Conservativity of ultrafilters over subsystems of second order arithmetic

Antonio Montalbán*
Department of Mathematics
University of California, Berkeley
Berkeley CA 94720 USA

Richard A. Shore**
Department of Mathematics
Cornell University
Ithaca NY 14853

October 10, 2017

Abstract

We extend the usual language of second order arithmetic to one in which we can discuss an ultrafilter over of the sets of a given model. The semantics are based on fixing a subclass of the sets in a structure for the basic language that corresponds to the intended ultrafilter. In this language we state axioms that express the notion that the subclass is an ultrafilter and additional ones that say it is idempotent or Ramsey. The axioms for idempotent ultrafilters prove, for example, Hindman's theorem and its generalizations such as the Galvin-Glazer theorem and iterated versions of these theorems (IHT and IGG). We prove that adding these axioms to IHT produce conservative extensions of ACA_0 (+IHT), ACA_0^+, ATR_0, Π^1_2-CA_0 and Π^1_2-CA_0 for all sentences of second order arithmetic and for full Z_2 for the class of Π^1_4 sentences. We also generalize and strengthen a metamathematical result of Wang [1984] to show, for example, that any Π^1_2 theorem $\forall X \exists Y \Theta(X, Y)$ provable in ACA_0 or ACA_0^+ there are $e, k \in \mathbb{N}$ such that ACA_0 or ACA_0^+ proves that $\forall X(\Theta(X, \Phi_e(J^{(k)}(X)))$ where Φ_e is the eth Turing reduction and $J^{(k)}$ is the kth iterate of the Turing or Arithmetic jump respectively. (A similar result is derived for Π^1_3 theorems of Π^1_1-CA_0 and the hyperjump.)

1 Introduction

An increasingly common phenomena in combinatorics is the use of higher order notions, objects and principles to prove combinatorial facts about the natural numbers $\mathbb{N}$ (or some

*Partially supported by NSF grant DMS-1363310 and by a Packard fellowship.
**Partially supported by NSF Grant DMS-1161175.

other countable set) and its subsets. One particularly fruitful such notion has been that of a (nonprincipal) ultrafilters on $\mathbb{N}$. (The definition of ultrafilters and most of the other notions mentioned in this introduction are given in §2. Hindman [2005] gives a brief survey of some examples. See Todorcevic [2010] and, especially, Hindman and Strauss [2012] for more extensive treatments.) These results provide a challenge to the analysis of the inherent complexity of such second order theorems in terms of the axioms (of second order arithmetic) needed to prove them and the computational complexity of the objects whose existence are asserted by such theorems. The study of such questions of complexity in general is the realm of reverse mathematics. (See the standard text on the subject, Simpson [2009] for a general introduction.)

One successful approach to this problem has been to replace the ultrafilters used in the proofs by countable approximations of various sorts that can be proven to exist in different axiomatic subsystem of second order arithmetic. (See for example Avigad [1998], Hirst [2004] and Towsner [2011], [2011a].) Another attractive idea is to attack the problem in a more wholesale fashion by expanding the language of second order arithmetic and the associated axiomatic systems in a way that enables one to carry out the proofs using ultrafilters but in a system that one can prove is conservative over the associated set of second order axioms. (For a set of sentences Γ in some language L and sets R and T of axioms the language L, we say that R is Γ-conservative over T if, for every $\Phi \in \Gamma$, if $R \vdash \Phi$ then $T \vdash \Phi$.) So, in our situation we want L to be an extension of the language of second order arithmetic in which we can talk about an ultrafilter is some reasonable way. We want Γ to be a natural subclass of the sentences of second order arithmetic containing theorems of interest. T should be a (standard) axiom system of second order arithmetic and R a natural extension of T to the language L. The system R should assert the existence of some type of ultrafilter and in it we should be able to carry out the proofs of the second order combinatorial theorems of interest. If we can then prove that R is Γ-conservative over T we will have shown that the theorems of interest are actually provable in the known second order theory T.

We first saw this approach carried out in Towsner [2014] who proved the desired results for (general) nonprincipal ultrafilters and several of the standard subsystems T of reverse mathematics (ACA_0, ATR_0 and Π^1_1-CA_0). His approach was to extend the language by adding a unary predicate (for the ultrafilter) as well as some technical term formation rules to allow various natural operations using the ultrafilter to be expressed in the extension. He then defined axiom systems R extending T in which one might be able to carry out arguments using the ultrafilter. Finally, he used forcing for this language and proved in T that all the axioms of R were all forced to be true (as then are all their consequences in second order arithmetic). In the standard way, this then shows that R is conservative over T for all sentences of second order arithmetic. (Towsner points out that this result for ACA_0, as well as the very strong theory $(\mathrm{Z}_2) + \Pi^1_\infty$-$\mathrm{DC}_0$, also follow from Enayat [2006] and that Kreuzer [2012] gives a proof for ACA_0 and Π^1_2 conservation using the functional interpretation.)

Unfortunately, most (if not all) of the classical and more recent examples of the applications of ultrafilters to second order combinatorics depend on the existence of ultrafilters with various special properties. Towsner provides two examples, Ramsey ultrafilters and divisible ultrafilters when the notion of divisibility is arithmetically definable, where his methods apply but no examples of proofs of known combinatorial theorems in the relevant systems. Historically, the basic and motivating example of uses of ultrafilters was the Galvin-Glazer proof of Hindman's theorem and its many generalizations. Hindman's theorem can be phrased in terms of a more complicated (Σ^1_1 complete) divisibility property (called IP). So the existence of an ultrafilter all of whose members are IP would prove Hindman's theorem. Towsner conjectures that such an assumption is conservative over ATR_0. We prove more below.

Hindman's original proof and the direct reverse mathematical analyses of his theorem are all long and difficult combinatorial arguments. The Galvin-Glazer proof of Hindman's theorem is a very short one based on the existence an algebraically defined type of ultrafilter called idempotent. (The algebra of ultrafilters on a countable set is based on an operation inherited from one in the underlying countable set such as $+$ on $\mathbb{N}$.) Towsner then asks for which (of the standard) theories T is the existence of an idempotent ultrafilter conservative. (The point here is that every member of an idempotent ultrafilter is easily seen to be IP.)

Our goal in this paper is to answer these questions positively except for the fundamental one of whether the existence of idempotent ultrafilters is conservative over ACA_0. (Such a result would show that IHT is provable in ACA_0 and so settle a long standing open problem.) We do prove, however, that it is conservative over a variation of Hindman's theorem (Iterated Hindman's theorem, IHT). Thus, in terms of our systems and languages, IHT is equiconsistent with the existence of an idempotent ultrafilter. Reverse mathematically, it is known that both Hindman's theorem (HT) and the iterated version, IHT, imply ACA_0 and are provable in ACA_0^+ (Blass, Hirst, Simpson [1987]). Our methods also show that the existence of an idempotent ultrafilter (§4) or a Ramsey ultrafilter (§5) is conservative (for all sentences of second order arithmetic) over ACA_0^+, ATR_0, Π^1_1-CA_0 and Π^1_2-CA_0 and conservative over full Z_2 for Π^1_4 sentences.

Of course, as did Towsner, we must also define an extension of the language of second order arithmetic to include ways of talking about and using the intended ultrafilters. We also use forcing notions designed so that the generic objects will be ultrafilters with the properties we want. Our proofs of conservativity take the semantic approach of analyzing the structure for our extended language defined from a generic object as the intended ultrafilters rather than the relation between provability and forcing.

As we were writing this paper, we found that Kreuzer [2015] had proved some of our results in a different (higher order) setting and by proof theoretic methods to get conservation results over systems like ACA_0 for Π^1_2 sentences and had extended them to minimal idempotent ultrafilters and so to applications to other well known combinatorial principles [2015a]. We present the standard ultrafilter proof of IHT (and so some generalizations as well) in our system ($\mathrm{ACA}_0^{I\mathcal{U}}$) in Theorem 3.3. We also give another example of a standard proof of the Milliken-Taylor theorem using ultrafilters in our system in Theorem 3.4. Our general conservation result then shows that it is provable from IHT (Hirst [2004]).

Most applications of ultrafilters to combinatorics, however, use ultrafilters with more properties than idempotence and/or considerations of extra structural features beyond addition on $\mathbb{N}$ or, more generally, multiplication in various semigroups. In future work, we hope to apply some of the ideas in this paper to another approach to conservation results that will handle additional properties such as minimality for ultrafilters and additional structure on the underlying algebra as well as a more robust language and setting for working with the ultrafilters. We then hope to apply this approach to well known combinatorial theorems such as Gowers' Fin_k Theorem and the Infinite Hales-Jewett Theorem.

2 Definitions and Notations

We use one of the common notations to view a single set X as a sequence of sets $\langle X^{[n]} \rangle$ where $X^{[n]} = \{i | \langle n, i \rangle \in X\}$. We let $X^{[<n]} = \{\langle j, i \rangle \in X | j < n\}$ and use $\bar{X}$ to denote the complement of X. The set of finite subsets of X is denoted by $[X]^{<\omega}$. We denote the standard natural numbers by $\mathbb{N}$.

Note that all sets and structures considered will be countable.

2.1 Subsystems of Second Order Arithmetic

We first briefly review the five standard systems of reverse mathematics as well as a couple of other systems that we will use. Details, general background and results can be found in Simpson [2009], the standard text on reverse mathematics. Each of the systems is given in the (two sorted) language $\mathcal{L}$ of second order arithmetic, that is, the usual first order language of arithmetic augmented by set variables with their quantifiers and the membership relation $\in$ between numbers and sets. A structure for this language is one of the form $\mathcal{M} = \langle M, S, +, \times, <, 0, 1, \in \rangle$ where M is a set (the set of "numbers" of $\mathcal{M}$) over which the first order quantifiers and variables of our language range; $S \subseteq 2^M$ is the collection of subsets of "numbers" in $\mathcal{M}$ over which the second order quantifiers and variables of our language range; $+$ and $\times$ are binary functions on M; $<$ is a binary relation on M while 0 and 1 are members of M. We also use ω to denote the set of numbers in such a structure in many familiar notations such as writing $2^{<\omega}$ to mean the set of all finite binary strings where we are taking for granted some standard coding of binary strings by numbers of the structure. Indeed, in general, when we talk about finite sets or sequences of elements in $\mathcal{M}$ we are assuming some standard coding in the usual reverse mathematical style. Thus $M^{<\omega}$ is the collection of all finite sequences in $\mathcal{M}$ and $[M]^{<\omega}$ is the collection of all finite sets in $\mathcal{M}$.

Each subsystem of second order arithmetic that we consider contains the standard basic axioms for $+$, $\cdot$, and $<$ (which say that M is an ordered semiring). In addition, they all include a form of induction that applies only to sets (that belong to the model):

(I_0) $\quad (0 \in X \,\&\, \forall n\,(n \in X \to n+1 \in X)) \to \forall n\,(n \in X)$.

We call the system consisting of I_0 and the basic axioms of ordered semirings P_0. All five standard systems are defined by adding various types of set existence axioms to P_0. All but one are determined by comprehension axioms for a standard class Γ of formulas. Note that these formulas may have free set or number variables. As usual the existence assertion $\exists X....$ of the axiom is taken to mean that for each instantiations of the free variables there is an X as described.

(RCA_0) Recursive Comprehension Axioms: This is a system just strong enough to prove the existence of the computable sets. In addition to P_0 its axioms include the schemes of Δ^0_1 comprehension and Σ^0_1 induction:

(Δ^0_1-CA_0) $\quad \forall n\,(\varphi(n) \leftrightarrow \psi(n)) \to \exists X\,\forall n\,(n \in X \leftrightarrow \varphi(n))$ for all Σ^0_1 formulas φ and Π^0_1 formulas ψ in which X is not free.

($\mathrm{I}\Sigma^0_1$) $\quad (\varphi(0) \,\&\, \forall n\,(\varphi(n) \to \varphi(n+1))) \to \forall n\,\varphi(n)$ for all Σ^0_1 formulas φ.

The next system says that every infinite binary tree has an infinite path.

(WKL_0) Weak König's Lemma: This system consists of RCA_0 plus the statement that every infinite subtree of $2^{<\omega}$ has an infinite path.

We next move up to arithmetic comprehension.

(ACA_0) Arithmetic Comprehension Axioms: This system consists of RCA_0 plus the axioms $\exists X\,\forall n\,(n \in X \leftrightarrow \Phi(n))$ for every arithmetic formula (i.e. Σ^0_n for some n) φ in which X is not free.

This system is equivalent to the totality of the Turing jump operator, i.e. for every X, X' exists. The next system says that arithmetic comprehension can be iterated along any countable well order.

(ATR_0) Arithmetical Transfinite Recursion: This system consists of RCA_0 plus the following axiom for each arithmetic formula $\Theta(n, Z)$: If $\langle X, <_X\rangle$ is (a code for) a well-ordering, then there is a set Y such that, for each $x \in X$, $\forall n(n \in Y^{[x]} \Leftrightarrow \Theta(n, Y^{[<x]})$.

The last of the standard systems is characterized by the comprehension axiom for Π^1_1 formulas:

(Π^1_1-CA_0) This system consists of RCA_0 plus the Π^1_1 comprehension axioms: $\exists X\,\forall k\,(k \in X \leftrightarrow \Phi(k))$ for every Π^1_1 formula Φ in which X is not free.

(Z_2) Full second order arithmetic, Z_2 consists of RCA_0 plus all the Π^1_n-CA$_0$, i.e. the comprehension axioms as above but for Π^1_n formulas φ for every $n \in \mathbb{N}$.

We also we refer to a system between ACA_0 and ATR_0 that corresponds to the ability to iterate the Turing jump ω many times, i.e. the totality of the ω-jump operator.

(ACA_0^+) This system consists of RCA_0 plus the axiom $\forall X \exists Y (Y^{[0]} = X \;\&\; \forall n > 0 (Y^{[n]} = (Y^{[n-1]})'))$. (Here we take some standard definition of the Turing jump in first order arithmetic.) The axiom says we can iterate the Turing jump through the numbers (of the model).

The above axiom systems are strictly ordered in the obvious way. We also need to use some choice principles whose relationships to the ones above are not so easily categorized. (See Simpson [2009].)

(strong Σ^1_n-DC$_0$) For every Σ^1_n formula $\Phi(k, X, Y)$, $\exists Z \forall k \forall Y (\Phi(k, Z^{[<n]}, Y) \to \Phi(k, Z^{[<n]}, Z^{[n]})$.

We note for later use that Π^1_1-CA$_0$ implies (indeed is equivalent to) strong Σ^1_1-DC$_0$ and Π^1_2-CA$_0$ implies (indeed is equivalent to) strong Σ^1_2-DC$_0$ (Simpson [2009, VII.6.9]).

2.2 Combinatorics

Definition 2.1. Let $\mathcal{M}$ be a structure for the language $\mathcal{L}$ of second order arithmetic with $\mathcal{M} \vDash RCA_0$. An *$\mathcal{M}$-filter* is a nonempty subset $\mathcal{F}$ of S which is closed upward and under intersection, i.e. if $X, Y \in S$, $X \subseteq Y$ and $X \in \mathcal{F}$, then $Y \in \mathcal{F}$; and if $X, Y \in \mathcal{F}$ then $X \cap Y \in \mathcal{F}$. An $\mathcal{M}$-filter is *principal* if there is an $A \in S$ such that $\mathcal{F} = \{X \in S | X \supseteq A\}$; otherwise it is nonprincipal. A nonprincipal $\mathcal{M}$-filter $\mathcal{U}$ is an *$\mathcal{M}$-ultrafilter* if for every $A \in S$, $A \in \mathcal{U}$ or $\bar{A} \in \mathcal{U}$.

