable words of type $> i$ are $Z_i$-invariant.
- In Step i.c, we will find some subset $W_i \subseteq Z_i$ and some $X_0 \cup \cdots \cup X_{i-1} \cup W_i \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,c} \in \mathsf{USub}_A^=(w^{i,b})$ such that the 1-variable words of type $i$ are $W_i$-invariant.

By induction, at the end of Step i.c, for every 1-variable word $u \in \mathsf{USub}_A^{1,=}(w^{i,c})$,

- if the type of $u$ is $< i$, then it is $W_i \cup \cdots \cup W_{k-1}$-invariant;
- if the type of $u$ is $\geqslant i$, then $f(u)$ only depends on the valuations of $X_0 \cup \cdots \cup X_{i-1}$ and on its type.

In particular, assuming that each $W_i$ is $\mathfrak{w}^n \cdot r(\min X_i)$-large$(\theta)$, at the end of Step 0.c, $w^{0,c} \in \mathsf{USub}_A^=(w)$ is a $W_0 \cup \cdots \cup W_{k-1}$-variable word satisfying the conclusion of Proposition 4.2.

**Formal proof:** Consider the extended alphabets

$$B := A \sqcup \{x_p : p \in X_0 \cup \cdots \cup X_{i-1}\}$$
$$C := A \sqcup \{x_p : p \in X_0 \cup \cdots \cup X_{i-1} \cup X_i\}$$

**Step i.a**. Assume $w^{i+1,c} = w^- \cdot w_i \cdot w^+$ has already been defined, where $w^-$ is an $X_0 \cup \cdots \cup X_{i-1}$-variable word over $A$, $w_i$ is an $X_i$-variable word over $B$ and $w^+$ is an $W_{i+1} \cup \cdots \cup W_{k-1}$-variable word over $C$. By construction, we have $w^- = w_0 \cdots w_{i-1}$.
For every $u^- \in \mathsf{USub}_A^{1,=}(w^-)$, let

$$g_{u^-} : \mathsf{USub}_B^{0,=}(w_i) \to \ell$$

be defined by

$$g_{u^-}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)])$$

for any $u^+ \in \mathsf{USub}_C^{0,=}(w^+)$. By Steps j.a for $j > i$, the coloring $g_{u^-}$ is well-defined. Finally, let

$$g : \mathsf{USub}_B^{0,=}(w_i) \to \ell^{|\mathsf{USub}_A^{1,=}(w^-)|}$$

be defined by

$$g(u_i) = \langle g_{u^-}(u_i) : u^- \in \mathsf{USub}_A^{1,=}(w^-)\rangle$$

We have $\min X_i > \ell^{|\mathsf{USub}_A^{1,=}(w^-)|}$. Indeed, in the case where $i = 0$ we have $\ell^{|\mathsf{USub}_A^{1,=}(w^-)|} = \ell$, and in the case where $i \neq 0$, we have $\ell^{|\mathsf{USub}_A^{1,=}(w^-)|} \leqslant \ell^{(\max X_{i-1})^{\max X_{i-1}}} < \min X_i$ by $(x \mapsto x^{x^x})$-sparsity of $X$. Hence, by definition of $f_{\mathsf{GR}^0}$, there exist some $\omega^{f_{\mathsf{GR}^0}(a)}$-large($\theta$) subset $Y_i \subseteq X_i$ and some $Y_i$-variable word $w_i^{i,a} \in \mathsf{USub}_B^{=}(w^i)$ such that $\mathsf{USub}_B^{0,=}(w_i^{i,a})$ is $g$-monochromatic. By letting $w^{i,a} := w^- \cdot w_i^{i,a} \cdot w^+$, we have $w^{i,a} \in \mathsf{USub}_A^{=}(w^{i+1,c})$ and all the 1-variable words of type $< i$ are $Y_i$-invariant.

**Step i.b**. Assume $w^{i,a} = w^- \cdot w_i^{i,a} \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$.
For every $u^- \in \mathsf{USub}_A^{0,=}(w^-)$ and $j \in \{i+1, \ldots, k-1\}$, let

$$h_{u^-,j} : \mathsf{USub}_B^{0,=}(w_i^{i,a}) \to \ell$$

be defined by

$$h_{u^-,j}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)])$$

for any $u^+ \in \mathsf{USub}_C^{1,=}(w^+)$ of type $j$. By Steps j.a, j.b and j.c, the coloring $h_{u^-,j}$ is well-defined. Finally, let

$$h : \mathsf{USub}_B^{0,=}(w_i^{i,a}) \to \ell^{|\mathsf{USub}_A^{0,=}(w^-)| \times (k-i-1)}$$

be defined by

$$h(u_i) = \langle h_{u^-,j}(u_i) : j \in \{i+1, \ldots, k-1\} \text{ and } u^- \in \mathsf{USub}_A^{0,=}(w^-) \text{ of type } j\rangle$$

By $(x \mapsto x^{x^x})$-sparsity of $X$ and from the fact that $\ell^k < \min X_0$, we have $\min Y_i > \ell^{|\mathsf{USub}_A^{1,=}(w^-)| \times k}$, hence, by definition of $f_{\mathsf{GR}^0}$, there exists an $\omega^a$-large($\theta$) subset $Z_i \subseteq Y_i$ and some $Z_i$-variable word $w_i^{i,b} \in \mathsf{USub}_B^{=}(w_i^{i,a})$ such that $\mathsf{USub}_B^{0,=}(w^{i,b})$ is $h$-monochromatic. By letting $w^{i,b} := w^- \cdot w_i^{i,b} \cdot w^+$, we have $w^{i,b} \in \mathsf{USub}_A^{=}(w^{i,a})$ and all the 1-variable words of type $> i$ are $Z_i$-invariant.

**Step i.c**. Assume $w^{i,b} = w^- \cdot w_i^{i,b} \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$.

For every $u^- \in \mathsf{USub}_A^{0,=}(w^-)$, let

$$t_{u^-} : \mathsf{USub}_B^{1,=}(w_i^{i,b}) \to \ell$$

be defined by

$$t_{u^-}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)]))$$

for any $u^+ \in \mathsf{USub}_C^{0,=}(w^+)$. By Step j.a for $j > i$, the coloring $t_{u^-}$ is well-defined. Finally, let

$$t : \mathsf{USub}_B^{1,=}(w_i^{i,b}) \to \ell^{|\mathsf{USub}_A^{0,=}(w^-)|}$$

be defined by

$$t(u_i) = \langle t_{u^-}(u_i) : u^- \in \mathsf{USub}_A^{0,=}(w^-)\rangle$$

By $(x \mapsto x^{x^x})$-sparsity of $X$, we have $\min Z_i > \ell^{|\mathsf{USub}_A^{0,=}(w^-)|}$, hence, by assumption, there exists some $\omega^n \cdot r(\min Z_i)$-large($\theta$) subset $W_i \subseteq Z_i$ (and therefore $\omega^n \cdot r(\min X_i)$-large($\theta$) as $r$ is non-decreasing) and some $W_i$-variable word $w_i^{i,c} \in \mathsf{USub}_B^{=}(w^{i,b})$ such that $\mathsf{USub}_B^{1,=}(w^{i,c})$ is $t$-monochromatic. By letting $w^{i,c} := w^- \cdot w_i^{i,c} \cdot w^+$, we have $w^{i,c} \in \mathsf{USub}_A^{=}(w^{i,b})$ and all the 1-variable words of type $i$ are $W_i$-invariant.

Finally, $v := w^{0,c}$ has the desired properties. □

We are now ready to prove the main theorem of our section:

**Theorem 4.1** ($\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{GR}^1} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\omega^{f_{\mathsf{GR}^1}(n)}$-large($\theta$) set is $\mathsf{GR}^1$-$\omega^n$-large($\theta$).*

*Proof.* Consider the sequences $(b_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ defined by mutual induction as follows:

$$b_0 = 0, c_n = f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(b_n)) + 2 \text{ and } b_{n+1} = f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(c_n)) + 1.$$

Then, let $f_{\mathsf{GR}^1}(n) = 2b_n + n_0$ with $n_0$ given by Lemma 3.7 is such that every $\omega^{f_{\mathsf{GR}^1}(n)}$-large($\theta$) set contains a $(x \mapsto x^{x^x})$-sparse $\omega^{b_n}$-large($\theta$) subset.

We prove by mutual induction on the definition of $(b_n)$ and $(c_n)$ that

(1) every $(x \mapsto x^{x^x})$-sparse $\omega^{b_n}$-large($\theta$) set $X$ is $\mathsf{GR}^1$-$\omega^n$-large($\theta$); and that

(2) every $(x \mapsto x^{x^x})$-sparse $\omega^{c_n}$-large($\theta$) set $X$ is $\mathsf{GR}^1$-$\omega^n \cdot \min X$-large($\theta$).

Every $\omega^{b_0}$-large($\theta$) set $X$ is $\mathsf{GR}^1$-$\omega^0$-large($\theta$), hence we have the base case.

