

# WHERE PIGEONHOLE PRINCIPLES MEET KÖNIG LEMMAS

DAVID BELANGER, C. T. CHONG, WEI WANG, TIN LOK WONG,  
AND YUE YANG

ABSTRACT. We study the pigeonhole principle for  $\Sigma_2$ -definable injections with domain twice as large as the codomain, and the weak König lemma for  $\Delta_2^0$ -definable trees in which every level has at least half of the possible nodes. We show that the latter implies the existence of 2-random reals, and is conservative over the former. We also show that the former is strictly weaker than the usual pigeonhole principle for  $\Sigma_2$ -definable injections.

## 1. INTRODUCTION

As Stephen Simpson maintained in his book [28], the goal of reverse mathematics is to investigate which set existence axioms are needed to prove theorems of ordinary mathematics. Since a great amount of ordinary mathematics can be formalized in the framework of second-order arithmetic through the process of coding, one may regard the set existence axioms to be those concerned with subsets of  $\mathbb{N}$ . Our focus in this paper is on the following slightly different question:

Which elementary or first-order properties of  $\mathbb{N}$  are needed to prove theorems of ordinary mathematics?

We study one instance of this very broad question.

To motivate this line of thought, we briefly review some recent developments in reverse mathematics. Traditionally the most prominent axiom systems about the existence of subsets of  $\mathbb{N}$  are the so-called *Big Five* systems, i.e.,  $\text{RCA}_0$ ,  $\text{WKL}_0$ ,  $\text{ACA}_0$ ,  $\text{ATR}_0$  and  $\Pi_1^1\text{-CA}_0$ . The weakest system  $\text{RCA}_0$ , whose principal constituents are the induction scheme for  $\Sigma_1^0$ -formulas and the comprehension scheme asserting the existence of all  $\Delta_1^0$ -definable sets, is usually taken—as we do in this

---

2010 *Mathematics Subject Classification.* 03B30, 03F35, 03F30, 03D32.

Belanger was partially supported by BOF grant number 01P01117. Chong's research was partially supported by NUS grants C-146-000-042-001 and WBS : R389-000-040-101. Wang was partially supported by China NSF Grants 11471342 and 11971501. Wong was financially supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE2016-T2-1-019 / R146-000-234-112 when this research was carried out. Yang was partially supported by Singapore MOE grant MOE-2019-t2-2-121. All the authors acknowledge the support of JSPS–NUS grants R146-000-192-133 and R146-000-192-733 during the course of the work.

paper—to be the base system. Over  $\text{RCA}_0$ , many important theorems in ordinary mathematics are known to be equivalent to one of the Big Five. However, in the last two decades a growing body of exceptions have appeared: a number of theorems in ordinary mathematics were found to be inequivalent to any of the Big Five, and among them some were proved to be independent of each other. Classical computability-theoretic methods over the standard natural numbers have been popular and fruitful in driving this development. They provide powerful tools for constructing models of the form  $(\mathbb{N}, \mathcal{S})$ , where  $\mathcal{S}$  is a subset of the power set of  $\mathbb{N}$ , to establish independence results.

Recently, the study of first-order strength of a given subsystem of second-order arithmetic has attracted much attention. In such studies, model-theoretic and proof-theoretic techniques come in naturally. This approach has provided insights into reverse mathematics that the computability-theoretic approach does not. It has introduced avenues for answering open questions in reverse mathematics, and in fact questions not answerable using standard models  $(\mathbb{N}, \mathcal{S})$ , by nature of the questions themselves. The best-known examples are all concerned with Ramsey’s theorem for pairs ( $\text{RT}_2^2$ ): first, the theorem of Chong, Slaman and Yang [4] which separates  $\text{RT}_2^2$  from its stable version ( $\text{SRT}_2^2$ ); second, the theorem by the same authors [5] that  $\text{RT}_2^2$  does not imply the induction scheme for  $\Sigma_2^0$ -formulas; and third, the theorem of Patey and Yokoyama [25] which says that all  $\Pi_3^0$ -consequences of  $\text{RT}_2^2$  are already provable in  $\text{RCA}_0$ .