Definition 2.2. An $\mathcal{M}$-ultrafilter $\mathcal{U}$ is *Ramsey* if for every partition $\langle X_n \rangle \in S$ of M into nonempty pairwise disjoint sets such that $X_n \notin \mathcal{U}$ for every n, there is a $Z \in \mathcal{U}$ such that $|X_n \cap Z| = 1$ for every n.

Before we can, even classically, define a notion of an idempotent ultrafilter we need a number of definitions.

Definition 2.3. A *(partial) semigroup (Z,$*$)* is a set Z with a (partial) binary operation $*$ which is associative. An *idempotent* is an element x such that $x * x = x$. A partial semigroup is *directed* if, for every finite sequence $x_0, \ldots, x_n$ of elements of Z, there is a $y \in Z$ such that $y \neq x_i$ for each $i \leq n$ and $x_0 * y, \ldots, x_n * y$ are all defined. It is *left-cancelative*, if, for every x, y, z with $x * y = x * z$, $y = z$. A *basic sequence* in a partial semigroup is a finite or infinite sequence $\langle x_i \rangle$ of pairwise distinct elements such that $x_{n_0} * x_{n_1} * \ldots * x_{n_k}$ is defined for every finite sequence $n_0 < \ldots < n_k$.

Notation 2.4. *For $X \subseteq M$,* FS(X) $= \{x_0 + x_1 + \ldots + x_n|$ *for any* $\langle x_i \rangle \in M^{<\omega}$ *with* $x_0 < x_1 < \ldots < x_n$ *each in* $X\}$*. We let* $X+n = \{x+n|x \in X\}$ *and* $X-n = \{m|n+m \in X\}$*. For* $X = \{D_i\} \subseteq [M]^{<\omega}$*,* $FU(X) = \{\cup\{D_i|i \in E\}|\ E \in [M]^{<\omega}\}$*. We define* $X * n$ *and* X/n *similarly for* $X \subseteq Z$ *and* $n \in Z$ *for a partial semigroup* $(Z, *)$*. We extend the operations* $FS(X)$ *and* $FU(X)$ *to basic sequences* $X = \langle x_i \rangle$ *in a semigroup* $(Z, *)$ *by setting* $[X] = \{x_{n_0} * \cdots * x_{n_k}|n_0, n_1, \cdots, n_k$ *an increasing sequence in* $X^{<\omega}\}$*.*

For background information and intuition for some of the standard terminology, we now describe some notions in the setting of classical mathematics.

One can extend the $*$ operation on a partial semigroup $(Z, *)$ to a total operation $*$ on a subclass γZ of the class βZ of all ultrafilters over Z: $\gamma Z = \{\mathcal{U} \in \beta Z|(\forall x \in Z)(\{y \in Z|x * y \text{ is defined}\} \in \mathcal{U}\}$. For $\mathcal{U}$, $\mathcal{V} \in \gamma Z$, $\mathcal{U} * \mathcal{V} = \{A \subseteq Z|\{x \in Z|\{y \in Z|x * y \in A\} \in \mathcal{V} \in \mathcal{U}\}$. (One can read this as the collection of subsets A of Z such that, for almost all x (w.r.t. to $\mathcal{U}$) and almost all y (w.r.t. $\mathcal{V}$), $x * y \in A$.) One can check that this operation is total and associative on γZ. (See e.g. Todorcevic [2010, 2.1 and 2.2] for all of this.) An ultrafilter $\mathcal{U} \in \gamma Z$ is then *idempotent* if $\mathcal{U} * \mathcal{U} = \mathcal{U}$.

After we introduce our extension of second order arithmetic in §3, we will give an axiom ($\mathcal{AU}5$) which is easily seen to be a translation of this definition into our language.

We now state several versions of, and variations on, Hindman's theorem in structures $\mathcal{M}$ for second order arithmetic without considering, at this point, the axiom systems needed to prove them. The first is what is usually called Hindman's Theorem (HT) (or the Finite Sums Theorem). The second is the Finite Union Theorem (FU) but also sometimes called Hindman's Theorem. The third is the analog for directed partial semigroups and is usually called the Galvin-Glazer Theorem (GG). Each one asserts the existence of a type of homogeneous set for various colorings. We then state the (obviously at least as strong) variations of these theorems for sequences of colorings. Finally, we prove the equivalences for the iterated versions and note that the proofs specialize to the ones for the ordinary versions of these theorems.

Theorem 2.5 (Hindman's Theorem (HT)). *If f is a finite coloring of M, i.e. $f : M \to k$, then there is an infinite $X \subseteq M$ such that $FS(X)$ is monochromatic, i.e. there is an $i < k$ such that for every finite sum z of distinct elements of X, $f(z) = i$. Such a set is called homogeneous for f.*

Theorem 2.6 (Finite Union Theorem (FU)). *If f is a finite coloring of $[M]^{<\omega}$, then there is an infinite subset X of $[M]^{<\omega}$ such that $FU(X)$ is monochromatic, i.e. all finite unions of members of X have the same color. Moreover, we may choose $X = \langle x_i \rangle$ to be a* block sequence*, i.e. $\forall i(\max x_i < \min x_{i+1})$. Such a set is called homogeneous for f.*

Of course, FU is equivalent (over RCA_0) to the same theorem for any countable set W in place of M.

Theorem 2.7 (Galvin-Glaser Theorem (GG)). *If $(Z, *)$ is a directed partial semigroup with no idempotents or which is left-cancelative and f is a finite coloring of Z, then* *there is an infinite basic sequence $X = \langle x_n \rangle$ in S such that $[X]$ is (necessarily infinite and) monochromatic. Such a set is called homogeneous for f.*

Theorem 2.8 (Iterated Hindman's Theorem (IHT)). *Given a sequence $\langle f_n, k_n \rangle$ where f_n is a k_n coloring of M, there is an infinite sequence $X = \langle x_i \rangle$ from M such that, for each n, $FS(\{x_i | i \geq n\})$ is monochromatic for f_n. Such a set is called homogeneous for $\langle f_n \rangle$.*

Note that this theorem is equivalent (over RCA_0) to the one where all $k_n = 2$.

Theorem 2.9 (Iterated Finite Union Theorem (IFU)). *Given a sequence $\langle f_n, k_n \rangle$ where f_n is a k_n coloring of $[M]^{<\omega}$, there is an infinite block sequence $X = \langle x_i \rangle$ from $[M]^{<\omega}$ such that, for each n, $FU(\{x_i | i \geq n\})$ is monochromatic for f_n. Such a set is called homogeneous for $\langle f_n \rangle$.*

Again this theorem is equivalent (over RCA_0) to the one where all $k_n = 2$. Of course, IFU implies the same theorem for any countable set W in place of M.

Theorem 2.10 (Iterated Galvin-Glaser Theorem (IGG)). *If $(Z, *)$ is as in Theorem 2.7, and we are given a sequence $\langle f_n, k_n \rangle$ where f_n is a k_n coloring of Z, then there is an infinite basic sequence $X = \langle x_n \rangle$ in Z such that for each n, $[\{x_i | i \geq n\}]$ is monochromatic for f_n. Such a set is called homogeneous for $\langle f_n \rangle$.*

We now wish to prove that HT, FU and GG are equivalent over RCA_0 as are the iterated versions. These equivalences were all essentially known to Hindman and others although they were not working in the framework of RCA_0. There are proofs of these equivalences for HT, FU and GG in Bergelson and Hindman [1993, Lemma 2.1] which are described there as "well known among aficionados" and are also outlined in Hindman and Strauss [2012, §5.5, Ex. 5.2.1-5.2.4] but the proofs in both cases for GG are only for semigroups. The ones for partial semigroups suggested in Ex 5.2.5 and Todorcevic [2010] for partial semigroups uses ultrafilters. We present similar proofs for the iterated versions in RCA_0 which also specialize to the noniterated versions. We begin with a simple Lemma.

Lemma 2.11 ($\mathbf{RCA}_0$ (Hindman [1972], Lemma 2.2)). *If $F \in [M]^{<\omega}$ and $|F| = 2^n$ then there is a $G \subseteq F$ such that $2^n | \Sigma G$.*

Proof. As Hindman [1972] points out this is a simple induction. It clearly works in RCA_0. □

Now for a bit more complicated Lemma that we need for the analysis of IGG.

Lemma 2.12. *If $(Z, *)$ is a directed partial semigroup it has an infinite basic sequence. If Z has no idempotents or is left-cancelative, then there is an infinite basic sequence $Y = \langle y_n \rangle$ in Z such that the products $y_{n_0} * \cdots * y_{n_{k-1}} = \Pi \bar{u}$ and $y_{m_0} * \cdots * y_{m_{l-1}} = \Pi \bar{v}$ are all distinct for $n_{k-1} < m_0$ which we write as $\bar{u} < \bar{v}$ where $\bar{u} = \left\langle y_{n_0}, \ldots, y_{n_{k-1}} \right\rangle$ and $\bar{v} = \left\langle y_{m_0}, \ldots, y_{m_{l-1}} \right\rangle$.*

Proof. If Z is directed we can easily build a basic sequence by induction: Given y_i, $i < n$, directedness shows that there is a z (indeed infinitely many) such that $\Pi\bar{u} * z$ is defined for all $\bar{u}$ with $n_{k-1} < n$. So we may take any such z as y_n.

If Z is left-cancelative we also want maintain the inductive hypothesis that $\Pi\bar{u} \neq \Pi\bar{v}$ for all $\bar{u} < \bar{v}$ with the last component of $\bar{v}$ having subscript less than n. Now, for each such $\bar{u} < \bar{v}$, there can be at most one z with $\bar{u} = \bar{v} * z$ as if $\bar{u} = \bar{v} * z_1 = \bar{v} * z_2$ then $z_1 = z_2$ since Z is left-cancellative. Thus we can choose a z as required to continue the block sequence and such that for all such $\bar{u} < \bar{v}$, $\bar{u} \neq \bar{v} * z$ and so maintain our inductive hypothesis by taking any such a z as y_n.

If Z has no idempotents let Y be any infinite basic sequence. If there is a z such that there are infinitely many k for which there are $\bar{u}$ with $y_{n_0} = k$ and $\Pi\bar{u} = z$, then we can find $\bar{u}_i$ with $\Pi\bar{u}_i = z$ and $\bar{u}_i < \bar{u}_j$ for $i < j$. Since Y is a basic sequence, all the products $\Pi u_0 * \cdots * \Pi_{u_n}$ i.e. all the z^n are defined. If they are all distinct then $\left\langle z^{2^i} \right\rangle$ is a basic sequence as desired. Otherwise there are n and m such that $z^n = z^m$ and so the subsemigroup $\hat{Z}$ generated by z is finite (of size at most m). We claim it has an idempotent for a contradiction: If not, Let $t \leq m$ be the size of the smallest subsemigroup Z' of $\hat{Z}$ generated by some $z^k = w$ which has no idempotent. Clearly $t > 1$ and so there are $p < q \leq t$ such $w^p = w^q$. Now $w^{q-p} * w^p = w^q = w^p$ and so by repeated multiplication by w^p, $w^{q-p} * w^{p \cdot l} = w^{p \cdot l}$ for every $l \leq t$. If the subsemigroup generated by w has size less than t it (and so $\hat{Z}$) would have an idempotent. Thus it has size t and is all of Z'. In particular, some $w^{p \cdot l} = w^{q-p}$ and so is our desired idempotent.

Finally, we have the case that for every z there are only finitely many k such that there are $\bar{u}$ with $y_{n_0} = k$ and $\Pi\bar{u} = z$. With this assumption we can easily thin out Y to satisfy the conclusions of the Lemma. Suppose we have $y_{m_0}, \ldots, y_{m_l}$ and a_l so that $\Pi\bar{u} \neq \Pi\bar{v}$ for all $\bar{u}$ from this sequence so far and all $\bar{v}$ from Y with first component v_i with $i \geq a_l$. By our assumption we may now take $m_{i+1} = a_l$ and choose a_{l+1} large enough to continue the induction. The new basic sequence clearly has the desired property. □

Proposition 2.13 (RCA$_0$). *IHT, IFU and IGG are all equivalent as are HT, FU and GG as are HT$_n$, FU$_n$ and GG$_n$ for each $n \in M$ where each of these subscripted versions is the restriction of the full theorem to n-colorings.*

Proof. We assume IHT and prove IFU. Let $\langle f_n, k_n \rangle$ be as in the hypothesis of IFU and define g_n from f_n by $g_n(x) = f_n(B_x)$ where B_x is the set of a_i which appear in the binary expansion of x, i.e. if $x = \Sigma 2^{a_i}$. Now let $X = \langle x_i \rangle$ be homogeneous for $\langle g_n, k_n \rangle$ as in IHT. We define a sequence $\langle y_i \rangle$ by recursion and show it is homogeneous for $\langle f_n, k_n \rangle$:

In addition to the y_i we define $z_i \in FS\{x_j | j \geq n\}$ such that $y_i = B_{z_i}$. We begin with $z_0 = x_0$ and $y_0 = B_{z_0}$. Assume for the recursion step that we have $y_i = B_{z_i}$ for $i \leq n$ defined as desired. By Lemma 2.11, we may choose $n < j_0 < \cdots < j_m$ so that if $z_{n+1} = \Sigma x_{j_k}$ and $y_{n+1} = B_{z_{n+1}}$ then $\max y_n < \min y_{n+1}$.

As for the required homogeneity, consider any $n \leq i_0 < \cdots < i_m$. As $\langle y_i \rangle$ is a block sequence and $y_i = B_{z_i}$, the definition of g_n says that $f_n(\cup\{y_{i_k} | k \leq m\}) = g_n(\Sigma\{z_{i_k} | k \leq m\})$.

As $z_{i_k} \in FS\{x_j | j \geq i_k\} \subseteq FS\{x_j | j \geq n\}$ the fact that $FS\{x_j | j \geq n\}$ is monochromatic for g_n implies that $FU(\{x_i | i \geq n\})$ is monochromatic for f_n as required.

That HT (HT_n) implies FU (FU_n) now follows by the same proof applied to a single coloring f of $M^{<\omega}$.

Next we assume IFU and prove IGG. Let an appropriate semigroup $(Z, *)$ be given along with a sequence $\langle f_n, k_n \rangle$ of colorings as in IGG. We begin with an infinite basic sequence Y as given by Lemma 2.12. For $F = \{y_{n_0}, \ldots, y_{n_{k-1}}\}$ let $g_n(F) = f_n(y_{n_0} * \cdots * y_{n_{k-1}})$. Now choose a block sequence $X = \langle x_i \rangle$ homogeneous for the g_n in the sense of IFU. The definition of the g_n and the properties of the sequences now show that if $x_i = \{y_{n_{i,0}}, \ldots, y_{n_{i,n_{k_i-1}}}\}$ then $\langle z_i \rangle$ is homogeneous for f_n in the sense of IGG where $z_i = y_{n_{i,0}} * \cdots * y_{n_{i,n_{k_i-1}}}$.