Assume that, for some $n \in \mathbb{N}$, every $(x \mapsto x^{x^x})$-sparse and $\omega^{b_n}$-large($\theta$) set is $\mathsf{GR}^1$-$\omega^n$-large($\theta$). Let $X$ be $(x \mapsto x^{x^x})$-sparse and $\omega^{c_n}$-large($\theta$) and let $f : \mathsf{USub}_A^{1,=}(w) \to \ell$ be a coloring for some $X$-variable word $w$, some $\ell < \min X$ and some alphabet $A$ with $|A| < \min X$. As $c_n = f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(b_n)) + 2$, $\hat{X} := X\backslash\{\min X\}$ is $\omega^{f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(b_n))+1}$-large($\theta$). Then, by $(x \mapsto x^{x^x})$-sparsity,

$$\min \hat{X} > \min X \times \ell$$

and there exist subsets $X_0 < \cdots < X_{\min X \times \ell - 1}$ of $\hat{X}$ that are pairwise $\theta$-apart and $\omega^{f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(b_n))}$-large($\theta$). We also have $\min X_0 > \ell^{\min X \times \ell}$, hence, by Proposition 4.2 with $r$ being the constant function 1, there exists some $\omega^n$-large($\theta$) subsets $W_i \subseteq X_i$ for every $i < \min X \times \ell$ and some $W_0 \cup \cdots \cup W_{\min X \times \ell - 1}$-variable word $v \in \mathsf{USub}_A^{=}(w)$ such that the color of any 1-variable word $u \in \mathsf{USub}_A^{1,=}(v)$ only depends on its type (i.e., in which $W_i$ does the first occurrence of the variable belong) and let $c_i$ be that color. By the finite pigeonhole principle, there exists some color $c < \ell$ such that $c_i = c$ for at least $\min X$ indices $i$. Let $i_0, \ldots, i_{\min X - 1}$ be those indices, and consider the $\omega^n \cdot (\min X)$-large($\theta$) set $W := W_{i_0} \cup \cdots \cup W_{i_{\min X - 1}}$. Then, by construction, any 1-variable word

$u \in \mathsf{USub}^{1,=}_A(v)$ whose first occurrence of the variable belongs to $W$ will have the same color. Hence, $X$ is $\mathsf{GR}^1$-$\omega^n \cdot (\min X)$-large($\theta$).

Assume that, for some $n \in \mathbb{N}$, every $(x \mapsto x^{x^x})$-sparse and $\omega^{c_n}$-large($\theta$) set $Y$ is $\mathsf{GR}^1$-$\omega^n \cdot (\min Y)$-large($\theta$). Let $X$ be $(x \mapsto x^{x^x})$-sparse and $\omega^{b_{n+1}}$-large($\theta$) and let $f : \mathsf{USub}^{1,=}_A(w) \to \ell$ be a coloring for some $X$-variable word $w$, some $\ell < \min X$ and some alphabet $A$ with $|A| < \min X$. Let $X_0 < \cdots < X_{\min X - 1}$ be pairwise $\theta$-apart and $\omega^{f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(c_n))}$-large($\theta$) subsets of $X \setminus \{\min X\}$. By $(x \mapsto x^{x^x})$-sparsity of $X$, we have $\ell^{\min X} < \min X_0$, hence, by Proposition 4.2 with $r$ being the identity function, there exists $\omega^n \cdot (\min X_i)$-large($\theta$) subsets $W_i \subseteq X_i$ for every $i < \min X$ and some $X_0 \cup \cdots \cup X_{\min X - 1}$-variable word $v \in \mathsf{USub}^{=}_A(w)$ such that the color of any 1-variable word $u \in \mathsf{USub}^{1,=}_A(v)$ only depends on its type and let $c_i$ be the color of the 1-variable words of type $i$. By the finite pigeonhole principle, there exist two indices $i_0 < i_1$ such that $c_{i_0} = c_{i_1}$. Let $x \in W_{i_0}$ and consider the set $W := \{x\} \cup W_{i_1}$. As $x < \min X_{i_1}$, $W$ is $\omega^{n+1}$-large($\theta$). Then, by construction, any 1-variable word $u \in \mathsf{USub}^{1,=}_A(v)$ whose first occurrence of the variable belongs to $W$ will have the same color. Hence, $X$ is $\mathsf{GR}^1$-$\omega^{n+1}$-large($\theta$). □

# 5 Carlson-Simpson lemma

The goal of this section is to prove the following theorem, where a set $X$ is $\mathsf{CSL}^n_\ell$-$\alpha$-large($\theta$) if it is $\mathsf{CSL}^n_\ell(k)$-$\alpha$-large($\theta$) for every $k < \min X$.

**Theorem 5.1** ($\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{CSL}^1_2} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\omega^{f_{\mathsf{CSL}^1_2}(n)}$-large($\theta$) set is $\mathsf{CSL}^1_2$-$\omega^n$-large($\theta$).*

By iterating Theorem 5.1, one can obtain bounds for $\mathsf{CSL}^1_\ell$-$\omega^n$-largeness($\theta$) for every $\ell \in \mathbb{N}$. On the other hand, there is no bound for $\mathsf{CSL}^1$-$\omega^n$-largeness($\theta$) in general by Corollary 2.7. Indeed, the existence of such a bound would yield that $\mathsf{RCA}_0 + \forall \ell, k \mathsf{CSL}^1_\ell(k)$ is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$, but $\mathsf{RCA}_0 \vdash \forall \ell \mathsf{CSL}^1_\ell(2) \to \mathsf{B}\Sigma^0_3$, and $\mathsf{B}\Sigma^0_3$ is not even $\Pi_1$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

However, one can decompose the statement $\forall \ell, k \mathsf{CSL}^1_\ell(k)$ into its "variable word part" and its "Ramsey part", as follows: Consider the projection function $\pi$, which maps any $n$-variable word $w$ over an alphabet $A$ to the $(n+1)$-tuple $\{p_0 < \cdots < p_n\}$ where $p_i$ is the position of the first occurence of the $i$th variable kind, and $p_n = |w|$. Note that $\pi(w) = \mathsf{set}(w) \cup \{|w|\}$. The following statement is a consequence of the Carlson-Simpson lemma:

**Statement 5.2** (Level $\mathsf{CSL}$)**.** *Fix $n \geqslant 0$ and $\ell \geqslant 1$. For every finite alphabet $A$ and every coloring $f : A^{<\omega, n} \to \ell$, there is an infinite $\omega$-variable word $W$ such that for every $u_0, u_1 \in \mathsf{USub}^{n,\star}_A(W)$, if $\pi(u_0) = \pi(u_1)$, then $f(u_0) = f(u_1)$.*

We denote by $\mathsf{LCSL}^n_\ell(k)$ the Level Carlson-Simpson lemma for $\ell$-colorings of $n$-variable words over finite alphabets of size $k$.

**Lemma 5.3** (Liu and Patey [43])**.** $\mathsf{RCA}_0 \vdash \forall n, k, \ell(\mathsf{CSL}^n_\ell(k) \leftrightarrow \mathsf{LCSL}^n_\ell(k) \wedge \mathsf{RT}^{n+1}_\ell)$.

We shall actually see that the whole induction strength of the statement $\mathsf{CSL}^1$ is contained in its Ramsey part, namely, $\mathsf{RT}^2$, by proving that its level version is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0+\mathsf{B}\Sigma^0_2$. One could define $\mathsf{LCSL}^1$-$\omega^n$-largeness($\theta$) and prove the existence of a primitive recursive function $f_{\mathsf{LCSL}^1}:\mathbb{N}\to\mathbb{N}$ such that every $\omega^{f_{\mathsf{LCSL}^1}(n)}$-large($\theta$) set is $\mathsf{LCSL}^1$-$\omega^n$-large($\theta$). However, as we shall see in the next section, neither $\mathsf{CSL}^1_2$-$\omega^n$-largeness($\theta$), nor $\mathsf{LCSL}^1$-$\omega^n$-largeness($\theta$) provides a good invariant for proving the existence of such bounds. We now define a notion in between $\mathsf{CSL}^1$ and $\mathsf{LCSL}^1$ which makes only sense in the finitary version of the statement, for $\omega^n$-largeness($\theta$). Given an $n$-variable word $w$ and $i<n$, we write $\pi_i(w)$ for the position of its first occurrence of $x_i$, and $\pi_n(w)$ for $|w|$.

**Definition 5.4.** *Let $X\subseteq_{\mathtt{fin}}\mathbb{N}$ be a set, $w$ be an $X$-variable word over an alphabet $A$, and $f:\mathsf{USub}^{1,\star}_A(w)\to\ell$ be a coloring. We define $f$-$\alpha$-block-homogeneity($\theta$) for some color $\sigma$ inductively on $n$ as follows:*

- *$X$ is $f$-$\omega^0$-block-homogeneous($\theta$) for color $\diamondsuit$ if $X\neq\varnothing$.*
- *$X$ is $f$-$\omega^n\cdot k$-block-homogeneous($\theta$) for color $\sigma\cdot c$ with $c<\ell$ if $X=X_0\cup\cdots\cup X_{k-1}$ for some there are pairwise $\theta$-apart subsets $X_0<\cdots<X_{k-1}$ such that*
  - *Each $X_i$ is $\omega^n$-block-homogeneous($\theta$) for color $\sigma$;*
  - *For every $u\in\mathsf{USub}^{1,\star}_A(w)$ with $\pi_0(u)\in X_i$ and $\pi_1(u)\in X_j$, if $i<j$ then $f(u)=c$.*
- *$X$ is $f$-$\omega^{n+1}$-block-homogeneous($\theta$) for color $\sigma\cdot c$ with $c<\ell$ if $X\backslash\{\min X\}$ is $f$-$\omega^n\cdot(\min X)$-block-homogeneous($\theta$) for color $\sigma$ and if $f(u)=c$ for every $u\in\mathsf{USub}^{1,\star}_A(w)$ with $\pi_0(u)=\min X$.*

Note that if $X$ is $f$-$\omega^n$-block-homogeneous($\theta$) for some color $\sigma$, then it is a solution to the Level 1-dimensional Carlson-Simpson lemma. The main advantage of this definition is that there are only $\ell^{2n}$ many possible colors of block-homogeneity, while for $\mathsf{LCSL}^1$, there would be $\ell^{|X|}$ possible colors. This difference will be used in an essential way in the next section.

**Definition 5.5.** *A set $X$ is $\mathsf{BCSL}^1$-$\alpha$-large($\theta$) if for every $X$-variable word $w$ over an alphabet $A$ and every coloring $f:\mathsf{USub}^{1,\star}_A(w)\to\ell$ with $|A|,\ell<\min X$, there is a subset $Y\subseteq X$ and a $Y$-variable word $v\in\mathsf{USub}_A(w)$ such that $Y$ is $f^v$-$\alpha$-block-homogeneous($\theta$) for some color, where $f^v$ is the restriction of $f$ to $\mathsf{USub}^{1,\star}_A(v)$.*

The goal is therefore to prove the following theorem:

**Theorem 5.6** ($\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{BCSL}^1}:\mathbb{N}\to\mathbb{N}$ such that for every $n\in\mathbb{N}$, every $\omega^{f_{\mathsf{BCSL}^1}(n)}$-large($\theta$) set is $\mathsf{BCSL}^1$-$\omega^n$-large($\theta$).*

In particular, any $\omega^{f_{\mathsf{BCSL}^1}(n)}$-large($\theta$) set is $\mathsf{LCSL}^1$-$\omega^n$-large($\theta$). As for the Graham-Rothschild theorem, the core of the argument lies in the following proposition which combines blocks of solutions to a larger solution:

**Proposition 5.7.** *Let $n \in \mathbb{N}$, let $r : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function, and suppose that the following induction hypothesis holds:*

> *There exists some $a \in \mathbb{N}$ such that every $(x \mapsto x^{x^x})$-sparse $\omega^a$-large($\theta$) set $X$ is $\mathsf{BCSL}^1$-$\omega^n \cdot r(\min X)$-large($\theta$).*

*Let $k, \ell \in \mathbb{N}$, and $X_0 < \cdots < X_{k-1}$ be $\theta$-apart and $\omega^{f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(a))))}$-large($\theta$) sets such that $\ell^{(2n+1)k^2} < \min X_0$. Let $X = X_0 \cup \cdots \cup X_{k-1}$ and let $w = w_0 \dots w_{k-1}$ be an $X$-variable word over an alphabet $A$ of size smaller than $\min X_0$.*

*Then, for every coloring $f : \mathsf{USub}_A^{1,\star}(w) \to \ell$, there exist $\omega^n \cdot r(\min X_i)$-large($\theta$) subsets $W_i \subseteq X_i$ and some $W_0 \cup \cdots \cup W_{k-1}$-variable word $v \in \mathsf{USub}_A^{=}(w)$ such that:*

*(1) for every $i < k$, there exists a color $\sigma_i$, such that, for every $W_i$-variable word $u \in \mathsf{USub}_A^{=}(v)$, $W_i$ is $f^u$-$\omega^n \cdot r(\min X_i)$-block-homogeneous($\theta$) for $\sigma_i$, where $f^u$ is the restriction of $f$ to $\mathsf{USub}_A^{1,\star}(u)$.*

*(2) the color of every 1-variable word $u \in \mathsf{USub}_A^{1,\star}(v)$ only depends on the pair $(i_0, i_1)$ such that $\pi_0(u) \in W_{i_0}$ and $\pi_1(u) \in W_{i_1}$ when $i_0 \neq i_1$.*

*Proof.* We say that some 1-variable word $u \in \mathsf{USub}_A^{1,\star}(w)$ is of *type* $(i_0, i_1)$ (for some $i_0 \leqslant i_1$) if $\pi_0(u) \in X_{i_0}$ and $\pi_1(u) \in X_{i_1}$. Given a set $Y \subseteq X$, the 1-variable words of type $(i_0, i_1)$ are *$Y$-invariant* if their color do not depend on the valuations of the variables $\{x_n : n \in Y\}$.

**Proof outline:** The construction will follow $5k$ steps, in the following order:

Step $(k-1)$.a, Step $(k-1)$.b, Step $(k-1)$.c, Step $(k-1)$.d, Step $(k-1)$.e, …, Step 0.a, Step 0.b, Step 0.c, Step 0.d, Step 0.e.

These steps will ensure that the 1-variable words $u \in \mathsf{USub}_A^{1,\star}(v)$ of any type $(i_0, i_1)$ are $W_i$-invariant for every $i < k$, except in the case where $i_0 = i_1 = i$, where we will find some color $\sigma_i$ and ensure the block-homogeneity of $W_i$ for that color.

Let $i < k$, and assume $W_j \subseteq X_j$ have been defined for $j \in \{i+1, \dots, k-1\}$.

- In Step i.a, we are going to find some subset $Y_{i,a} \subseteq X_i$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Y_{i,a} \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,a} \in \mathsf{USub}_A^{=}(w^{i+1,e})$ (or $w^{i,a} \in \mathsf{USub}_A^{=}(w)$ in the case where $i = k-1$) such that the 1-variable words of type $(i_0, i)$ with $i_0 < i$ are $Y_{i,a}$-invariant.

- In Step i.b, we are going to find some subset $Y_{i,b} \subseteq Y_{i,a}$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Y_{i,b} \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,b} \in \mathsf{USub}_A^{=}(w^{i,a})$ such that the 1-variable words of type $(i_0, i_1)$ with $i_0, i_1 > i$ are $Y_{i,b}$-invariant.

- In Step i.c, we are going to find some subset $Y_{i,c} \subseteq Y_{i,b}$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Y_{i,c} \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,c} \in \mathsf{USub}_A^{=}(w^{i,b})$ such that the 1-variable words of type $(i_0, i_1)$ with $i_0 < i < i_1$ are $Y_{i,c}$-invariant.

- In Step i.d, we are going to find some subset $Y_{i,d} \subseteq Y_{i,c}$ and some $X_0 \cup \cdots \cup X_{i-1} \cup Y_{i,d} \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,d} \in \mathsf{USub}_A^=(w^{i,c})$ such that the 1-variable words of type $(i, i_1)$ with $i < i_1$ are $Y_{i,d}$-invariant.

- In Step i.e, we are going to find some subset $W_i \subseteq Y_{i,d}$ and some $X_0 \cup \cdots \cup X_{i-1} \cup W_i \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,e} \in \mathsf{USub}_A^=(w^{i,d})$ such that, for every $W_i$-variable word $u \in \mathsf{USub}_A^=(w^{i,e})$, $W_i$ is $f^u$-$\mathfrak{w}^n \cdot r(\min X_i)$-block-homogeneous($\theta$) for some $\sigma_i \in \ell^{2n+1}$, where $f^u$ is the restriction of $f$ to $\mathsf{USub}_A^{1,\star}(u)$.

Notice that we already have that the 1-variable words of type $(i_0, i_1)$ are $X_i$ invariant when $i_0, i_1 < i$ as the 1-variable words of type $(i_0, i_1)$ are of length $< \min X_i$.

**Formal proof:** Consider the extended alphabets

$$B := A \sqcup \{x_p : p \in X_0 \cup \cdots \cup X_{i-1}\}$$
$$C := A \sqcup \{x_p : p \in X_0 \cup \cdots \cup X_{i-1} \cup X_i\}$$

At each step of the construction, we will be defining a coloring as follows:

**Step i.a**. Assume $w^{i+1,e} = w^- \cdot w_i \cdot w^+$ has already been defined, where $w^-$ is an $X_0 \cup \cdots \cup X_{i-1}$-variable word over $A$, $w_i$ is an $X_i$-variable word over $B$ and $w^+$ is an $W_{i+1} \cup \cdots \cup W_{k-1}$-variable word over $C$. By convention, $w^{i+1,e} = w$ if $i = k - 1$. By construction, we have $w^- = w_0 \cdots w_{i-1}$.

For every $u^- \in \mathsf{USub}_A^{1,=}(w^-)$, let

$$g_{i.a}^{u^-} : \mathsf{USub}_B^{0,\star}(w_i) \to \ell$$

be defined by

$$g_{i.a}^{u^-}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)])$$

Then, let

$$g_{i.a} : \mathsf{USub}_B^{0,\star}(w_i) \to \ell^{|\mathsf{USub}_A^{1,=}(w^-)|}$$

be defined by

$$g_{i.a}(u_i) = \langle g_{i.a}^{u^-}(u_i) : u^- \in \mathsf{USub}_A^{1,=}(w^-) \rangle$$

**Step i.b**. Assume $w^{i,a} = w^- \cdot w_i^{i,a} \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$.

For every $u^- \in \mathsf{USub}_A^{0,=}(w^-)$ and $i_0, i_1 \in \mathbb{N}$ with $i < i_0 < i_1 < k$, let

$$g_{i.b}^{u^-,i_0,i_1} : \mathsf{USub}_B^{0,=}(w_i^{i,a}) \to \ell$$

be defined by

$$g_{i.b}^{u^-,i_0,i_1}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)])$$

for any $u^+ \in \mathsf{USub}^{1,\star}_C(w^+)$ of type $(i_0, i_1)$. By Steps $j$.a, $j$.b, $j$.c and $j$.d for $j > i$, the coloring $g^{u^-,i_0,i_1}_{i.b}$ is well-defined.