As above, this proof specializes to show that FU (FU_n) implies GG (GG_n).

Finally, for the implication from IGG to IHT, just note that if we take our semigroup to be $\langle M, + \rangle$ (which is left-cancelative) and consider any $\langle f_n, k_n \rangle$ as in the hypothesis of IHT then it is also a sequence of colorings in the sense of IGG. It is then clear that any homogeneous set in the sense of IGG is also homogeneous in the sense of IHT. The same observation shows that GG (GG_n) implies HT (HT_n). $\square$

3 Extended Language

To be able to work with ultrafilters in the setting of reverse mathematics and second order arithmetic, we must extend our language $\mathcal{L}$ of second order arithmetic. The expressiveness that we need introduces, at least, a way of talking about an individual third order object such as a specified ultrafilters $\mathcal{U}$. To do this we extend $\mathcal{L}$ by adding on a unary function symbol $\delta_{\mathcal{U}}$ from sets to sets. This introduces new terms such as $\delta_{\mathcal{U}}(X)$ (or $\delta^k_{\mathcal{U}}(X)$ for an k-fold iteration with $k \in \mathbb{N}$) and so we have new atomic formulas such as $n \in \delta_{\mathcal{U}}(X)$, $Y = \delta_{\mathcal{U}}(X)$ and $\delta_{\mathcal{U}}(X) = \delta_{\mathcal{U}}(Y)$.

The structures for this language $\mathcal{L}^{\mathcal{U}}$ are formed from ones $\mathcal{M} = \langle M, S, +, \times, <, 0, 1, \in \rangle$ for $\mathcal{L}$ and a subset $\mathcal{U}$ of S to get a structure $\mathcal{M}^{\mathcal{U}} = \langle M, S, \mathcal{U}, +, \times, <, 0, 1, \in \rangle$. The semantics for $\mathcal{L}^{\mathcal{U}}$ are determined by fixing a $\mathcal{U} \subseteq S$ and interpreting $\delta_{\mathcal{U}}(X)$ as the subset of $\mathcal{M}$ given by $0 \in \delta_{\mathcal{U}}(X) \Leftrightarrow X \in \mathcal{U}$ and $n + 1 \in M \Leftrightarrow X^{[n]} \in \mathcal{U}$. Of course, we must make sure that this operation is well-defined and the the class S of subsets of M which form the sets of $\mathcal{M}$ is closed under $\delta_{\mathcal{U}}(X)$ to get a structure for $\mathcal{L}^{\mathcal{U}}$.

We can now say in $\mathcal{L}^{\mathcal{U}}$, for example, that $\mathcal{U}$ is an $\mathcal{M}$-ultrafilter or is any of a variety of special $\mathcal{M}$-ultrafilters (or a Z-ultrafilter for any $Z \in S$). We incorporate these sentences as axioms for our extensions R of theories T of second order arithmetic that essentially assert the existence of an ultrafilter (with some additional properties) to form the desired conservative extensions. The following sentences of $\mathcal{L}^{\mathcal{U}}$ say that $\mathcal{U}$ is a nonprincipal ultrafilter:

$\mathcal{AU}1$ $\forall X(0 \in \delta_{\mathcal{U}}(X) \to \forall x \exists y(y > x \;\&\; y \in X)$. (Every set in $\mathcal{U}$ is infinite.)

$\mathcal{AU}2$ $\forall X \forall Y (X \subseteq Y \ \& \ 0 \in \delta_{\mathcal{U}}(X) \to 0 \in \delta_{\mathcal{U}}(Y))$. ($\mathcal{U}$ is closed under supersets.)

$\mathcal{AU}3$ $\forall X \forall Y \forall Z (0 \in \delta_{\mathcal{U}}(X) \ \& \ 0 \in \delta_{\mathcal{U}}(Y) \ \& \ X \cap Y = Z \to 0 \in \delta_{\mathcal{U}}(Z))$. ($\mathcal{U}$ is closed under intersections.)

$\mathcal{AU}4$ $\forall X \forall Y (Y = \bar{X} \to (0 \in \delta_{\mathcal{U}}(X) \ \vee \ 0 \in \delta_{\mathcal{U}}(Y)$. (For every X, X or $\bar{X}$ is in $\mathcal{U}$.)

The next axiom guarantees that $\mathcal{U}$ satisfies the translation of the classical definition of an ultrafilter being idempotent as given after Notation 2.4.

$\mathcal{AU}5$ $\forall X \forall Y \forall Z (0 \in \delta_{\mathcal{U}}(X) \ \& \ \forall n (Y^{[n]} = X - n)) \ \& \ \forall i (i \in Z \Leftrightarrow i+1 \in \delta_{\mathcal{U}}(Y)) \to 0 \in \delta_{\mathcal{U}}(Z))$. (If $X \in \mathcal{U}$ then $\{\{n | (X-n) \in \mathcal{U}\} = \{n | \{y | \{n+y\} \in \mathcal{U}\} \in \mathcal{U}$.)

We now explain how to extend theories T of second order arithmetic.

Definition 3.1. If T is theory of second order arithmetic specified by a set of axioms as well as axiom schemes giving an axiom involving φ for each $\varphi \in \Gamma$ for some syntactically defined class Γ (in the usual way as for comprehension axioms), then $T^{I\mathcal{U}}$ ($T^{\mathcal{U}}$) is the same set of axioms as T but we replace the class Γ by the class of formulas of $\mathcal{L}^{\mathcal{U}}$ with the same syntactic specification with the understanding, of course, that formulas of the form $n \in \delta_{\mathcal{U}}(X)$, $Y = \delta_{\mathcal{U}}(x)$ and $\delta_{\mathcal{U}}(X) = \delta_{\mathcal{U}}(Y)$ are atomic. (For example, For $\mathrm{ACA}_0^{I\mathcal{U}}$ we include comprehension axioms for all arithmetic formulas in the language $\mathcal{L}^{\mathcal{U}}$.) In addition, $T^{I\mathcal{U}}$ ($T^{\mathcal{U}}$) has axioms $\mathcal{AU}1 - \mathcal{AU}5$ ($\mathcal{AU}1 - \mathcal{AU}4$).

Remark 3.2. For those familiar with the extended language of Towsner [2014], it is not hard to see that ours is at least as expressive as his. His language introduces second order terms T_t for each second order term T and each first order term t and a unary predicate $\mathcal{U}$ on second order terms written $T \in \mathcal{U}$. The intended semantics for the T_t are expressed by the added axioms saying that $n \in T_t \Leftrightarrow \langle t, n \rangle \in T$. We can convert his formulas to ours by replacing an arbitrary new atomic formula of the form $X \in \mathcal{U}$ by $0 \in \delta_{\mathcal{U}}(X)$, ones of the form $(\ldots(X_{t_1})_{t_2})_{t_3}\ldots)_{t_n} \in \mathcal{U}$ by $\langle t_1, \ldots, t_n \rangle + 1 \in \delta_{\mathcal{U}}(X)$ and ones of the form $s \in (\ldots(X_{t_1})_{t_2})_{t_3}\ldots)_{t_n}$ by $\langle t_1, \ldots, t_n, s \rangle \in X$. It is not immediately obvious that (in ACA_0) one can express formulas such as $n \in \delta^k_{\mathcal{U}}(X)$ in his language.

In the next section we will prove conservation results for $T^{I\mathcal{U}}$ over T for many of the standard systems of reverse mathematics. To see that this is meaningful in the sense of possibly showing that proofs using idempotent ultrafilters can be replaced by ones in second order arithmetic otherwise at the same level of complexity, we show how to convert the standard proof of IHT using an idempotent ultrafilters to one in $\mathrm{ACA}_0^{I\mathcal{U}}$.

We mimic the standard Galvin-Glazer proof using an idempotent ultrafilters on $\langle M, + \rangle$ (as in Todorcevic [2010, Theorem 2.20]). Like many other proofs using ultrafilters, this proof uses an inductive construction choosing sets in the ultrafilter in an inductive fashion that depends at each step on the previously chosen sets. In general this is not a procedure that can be done in $\mathrm{ACA}_0^{I\mathcal{U}}$ (even if the inductive step is arithmetic). (On its face, such a construction may use something like strong Σ_1^1-DC_0 and certainly some are provably beyond the reach of $\mathrm{ACA}_0^{I\mathcal{U}}$.) We make one simple change for this construction. Instead of choosing the successive sets in the ultrafilter, we describe the results of all possible paths through the inductive argument in advance. This allows us to work inside ACA_0 after using only one instance of our operator $\delta_{\mathcal{U}}$ instead of something that seems like a dependent sequence of choices. This type of argument will play a crucial role in the analysis of our notion of forcing and our conservation results in §4. Two different approaches to this problem appear in our treatment of another example (the Milliken-Taylor theorem) in Theorem 3.4 in $\mathrm{ACA}_0^{I\mathcal{U}}$ and in Theorem 4.16 in stronger systems but ones still weaker than Σ_1^1-DC_0.

Theorem 3.3. *$ACA_0^{I\mathcal{U}} \vdash IHT$.*

Proof. We work in $\mathrm{ACA}_0^{I\mathcal{U}}$. Let $\langle f_k, n_k\rangle$ be a sequence of colorings of M as in the statement of IHT (Theorem 2.8). We recursively define by induction on levels a tree of sequences σ labeled with sets X_σ which will be downward nested with increasing σ. We begin with the root $\emptyset$ labeled with M. Given σ of length $2k$ the immediate successors of σ are $\sigma\hat{\ }i$ for $i < n_k$ and we set $X_{\sigma\hat{\ }i} = \{x \in X_\sigma | f_k(x) = i\}$. If σ is of length $2k+1$, its immediate successors are $\sigma\hat{\ }x$ with $x \in X_\sigma$ and $X_{\sigma\hat{\ }x} = \{y | y, x+y \in X_\sigma\}$. We now let X be the code of the sets X_σ. We suppress the recursive coding of strings as numbers as usual but note that for convenience we assume that the numbers which are codes for strings σ begin with 1 rather than 0. (This allows us to avoid the notational problems engendered by reserving $\delta_{\mathcal{U}}(X)(0)$ to code the membership of X in $\mathcal{U}$ and using the numbers larger than 0 for the columns of X.)

Using $\delta_{\mathcal{U}}(X)$ and the tree above, we now arithmetically (or indeed as we shall see recursively) define a path in the tree with nodes σ_n at level n and a sequence $x_{2n+2} \in X_{\sigma_{2n+1}}$ that will satisfy the requirements for the given instance of IHT. We begin with $\sigma_0 = \emptyset$. Given σ_{2k} we let $\sigma_{2k+1} = \sigma_{2k}\hat{\ }i$ for the least $i < n_k$ such that $\sigma_{2k}\hat{\ }i \in \delta_{\mathcal{U}}(X)$ (i.e. $X_{\sigma_{2k}\hat{\ }i} \in \mathcal{U}$). Given σ_{2k+1} we let $\sigma_{2k+2} = \sigma_{2k+1}\hat{\ }x$ for the least $x > 0, x_{2k+1}$ (necessarily in $X_{\sigma_{2k+1}}$) such that $\sigma_{2k+1}\hat{\ }x \in \delta_{\mathcal{U}}(X)$ (i.e. $X_{\sigma_{2k+1}\hat{\ }x} \in \mathcal{U}$) and let $x_{2k+2} = x$. In either case, if there is no such extension the sequences terminates.

We now prove by induction on n that $\sigma_n \in \delta_{\mathcal{U}}(X)$ (i.e. $X_{\sigma_n} \in \mathcal{U}$) and the sequences of strings σ and sets X_σ never terminate. Given $\sigma_{2k} \in \delta_{\mathcal{U}}(X)$, there are only n_k many immediate successors $\sigma_{2k}\hat{\ }i$ and the corresponding $X_{\sigma_{2k}\hat{\ }i}$ partition $X_{\sigma_{2k}}$. As $\mathcal{U}$ is an ultrafilter, one of them must be in $\mathcal{U}$. More formally, one of the $\sigma_{2k}\hat{\ }i$ is in $\delta_{\mathcal{U}}(X)$ by Axiom $\mathcal{AU}3$ and so there is a least one i which gives us $\sigma_{2n+1} = \sigma_{2n}\hat{\ }i$ with $\sigma_{2n+1} \in \delta_{\mathcal{U}}(X)$ as required. (Note that arithmetic comprehension for formulas of $\mathcal{L}^{\mathcal{U}}$ and open induction easily imply that $\mathcal{U}$ is closed under $\mathcal{M}$-finite intersections.) Given $\sigma_{2k+1} \in \delta_{\mathcal{U}}(X)$, the axiom $\mathcal{AU}5$ for $\mathcal{U}$ being an idempotent ultrafilters, guarantees that for some (indeed $\mathcal{U}$ many) $x \in X_{\sigma_{2n+1}}$, $\sigma_{2n+1}\hat{\ }x \in \delta_{\mathcal{U}}(X)$ (i.e. $X_{\sigma_{2n+1}\hat{\ }x} \in \mathcal{U}$). (The axiom says that, as $X_{\sigma_{2n+1}} \in \mathcal{U}$, $\{x | X_{\sigma_{2n+1}} - x \in \mathcal{U}\} = \{x | \{y | x+y \in X_{\sigma_{2n+1}}\} \in \mathcal{U}\} \in \mathcal{U}$ and so by $\mathcal{AU}3$ $\{x | \{y \in X_{\sigma_{2n+1}} | x+y \in X_{\sigma_{2n+1}}\} \in \mathcal{U}\}\} \in \mathcal{U}$. For any x in this last set, $X_{\sigma_{2n+1}\hat{\ }x} =$

$\{y|y, x+y \in X_{\sigma_{2n+1}}\} \in \mathcal{U}$ as required.) Thus there is a least one x and $\sigma_{2n+2} = \sigma_{2n+1}\hat{\ }x$ and $x_{2n+2} = x$ for the least such x.

Finally we prove that the sequence of x_{2n+2} witnesses the desired conclusion of IHT. Consider any m. We want to prove that $FS(\{x_{2n}+2|n \geq m\}) \subseteq X_{\sigma_{2m+1}}$ and so is monochromatic for f_m as if $\sigma_{2m\hat{\ }i} = \sigma_{2m+1}$, $f_m(x) = i$ for all $x \in X_{\sigma_{2m+1}}$ by the definition of σ_{2m+1}. We claim that $FS(\{x_{2n+2}|n \geq m\})$ is contained in $X_{\sigma_{2m+1}}$ and so monochromatic to i for f_m. Consider any sum $x = x_{2n_0+2} + x_{2n_1+2} + \ldots + x_{2n_k+2}$ with $m \leq n_0 < \ldots < n_k$. We prove by induction on k that $x \in X_{\sigma_{2n_0+1}}$ and so $x \in X_{\sigma_{2m+1}}$ as $n_0 \geq m$ and the X_{σ_n} are downward nested. We have the base case $k = 0$ by the argument above taking $m = n_0$. Consider now the case that $k > 0$ and let $z = x_{2n_1+2} + \ldots + x_{2n_k+2}$. By induction $z \in X_{\sigma_{2n_1+2}} \subseteq X_{\sigma_{2n_0+1}\hat{\ }x_{2n_0+2}} = \{y|y, x_{2n_0+2}+y \in X_{\sigma_{2n_0+1}}\}$. Thus $x = x_{2n_0+2}+z \in X_{\sigma_{2n_0+1}}$ as required. □

As another example of a common ultrafilter proof of a standard combinatorial principle that can be done in $\mathrm{ACA}_0^{I\mathcal{U}}$, we consider the proof of the Milliken-Taylor theorem in Bergelson and Hindman [1989, Corollary 2.4]. This theorem is a combination of Ramsey's theorem for r-colorings of size k sets of natural numbers and Hindman's theorem for finite sums.