For every $u^- \in \mathsf{USub}^{0,=}_A(w^-)$ and $i_0 \in \mathbb{N}$ with $i < i_0 < k$, let

$$g^{u^-,i_0}_{i.b} : \mathsf{USub}^{0,=}_B(w^{i,a}_i) \to \ell^{2n+1}$$

be defined by $g^{u^-,i_0}_{i.b}(u_i) = \sigma$ where $f^{u^-,u_i}$ is the restriction of $f$ to $\mathsf{USub}^{1,\star}_A(u^- \cdot u_i[x_p \mapsto u^-(p)])$ and $\sigma$ is the color of $f^{u^-,u_i}$-$\mathfrak{w}^n \cdot r(\min X_{i_0})$-block-homogeneity$(\theta)$ for $W_{i_0}$. By Steps $j$.e for $j > i$, the coloring $g^{u^-,i_0}_{i.b}$ is well-defined.

Finally, let

$$g_{i.b} : \mathsf{USub}^{0,=}_B(w^{i,a}_i) \to \ell^{|\mathsf{USub}^{0,=}_A(w^-)| \times \left((2n+1)(k-i-1) + \frac{(k-i-1)(k-i-2)}{2}\right)}$$

be defined by

$$g_{i.b}(u_i) = \left\langle \begin{array}{ll} g^{u^-,i_0,i_1}_{i.b}(u_i) & : i_0 < i_1 \in \{i+1, \dots, k-1\} \text{ and } u^- \in \mathsf{USub}^{0,=}_A(w^-) \\ g^{u^-,i_0}_{i.b}(u_i) & : i_0 \in \{i+1, \dots, k-1\} \text{ and } u^- \in \mathsf{USub}^{0,=}_A(w^-) \end{array} \right\rangle$$

**Step i.c**. Assume $w^{i,b} = w^- \cdot w^{i,b}_i \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$.

For every $u^- \in \mathsf{USub}^{1,=}_A(w^-)$ and $i_1 \in \{i+1, \dots, k-1\}$, let

$$g^{u^-,i_1}_{i.c} : \mathsf{USub}^{0,=}_B(w^{i,b}_i) \to \ell$$

be defined by

$$g^{u^-,i_1}_{i.c}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)])$$

for any $u^+ \in \mathsf{USub}^{0,\star}_C(w^+)$ with $|u^+| \in W_{i_1}$. Recall that we consider $u^+$ as a located variable word, in which case $|u^+|$ denotes the least upper bound of its domain. By Steps j.a and j.c for $j > i$, the coloring $g^{u^-,i_1}_{i.c}$ is well-defined.

Finally, let

$$g_{i.c} : \mathsf{USub}^{0,=}_B(w^{i,b}_i) \to \ell^{|\mathsf{USub}^{1,=}_A(w^-)| \times (k-i-1)}$$

be defined by

$$g_{i.c}(u_i) = \langle g^{u^-,i_1}_{i.c}(u_i) : i_1 \in \{i+1, \dots, k-1\} \text{ and } u^- \in \mathsf{USub}^{1,=}_A(w^-) \rangle$$

**Step i.d**. Assume $w^{i,c} = w^- \cdot w^{i,c}_i \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$.

For every $u^- \in \mathsf{USub}^{0,=}_A(w^-)$ and $i_1 \in \{i+1, \dots, k-1\}$, let

$$g^{u^-,i_1}_{i.d} : \mathsf{USub}^{1,=}_B(w^{i,c}_i) \to \ell$$

be defined by

$$g_{i.d}^{u^-,i_1}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)] \cdot u^+[x_p \mapsto (u^- \cdot u_i)(p)])$$

for any $u^+ \in \mathsf{USub}_C^{0,\star}$ with $|u^+| \in W_{i_1}$. By Steps j.a and j.c for $j > i$, the coloring $g_{i.d}^{u^-,i_1}$ is well-defined.

Finally, let

$$g_{i.d} : \mathsf{USub}_B^{1,=}(w_i^{i,c}) \to \ell^{|\mathsf{USub}_A^{0,=}(w^-)| \times (k-i-1)}$$

be defined by

$$g_{i.d}(u_i) = \langle g_{i.d}^{u^-,i_1}(u_i) : i_1 \in \{i+1, \ldots, k-1\} \text{ and } u^- \in \mathsf{USub}_A^{0,=}(w^-) \rangle$$

**Step i.e**. Assume $w^{i,d} = w^- \cdot w_i^{i,d} \cdot w^+$ have already been defined. By construction, we have $w^- = w_0 \cdots w_{i-1}$. For every $u^- \in \mathsf{USub}_A^{0,=}(w^-)$, let

$$g_{i.e}^{u^-} : \mathsf{USub}_B^{1,\star}(w_i^{i,d}) \to \ell$$

be defined by

$$g_{i.e}^{u^-}(u_i) = f(u^- \cdot u_i[x_p \mapsto u^-(p)])$$

Then, let

$$g_{i.e} : \mathsf{USub}_B^{1,\star}(w_i^{i,d}) \to \ell^{|\mathsf{USub}_A^{0,=}(w^-)|}$$

be defined by

$$g_{i.e}(u_i) = \langle g_{i.e}^{u^-}(u_i) : u^- \in \mathsf{USub}_A^{0,=}(w^-) \rangle$$

In Steps i.a through i.e, we have considered successively: an instance of $\mathsf{CSL}^0$, an instance of $\mathsf{GR}^0$, an instance of $\mathsf{GR}^0$, an instance of $\mathsf{GR}^1$ and an instance of $\mathsf{CSL}^1$. As in the proof of Proposition 4.2, we can use the bounds $f_{\mathsf{GR}^0}, f_{\mathsf{GR}^1}$ and $f_{\mathsf{CSL}^0}$ and the definition of $a$, to obtain that it is sufficient for $X_i$ to be $(x \mapsto x^{x^x})$-sparse and $\omega^{f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(a))))}$-large$(\theta)$ in order to find a sequence of subsets $W_i \subseteq Y_{i,d} \subseteq Y_{i,c} \subseteq Y_{i,b} \subseteq Y_{i,a} \subseteq X_i$, some $X_0 \cup \ldots X_{i-1} \cup W_i \cup W_{i+1} \cup \cdots \cup W_{k-1}$-variable word $w^{i,e} \in \mathsf{USub}_A^{=}(w)$, and some color $\sigma_i$, so that:

- for every $W_i$-variable word $u \in \mathsf{USub}_A^{=}(w^{i,e})$, $W_i$ is $f^u$-$\omega^n \cdot r(\min X_i)$-block-homogeneous$(\theta)$ for $\sigma_i$, where $f^u$ is the restriction of $f$ to $\mathsf{USub}_A^{1,\star}(u)$.
- the 1-variable word $u \in \mathsf{USub}_A^{1,\star}(w^{i,e})$ of type $(i_0, i_1)$ are $W_i$-invariant when $(i_0, i_1) \neq (i, i)$.

The assumption that $\ell^{(2n+1)k^2} < \min X_0$ is used in Step 0.b to ensure that

$$\ell^{\left((2n+1)(k-1)+\frac{(k-1)(k-2)}{2}\right)} \leqslant \ell^{(2n+1)k^2} < \min X_0$$

so that the bound $f_{\mathsf{GR}^0}$ can be used. Thus, once all the steps have been done, the set $W_0 \cup \cdots \cup W_{k-1}$ and the resulting $W_0 \cup \cdots \cup W_{k-1}$-variable word $w^{0,e}$ satisfy the desired properties. □

In what follows, for $k, \ell \in \mathbb{N}$, let $R_\ell(k) \in \mathbb{N}$ denote the smallest bound $b \in \mathbb{N}$ such that every coloring $f : [b]^2 \to \ell$ admits a subset $H \subseteq b$ of cardinality $k$ satisfying $f([H]^2) = 1$. The standard upper bound for the finite Ramsey's theorem for pairs and two colors, obtained by Erdős and Szekeres in [24], states that

$$R_2(d) \leqslant \binom{2d-1}{d-1} \leqslant 4^d.$$

This upper bound generalizes to $R_k(d) \leqslant k^{kd}$ (see Graham, Rothschild and Spencer [26, Section 1.1]).

**Theorem 5.6** ($\mathsf{RCA}_0$)**.** *There exists a primitive recursive function $f_{\mathsf{BCSL}^1} : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, every $\mathfrak{w}^{f_{\mathsf{BCSL}^1}(n)}$-large($\theta$) set is $\mathsf{BCSL}^1$-$\mathfrak{w}^n$-large($\theta$).*

*Proof.* Consider the sequences $(b_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ defined by mutual induction as follows:

$$\begin{cases} b_0 = 0, \\ c_n = f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(b_n)))) + 2, \\ b_{n+1} = f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(c_n)))) + 2 \end{cases}$$

Then, let $f_{\mathsf{BCSL}^1}(n) = 2b_n + n_0$ with $n_0$ given by Lemma 3.7 is such that every $\mathfrak{w}^{f_{\mathsf{BCSL}^1}(n)}$-large($\theta$) set contains a $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^{b_n}$-large($\theta$) subset.