Theorem 3.4 (Milliken-Taylor). *Let $k, r \in \mathbb{N}$ and let $[\mathbb{N}]^k = \cup\{A_i|i < r\}$. Then there exists an $i < r$ and an increasing sequence $\langle x_i \rangle$ such that whenever $-1 = n(0) < n(1) < \cdots < n(k)$ and, for $1 \leq t \leq k$, $a_t \in FS(\langle x_m|n(t-1)+1 \leq m \leq n(t)\rangle)$, $\{a_1, \ldots, a_k\} \in A_i$.*

Proof (in $ACA_0^{I\mathcal{U}}$). We denote by $\mathrm{MT}(k)$ the assertion of the theorem for fixed k (but for all r). We indicate how to read the proof in Bergelson and Hindman [1989] as one for a model $\mathcal{M}^{\mathcal{U}}$ of $\mathrm{ACA}_0^{I\mathcal{U}}$. Note that while we may take r to be any number in the model, we restrict ourselves to proving $\mathrm{MT}(k)$ for standard k so that we can handle the backward induction on k in Bergelson and Hindman [1989], Theorem 2.3] in $\mathrm{ACA}_0^{I\mathcal{U}}$. We fix both $k \in \mathbb{N}$ and $r \in M$ for the argument.

The basic tool in the proof is a form of Ramsey's theorem for k-tuples (and r colors) where the homogeneous set is constructed from a descending sequence D_n of sets in the given ultrafilter $\mathcal{U}$ (called p in their notation) with certain extension properties. As Bergelson and Hindman point out, the proof of their Theorem 2.3 does not use the idempotence of the ultrafilter and so we carry it out in $\mathrm{ACA}_0^{\mathcal{U}}$. The main issue from the viewpoint of $\mathrm{ACA}_0^{\mathcal{U}}$ is the construction of the sets $B_t(E, i)$ for $t \leq k$, $i < r$ and $E \in [M]^{t-1}$. The definition begins innocuously enough with $B_k(E, i) = \{y \in M \backslash E|\ E \cup \{y\} \in A_i\}$. It then proceeds by backward induction on $t < k$ by setting $B_t(E, i) = \{y \in M \backslash E|\ B_{t+1}(E \cup \{y\}, i) \in \mathcal{U}\}$. Assuming we have (a real in our model coding) which of all the $B_{t+1}(E, i)$ for $E \in [M]^t$ are in $\mathcal{U}$ (which we certainly do for k itself by $\mathrm{ACA}_0^{\mathcal{U}}$), we can clearly define (a real coding) all the $B_t(E, i)$ and so there would be a real in our model coding which of all the $B_t(E, i)$ are in $\mathcal{U}$ again by $\mathrm{ACA}_0^{\mathcal{U}}$. Thus as long as k is standard we can show (by an external induction through k many steps) that there is a real in the model coding all the $B_t(E,i)$ for $t \leq k$ and which of them are in $\mathcal{U}$. Equipped with this real as a parameter the construction of the D_n in Bergelson and Hindman [1989, Theorem 2.3] is carried out routinely in $\mathrm{ACA}_0^{\mathcal{U}}$ using just the properties for $\mathcal{U}$ being an ultrafilter.

Bergelson and Hindman [1989] then deduce our Theorem 3.4 as their Corollary 2.4 using the fact that $\mathcal{U}$ is idempotent. The procedure is much like the Galvin-Glazer proof of Hindman's theorem and can be modified to fit $\mathrm{ACA}_0^{I\mathcal{U}}$ in the same way as in our proof of Theorem 3.3, by building a tree of all possible choices of the x_i and corresponding sets C_i with $C = D_0$, $x_0 \in C_0$ and $C_0 - x_0 \in \mathcal{U}$. Inductively if we have the sequences up to C_n we consider the $x \in C_n$ along with $C_n - x$ and intend to choose as x_{n+1} an x larger than $l(n) =$ the sum of all the previous x_i with $C_n - x \in \mathcal{U}$ and set $C_{n+1} = C_n \cap C_n - x \cap D_{l(n+1)}$. As before, we can recursively lay out the tree of all possible constructions with a single real coding all the possible choices of the x_n and C_n along each path. Asking the appropriate array of questions about membership in $\mathcal{U}$ then produces a real which guides the choice of an infinite path in this tree to construct a specific sequences which has all the desired sets in $\mathcal{U}$. (The proof of nontermination is just the proof that the inductive step in Bergelson and Hindman [1989, Corollary 2.4] works, i.e. that $\mathcal{U}$ is idempotent.) Armed with this sequence the proof that $\langle x_n \rangle$ is as desired, is as they say routine and we note that it works in $\mathrm{ACA}_0^{I\mathcal{U}}$ without any concerns . $\square$

We note that essentially the same ultrafilter proof of Theorem 3.4 appears in Hindman and Strauss [2012, Ch. 18] along with two other versions of the theorem bearing the same relationships to the Milliken-Taylor theorem that FU and GG bear to HT. The derivations of these results there from the Milliken-Taylor Theorem are purely combinatorial and work in ACA_0. Thus they too are consequences of IHT. Thus all of these theorems are provable in IHT and so ACA_0^+.

We also note that Hirst [2004] has shown this result for MT(k) for each $k \in \mathbb{N}$ by a much more direct proof. We discuss the proof theoretic strengths of $\forall k MT(k)$ in the explication of Theorem 4.16 after the analysis of conservation results over ATR_0 (Theorem 4.14 and Corollary 4.15) and its recursion theoretic strength in the appendix after Corollary 6.2.

4 Notions of Forcing

We assume from now on that all structures $\mathcal{M}$ for second order arithmetic are models of ACA_0. We use the language of forcing to construct the models of $T^{I\mathcal{U}}$ (for various theories T that include at least IHT and so ACA_0) but need very little of even the basic lemmas on forcing for the theories considered in Towsner [2014].

Recall that our language is that of second order arithmetic augmented by a unary function symbol $\delta_{\mathcal{U}}$ taking sets to sets. Our goal is, given a model $\mathcal{M}$ of T, to build a subset $\mathcal{U}$ of the "sets" S of $\mathcal{M}$ and to interpret $\delta_{\mathcal{U}}(X)$ so that $0 \in \delta_{\mathcal{U}}(X) \Leftrightarrow X \in \mathcal{U}$ and $n+1 \in M \Leftrightarrow X^{[n]} \in \mathcal{U}$. We define a notion of forcing $\Vdash_{I\mathcal{U}}$ with a definition of the forcing relation which differs slightly from the usual one. It is defined globally and directly for all arithmetic sentences to correspond to being true for every sufficiently generic filter extending the condition. We take $\mathcal{U}$ to be canonically defined from a generic (over $\mathcal{M}$) filter for this forcing.

Definition 4.1. Our notion $\mathbb{P}_{I\mathcal{U}}$ of forcing (really this should be $\mathbb{P}_{I\mathcal{U}}^{\mathcal{M}}$ but we omit the superscript when no confusion should arise) has as conditions sequences u in $\mathcal{M}$ of the form $\langle U_0, U_1, \ldots, U_i, \ldots \rangle_{i \in M}$ such that there is a sequence $y_0 < y_1 < y_2 < \ldots$ (in $\mathcal{M}$) such that, for every i, $U_i = FS(y_i, y_{i+1}, \ldots)$. We define extension for $\mathbb{P}_{I\mathcal{U}}$ by $v = \langle V_i \rangle \le \langle U_i \rangle = u \Leftrightarrow \forall i \exists j (U_i \supseteq V_j)$.

Remark 4.2. This definition clearly implies that, if $u = \langle U_0, U_1, \ldots \rangle \in \mathbb{P}_{I\mathcal{U}}$ then, for each i, U_i is infinite, $U_i \supseteq U_{i+1}$ and $\cap U_i = \emptyset$. Moreover, we can uniformly arithmetically translate between the representations of u as $\langle U_i \rangle$ and $\langle y_i \rangle$: Clearly, U_i is uniformly arithmetic in $\langle y_i \rangle$ as $U_i = FS(y_i, y_{i+1}, \ldots)$. In the other direction y_i is the least element of $U_i - U_{i+1}$. We are thinking of u as representing the filter $F_u = \{A \in S | \exists i (A \supseteq U_i)\}$. It is easy to see that this set is an $\mathcal{M}$-filter for any $u \in \mathbb{P}_{I\mathcal{U}}$ and, for example, if $v \le u$ then $F_v \supseteq F_u$. So we abuse notation by using $A \in u$ (for $A \in S$ and u a condition in $\mathbb{P}_{I\mathcal{U}}$) to mean that $\exists i (A \supseteq U_i)$. Given any ($\mathcal{M}$-generic) filter $\mathbb{U}$ on $\mathbb{P}_{I\mathcal{U}}$ (i.e. a filter on $\mathbb{P}_{I\mathcal{U}}$ which meets all dense sets definable in $\mathcal{M}$), the corresponding ($\mathcal{M}$-generic) object $\mathcal{U}$ is $\{A \in S | (\exists u \in \mathbb{U})(A \in u)\}$. Note that both membership and extension in $\mathbb{P}_{I\mathcal{U}}$ are arithmetic relations in $\mathcal{M}$.

Lemma 4.3. *If $\langle u_i \rangle$ is a descending sequence in $\mathbb{P}_{I\mathcal{U}}$ in $\mathcal{M}$ then there is a condition u extending every u_i.*

Proof. Let the sequences $y_{i,j}$ and $U_{i,j}$ be associated with u_i as in Definition 4.1. By Remark 4.2 and the definition of extension there is a function h defined by the following recursion:

$h(0) = 0$

$h(i+1) = \mu k (k > h(i) \;\&\; (\forall n, m \le h(i))(U_{n,m} \supseteq U_{h(i),k})$.

We now let $z_i = y_{h(i),h(i+1)}$ and define v to be the associated condition. We claim that $v \le u_i$ for every i. Indeed, for any $U_{n,m}$ choose an i such that $n, m < h(i)$ and note that, by induction on j, for every $j > i$, $U_{n,m} \supseteq U_{h(j),h(j+1)}$ and so $z_j = y_{h(j),h(j+1)} \in U_{n,m}$ and $U_{n,m} \supseteq V_j$ as required. □

To get the forcing relation started (for arithmetic sentences of $\mathcal{L}^{\mathcal{U}}$ with set parameters) we begin with the idea of deciding a set X.

Definition 4.4. A condition $u \in \mathbb{P}_{I\mathcal{U}}$ *decides a set* $X \in S$ if $(X \in u \vee \bar{X} \in u)$ and $\forall n (X^{[n]} \in u \vee \overline{X^{[n]}} \in u)$.

Proposition 4.5. *If u decides X, then there is a set $Y \in S$ such that $0 \in Y \Leftrightarrow X \in u$ and $n+1 \in Y \Leftrightarrow X^{[n]} \in u$. We then say that $u \Vdash \delta_{\mathcal{U}}(X) = Y$. The relation u decides*

X, the set Y such that $u \Vdash \delta_{\mathcal{U}}(X) = Y$ and the relation $u \Vdash \delta_{\mathcal{U}}(X) = Y$ are uniformly arithmetic (in u, X and Y as relevant).

Proof. That the relation u decides X is uniformly arithmetic is immediate from the definition of deciding. (Note, for example that, $X^{[n]} \in u$ means that $(\exists i)(\forall m)(m \in U_i \rightarrow \langle n, m\rangle \in X.)$ The first assertion then follows immediately from the definitions and the fact that $\mathcal{M} \vDash ACA_0$. The other assertions then follow from the definitions. $\square$

Definition 4.6. As a matter of notation at this point, we let $\delta^0_{\mathcal{U}}(X) = X$ and $\delta^{n+1}_{\mathcal{U}}(X) = \delta_{\mathcal{U}}(\delta^n_{\mathcal{U}}(X))$. We say u *decides* $\delta^{n+1}_{\mathcal{U}}(X)$ if u decides Y for the Y such that $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$ where we are also defining $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$ by the obvious induction.

Proposition 4.7. *By iterating the previous Proposition, the relations u decides $\delta^{n+1}_{\mathcal{U}}(X)$, $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$ and the set $Y \in S$ such that $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$ are uniformly arithmetic in the sense that there is a recursive function that, for each n, supplies arithmetic formulas φ^n_1, φ^n_2 and φ^n_3 (whose complexity depends on n) such that $\varphi^n_1(u, X)$ holds if and only if u decides $\delta^{n+1}_{\mathcal{U}}(X)$; $\varphi^n_2(u, X, Y)$ holds if and only if $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$ and $\varphi^n_3(u, X, m)$ holds if and only if m is a member of the set Y such that $\varphi^n_2(u, X, Y)$, i.e. $u \Vdash \delta^n_{\mathcal{U}}(X) = Y$.*

We define the forcing relation between conditions and sentences of $\mathcal{L}^{\mathcal{U}}$ for arithmetic sentences in terms of these relations.

Definition 4.8. Let $\Phi(Z_1, \dots Z_m)$ be an arithmetic formula of $\mathcal{L}^{\mathcal{U}}$ with the free second order variables displayed. Consider an instance of $\Phi(X_1, \dots X_m)$ specified by choosing values $X_i \in S$ for the Z_i and a condition u. We say that $u \Vdash \Phi(X_1, \dots X_m)$ if u decides every $\delta^{n_j}_{\mathcal{U}}(X_i)$ occurring in Φ and $\mathcal{M}$ satisfies the sentence gotten from $\Phi(X_1, \dots X_m)$ by substituting the formula saying that $k \in Y_{i,j}$ for the $Y_{i,j}$ such that $u \Vdash \delta^{n_j}_{\mathcal{U}}(X_i) = Y_{i,j}$ for the new atomic formulas $k \in \delta^{n_j}_{\mathcal{U}}(X_i)$, $X_i = Y_{j,k}$ for $X_i = \delta^{n_k}_{\mathcal{U}}(X_j)$ and $Y_{i,j} = Y_{l,k}$ for $\delta^{n_j}_{\mathcal{U}}(X_i) = \delta^{n_k}_{\mathcal{U}}(X_l)$. We now proceed by induction to define forcing for second order sentences of $\mathcal{L}^{\mathcal{U}}$ (which we assume to be in prenex normal form) in the obvious way. We say $u \Vdash \exists Z \Phi((X_1, \dots X_m, Z)$ if there is an $X_{m+1} \in S$ such that $u \Vdash \Phi((X_1, \dots X_m, X_{m+1})$. On the $\forall$ side we say $u \Vdash \forall Z \Phi((X_1, \dots X_m, Z)$ if there is no $v \leq u$ and no $X_{m+1} \in S$ such that $v \Vdash \neg\Phi((X_1, \dots X_m, X_{m+1})$ where we use $\neg\Phi$ as an abbreviation of the standard prenex normal form equivalent of the negation of Φ.

Proposition 4.9. *For arithmetic sentences $\Phi(X_1, \dots X_m)$ of $\mathcal{L}^{\mathcal{U}}$, the relation $u \Vdash \Phi(X_1, \dots X_m)$ is uniformly arithmetic in u and $X_1, \dots X_m$. For Σ^1_n (Π^1_n) sentences $\Phi(X_1, \dots X_m)$ of $\mathcal{L}^{\mathcal{U}}$ the relation $u \Vdash \Phi(X_1, \dots X_m)$ is uniformly Σ^1_n (Π^1_n).*

Proof. The arithmetic case follows directly from the definition of forcing and Proposition 4.7. The second order cases then follow by the obvious induction. $\square$

Proposition 4.10 (IHT). *Assume $\mathcal{M} \vDash IHT$. For each sequence $\langle X_m\rangle \in S$, the set $\{u | \forall m(u \text{ decides } X_m)\}$ is dense in $\mathbb{P}_{I\mathcal{U}}$. For each sentence Φ of $\mathcal{L}^{\mathcal{U}}$ (with set parameters), the set $\{u | u \text{ decides } \Phi\}$ is also dense in $\mathbb{P}_{I\mathcal{U}}$. (We say u decides Φ if $u \Vdash \Phi$ or $u \Vdash \neg\Phi$ as usual.)*

Proof. Consider any $\langle X_m \rangle \in S$ and $v \in \mathbb{P}_{I\mathcal{U}}$ generated by the sequence $\langle z_i \rangle$ as in Definition 4.1. By Lemma 2.11 we may, by a refinement gotten by replacing z_i by sums of z_j for $j \geq i$, assume that there is an increasing sequence s_i such that $2^{s_i} | z_i$ but $2^{s_i+1} \nmid z_i$. So for $z = z_{i_0} + \ldots + z_{i_n}$ with $i_0 < \ldots < i_n$, $z \in V_{i_0} - V_{i_0+1}$ and i_0 is the maximum j such that $z \in V_j$. Now consider the subsemigroup Z of M generated by the z_i and apply IGG for the sequence of colorings given by $f_{\langle 0,m \rangle}(x) = 1$ if $x \in X_m$, $f_{\langle 0,m \rangle}(x) = 0$ if $x \notin X_m$, $f_{\langle n+1,m \rangle}(x) = 1$ if $x \in X_m^{[n]}$ and $f_{\langle n+1,m \rangle}(x) = 0$ if $x \notin X_m^{[n]}$ to get a homogeneous subset q_i of Z. We may now refine the condition w generated by the q_i (which are themselves finite sums of z_i) as before to get one u generated by y_i with an increasing sequence $t_i > s_i$ such that $2^{t_i} | y_i$ but $2^{t_i+1} \nmid y_i$ but that is still homogeneous in the sense of IGG. Thus for each y_i there is a maximal j such that $y_i \in V_j$. (As with the sum z above, this j is determined by the positions of the t_k with respect to the s_l.) For this j, $U_i \supseteq V_j$ as required for u to be an extension of v. The IGG homogeneity of the sequence y_i guarantees that for each m, $X_m \supseteq U_i$ or $\overline{X_m} \supseteq U_i$ and for each $\langle n+1, m \rangle$, $X_m^{[n]} \supseteq U_i$ or $\overline{X_m^{[n]}} \supseteq U_i$ for some i as required.

The second assertion now follows easily from the definitions of deciding sentences Φ and terms $\delta^n_{\mathcal{U}}(X)$ by induction on both the complexity of Φ and the depth of $\delta_{\mathcal{U}}$ terms occurring in Φ. □

Theorem 4.11. *If $\mathcal{M} \models IHT$ and both $\mathbb{U}$ and the associated $\mathcal{U}$ are $\mathcal{M}$-generic for $\mathbb{P}_{I\mathcal{U}}$, then $\mathcal{M}^{\mathcal{U}}$ is an $\mathcal{L}^{\mathcal{U}}$ structure and a model of $ACA_0^{I\mathcal{U}}$. Moreover, any sentence Φ of $\mathcal{L}^{\mathcal{U}}$ (with second order parameters) is true in $\mathcal{M}^{\mathcal{U}}$ if and only if there is a $u \in \mathbb{U}$ such that $u \Vdash \Phi$. Finally, $\mathcal{U}$ is an idempotent $\mathcal{M}$-ultrafilter.*

Proof. Consider any $\mathcal{M}$-generic $\mathcal{U}$ and any $X \in S$. By Proposition 4.10, there is a $u \in \mathbb{U}$ (which decides X) and a (necessarily unique) $Y \in S$ such that $u \Vdash \delta_{\mathcal{U}}(X) = Y$. Moreover, for any $v \in \mathcal{U}$ and $Z \in S$ such that $v \Vdash \delta_{\mathcal{U}}(X) = Z$, $Y = Z$. If not, then, without loss of generality, there is some n such that $X^{[n]} \in u$ but $X^{[n]} \notin v$ (or $X \in u$ but $X \notin v$) and so $\overline{X^{[n]}} \in v$ ($\bar{X} \in v$). So there are i and j such that $X^{[n]} \supseteq U_i$ but $\overline{X^{[n]}} \supseteq V_j$ (or $X \supseteq U_i$ but $\overline{X} \supseteq V_j$). This would contradict the compatibility of all $u, v \in \mathbb{U}$ as any $w \leq u, v$ would have some $W_k \subseteq U_i \subseteq X^{[n]}$ (or X) and some $W_l \subseteq V_j \subseteq \overline{X^{[n]}}$ (or $\bar{X}$) and so $W_k \cap W_l = \emptyset$. Thus there is a well-defined function specified on S by $\delta_{\mathcal{U}}(X) = Y$ where $u \Vdash \delta_{\mathcal{U}}(X) = Y$ for some $u \in \mathbb{U}$. This function makes $\mathcal{M}^{\mathcal{U}}$ into an $\mathcal{L}^{\mathcal{U}}$ structure as desired.

Next consider any arithmetic sentence $\Phi(Y_1, \ldots Y_k)$ of $\mathcal{L}^{\mathcal{U}}$ with second order parameters displayed. By Proposition 4.10, there is a $u \in \mathbb{U}$ which decides $\Phi(Y_1, \ldots Y_k)$. It is clear from the definition of forcing and our interpretation of $\delta_{\mathcal{U}}(X)$ in $\mathcal{M}^{\mathcal{U}}$ that $\mathcal{M}^{\mathcal{U}} \models \Phi$ if $u \Vdash \Phi$ and $\mathcal{M}^{\mathcal{U}} \models \neg\Phi$ if $u \Vdash \neg\Phi$. It is now an easy induction to show that every sentence Φ of $\mathcal{L}^{\mathcal{U}}$ is true in $\mathcal{M}^{\mathcal{U}}$ if and only if it is forced by some $u \in \mathbb{U}$.

Now consider any instance of a comprehension axiom of $\mathrm{ACA}_0^{I\mathcal{U}}$ given by $\Phi(Y_1, \ldots Y_k, x)$ with Φ as above except that it now has one free number variable x. By Proposition 4.10, there is a $u \in \mathbb{U}$ which decides every $\delta^n_{\mathcal{U}}(Y_i)$ occurring in Φ and so for each a, u forces

$\Phi(Y_1, \dots Y_k, a)$ or its negation and the corresponding sentence is true in $\mathcal{M}^{\mathcal{U}}$. As u forcing $\Phi(Y_1, \dots Y_k, a)$ and forcing its negation are arithmetic relations in $\mathcal{M}$, the set of $a \in \mathcal{M}$ such that $\mathcal{M}^{\mathcal{U}} \vDash \Phi(Y_1, \dots Y_k, a)$ is arithmetically definable in $\mathcal{M}$ (using u) and so a member of S as required (as $\mathcal{M} \vDash ACA_0$).

Finally, we show that $\mathcal{M}^{\mathcal{U}}$ satisfies the special axioms for $\mathcal{U}$ being an idempotent ultrafilter. If $X \in \mathcal{U}$, i.e. $0 \in \delta_{\mathcal{U}}(X)$, then by definition $X \in u$ for some $u \in \mathbb{U}$, i.e. $X \supseteq U_i$ for some i and so is infinite by Definition 4.1 and $\mathcal{M}^{\mathcal{U}}$ satisfies $\mathcal{AU}1$. If $Y \supseteq X \in \mathcal{U}$ then for some $u \in \mathbb{U}$, $Y \supseteq X \supseteq U_i$ and so $Y \in \mathcal{U}$ and $\mathcal{M}^{\mathcal{U}}$ satisfies $\mathcal{AU}2$. If $X, Y \in \mathcal{U}$ then there are $u, v \in \mathbb{U}$ and i and j such that $X \supseteq U_i$ and $Y \supseteq V_j$. As $\mathbb{U}$ is a generic filter it contains a condition $w \leq u, v$. The definition of extension provides k and l such that $U_i \supseteq W_k$ and $V_j \supseteq W_l$. Now clearly, $X \cap Y \supseteq W_k \cap W_l = W_{\max\{k.l\}}$ by Definition 4.1. So $X \cap Y \in \mathcal{U}$ and $\mathcal{M}^{\mathcal{U}}$ satisfies $\mathcal{AU}3$. Next for any X there is a $u \in \mathbb{U}$ which decides X and so an i such that $X \supseteq U_i$ or $\bar{X} \supseteq U_i$, i.e. $X \in \mathcal{U}$ or $\bar{X} \in \mathcal{U}$ and $\mathcal{M}^{\mathcal{U}}$ satisfies $\mathcal{AU}4$.

For the last axiom, $\mathcal{AU}5$, suppose $X \in \mathcal{U}$ and so $X \supseteq U_i$ for some $u \in \mathbb{U}$ (determined by $\langle y_i \rangle$). We also suppose that $\forall n(Y^{[n]} = X - n))$ and $\forall i(i \in Z \Leftrightarrow i + 1 \in \delta_{\mathcal{U}}(Y))$. By Definition 4.1, $U_i = FS\{y_j | j \geq i\}$ and if we take any $z \in U_i$ with $z = y_{i_0} + \cdots + y_{i_m}$ $(i \leq i_0 < \cdots < i_m)$ then $X - z \supseteq U_i - z \supseteq U_j$ for any $j > i_m$ and so $X - z \in \mathcal{U}$ for every $z \in U_{i_m+1}$ and we satisfy $\mathcal{AU}5$. □

Corollary 4.12. *$ACA_0^{I\mathcal{U}}$ is a conservative extension of IHT for all sentences of second order arithmetic.*

Proof. Note that by Theorem 3.3, $ACA_0^{I\mathcal{U}} \vdash IHT$. On the other hand, suppose $ACA_0^{I\mathcal{U}} \vdash \Phi$. Theorem 4.11 shows that every model of IHT can be extended to one $\mathcal{M}^{\mathcal{U}}$ of $\mathrm{ACA}_0^{I\mathcal{U}}$ with the same M and S. So, if for any sentence Φ of second order logic there is an $\mathcal{M} \vDash IHT \ \& \ \neg\Phi$, then for any $\mathcal{M}$-generic $\mathcal{U}$, $\mathcal{M}^{\mathcal{U}} \vDash \neg\Phi \ \& \ \mathrm{ACA}_0^{I\mathcal{U}}$ for the desired contradiction. □

As noted in Towsner [2014] for his language in the context of $\mathrm{ACA}_0^{\mathcal{U}}$, it does not make sense to consider weaker systems than ACA_0 as $\mathrm{ACA}_0^{I\mathcal{U}}$ is a conservative extension of $\mathrm{RCA}_0^{I\mathcal{U}}$ for all sentences of second order arithmetic. Indeed the axioms of the apparently stronger theory are provable in the weaker. (He says this follows from results of Kirby [1984].) We give a easy proof.

Proposition 4.13. *$RCA_0^{\mathcal{U}} \vdash ACA_0^{\mathcal{U}}$.*

Proof. We consider definitions of sets by arithmetic formulas in prenex normal form $\{n | \varphi(n, \bar{x}, \bar{W})\}$ (with additional free number and set variables $\bar{x}$, $\bar{W}$). We prove by (external) induction on the number of (first order) quantifiers in $\varphi \in \mathcal{L}^{\mathcal{U}}$ that a set as defined by the formula exists, i.e. $\forall \bar{W} \forall \bar{x} \exists Y(n \in Y \leftrightarrow \varphi(n, \bar{x}, \bar{W}))$. (As usual we can suppress the $\bar{W}$ which gets carried along for the ride in the proof.) Consider the case that φ has one existential quantifier $\varphi = \exists x_1 \varphi_1(n, x_1, x_2, \dots, x_k)$. In $\mathrm{RCA}_0^{\mathcal{U}}$ we see that there is a set Z such that $Z^{[\langle n,x_2,\ldots,x_k\rangle]} = \{s|(\exists x_1 < s)(\varphi_1(n,x_1,x_2,\ldots,x_k)$. It is clear that each such column of Z is either empty or cofinite. Moreover, for every $\langle x_2,\ldots,x_k\rangle$, the set $\{n|\,\langle n,x_2,\ldots,x_k\rangle + 1 \in \delta_{\mathcal{U}}((Z)\}$ exists (by $\mathrm{RCA}_0^{\mathcal{U}}$) and supplies the witnesses needed for this φ. Indeed one has Y_1 with $Y_1^{\langle x_2,\ldots,x_k\rangle} = \{n|\exists x_1 < t)(\varphi_1(n,x_1,x_2,\ldots,x_k)\}$. We can now prove the existence of Y_2 with $Y_1^{\langle x_3,\ldots,x_k\rangle} = \{n|\forall x_2 \exists x_1 < t)(\varphi_1(n,x_1,x_2,\ldots,x_k)\}$ in the same way using Y_1 as an additional parameter and then proceed by induction. $\square$

On the other hand, it makes perfect sense to ask about theories stronger than ACA_0.

Theorem 4.14. *If T is any one of ACA_0^+, ATR_0, Π_1^1-CA_0 or Π_2^1-CA_0, $\mathcal{M} \vDash T$ and $\mathcal{U}$ is $\mathcal{M}$-generic for $\mathbb{P}_{I\mathcal{U}}$ then $\mathcal{M}^{\mathcal{U}} \vDash T^{I\mathcal{U}}$.*

Proof. ACA_0^+: The desired result follows immediately from Theorem 4.11 and the definition of ACA_0^+: Every $X \in \mathcal{M}^{\mathcal{U}}$ is in $\mathcal{M}$ by the definition of $\mathcal{M}^{\mathcal{U}}$. If $\mathcal{M} \vDash ACA_0^+$ then there is a (unique) set Y in $\mathcal{M}$ satisfying (in $\mathcal{M}$) the implicit arithmetic definition of Y from X given by the defining axiom of ACA_0^+. Again, by the definition of $\mathcal{M}^{\mathcal{U}}$, this Y is in $\mathcal{M}^{\mathcal{U}}$. As the numbers of $\mathcal{M}^{\mathcal{U}}$ are the same as those of $\mathcal{M}$, X and Y satisfy the implicit definition of the defining axiom in $\mathcal{M}^{\mathcal{U}}$ as required.