We prove by mutual induction on the definition of $(b_n)$ and $(c_n)$ that

(1) every $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^{b_n}$-large($\theta$) set $X$ is $\mathsf{BCSL}^1$-$\mathfrak{w}^n$-large($\theta$); and that

(2) every $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^{c_n}$-large($\theta$) set $X$ is $\mathsf{BCSL}^1$-$\mathfrak{w}^n \cdot \min X$-large($\theta$).

Every $\mathfrak{w}^{b_0}$-large($\theta$) set $X$ is $\mathsf{BCSL}^1$-$\mathfrak{w}^0$-large($\theta$), hence we have the base case.

Assume that every $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^{b_n}$-large($\theta$) set is $\mathsf{BCSL}^1$-$\mathfrak{w}^n$-large($\theta$) and let $X$ be $\mathfrak{w}^{c_n}$-large($\theta$) and $(x \mapsto x^{x^x})$-sparse. Let $w$ be an $X$-variable word over an alphabet $A$ with $|A| < \min X$ and let $f : \mathsf{USub}^{1,\star}_A(w) \to \ell$ be a coloring for some $\ell < \min X$.

Let $a = \ell^{2n} \times R_\ell(\min X)$, we therefore have $a \leqslant \ell^{2n} \times \ell^{\ell \times \min X}$. As $X$ is $\mathfrak{w}^{c_n}$-large($\theta$), there exists some $\mathfrak{w}^{f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(b_n))))+1}$-large($\theta$) subsets $X' < X''$ of $X$. Then, by $(x \mapsto x^{x^x})$-sparsity of $X$ and the fact that $f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(b_n)))) + 1 > n$, we get that $\min X'' \geqslant a$, hence that there exist $X_0 < \cdots < X_{a-1}$ pairwise $\theta$-apart $\mathfrak{w}^{f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(b_n))))}$-large($\theta$) subsets of $X''$.

We also have $\min X'' > \ell^{(2n+1)a^2}$, hence, by Proposition 5.7, there exist $\mathfrak{w}^n$-large($\theta$) subsets $Y_i \subseteq X_i$ and some $Y_0 \cup \cdots \cup Y_{a-1}$-variable word $v \in \mathsf{USub}^=_A(w)$ such that:

(1) for every $i < a$, there exists a color $\sigma_i$, such that, for every $W_i$-variable word $u \in \mathsf{USub}^=_A(v)$, $W_i$ is $f^u$-$\mathfrak{w}^n$-block-homogeneous($\theta$) for $\sigma_i$, where $f^u$ is the restriction of $f$ to $\mathsf{USub}^{1,\star}_A(u)$.

(2) the color of every 1-variable word $u \in \mathsf{USub}_A^{1,\star}(v)$ only depends on the pair $(i_0, i_1)$ such that $\pi_0(u) \in W_{i_0}$ and $\pi_1(u) \in W_{i_1}$ when $i_0 \neq i_1$.

By the finite pigeonhole principle, there exists some $\sigma \in \ell^{2n}$ for which $\sigma_i = \sigma$ for at least $R_\ell(\min X)$ values of $i < a$. Let $V_0 < \cdots < V_{R_\ell(\min X)-1}$ be the corresponding blocks and pick some $V_0 \cup \cdots \cup V_{R_\ell(\min X)-1}$-variable word $v' \in \mathsf{USub}_A^{=}(v)$. By our application of Proposition 5.7, the color of any $u \in \mathsf{USub}_A^{1,\star}(v')$ only depends on the pair $(i_0, i_1)$ such that $\pi_0(u) \in V_{i_0}$ and $\pi_0(u) \in V_{i_1}$ when $i_0 \neq i_1$. Hence, by definition of $R_\ell(\min X)$, there exists $\min X$ blocks $K_0 < \cdots < K_{\min X-1}$ among $V_0, \ldots, V_{R_\ell(\min X)-1}$ and some $c < \ell$ such that $f(u) = c$ for any $u \in \mathsf{USub}_A^{1,\star}(v')$ for which $\pi_0(u) \in K_i$ and $\pi_0(u) \in K_j$ for some $i < j$. Thus, the set $K_0 \cup \cdots \cup K_{\min X-1}$ is $f$-$\mathfrak{w}^n \cdot \min X$-block-homogeneous($\theta$) for color $\sigma \cdot c$.

Assume that every $(x \mapsto x^{x^x})$-sparse $\mathfrak{w}^{c_n}$-large($\theta$) set is $\mathsf{BCSL}^1$-$\mathfrak{w}^n \cdot \min X$-large($\theta$) and let $X$ be $\mathfrak{w}^{b_{n+1}}$-large($\theta$) and $(x \mapsto x^{x^x})$-sparse. Let $w$ be an $X$-variable word over an alphabet $A$ with $|A| < \min X$ and let $f : \mathsf{USub}_A^{1,\star}(w) \to \ell$ be a coloring for some $\ell < \min X$.

As $X$ is $\mathfrak{w}^{b_{n+1}}$-large($\theta$) and $(x \mapsto x^{x^x})$-sparse, there exist $X_0 < \cdots < X_{\ell^{2n+1}}$ pairwise $\theta$-apart $\mathfrak{w}^{f_{\mathsf{CSL}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^0}(f_{\mathsf{GR}^1}(b_n))))}$-large($\theta$) subsets of $X$ by a reasoning similar to the previous case. Moreover, $\ell^{(2n+1)\ell^{4n+2}} < \min X_0$, so by Proposition 5.7, there exist $\mathfrak{w}^n \cdot (\min X_i)$-large($\theta$) subsets $Y_i \subseteq X_i$ and some $Y_0 \cup \cdots \cup Y_{\ell^{2n+1}}$-variable word $v \in \mathsf{USub}_A^{=}(w)$ such that:

(1) for every $i \leqslant \ell^{2n+1}$, there exists a color $\sigma_i$, such that, for every $W_i$-variable word $u \in \mathsf{USub}_A^{=}(v)$, $W_i$ is $f^u$-$\mathfrak{w}^n \cdot \min X_i$-block-homogeneous($\theta$) for $\sigma_i$, where $f^u$ is the restriction of $f$ to $\mathsf{USub}_A^{1,\star}(u)$.

(2) the color of every 1-variable word $u \in \mathsf{USub}_A^{1,\star}(v)$ only depends on the pair $(i_0, i_1)$ such that $\pi_0(u) \in W_{i_0}$ and $\pi_1(u) \in W_{i_1}$ when $i_0 \neq i_1$.

By the finite pigeonhole principle, there exists some indices $i_0 < i_1 \leqslant \ell^{2n+1}$ such that $\sigma_{i_0} = \sigma_{i_1}$. Let $x \in Z_{i_0}$, then, as $x \leqslant \max Z_{i_0} < \min Y_{i_1}$, the set $\{x\} \cup Z_{i_1}$ is $f$-$\mathfrak{w}^{n+1}$-block-homogeneous($\theta$) for color $\sigma_{i_0} \cdot c$ with $c = f(u)$ for any $u \in \mathsf{USub}_A^{1,\star}(v)$ such that $\pi_0(u) \in Y_{i_0}$ and $\pi_0(u) \in Y_{i_1}$. □

# 6 Conservation theorems

We now apply the previous closure theorems under $\alpha$-largeness($\theta$) to deduce $\forall\Pi^0_4$-conservation of $\mathsf{CSL}^1_2$ over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$. Such partial conservation results are proven as follows: given a $\forall\Pi^0_4$-sentence $\varphi$ and a countable model $\mathcal{M} \models \mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \neg\varphi$, we will construct another model $\mathcal{N} \models \mathsf{RCA}_0 + \mathsf{CSL}^1_2 + \neg\varphi$. The first-order part of our model $\mathcal{N}$ will actually be an initial segment of the first-order part of $\mathcal{M}$.

**Definition 6.1** (Cut)**.** *A* cut *in a model $M$ of first-order arithmetic is a nonempty subset $I \subseteq M$ which is closed under successor (if $a \in I$, then $a+1 \in I$),*

*and is an initial segment of $M$ (if $a \in I$ and $b \leqslant a$, then $b \in I$). A cut $I \subseteq M$ is* proper *if $I \neq M$.*

The model $\mathcal{N}$ that we shall construct will be fully characterized the cut forming its first-order part. Indeed, its second-order part of $\mathcal{N}$ will the collection of all its coded sets:

**Definition 6.2** (Coded sets)**.** *For $I \subseteq M$ a proper cut, a set $X \subseteq I$ is said to be $M$-*coded *if $X = I \cap \hat{X}$ for some $M$-finite set $\hat{X}$. Let $\mathsf{Cod}(M/I)$ be the set of all $M$-coded sets.*

There exists a neat characterization of the cuts $I \subseteq_e M$ such that $\mathsf{Cod}(M, /I)$ forms a model of $\mathsf{WKL}_0$. These cuts are called *semi-regular*, a notion related to regularity in set theory:

**Definition 6.3** (Semi-regular cut [36])**.** *A proper cut $I \subseteq M$ is said to be* semi-regular *if for every $M$-finite set $E \subseteq M$ such that $|E| \in I$, $E \cap I$ is bounded in $I$ (i.e., there exists $a \in I$ such that $E \cap I < a$).*

**Proposition 6.4** (Scott [49] and Kirby and Paris [36, Proposition 1]. See also Simpson [51, Lemma IX.3.11])**.** *Let $M \models \mathsf{PRA}$ and let $I \subseteq M$ be a proper cut. Then, $I$ is semi-regular if and only if $(I, \mathsf{Cod}(M/I)) \models \mathsf{WKL}_0$.*

We now turn to the actual proof of $\forall\Pi^0_4$-conservation, from which all our results will follow:

**Theorem 6.5.** $\mathsf{WKL}_0 + \mathsf{RT}^2_2 + \mathsf{LCSL}^1$ *is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.*

*Proof.* Let $\phi(X\restriction_z, t, x, y, z)$ be a $\Delta^0_0$-formula and assume that $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \not\vdash \forall X \forall t \exists x \forall y \exists z \phi(X\restriction_z, t, x, y, z)$. In what follows, we enrich the language with two constant symbols $c$ and $C$, of first and second order, respectively.

By completeness and the Löwenheim-Skolem theorem, there exists some countable model

$$\mathcal{M} = (M, S, c^{\mathcal{M}}, C^{\mathcal{M}}) \models \mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \forall x \exists y \forall z \neg\phi(C\restriction_z, c, x, y, z)$$

Let $\theta(x, y, z)$ be the $\Delta^{0,C,c}_0$ formula $\neg\phi(C\restriction_z, c, x, y, z)$.

By Proposition 3.3, for all standard $n$, there exists some $\omega^n$-large($\theta$) subset $X$ of $M$ such that $c \leqslant \min X$. So, by compactness, we can assume that $\mathcal{M}$ contains an $\omega^d$-large($\theta$) set $X$ with $c \leqslant \min X$ for some non-standard integer $d$.

By a standard argument, using Theorem 3.4 for (4) and Theorem 5.6 for (5), we can consider a decreasing sequence $X = X_0 \supseteq X_1 \supseteq \dots$ such that for every $i < \omega$:

(1) $X_i$ is $\omega^{d_i}$-large($\theta$) for some non-standard $d_i \in M$;

(2) $\min X_{i+1} > \min X_i$;

(3) For every $M$-finite set $E$ such that $|E| < \min X_i$, there exists some $j > i$ such that $[\min X_j, \max X_j] \cap E = \varnothing$;

(4) For every coloring $f : [M]^2 \to 2$ in $\mathcal{M}$, there exists some $j > i$ such that $[X_j]^2$ is $f$-monochromatic;

(5) For every coloring $f : A^{<M,1} \to \ell$ for some $\ell < \min X_i$ and some $M$-finite alphabet $A$ with $|A| < \min X_i$ with $f \in \mathcal{M}$, there exists some $j > i$ and an $X_j$-variable word $W \in \mathcal{M}$ such that for every $u_0, u_1 \in \mathsf{USub}^{1,\star}_A(W)$ for which $\pi(u_0) = \pi(u_1)$, $f(u_0) = f(u_1)$;

(6) There exists some $j > i$ such that

$$\forall x < \min X_j \exists y < \min X_{j+1} \forall z < \max X_{j+1} \theta(x, y, z)$$

Then, consider $I = \sup\{\min X_i | i \in \omega\}$. $I$ is a semi-regular cut of $M$ by (3), therefore $(I, \mathsf{Cod}(M/I)) \models \mathsf{WKL}_0$. The constraint $\min X \geqslant c$ imposed on largeness($\theta$) ensures that $c$ is in $I$. The condition $\min X_{i+1} > \min X_i$ implies that every $X_i \cap I$ is infinite in $I$.

Let $f : [I]^2 \to 2$ be a coloring in $(I, \mathsf{Cod}(M/I))$. There exists some coloring $g : [M]^2 \to 2$ in $\mathcal{M}$ such that $f = g \cap [I]^2$. By (4), there is some $j > i$ such that $[X_j]^2$ is $g$-monochromatic. Since $X_j \cap I$ is infinite in $I$, $(I, \mathsf{Cod}(M/I)) \models \mathsf{RT}^2_2$.

Let $\ell \in I$ and $A$ a finite alphabet with $|A| \in I$, let $f : A^{<I,1} \to \ell$ be a coloring in $(I, \mathsf{Cod}(M/I))$. There exists some coloring $g : A^{<M,1} \to \ell$ in $\mathcal{M}$ such that $f = g \restriction I$ and let $i$ such that $\ell, |A| < \min X_i$. By (5), there is some $j > i$ and an $X_j$-variable word $W \in \mathcal{M}$ such that for every $u_0, u_1 \in \mathsf{USub}^{1,\star}_A(W)$ for which $\pi(u_0) = \pi(u_1)$, $g(u_0) = g(u_1)$; then $W \restriction I$ is an $X_j \cap I$-variable word such that for every $u_0, u_1 \in \mathsf{USub}^{1,\star}_A(W \restriction I)$ for which $\pi(u_0) = \pi(u_1)$, $f(u_0) = f(u_1)$. Since $X_j \cap I$ is infinite in $I$, $W \restriction I$ is an $\omega$-variable word in $(I, \mathsf{Cod}(M/I))$, so $(I, \mathsf{Cod}(M/I)) \models \mathsf{LCSL}^1$.

Finally, $(I, \mathsf{Cod}(M/I), c^{\mathcal{M}}, C^{\mathcal{M}} \cap I) \models \forall x \exists y \forall z \theta(x, y, z)$ as for every $k \in I$, there exists an index $i \in \omega$ such that $k < \min X_i$ and therefore by (6), there is some $j > i$ such that $\forall x < k \exists y < \min X_{j+1} \forall z < \max X_{j+1}\ \theta(x, y, z)$, so $(I, \mathsf{Cod}(M/I), c^{\mathcal{M}}, C^{\mathcal{M}} \cap I) \models \forall x < k \exists y \forall z\ \theta(x, y, z)$ (since $\max X_{j+1} > I$). Hence

$$(I, \mathsf{Cod}(M/I)) \models \mathsf{WKL}_0 + \mathsf{RT}^2_2 + \mathsf{LCSL}^1 + \neg \forall X \forall t \exists x \forall y \exists z \phi(X\restriction_z, t, x, y, z)$$

□

This implies in particular our main theorem:

**Main Theorem 1.1.** $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ *is* $\forall\Pi^0_4$*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

*Proof.* Immediate by Theorem 6.5 and Lemma 5.3. □

**Corollary 6.6.** $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ *is* $\forall\Pi^0_3$*-conservative over* $\mathsf{RCA}_0$ *and* $\Pi^0_2$*-conservative over* $\mathsf{PRA}$.

*Proof.* By a parameterized version of the Parsons, Paris and Friedman conservation theorem (see [27, 34]), $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ is $\forall\Pi^0_3$-conservative over $\mathsf{RCA}_0$. By Friedman [25] (see [51]), $\mathsf{RCA}_0$ is $\Pi^0_2$-conservative over $\mathsf{PRA}$. □

**Corollary 6.7** (Liu and Patey [43])**.** $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ *does not imply* $\mathsf{ACA}_0$.

*Proof.* By Simpson [51, Corollary VIII.1.7], $\mathsf{ACA}_0$ proves the consistency of $\mathsf{RCA}_0$, which is a $\Pi^0_1$ statement, while $\mathsf{RCA}_0$ does not by the second Gödel incompleteness theorem. Thus, by Corollary 6.6, $\mathsf{WKL}_0 + \mathsf{CSL}^1_2$ does not imply the consistency of $\mathsf{RCA}_0$. □

## 6.1 Universal triangle-free graph

We now present an application in Structural Ramsey theory. Let $\mathbb{H}_3$ be the universal triangle-free Henson graph, which is the Fraissé limit of the class of all finite triangle-free graphs. Given two structures $\mathbf{F}, \mathbf{G}$, we write $\binom{\mathbf{G}}{\mathbf{F}}$ for the set of all embeddings from $\mathbf{F}$ to $\mathbf{G}$. We write $\mathbf{F} \cong \mathbf{G}$ iff $\mathbf{F}$ is isomorphic to $\mathbf{G}$. Given integers $k, \ell \in \omega$, structures $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ we write $\mathbf{C} \longrightarrow (\mathbf{A})^{\mathbf{B}}_{k,\ell}$ for the following statement:

**Definition 6.8.** $\mathbf{C} \longrightarrow (\mathbf{A})^{\mathbf{B}}_{k,\ell}$*: For any coloring* $\chi : \binom{\mathbf{A}}{\mathbf{B}} \to k$*, there is an embedding* $g \in \binom{\mathbf{A}}{\mathbf{C}}$ *such that* $\binom{g(\mathbf{C})}{\mathbf{B}}$ *uses at most* $\ell$ *many colors.*

The *big Ramsey degree* of $\mathbf{B}$ in $\mathbf{A}$ is the smallest $L \in \omega \cup \{\omega\}$ such that $\mathbf{A} \longrightarrow (\mathbf{A})^{\mathbf{B}}_{k,L}$ for every $k \in \omega$, and a infinite structure $\mathbf{A}$ is said to have *finite big Ramsey degrees* if for all finite substructure of $\mathbf{B}$, $\mathbf{B}$ has a finite big Ramsey degree in $\mathbf{A}$.

Dobrinen [18] proved that the triangle-free Henson graph admits finite big Ramsey degrees using a notion of forcing with coding trees. More recently, Hubička [33] gave an alternative proof of Dobrinen's theorem using $\mathsf{CSL}$ as its combinatorial core, but with no optimal big Ramsey degree bounds. Anglès d'Auriac, Liu, Mignoty and Patey [1] noticed that his proof could be formalized in $\mathsf{RCA}_0$.