ATR_0: We just have to prove the comprehension axioms for $\mathrm{ATR}_0^{I\mathcal{U}}$ and so, for each instance of the axiom and each condition u (forcing the hypotheses of the axiom) we need an extension v of u which forces the conclusion. Note that a set codes a well ordering in $\mathcal{M}$ if and only if it codes one in $\mathcal{M}^{\mathcal{U}}$ by Theorem 4.11. Suppose we are given an instance of the axioms specified by a well ordering $\langle X, <_X\rangle$ in $\mathcal{M}^{\mathcal{U}}$ and an arithmetic formulas $\Phi(z,Z)$ of $\mathcal{L}^{\mathcal{U}}$ possibly with number and set parameters. For convenience we assume that $\langle X, <_X\rangle$ is one of the set parameters and x one of the number parameters.

The given instance of $\mathrm{ATR}_0^{I\mathcal{U}}$ is then specified by an arithmetic formula $\Theta(n,Z)$ of $\mathcal{L}^{\mathcal{U}}$ containing these parameters with n and Z free. We want a set Y such that, for each $x \in X$, $\forall n(n \in Y^{[x]} \Leftrightarrow \Theta(n, Y^{[<x]})$. The basic idea is that we want to prove the existence of a set $\langle Y, Z\rangle$ such that, for each $x \in X$, $Z^{[x]} \Vdash \forall n(n \in Y^{[x]} \Leftrightarrow \Theta(n, Y^{[<x]}))$ and, for $x <_X y$, $Z^{[y]} \le Z^{[x]} \le u$. Our plan is to expand $<_X$ by inserting additional steps so that a fixed $\hat{\Theta}$ will define the desired sequence of pairs. For each set $\left\langle Y^{[<x]}, Z^{[<x]}\right\rangle$, and $x \in X$ we know that there is a $\left\langle \hat{Y}, \hat{Z}\right\rangle$ with $\left\langle \hat{Y}^{[<x]}, \hat{Z}^{[<x]}\right\rangle = \left\langle Y^{[<x]}, Z^{[<x]}\right\rangle$ such that $Z^{[x]} \Vdash \forall n(n \in \hat{Y}^{[x]} \Leftrightarrow \Theta(n, \hat{Y}^{[<x]}))$ and $Z^{[x]} \le Z^{[y]}$ for every $y < x$. (For notational simplicity, we are assuming that $\langle Y, Z\rangle^{[x]} = \left\langle Y^{[x]}, Z^{[x]}\right\rangle$.) We want to specify a well-ordered sequence of arithmetic operators that will produce such a result. One could do this by hand but we use a metatheorem that relies on the fact that this Π_2^1 assertion is provable in IHT and so in ACA_0^+ (Blass, Hirst and Simpson [1987,4.13]). Theorem 6.1 says that there are $e, k \in \mathbb{N}$ such that $\forall X \forall x \forall \langle Y, Z\rangle \exists \left\langle \hat{Y}, \hat{Z}\right\rangle ((\left\langle \hat{Y}^{[<x]}, \hat{Z}^{[<x]}\right\rangle = \left\langle Y^{[<x]}, Z^{[<x]}\right\rangle$ & $Z^{[x]} \Vdash \forall n(n \in \hat{Y}^{[x]} \Leftrightarrow \Theta(n, \hat{Y}^{[<x]}))$ &$(\forall y <_X x)(Z^{[x]} \le Z^{[y]})$ & $\left\langle \hat{Y}, \hat{Z}\right\rangle = \Phi_e(\left\langle \hat{Y}^{[<x]}, \hat{Z}^{[<x]}, \langle X, <_X\rangle, x\right\rangle^{(\omega\cdot k)})$. So we can form a new well ordering which inserts $\omega \cdot k$ steps between each $x \in X$ and its immediate successor $x+1$ each of which produces the Turing jump of its input and at the step at the end corresponding to $x+1$ in $<_X$ applies Φ_e to $\left\langle \hat{Y}^{[<x]}, \hat{Z}^{[<x]} \langle X, <_X \rangle, x \right\rangle^{(\omega \cdot k)}$ which is arithmetic in the new sequence inserted between x and $x+1$. (In the usual way, we are using $Z^{(\omega \cdot k)}$ to denote the result of iterating the Turing jump $\omega \cdot k$ many times beginning with Z.)

The fact that the relations of extension and forcing the relevant sentences are arithmetic follow from Remark 4.2 and Proposition 4.9. The existence of some $\left\langle Y^{[x]}, Z^{[x]} \right\rangle$ as required follows from Proposition 4.10 at successor steps in $<_X$. For x a limit point, one would first arithmetically define a strictly ascending sequence in $<_X$ with limit x and apply Lemma 4.3 before using Proposition 4.10. Thus the desired Π^1_2 assertion is provable in IHT and so ACA_0^+ as required for Theorem 6.1

Now given the $\langle Y, Z \rangle$ witnessing the truth of this new instance of ATR_0 in $\mathcal{M}$, we can define $\hat{Y}$ by the arithmetic formula $n \in \hat{Y}^{[x]} \Leftrightarrow Z^{[x]} \Vdash \Theta(n, \hat{Y}^{[<x]})$. A standard transfinite induction together with the representation of truth in $\mathcal{M}^{\mathcal{U}}$ by forcing conditions in $\mathbb{P}_{I\mathcal{U}}$ now shows that the $\hat{Y}$ so defined is the one required for the instance of $\mathrm{ATR}_0^{I\mathcal{U}}$ with which we began and one more application of Lemma 4.3 gives the $z \leq Z^{[x]} \leq u$ for every $x \in X$ that forces everything and gives the density result needed to show that there is a $\hat{Y}$ as required in every $\mathcal{M}^{\mathcal{U}}$.

Π^1_1-CA_0: We prove Σ^1_1-CA_0 which is obviously equivalent to Π^1_1-CA_0. Suppose $\exists X \Phi(X, i)$ is a Σ^1_1 formula of $\mathcal{M}^{\mathcal{U}}$. As $u \Vdash \Phi(Y, n)$ and being an extension in $\mathbb{P}_{I\mathcal{U}}$ are arithmetic predicates (of Y, u and n), we can apply strong Σ^1_1-DC_0 (which is a consequence of Π^1_1-CA_0) to get a sequence $\langle Y_i, u_i \rangle$ with $u_0 \in \mathbb{P}_{I\mathcal{U}}$ arbitrary, $u_{i+1} \leq u_i$ and $u_{i+1} \Vdash \Phi(Y_i, i)$ if there is any $\langle Y, v \rangle$ with $v \leq u_i$ such that $v \Vdash \Phi(Y, i)$. By Lemma 4.3, there is a condition u extending all the u_i. By genericity, then there is such a $u \in U$. We claim that $\mathcal{M}^{\mathcal{U}} \vDash \exists X \Phi(X, i) \Leftrightarrow u \Vdash \Phi(Y_i, i)$. Of course, by our previous argument, if $u_{i+1} \Vdash \Phi(Y_i, i)$, $\mathcal{M}^{\mathcal{U}} \vDash \Phi(Y_i, i)$ (and so $u \Vdash \Phi(Y_i, i)$) and $\mathcal{M}^{\mathcal{U}} \vDash \exists X \Phi(X, i)$. For the other direction, if $u_{i+1} \nVdash \Phi(Y_i, i)$, then, by strong Σ^1_1-DC_0, no $v \leq u_i$ forces $\exists X \Phi(X, i)$ so, in particular, $u \nVdash \exists X \Phi(X, i)$. In this case, $\mathcal{M}^{\mathcal{U}} \nvDash \exists X \Phi(X, i)$ by the usual facts about forcing. (In this case, we could argue for this directly from the definitions and above results.)

Π^1_2-CA_0: The argument is essentially the same as for the previous case since Π^1_1-CA_0 proves strong Σ^1_2-DC_0. □

Corollary 4.15. *For T any of the theories mentioned in Theorem 4.14, $T^{I\mathcal{U}}$ is a conservative extension of T for all sentences of second order arithmetic.*

We now return to the example of the Milliken-Taylor theorem, $\mathrm{MT}(k)$, for all k and the problem of carrying out constructions that inductively define sets using the ultrafilter at each step. Using the analysis in Theorem 4.14 for special cases of ATR_0 we can now handle situations in which the sets required are defined from the previous ones by an arithmetic formula including the ultrafilter (in the disguise of $\delta_{\mathcal{U}}$). Of course, we must now deal with all k in our model $\mathcal{M}^{\mathcal{U}}$ not just the standard ones in $\mathbb{N}$. Note that $\mathrm{MT}(k)$

clearly implies RT^k_r, Ramsey's theorem for k-tuples and all colors r by taking $a_t = x_{n(t)}$. As $\forall k \forall r RT^k_r$ is not provable in ACA_0 (see Hirschfeldt [2104, Corollary 6.24]), $\forall k MT(k)$ is not provable in $\mathrm{ACA}_0^{I\mathcal{U}}$ by Theorem 4.15.

The only part of the proof of MT(k) for standard k outlined in Theorem 3.4 that used the fact that $k \in \mathbb{N}$ was in the construction of the sets $B_t(E,i)$ for $t \leq k$. Note that the definition of $B_t(E,i)$ in Bergelson and Hindman [1989, Theorem 2.3] given above is arithmetic in the real coding all the $B_{t+1}(E,i)$ for $E \in [\mathbb{N}]^t$ in $\mathcal{L}^{\mathcal{U}}$. Thus for each $k \in M_0$ we can represent this definition as an example of $\mathrm{ATR}_0^{I\mathcal{U}}$ over an ordering of length k. Thus we immediately have that $\forall k MT(k)$ is provable in ATR_0. Of course, much less is needed as we only need instance of orderings of length k for each $k \in M_0$. Looking at the proof for ATR_0 in Theorem 4.14, one sees that all one needs in the ground model are instances of ATR_0 of length $k(\omega \cdot n)$ for some fixed $n \in \mathbb{N}$ (the k in that proof). Moreover, if IHT is provable in ACA_0 then (still by Theorem 6.1) we can replace the ω-jump in this calculation by the Turing jump and so instances of ATR_0 of length $k \cdot n$ (for some fixed $n \in \mathbb{N}$) in the model M would suffices.

It is clear that all instances of ATR_0 for orderings of length a number in M_0 are provable in the system usually denoted ACA_0' which asserts the existence of $X^{(k)}$, the kth Turing jump of X, for each k and X. (Indeed these two systems are clearly equivalent.) By analogy one could label the system which asserts the existence of $X^{\omega \cdot k}$, the, the $\omega \cdot k$ jump of X, for each k and X, $(\mathrm{ACA}_0^+)'$. It is equivalent to the axioms for ATR_0 restricted to orderings of type $\omega \cdot k$, so an alternative notation for ACA_0' and $(\mathrm{ACA}_0^+)'$ might be $\mathrm{ATR}_0^{<\omega}$ and $\mathrm{ATR}_0^{<\omega \cdot \omega}$ for instances of ATR_0 of length less than ω and $\omega \cdot \omega$, respectively. We thus have good bounds on the proof theoretic strength of $\forall k MT(k)$.

Theorem 4.16. *$(ACA_0^+)' \vdash \forall k MT(k)$. If $ACA_0 \vdash IHT$ then $ACA_0' \vdash \forall k MT(k)$ and indeed ACA_0' and $\forall k MT(k)$ are equivalent over RCA_0.*

Proof. We have already argued for everything except the last claimed equivalence. This follows from our previous observation that $\forall k MT(k) \vdash \forall k, r RT^k_r$ while McAloon [1985, Theorem B] (or see Hirschfeldt [2104, Theorem 6.27]) proves that $\forall k, r RT^k_r \vdash ACA_0'$. $\square$

Note that even if ACA_0 does not prove IHT, we are still quite close to the best possible result as even MT(3) proves IHT (Hirst [2004, Theorem 6]).

If we are interested essentially in the divide between second and third order methods, then we would like results analogous to Theorem 4.14 and Corollary 4.15 for Z_2. The same proof does not work as the relevant instances of strong dependent choice are not provable in Z_2 (see Simpson [2009, VII.6.3]). Nonetheless, our methods applied to theories satisfying $\exists X(V = L[X])$ yield a conservation result for $\mathrm{Z}_2^{I\mathcal{U}}$ over Z_2 for essentially all sentences of combinatorial interest.

Proposition 4.17. *If $T = Z_2 + \exists X(V = L[X])$, $\mathcal{M} \vDash T$ and $\mathcal{U}$ is $\mathcal{M}$-generic for $\mathbb{P}_{I\mathcal{U}}$ then $\mathcal{M}^{\mathcal{U}} \vDash T^{I\mathcal{U}}$.*

Proof. Proceed as in the case for Π^1_1-CA_0 in Theorem 4.14 but use strong Σ^1_n-DC_0 for each n as needed. (By Simpson [2009, VII.6.5], strong Σ^1_n-DC_0 is a theorem of $Z_2 + \exists X(V = L[X])$ for each n $\square$

Corollary 4.18. *$Z_2^{I\mathcal{U}}$ is Π^1_4-conservative over Z_2.*

Proof. Suppose, for the sake of a contradiction, that $\Phi(X, Y, Z, W)$ is arithmetic and $\mathrm{Z}_2^{I\mathcal{U}} \vdash \forall X \exists Y \forall Z \exists W \Phi(X, Y, Z, W)$ but $\mathcal{M} = \langle M, S, +, \times, <, 0, 1, \in\rangle \vDash \exists X \forall Y \exists Z \forall W \neg\Phi(X, Y, Z, W)$ and $\mathcal{M} \vDash Z_2$. Let $X \in S$ be a witness to the failure of our Π^1_4 sentence in $\mathcal{M}$ and let $\hat{S} = (L[X])^{\mathcal{M}}$ and $\mathcal{N} = \left\langle M, \hat{S}, +, \times, <, 0, 1, \in\right\rangle$. Note that, as usual, $\mathcal{N} \vDash Z_2$ & $V = L[X]$. We claim that $\mathcal{N} \vDash \forall Y \exists Z \forall W \neg\Phi(X, Y, Z, W)$: Consider any $Y \in \hat{S} \subseteq S$. By our assumptions, $\mathcal{M} \vDash \exists Z \forall W \neg\Phi(X, Y, Z, W)$ and so, by Shoenfield absoluteness (see Simpson [2104, VII.4.14]), $\mathcal{N} \vDash \exists Z \forall W \neg\Phi(X, Y, Z, W)$. Now by Proposition 4.17, we can extend $\mathcal{N}$ to $\mathcal{N}^{I\mathcal{U}}$ while maintaining the same numbers M and the same class of sets $\hat{S}$ so that $\mathcal{N}^{I\mathcal{U}} \vDash Z_2^{I\mathcal{U}} + V = L[X]$. As we also have $\mathcal{N}^{I\mathcal{U}} \vDash \forall Y \exists Z \forall W \neg\Phi(X, Y, Z, W)$ we have contradicted our assumption that $\mathrm{Z}_2^{I\mathcal{U}} \vdash \forall X \exists Y \forall Z \exists W \Phi(X, Y, Z, W)$ as desired $\square$

5 Ramsey Ultrafilters

We follow the same course as for idempotent ultrafilters with the natural changes. In place of $\mathcal{AU}5$ we have the following axiom saying that $\mathcal{U}$ is Ramsey:

$\mathcal{AU}5R$ $\forall X(\langle X^{[n]}\rangle$ partitions M into pairwise disjoint nonempty sets & $\forall n(n+1 \notin \delta_{\mathcal{U}}(X))$
$\rightarrow \exists Z(0 \in \delta_{\mathcal{U}}(Z)$ & $\forall n|X^{[n]} \cap Z| = 1))$

The language is the same as before. We denote by $T^{R\mathcal{U}}$ the same theory as $T^{I\mathcal{U}}$ except that we replace $\mathcal{AU}5$ by $\mathcal{AU}5R$. The notion of forcing $\mathbb{P}_{R\mathcal{U}}$ is the same as $\mathbb{P}_{I\mathcal{U}}$ except that the only restrictions on the U_i making up a condition u are that the U_i are infinite and nested with empty intersection. We now adopt the same definitions as for $\mathbb{P}_{I\mathcal{U}}$ (4.4, 4.6 and 4.8) for $\mathbb{P}_{R\mathcal{U}}$. All the same results hold with minor simplifications of some of the proofs as we no longer need to guarantee the extra requirements for conditions being in $\mathbb{P}_{I\mathcal{U}}$. We give the proofs only for those results where changes are needed. Note that we only need ACA_0 as our base theory rather than IHT.