**Definition 6.9.** *Let* $\mathbb{G} = (\{0\}^{<\omega,1}, E)$ *be the graph where* $E$ *is the symmetric binary relation defined as follows: for* $v, w \in \{0\}^{<\omega,1}$*,* $vEw$ *if* $|v| \neq |w|$ *and, assuming* $|v| < |w|$*, the following holds:*

*(1) Passing number property:* $w(|v|) = x_0$*;*

*(2) Triangle-freeness condition: there is no* $i < |v|$ *with* $v(i) = w(i) = x_0$*.*

Hubička [33] proved that for every unordered $\omega$-variable word $W$ over $\{0\}$, $(\mathsf{USub}^{1,\star}(W), E)$ is computably isomorphic to the triangle-free Henson graph $\mathbb{H}_3$. Moreover, these isomorphisms hold provably over $\mathsf{RCA}_0$, so the indivisibility of the triangle-free Henson graph follows from Carlson-Simpson lemma for 1-variable words over $\mathsf{RCA}_0$. In what follows, we identify the graph $\mathbb{H}_3$ with its set of vertices:

**Theorem 6.10** ($\mathsf{RCA}_0 + \mathsf{CSL}^1(1)$)**.** *For any integer $k > 0$ and any finite coloring $\chi : \mathbb{H}_3 \to k$, there exists $f \in \binom{\mathbb{H}_3}{\mathbb{H}_3}$ such that $\chi$ uses exactly one color on $f(\mathbb{H}_3)$.*

By Main Theorem 1.1 and Corollary 6.7, it follows that the previous theorem is $\forall\Pi^0_4$-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ and does not imply $\mathsf{ACA}_0$. The latter fact was recently proven by Liu and Patey [43].

## 6.2 Tree theorem for pairs

Chubb, Hirst, and McNicholl [15] introduced a partition theorem for comparable nodes over perfect binary trees, which was extensively studied in reverse mathematics. Given a set $S \subseteq 2^{<\mathbb{N}}$ and $n \in \mathbb{N}$, we write $[S]^n$ for the set of all pairwise comparable $n$-subsets of $S$, that is, the subsets of the form $\{\sigma_0, \dots, \sigma_{n-1}\}$ such that $\sigma_i \preceq \sigma_{i+1}$.

**Definition 6.11.** *A set $T \subseteq 2^{<\mathbb{N}}$ is* order-isomorphic to $2^{<\mathbb{N}}$ *(written $T \cong 2^{<\mathbb{N}}$) if there is a bijection $g : T \to 2^{<\mathbb{N}}$ such that for every $\sigma, \tau \in T$, $\sigma \preceq \tau$ iff $g(\sigma) \preceq g(\tau)$.*

**Statement 6.12.** $\mathsf{TT}^n_k$ *is the statement "For every coloring $f : [2^{<\mathbb{N}}]^n \to k$, there is a set $T \cong 2^{<\mathbb{N}}$ such that $[T]^n$ is $f$-monochromatic".*

Chubb, Hirst, and McNicholl [15] proved that $\mathsf{TT}^n_2$ is between $\mathsf{ACA}_0$ and $\mathsf{RT}^n_2$ over $\mathsf{RCA}_0$. In particular, for $n \geqslant 3$, $\mathsf{TT}^n_2$ is equivalent to $\mathsf{ACA}_0$. Patey [46] and Dzhafarov and Patey [23] proved that $\mathsf{TT}^2_2$ is strictly in-between $\mathsf{ACA}_0$ and $\mathsf{RT}^2_2$ over $\omega$-models. Chong, Li, Liu and Yang [11, 12, 13] published a series of article and proved, among other things, that $\mathsf{TT}^2_2$ does not imply $\mathsf{WKL}_0$ over $\mathsf{RCA}_0$.

From a proof-theoretic viewpoint, Corduan, Groszek and Mileti [16] showed that $\mathsf{B}\Sigma^0_2$ is strictly weaker than $\mathsf{TT}^1$ over $\mathsf{RCA}_0$, while Chong, Li, Wang and Yang [10] proved that $\mathsf{TT}^1$ is $\Pi^1_1$-conservative over $\mathsf{B}\Sigma^0_2 + \mathsf{P}\Sigma^0_1$, where $\mathsf{P}\Sigma^0_1$ is equivalent to the well-foundedness of $\omega^\omega$. In particular, $\mathsf{TT}^1$ does not imply $\mathsf{I}\Sigma^0_2$ over $\mathsf{RCA}_0$. Last, Chong, Wang and Yang [14] proved that $\mathsf{TT}^1$ is $\forall\Pi^0_3$-conservative over $\mathsf{RCA}_0$. Note that $\mathsf{TT}^2_2$ implies $\mathsf{TT}^1$ over $\mathsf{RCA}_0$.

**Proposition 6.13.** $\mathsf{RCA}_0 \vdash \forall k(\mathsf{CSL}^1_k \to \mathsf{TT}^2_k)$.

*Proof.* Let $f : 2^{<\mathbb{N}} \to k$ be an instance of $\mathsf{TT}^2_k$ for some $k \geqslant 1$. Let $A = \{0, 1\}$ and $h : A^{<\omega, 1} \to k$ be defined by

$$h(v) = f(v[\epsilon], v[0])$$

By $\mathsf{CSL}^1_k$, there is an infinite $\omega$-variable word $W$ over $A$ such that $\mathsf{USub}^{1,\star}_A(W)$ is $h$-monochromatic for some color $c$. Let

$$T = \{\sigma \in 2^{<\mathbb{N}} : |\sigma| \text{ is even } \wedge \forall i < \frac{|\sigma|}{2}, \sigma(2i) = 0\}$$

Note that $T \cong 2^{<\mathbb{N}}$. Finally, let $S = \{W[\sigma] : \sigma \in T\}$. Since $\sigma \mapsto W[\sigma]$ is order-preserving, $S \cong 2^{<\mathbb{N}}$. We claim that $[S]^2$ is $f$-monochromatic for color $c$.

Fix some $\hat{\sigma}, \hat{\tau} \in S$ such that $\hat{\sigma} \prec \hat{\tau}$. Let $\sigma, \tau \in T$ be such that $\hat{\sigma} = W[\sigma]$ and $\hat{\tau} = W[\tau]$. In particular, $\sigma \prec \tau$. By definition of $T$, there is some $\tau' \in 2^{<\mathbb{N}}$ such that $\tau = \sigma \cdot 0 \cdot \tau'$. Let $u = \sigma \cdot \star \cdot \tau'$. By choice of $W$, $h(W[u]) = c$. By definition of $h$,

$$f(\hat{\sigma}, \hat{\tau}) = f(W[\sigma], W[\tau]) = f(W[u][\epsilon], W[u][0]) = h(W[u]) = c$$

□

Chong, Li, Wang and Yang [10] asked whether $\mathsf{TT}^2_2$ implies $\mathsf{I}\Sigma^0_2$ over $\mathsf{RCA}_0$. We answer negatively with the following corollary:

**Corollary 6.14.** $\mathsf{WKL}_0 + \mathsf{TT}^2_2$ *is* $\forall\Pi^0_4$*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

# 7 Conclusion and open questions

The study of variable word theorems from the viewpoint of reverse mathematics remains a subject with a relative lack of understanding, as witnessed by the big gap between lower bounds and upper bounds. For instance, Hindman's theorem, $\mathsf{OVW}^1$ and $\mathsf{CSL}^2$ are known to imply $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$, but the best known upper bounds are $\mathsf{ACA}^+_0$ for $\mathsf{HT}$, and $\Pi^1_1\text{-}\mathsf{CA}_0$ for $\mathsf{OVW}^1$ (Liu, unpublished). The following questions are of fundamental importance:

**Question 7.1.** *Does* $\mathsf{ACA}_0 \vdash \mathsf{HT}$*?*

**Question 7.2.** *For every* $n \in \omega$*, does* $\mathsf{ACA}_0 \vdash \mathsf{OVW}^n$*?*

The framework of $\alpha$-largeness has been shown to be successful in proving separations from theories, where techniques over $\omega$-models failed [31]. The following questions might be more tractable and would still yield significant insights on the reverse-mathematical strength of variable word theorems:

**Question 7.3.** *Is* $\mathsf{RCA}_0 + \mathsf{HT}$ $\forall\Pi^0_3$*-conservative over* $\mathsf{ACA}_0$*?*

**Question 7.4.** *For every* $n \in \omega$*, is* $\mathsf{RCA}_0 + \mathsf{OVW}^n$ $\forall\Pi^0_4$*-conservative over* $\mathsf{ACA}_0$*?*

Last, the known lower bounds to $\mathsf{CSL}^n$ are all coming from the fact that $\mathsf{RCA}_0 \vdash \forall\ell(\mathsf{CSL}^n_\ell \to \mathsf{RT}^{n+1}_\ell)$ (see Proposition 2.6). It follows that its level version, $\mathsf{LCSL}^n$, has no known non-trivial lower bound.

**Question 7.5.** *For every* $n \in \omega$*, is* $\mathsf{RCA}_0 + \mathsf{LCSL}^n$ $\Pi^1_1$*-conservative over* $\mathsf{RCA}_0$*?*

# References


[1] Paul-Elliot Angles d'Auriac, Lu Liu, Bastien Mignoty, and Ludovic Patey. Carlson-Simpson's lemma and applications in reverse mathematics. *Ann. Pure Appl. Logic*, 174(9):Paper No. 103287, 16, 2023.

[2] Paul-Elliot Angles d'Auriac, Peter A Cholak, Damir D Dzhafarov, Benoˆıt Monin, and Ludovic Patey. Milliken's tree theorem and its applications: a computability-theoretic perspective. *arXiv preprint arXiv:2007.09739*, page 141, 2020.

[3] James E. Baumgartner. A short proof of Hindman's theorem. *Journal of Combinatorial Theory, Series A*, 17(3):384–386, 1974.

[4] Andreas R. Blass, Jeffry L. Hirst, and Stephen G. Simpson. Logical analysis of some theorems of combinatorics and topological dynamics. In *Logic and combinatorics (Arcata, Calif., 1985)*, volume 65 of *Contemp. Math.*, pages 125–156. Amer. Math. Soc., Providence, RI, 1987.

[5] Tristan Bompard, Lu Liu, and Ludovic Levy Patey. The reverse mathematics of Carlson's theorem for located words. *J. Comb.*, 16(2):251–279, 2025.

[6] Timothy J. Carlson. Some unifying principles in Ramsey theory. *Discrete Math.*, 68(2-3):117–169, 1988.

[7] Timothy J. Carlson and Stephen G. Simpson. A dual form of Ramsey's theorem. *Adv. in Math.*, 53(3):265–290, 1984.

[8] Peter Cholak, Natasha Dobrinen, and Charlie McCoy. The henson graphs: colorings and codings, 2026.

[9] Peter Cholak, Natasha Dobrinen, and Henry Towsner. The finite big ramsey degrees of henson graphs are provable in $\mathrm{ACA}_0$, 2026.

[10] C. T. Chong, Wei Li, Wei Wang, and Yue Yang. On the strength of Ramsey's theorem for trees. *Adv. Math.*, 369:107180, 39, 2020.

[11] Chi Tat Chong, Wei Li, Lu Liu, and Yue Yang. The Strength of Ramsey's Theorem For Pairs over trees: II. Chain-Anti-Chain and Ascending Descending Sequence. 2019. _eprint: arXiv:1912.09049.

[12] Chi Tat Chong, Wei Li, Lu Liu, and Yue Yang. The Strength of Ramsey's Theorem For Pairs over trees: III. low2 construction. 2019. _eprint: arXiv:1912.09049.

[13] Chi Tat Chong, Wei Li, Lu Liu, and Yue Yang. The Strength of Ramsey's Theorem For Pairs over trees: IV. Erdos-Rado's theorem for rationals. 2019. _eprint: arXiv:1912.09049.

[14] Chi Tat Chong, Wei Wang, and Yue Yang. Conservation strength of the infinite pigeonhole principle for trees. *Israel Journal of Mathematics*, pages 1–24, 2023.

[15] Jennifer Chubb, Jeffry L. Hirst, and Timothy H. McNicholl. Reverse mathematics, computability, and partitions of trees. *J. Symbolic Logic*, 74(1):201–215, 2009.

[16] Jared Corduan, Marcia J. Groszek, and Joseph R. Mileti. Reverse mathematics and Ramsey's property for trees. *J. Symbolic Logic*, 75(3):945–954, 2010.

[17] Denis Campau Devlin. *SOME PARTITION THEOREMS AND ULTRAFILTERS ON OMEGA*. ProQuest LLC, Ann Arbor, MI, 1980. Thesis (Ph.D.)–Dartmouth College.

[18] Natasha Dobrinen. The Ramsey theory of the universal homogeneous triangle-free graph. *Journal of Mathematical Logic*, 20(2):2050012, 75, 2020. tex.fjournal: Journal of Mathematical Logic tex.mrclass: 03E02 (03C15 03E05 03E40 03E75 05C05 05C55) tex.mrnumber: 4128725 tex.mrreviewer: G.Cherlin.

[19] Pandelis Dodos and Vassilis Kanellopoulos. *Ramsey theory for product spaces*, volume 212 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2016.

[20] Pandelis Dodos, Vassilis Kanellopoulos, and Konstantinos Tyros. A density version of the Carlson-Simpson theorem. *J. Eur. Math. Soc. (JEMS)*, 16(10):2097–2164, 2014.

[21] Damir Dzhafarov, Stephen Flood, Reed Solomon, and Linda Westrick. Effectiveness for the dual ramsey theorem. *Notre Dame Journal of Formal Logic*, 62(3):455–490, 2021.

[22] Damir D. Dzhafarov and Carl Mummert. *Reverse mathematics—problems, reductions, and proofs*. Theory and Applications of Computability. Springer, Cham, [2022] ©2022.

[23] Damir D. Dzhafarov and Ludovic Patey. Coloring trees in reverse mathematics. *Adv. Math.*, 318:497–514, 2017.

[24] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[25] Harvey Friedman. Personal communication to L. Harrington, 1977.

[26] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2013. Paperback edition of the second (1990) edition [MR1044995].

[27] Petr Hájek and Pavel Pudlák. *Metamathematics of first-order arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1998. Second printing.

[28] Neil Hindman. Finite sums from sequences within cells of a partition of $N$. *J. Combinatorial Theory Ser. A*, 17:1–11, 1974.

[29] Neil Hindman and Dona Strauss. *Algebra in the Stone-Čech compactification: theory and applications*. De Gruyter textbook. De Gruyter, Berlin ; Boston, 2nd revised and extended ed edition, 2012.

[30] Jeffry Lynn Hirst. *COMBINATORICS IN SUBSYSTEMS OF SECOND ORDER ARITHMETIC*. ProQuest LLC, Ann Arbor, MI, 1987. Thesis (Ph.D.)–The Pennsylvania State University.

[31] Quentin Le Houérou and Ludovic Levy Patey. $\Pi^0_4$ conservation of the Ordered Variable Word theorem. *arXiv preprint arXiv:2404.18749*, 2024.

[32] Quentin Le Houérou and Ludovic Patey. Largeness notions and polytime translation for $\forall\Sigma^0_3$-consequences of $\mathsf{RT}^2_2$, 2026.

[33] Jan Hubička. Big ramsey degrees using parameter spaces. *arXiv preprint arXiv:2009.00967*, page 20, 2020.

[34] Richard Kaye. *Models of Peano arithmetic*, volume 15 of *Oxford Logic Guides*. The Clarendon Press, Oxford University Press, New York, 1991. Oxford Science Publications.

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

[36] L. A. S. Kirby and J. B. Paris. Initial segments of models of Peano's axioms. In *Set theory and hierarchy theory, V (Proc. Third Conf., Bierutowice, 1976)*, Lecture Notes in Math., Vol. 619, pages 211–226. Springer, Berlin, 1977.

[37] Leszek Aleksander Kołodziejczyk, Tin Lok Wong, and Keita Yokoyama. Ramsey's theorem for pairs, collection, and proof size. *Journal of Mathematical Logic*, page 2350007, 2023.

[38] Leszek Aleksander Kołodziejczyk and Keita Yokoyama. Some upper bounds on ordinal-valued Ramsey numbers for colourings of pairs. *Selecta Math. (N.S.)*, 26(4):Paper No. 56, 18, 2020.

[39] Quentin Le Houérou, Ludovic Levy Patey, and Keita Yokoyama. $\Pi^0_4$ conservation of Ramsey's theorem for pairs. In preparation.

[40] Quentin Le Houérou, Ludovic Patey, and Keita Yokoyama. Corrigendum to: $\Pi^0_4$ conservation of Ramsey's theorem for pairs. 2026.

[41] Quentin Le Houérou, Ludovic Levy Patey, and Keita Yokoyama. $\Pi^0_4$ conservation of ramsey's theorem for pairs. *Journal of the London Mathematical Society*, 113(1):e70419, 2026.

[42] Lu Liu, Benoit Monin, and Ludovic Patey. A computable analysis of variable words theorems. *Proc. Amer. Math. Soc.*, 147(2):823–834, 2019.

[43] Lu Liu and Ludovic Patey. The reverse mathematics of the ordered variable word theorem. June 2026.

[44] Joseph S. Miller and Reed Solomon. Effectiveness for infinite variable words and the dual Ramsey theorem. *Arch. Math. Logic*, 43(4):543–555, 2004.

[45] Keith R. Milliken. A Ramsey theorem for trees. *J. Combin. Theory Ser. A*, 26(3):215–237, 1979.

[46] Ludovic Patey. The strength of the tree theorem for pairs in reverse mathematics. *J. Symb. Log.*, 81(4):1481–1499, 2016.

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

[48] N. W. Sauer. Coloring subgraphs of the Rado graph. *Combinatorica*, 26(2):231–253, 2006.

[49] Dana Scott. Algebras of sets binumerable in complete extensions of arithmetic. In *Proc. Sympos. Pure Math*, volume 5, pages 117–121, 1962.

[50] Saharon Shelah. Primitive recursive bounds for van der Waerden numbers. *J. Amer. Math. Soc.*, 1(3):683–697, 1988.

[51] Stephen G. Simpson. *Subsystems of second order arithmetic*. Perspectives in Logic. Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, second edition, 2009.

[52] Stevo Todorcevic. *Introduction to Ramsey spaces*, volume 174 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2010.

[53] Henry Towsner. Hindman's theorem: an ultrafilter argument in second order arithmetic. *J. Symbolic Logic*, 76(1):353–360, 2011.

[54] Henry Towsner. A simple proof and some difficult examples for Hindman's theorem. *Notre Dame J. Form. Log.*, 53(1):53–65, 2012.

[55] Henry Towsner and Keita Yokoyama. Erdős-moser and $I\Sigma_2$. *Israel J. Math.*, 263(2):843–870, 2024.