Lemma 5.1 (4.3). *If $\mathcal{M} \vDash ACA_0$ and u_i is a descending sequence in $\mathbb{P}_{\mathbb{R}\mathcal{U}}$ then there is a condition u extending every u_i.*

Proof. Let u_i as specified by $U_{i,j}$ be a descending sequence in $\mathbb{P}_{R\mathcal{U}}$. As in Lemma 4.3 there is a function h defined by the following recursion:

$h(0) = 0$

$h(i+1) = \mu k(k > h(i)$ & $(\forall n, m \leq h(i))(U_{n,m} \supseteq U_{h(i),k})$.

We now set $V_i = U_{h(i),h(i+1)}$ and verify as before that the associated condition $v \leq u_i$ for every i. $\square$

Proposition 5.2 (4.10 in ACA$_0$). *For each sequence $X_m \in S$, the set $\{u|\forall m(u$ decides $X_m)\}$ is dense in $\mathbb{P}_{R\mathcal{U}}$. For each sentence Φ of $\mathcal{L}^{\mathcal{U}}$ (with set parameters), the set $\{u|u$ decides $\Phi\}$ is also dense in $\mathbb{P}_{R\mathcal{U}}$.*

Proof. It is clear that by coding the X_m and $X_m^{[n]}$ as the columns $X^{[k]}$ of a single X that it suffices to find, for any condition $u = \langle U_i \rangle$, an extension $v = \langle V_k \rangle$ such that, for each k, $X^{[k]} \in v$ or $\overline{X^{[k]}} \in v$. Begin with $X^{[0]}$. It is clear that either there are infinitely many (and so all as the U_i are nested) i such that $U_i \cap X^{[0]} \neq \emptyset$ or there are infinitely many (and so all) such that $U_i \cap \overline{X^{[0]}} \neq \emptyset$. Let Z_0 be $X^{[0]}$ or $\overline{X^{[0]}}$ accordingly and $V_0 = U_0 \cap Z_0$. We now define Z_k and V_k by induction: Z_{k+1} is either $X^{[k]}$ or $\overline{X^{[k]}}$ whichever has nonempty intersection with all the $U_j \cap Z_0 \cap \cdots \cap Z_k$. We then set $V_{k+1} = U_{k+1} \cap Z_0 \cap \cdots \cap Z_k \cap Z_{k+1}$. It is clear that the V_k are nested, nonempty with empty intersection and so form a condition v, As $U_k \supseteq V_k$, $v \leq u$. It is also clear by the definition of the Z_k that $X^{[k]}$ or $\overline{X^{[k]}}$ contains V_k and so v decides X as required. □

Theorem 5.3 (4.11). *If $\mathcal{M} \vDash ACA_0$ and $\mathbb{U}$ and the associated $\mathcal{U}$ are $\mathcal{M}$-generic for $\mathbb{P}_{R\mathcal{U}}$ then $\mathcal{M}^{\mathcal{U}}$ is an $\mathcal{L}^{\mathcal{U}}$ structure and a model of $ACA_0^{R\mathcal{U}}$. Moreover, any sentence Φ of $\mathcal{L}^{\mathcal{U}}$ (with second order parameters) is true in $\mathcal{M}^{\mathcal{U}}$ if and only if there is a $u \in \mathbb{U}$ such that $u \Vdash \Phi$. Finally, $\mathcal{U}$ is a Ramsey $\mathcal{M}$-ultrafilter.*

Proof. No changes (other than replacing $I\mathcal{U}$ by $R\mathcal{U}$) are needed except for the verification that $\mathcal{U}$ is a Ramsey$\mathcal{M}$-ultrafilter, i.e. $\mathcal{M}^{\mathcal{U}} \vDash \mathcal{AU}5R$ (rather than $\mathcal{AU}5$). By what we know at this point it suffices to show that for every X such that $\langle X^{[n]} \rangle$ partitions M into pairwise disjoint nonempty sets and every condition u which decides X and W where $W^{[m]} = \cup\{X^{[n]}|n \geq m\}$ with $n+1 \in \delta_{\mathcal{U}}(W)$ and $n+1 \notin \delta_{\mathcal{U}}(X)$ for all n that there is a $v \leq u$ which forces the conclusion of $\mathcal{AU}5R$: $\exists Z(0 \in \delta_{\mathcal{U}}(Z)\ \&\ \forall n|X^{[n]} \cap Z| = 1$.

By definition, $\forall m \exists i(W^{[m]} \supseteq U_i)$ so, by the properties of the U_i determining a condition and the fact that the $X^{[n]}$ partition M, there is an increasing function g such that $\forall i(U_i \cap X^{[g(i)]} \neq \emptyset)$. We let r_i be the least element of $U_i \cap X^{[g(i)]}$ and $R = \{r_i|i \in M\}$ so $V_i = U_i \cap R$ is infinite for every i and defines a condition $v \leq u$ with $R \supseteq V_0$. Now choose $t_j \in X^{[j]}$ for each $j \notin rg(g)$ and set $T = \{t_j|j \notin rg(g)\}$. Let $Z = R \cup T$. Clearly $|Z \cap X^{[n]}| = 1$ for all n (the set consists of precisely r_n or t_n depending on whether $n \in rg(g)$ or not. Moreover, $Z \supseteq R \supseteq V_0$ and so v forces that Z is a witness to the conclusion of $\mathcal{AU}5R$ as required. □

The same arguments as for the remaining results in §4 with $R\mathcal{U}$ replacing $I\mathcal{U}$ and the corresponding versions of all the previous results of §4 now give all the same results for Ramsey ultrafilters as we had for idempotent ultrafilters.

Corollary 5.4. *$ACA_0^{R\mathcal{U}}$ is a conservative extension of ACA_0 for all sentences of second order arithmetic.*

As for $I\mathcal{U}$, it does not make sense to consider theories weaker than ACA$_0$ for such conservation results (as the same proof as for Proposition 4.13 shows that ACA$_0^{R\mathcal{U}}$ is a conservative extension of $\mathrm{RCA}_0^{R\mathcal{U}}$). However, the analogous theorems hold for the stronger theories.

Theorem 5.5. *If T is any one of ACA_0^+, ATR_0, Π^1_1-CA_0 or Π^1_2-CA_0, $\mathcal{M} \vDash T$ and $\mathcal{U}$ is $\mathcal{M}$-generic for $\mathbb{P}_{R\mathcal{U}}$ then $\mathcal{M}^{\mathcal{U}} \vDash T^{R\mathcal{U}}$.*

Corollary 5.6. *For T any of the theories mentioned in Theorem 5.5, $T^{R\mathcal{U}}$ is a conservative extension of T for all sentences of second order arithmetic.*

Proposition 5.7. *If $T = Z_2 + \exists X(V = L[X]$, $\mathcal{M} \vDash T$ and $\mathcal{U}$ is $\mathcal{M}$-generic for $\mathbb{P}_{R\mathcal{U}}$ then $\mathcal{M}^{\mathcal{U}} \vDash T^{R\mathcal{U}}$.*

Corollary 5.8. *$Z_2^{R\mathcal{U}}$ is Π^1_4-conservative over Z_2.*

6 Appendix

Hirschfeldt [2014, §6.3] mentions an old proof theoretic result of Wang [1993] as saying that if a Π^1_2 formula $\forall X(\Theta(X) \to \exists Y \Psi(X,Y))$ is provable in ACA_0 then there is a $k \in \mathbb{N}$ such that ACA_0 proves that for every X there is a $Y \in \Sigma^{0,X}_k$ such that $\Psi(X,Y)$. (Note that this means that in a nonstandard model the Σ^0_k formula may have nonstandard parameters.) We strengthen this result and extend it to ACA_0^+ as we use this new version to prove our preservation and so conservation result over ATR_0 in Theorem 4.14 and Corollary 4.15. We also derive some recursion theoretic bounds on results provable from IHT and so also by the ultrafilter arguments that are conservative over it.

The basic theorem can be generalized to systems containing RCA_0 and specified by additional Π^1_2 axioms. The strengthened version applies to systems which, like ACA_0 and ACA_0^+, contain ACA_0 and have axioms which define operators which preserve the order of Turing reducibility. The basic proof (for ACA_0^+ and the general case) follows the compactness style argument given in Hirschfeldt [2014, Theorem 6.23] and attributed there to Jockusch. We then apply a simple observation for the strengthening. Rather than carry out the general proofs we simplify notation by considering just ACA_0 and ACA_0^+. We then state the more general results which follow by the same arguments.

Theorem 6.1. *Let T be ACA_0 or ACA_0^+, J be the (implicitly arithmetically defined) Turing or arithmetic jump operator, respectively, with its iterates $J^{(k)}$ and let R be a Π^1_2 sentence $\forall X \exists Y(\Psi(X,Y))$ (with Ψ arithmetic) such that $T \vdash R$. Then $\exists e, k \in \mathbb{N}$ such that, $T \vdash \forall X \exists Y(Y = \Phi_e(J^{(k)}(X))$ & $\Psi(X,Y))$, i.e. $T \vdash \forall X((\exists Y, Z)(Z, X$ satisfy the Π^0_2 formula implicitly defining $J^{(k)}(X)$ & $Y = \Phi_e(Z)$ & $\Psi(X,Y))$ where $\Phi_e(Z)$ is the function given by the eth Turing machine with oracle Z.*

Proof. We first prove the apparently weaker conclusion analogous to the result of Wang that, for some $k \in \mathbb{N}$, $T \vdash \forall X \exists i \exists Y(Y = \Phi_i(J^{(k)}(X)$ & $\Psi(X,Y))$, i.e. $T \vdash \forall X \exists i((\exists Y, Z)(Z, X$

satisfy the Π^0_2 formula implicitly defining $J^{(k)}(X)$ & $Y = \Phi_i(Z)$ & $\Psi(X,Y)$). (Note that, for ACA_0 this is the same as saying that there is a solution to Ψ for X which is Δ^X_{k+1} (with possibly nonstandard parameters) and we have Wang's theorem above.) Consider the parameterless arithmetic formulas $\Theta_k(X)$, for $k \in \mathbb{N}$, which say that $\exists i \Psi(X, \Phi_i(X^{(k)}))$. (Formally, $\Psi(X, \Phi_i(X^{(k)}))$ denotes the formula saying $\Phi_i(X^{(k)})$ is a total characteristic function and that the formula gotten by replacing, as usual, appearances of $m \in Y$ in Ψ by $\Phi_i(X^{(k)}))(m) = 1$.) If our desired result fails then $T \cup \{\neg\Theta_k(C)\}$ is consistent for each k. Indeed, as it is obvious that $T \vdash \Theta_k(C) \to \Theta_{k+1}(C)$, $T \cup \{\neg\Theta_k(C) | k \in \mathbb{N}\}$ is also consistent by compactness. So we may suppose that we have a countable model $\mathcal{M}$ of T with a set C such that $\mathcal{M} \vDash \neg\Theta_k(C)$ for each $k \in \mathbb{N}$. Now form the subset $\hat{S}$ of the sets S of $\mathcal{M}$ generated from C by J and closure under Turing reductions in $\mathcal{M}$. Define $\hat{\mathcal{M}}$ as the substructure of $\mathcal{M}$ with first order part that of $\mathcal{M}$ (i.e. M) and second order part $\hat{S}$. As the first order parts of $\mathcal{M}$ and $\hat{\mathcal{M}}$ are the same, truth of first order sentences with second order parameters in $\hat{S}$ are the same in $\mathcal{M}$ and $\hat{\mathcal{M}}$. Thus, $\hat{\mathcal{M}}$ is also a model of T and the J operator in $\hat{\mathcal{M}}$ is the restrictions of that in $\mathcal{M}$. Now, by the usual (external) construction of the closure in $\mathcal{M}$ of $\{C\}$ under J and Turing reductions (in ω many steps) and the fact that J preserves Turing reducibility, every $Y \in \hat{S}$ is of the form $\Phi_i(C^{(k)})$ for some $i \in M$ and $k \in \mathbb{N}$ and none of them satisfy $\Psi(X,Y)$ in $\mathcal{M}$ by assumption and so none do in $\hat{\mathcal{M}}$. Thus $\hat{\mathcal{M}}$ is a model of T but not of R, for the desired contradiction.

Now we know that $T \vdash \forall X \exists i \Psi(X, \Phi_i(X^{(\hat{k})}))$ for some fixed $\hat{k} \in \mathbb{N}$. Suppose that Ψ is Σ^0_n for some $n \in \mathbb{N}$. By standard arguments, we can choose $e, k \in \mathbb{N}$ (uniformly recursively in $\hat{k}$, the quantifier rank of Ψ and whether J is the Turing or arithmetic jump) such that (provably in T), for every X, m: $m \in \Phi_e(J^{(k)}(X)) \Leftrightarrow \exists i \Psi(X, \Phi_i(J^{(\hat{k})}(X)))$ & $m \in \Phi_j((J^{(\hat{k})}(X))$ for j the least such i). A point to notice here is that as $k > \hat{k}$, $J^{(\hat{k})}(X)$ is uniformly recursive in $J^{(k)}(X)$ and the properties we need can all be expressed in terms of indices relative to $J^{(\hat{k})}(X)$ while k is taken to be sufficiently larger than $\hat{k}$ that we can decide all the relevant sentences about $J^{(\hat{k})}(X)$ uniformly recursively in $J^{(k)}(X)$. It is now immediate that $T \vdash \forall X \Psi(X, \Phi_e(X^{(k)}))$. □

We can now draw some recursion theoretic consequences which bound the complexity of witness to Π^1_2 theorems of T (ACA_0 or ACA_0^+) in the standard model of arithmetic. The first is the obvious one that if $T \vdash \forall X \exists Y(\Psi(X,Y))$ then there is a k such that for each X such that $Y \leq_T J^{(k)}(X)$ (and indeed there is a fixed e such $\Phi_e(J^{(k)}(X))$ is such a witness). Thus, for example, as MT(n) is provable in IHT for each $n \in \mathbb{N}$ (Theorem 3.4) there is, for each $n \in \mathbb{N}$ a $k \in \mathbb{N}$ and and $e \in \mathbb{N}$ such that for every instance X of MT(n), $\Phi_e(J^{(k)}(X)$ is the desired homogenous set. Here J being the ω-jump works as IHT is provable in ACA_0^+. If it turns out that IHT is provable in ACA_0 then the Turing jump will work for J. Now the proof only works for standard n and so it does not say much on purely general grounds about the strength of the assertion $\forall n MT(n)$ which we analyzed proof theoretically in Theorem 4.16. We can, however, use our Theorem 6.1 to derive recursion theoretic bounds. These comments and the Corollary we now give will apply just as well to the extensions of Theorem 6.1 in Corollary 6.6 and Theorem 6.8.

Corollary 6.2. *Let T, J and $R(n) = \forall X \exists Y(\Psi(X, Y, n))$ be as in Theorem 6.1. If for each $n \in \mathbb{N}$, $T \vdash R(n)$ then there is are recursive functions $k(n)$ and $e(n)$ such that for every n and X, $\Psi(X, \Phi_{e(n)}(J^{k(n)}(X), n)$.*

Proof. We define the functions $k(n)$ and $e(n)$ by searching for the least triple consisting of a k, n and a proof in T of $\forall X \exists Y(Y = \Phi_e(J^{(k)}(X))\ \&\ \Psi(X, Y))$. This recursive search terminates by Theorem 6.1 and the functions so defined clearly satisfy the conclusion of our Corollary.

Now in some cases the numbers e and k of Theorem 6.1 can be read off (easily or with a bit of work) from the proof of R in T and it may be that the dependence on n of the proofs used in Corollary 6.2 can also be directly determined but in others it may not be at all clear how to produce this information. This is particularly true for Corollary 6.2 when the proof that $\forall n(T \vdash R(n))$ is produced by an external induction on $n \in \mathbb{N}$. For MT(n) our proof is quite indirect and recovering these numbers not at all simple. On the other hand, the proofs in Hirst [2004] are much more amenable to such an analysis. In future work (on Gowers' Fin_k theorem), we expect to see examples where recovering the number of jumps needed seems much more difficult and so Corollary 6.2 will give results that are far from obvious. □

We now provide a generalization of Wang's result that is proved by the same method.

Definition 6.3. We say that a proof theoretic system T (of second order arithmetic) is a Π^1_2 *system* if it is specified by adding a set of Π^1_2 axioms $A_i = \forall X \exists Y \Gamma_i(X, Y)$ to RCA_0.

Examples include ACA_0 and ACA_0^+ as they can be specified as the closure under the Turing and arithmetic jumps, respectively, and these operators are defined by arithmetic (even Π^0_2) formulas $\Phi(X, Y)$ with unique solutions Y for every X. Most of the systems weaker than ACA_0^+ studied in reverse mathematics are of this form as are many stronger ones.

Theorem 6.4. *Let T be a Π^1_2 system generated over RCA_0 by the Π^1_2 axioms $A_i = \forall X \exists Y \Gamma_i(X, Y)$. Let R be a Π^1_2 sentence $\forall X \exists Y(\Psi(X, Y))$ such that $T \vdash R$. Then $\exists k \in \mathbb{N}$ such that $T \vdash \forall X \exists l \leq k \exists i_0, \ldots, i_l \leq k(\exists Z_0, \ldots Z_l, Y_0, \ldots Y_l, e_0, \ldots e_l)(X = Z_0\ \&\ (\forall j)_{0 \leq j \leq l}(Y_j = \Phi_{e_j}(Z_j))\ \&\ (\forall j)_{0 \leq j < l}(\Gamma_{i_j}(Y_j, Z_{j+1}))\ \&\ \Psi(X, Y_l))$.*

Proof. The proof is essentially the same as the proof above for the weak version of Theorem 6.1. The sentences $\Theta_k(X)$, for $k \in \mathbb{N}$, now assert that there are $l \leq k, i_0, \ldots, i_l \leq k, Z_0, \ldots Z_l, Y_0, \ldots Y_l, e_0, \ldots e_l$ with the desired property. As above, if our conclusion fails $T \cup \{\neg\Theta_k(C)\}$ is consistent for each k. Indeed, as it is again obvious that $T \vdash \Theta_k(C) \to \Theta_{k+1}(C)$, $T \cup \{\neg\Theta_k(C) | k \in \mathbb{N}\}$ is also consistent by compactness. So we may suppose that we have a countable model $\mathcal{M}$ of T with a set C such that $\mathcal{M} \vDash \neg\Theta_k(C)$ for each $k \in \mathbb{N}$.

The generation process for a submodel $\hat{\mathcal{M}}$ of $\mathcal{M} \vDash T$ from C proceeds in ω many steps starting with C by alternately applying some Turing reduction Φ_i for $i \in M$ and adding a set Z of $\mathcal{M}$ such that $\mathcal{M} \models \Gamma_i(Y, Z)$ for some Y already added to our list and some $i \in \mathbb{N}$. As before $\hat{\mathcal{M}}$ is absolute with respect to $\mathcal{M}$ for arithmetic formulas with set parameters from $\hat{\mathcal{M}}$ and, moreover, is a model of T. By our our choice of C, $\mathcal{M} \models \neg\Theta_k(C)$ for every $k \in \mathbb{N}$. On the other hand, every set Y in $\hat{\mathcal{M}}$ is, for some $l, i_0, \dots i_l \in \mathbb{N}$, the Y_l of a sequence $Z_0, \dots Z_l, Y_0, \dots Y_l, e_0, \dots e_l$ such that $(X = Z_0 \;\&\; (\forall j)_{0 \leq j \leq l}(Y_j = \Phi_{e_j}(Z_j)) \;\&\; (\forall j)_{0 \leq j < l}(\Gamma_{i_j}(Y_j, Z_{j+1})))$and so $\mathcal{M} \models \neg\Psi(C, Y)$ for every $Y \in \hat{\mathcal{M}}$. Thus $\hat{\mathcal{M}} \models \forall Y \neg\Psi(C, Y)$, i.e. $\hat{\mathcal{M}}$ is a model of T but not of R for the desired contradiction. □

We now turn to the strengthening of Theorem 6.4 analogous to Theorem 6.1. Here we need additional assumptions on the axioms of our system T corresponding to the properties of ACA_0 and ACA_0^+ used to go from the weaker version above to the full theorem.

Definition 6.5. We say that the axiom $A = \forall X \exists Y \Gamma(X, Y)$ of a Π^1_2 system *corresponds to an operator* $J_\Gamma(X)$ if $T \vdash \forall X \exists! Y \Gamma(X, Y)$. For $k \in \mathbb{N}$, we denote the *k-fold iteration* of such a J_Γ by $J^{(k)}$: $J^{(0)}(X) = X$ and $J^{(k+1)}(X) = JJ^{(k)}(X)$. We say that such an operator J_Γ is *monotonic* if $\forall X \forall Y \forall Z \forall W (X \leq_T Y \;\&\; \Gamma(X, W) \;\&\; \Gamma(Y, Z) \rightarrow W \leq_T Z)$.

In the setting of Theorem 6.4, if the axioms A_i correspond to monotonic operators J_i (such as the Turing or arithmetic jump) one of which is the Turing jump (so $T \vdash ACA_0$), then we can deduce additional uniformities that provide a generalization of the full Theorem 6.1.

Corollary 6.6. *Let T be a Π^1_2 system all of whose axioms A_i, $i \in \mathbb{N}$, correspond to monotonic operators J_i one of which, say J_0, is the Turing jump and let R be a Π^1_2 sentence $\forall X \exists Y (\Psi(X, Y))$ such that $T \vdash R$. Then there are $e, k \in \mathbb{N}$ such that $T \vdash \forall X \Psi(X, \Phi_e(J^{(k)}(X)))$ where $J(X) = J_k J_{k-1} \dots J_0(X)$.*

Proof. By Theorem 6.4, we have a $k \in \mathbb{N}$ such that
$T \vdash \forall X \exists l \leq k \exists i_0, \dots, i_l \leq k (\exists Z_0, \dots Z_l, Y_0, \dots Y_l, e_0, \dots e_l)(X = Z_0 \;\&\; (\forall j)_{0 \leq j \leq l}(Y_j = \Phi_{e_j}(Z_j)) \;\&\; (\forall j)_{0 \leq j < l}(\Gamma_{i_j}(Y_j, Z_{j+1})) \;\&\; \Psi(X, Y_l))$.

Let $J(X) = J_k J_{k-1} \dots J_0(X)$. As the axioms A_j all define monotonic operators J_j, we see for any X, $l, i_0, \dots, i_l \leq k$ and sequences $Z_0, \dots Z_l, Y_0, \dots Y_l, e_0, \dots e_l$ as above, $Z_j, Y_j \leq_T J_{i_j} J_{i_{j-1}} \dots J_{i_0}(X) \leq_T J^{(k)}(X)$ for all $j \leq k$. So all possible such sequences can be represented by indices relative to $J^{(k)}(X)$. Moreover, for some n larger than the quantifier complexities of Ψ and the Γ_i for $i \leq k$, the assertions about sequences of indices satisfying the conditions described can all be decided uniformly recursively in $(J^{(k)}(X))^{(n)} = J_0^{(n)} J^{(k)}(X) \leq_T J^{(k+n)}(X)$. Thus as before, by minimizing the sequence of indices defining the sequence of sets desired, we can calculate, uniformly recursively in k, Ψ and the Γ_i, $i \leq k$, an index e which relative to $J^{(k+n)}(X)$ calculates (provably in T) a Y such that $\Psi(X, Y)$. This e along with $k + n$ are the standard numbers required for the Corollary. □

Theorems 6.1 and 6.4 as well as Corollary 6.6 all have analogous versions for Π^1_1-CA$_0$ and the hyperjump operator replacing ACA$_0$ and the Turing jump and Π^1_3 axioms A_1 and sentences R replacing the Π^1_2 ones. For definiteness and simplicity we define the hyperjump by $HJ(X) = \{e|\Phi_e(X)$ is the characteristic function of a well ordering$\}$. This is one of the standard versions of the complete $\Pi^1_1(X)$ set but any of the usual ones would do just as well. The crucial standard fact that we need to note is an absoluteness result for well-foundednes or, equivalently, Π^1_1 sentences. We denote the finite iterates of the hyperjump by $HJ^{(n)}(X)$ for $n \in \mathbb{N}$.

Proposition 6.7. *If $\mathcal{M} \vDash \Pi^1_1$-$CA_0$ and $\hat{\mathcal{M}}$ is an ω-submodel of $\mathcal{M}$, i.e. it has the same first order part as $\mathcal{M}$, and is closed under hyperjump then it is a β-submodel, i.e. for any Π^1_1 sentence Φ with parameters in $\hat{\mathcal{M}}$, $\mathcal{M} \vDash \Phi \Leftrightarrow \hat{\mathcal{M}} \vDash \Phi$.*

Proof. See Simpson [2009, VII,1.8]. □

We explicitly state the basic case of our metatheorem for Π^1_1-CA$_0$ and leave the other analogs to the reader.

Theorem 6.8. *If $R = \forall X \exists y \forall Z \Phi(X, Y, Z)$ is a Π^1_3 sentence, $T = \Pi^1_1$-CA_0 and $T \vdash R$ then there are $e, k \in \mathbb{N}$ such that $T \vdash \forall X \exists Y (Y = \Phi_e(HJ^{(k)}(X)$ & $\forall Z \Phi(X, Y, Z))$.*

Proof. The proof is essentially the same as Theorem 6.1 with hyperjump replacing the Turing or arithmetic jump J. Now, however, there are Π^1_1 formulas defining each of the nth hyperjumps for $n \in \mathbb{N}$ and one uses Proposition 6.7 to show that Π^1_1 formulas with parameters in $\hat{\mathcal{M}}$ are absolute to $\mathcal{M}$, i.e. $\hat{\mathcal{M}}$ is a β-submodel of $\mathcal{M}$. □

7 Bibliography

Avigad, Jeremy [1998], An effective proof that open sets are Ramsey, *Arch. Math. Logic* **37**, 235–240.

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

Bergelson, V. and Hindman, N. [1989], Ultrafilters and Multidimensional Ramsey Theorems, *Combinatorica* **9**, 1-7.

Enayat. Ali [2006], From bounded arithmetic to second order arithmetic via automorphisms in *Logic in Tehran*, Ali Enayat, Iraj Kalantari, Mojtaba Moniri EDS., *LNL* **26**, Association for Symbolic Logic (La Jolla CA), distributed by A.K. Peters (Wellesley MA).

Hindman, Neil [1972], The existence of certain ultrafilters on $\mathbb{N}$ and a conjecture of Graham and Rothschild. *Proc. Amer. Math. Soc.* **36,** 341-346.

Hindman, Neil [2005], Algebra in the Stone-Čech Compactification and its Applications to Ramsey Theory, *Scientiae Mathematicae Japonicae* **62**, 321-329.

Hindman, Neil and Strauss, Dona [2012], *Algebra in the Stone-Čech Compactification, Theory and Applications* 2nd ed., De Gruyter, Berlin/Boston,

Hirschfeldt, Denis [2014], *Slicing the Truth: On the Computable and Reverse Mathematics of Combinatorial Principles*, Chitat Chong, Qi Feng, Theodore A Slaman, W. Hugh Woodin, and Yue Yang eds., World Scientific, Singapore.

Hirst, Jeffrey L. [2004], Hindman's Theorem, Ultrafilters, and Reverse Mathematics, *J. Sym. Logic* **69**, 65-72.

Kirby. l. A. S. [1984], Ultrafilters and types on models of arithmetic, *Ann. Pure and Applied Math.* **27**, 215-252.

Kreuzer, Alexander [2012], Non-principal ultrafilters, program extractions and higher order reverse mathematics, *J. Math. Logic* **12**, 1250002 (16 pp.).

Kreuzer, Alexander [2015], On Idempotent Ultrafilters in Higher-order Reverse Mathematics, *J. Sym. Logic* **80**, 179-193.

Kreuzer, Alexander [2015a] Minimal Idempotent Ultrafilters and the Auslander-Ellis Theorem, Preprint. arXiv:1305.6530.

McAloon, K. [1985], Paris-Harrington incompleteness and progressions of theories in *Recursion Theory* (Ithaca, N.Y., 1982), Anil Nerode and Richard A. Shore, eds., *Proc. Sympos. Pure Math.* **42**, Amer. Math. Soc., Providence, RI.

Simpson, Stephen G. [2009] *Subsystems of Second Order Arithmetic* 2^{nd} ed., *Perspectives in Logic*, ASL and Cambridge University Press, New York.

Todorcevic, Stevo [2010] *Introduction to Ramsey Spaces*, Annals of Mathematical Studies **174**, Princeton University Press, Princeton NJ.

Towsner, Henry [2011], A combinatorial proof of the dense Hindman's theorem, *Discrete Math.* **311**, 1380–1384.

Towsner, Henry [2011a], Hindman's theorem: An ultrafilter argument in second order arithmetic, *J. Sym. Logic* **76**, 353–360.

Towsner, Henry [2104] Ultrafilters in reverse mathematics, *J., Math. Logic* **14**, 145001 (11 pp.),

Wang, Hao [1993], *Popular Lectures on Mathematical Logic*, Revised reprint of the 1981 second edition, Dover Publications, Inc., New York.