A particularly interesting aspect of the first example above is that the statements of  $\text{RT}_2^2$  and  $\text{SRT}_2^2$  make no direct reference to the first-order properties of  $\mathbb{N}$ . Conceivably, this independence result can be reproduced using classical computability-theoretic techniques. On the other hand, there are a number of second-order statements whose strengths can only be understood by studying their first-order consequences. An example is  $\text{TT}^1$ , which is a tree version of the infinite pigeonhole principle asserting, for every partition of  $2^{<\mathbb{N}}$  into finitely many parts, the existence of a monochromatic subtree isomorphic to the (infinite) perfect binary tree  $2^{<\mathbb{N}}$ . Any standard model of  $\text{RCA}_0$  is trivially a model of  $\text{TT}^1$ , because in this case a monochromatic subtree can be computed from the partition. Indeed,  $\text{RCA}_0 + \text{I}\Sigma_2^0 \vdash \text{TT}^1$ . However, Corduan et al. [8] proved that in the absence of  $\text{I}\Sigma_2^0$  a finite partition of  $2^{<\mathbb{N}}$  may fail to compute a monochromatic subtree isomorphic to  $2^{<\mathbb{N}}$ . As  $\text{RCA}_0 + \text{TT}^1 \vdash \text{B}\Sigma_2^0$  easily, the strength of  $\text{TT}^1$  lies between  $\text{B}\Sigma_2^0$  and  $\text{I}\Sigma_2^0$ , and thus cannot be calibrated in the classical computability-theoretic setting. Recently, Chong, Li, Wang and Yang [6] proved that  $\text{RCA}_0 + \text{TT}^1 \not\vdash \text{I}\Sigma_2^0$ .

Another example is  $2\text{-WWKL}_0(1/2)$ , which we investigate in this paper. It is related to measure theory and algorithmic randomness. Recall the principle  $\text{WWKL}_0$ , introduced by Yu and Simpson [31], which states

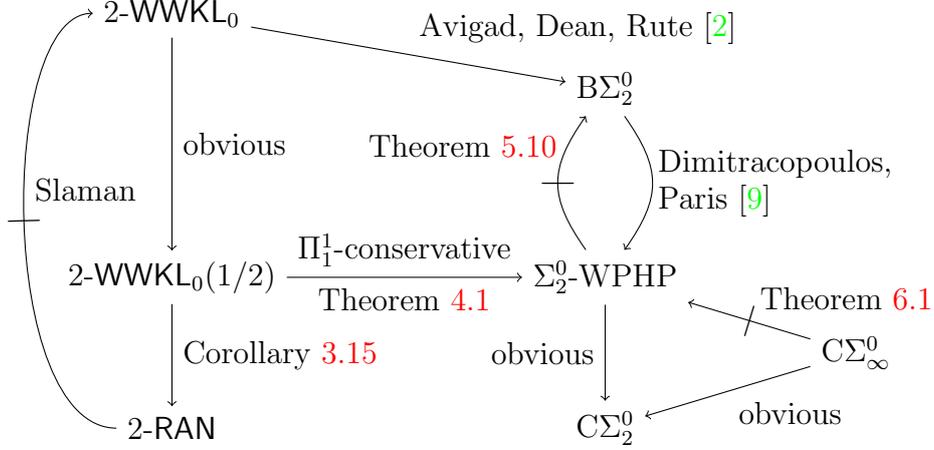


FIGURE 1. Some relationships between various subsystems of second-order arithmetic investigated in this paper, over  $\text{RCA}_0$

that if a binary tree  $T$  satisfies

$$(1.1) \quad \exists m > 0 \forall n (|T \cap 2^n| > 2^n/m),$$

where  $2^n$  denotes both a natural number (on the right-hand side of the inequality) and the set of all binary strings of length  $n$  (on the left-hand side) then there exists an  $X \in [T]$ , meaning  $X$  is an infinite path through  $T$ . This principle is known to be strictly weaker than  $\text{WKL}_0$ , but independent of  $\text{RCA}_0$ . Avigad et al. considered in a more recent paper [2] the analogue of  $\text{WWKL}_0$  for  $\Delta_n^0$ -definable trees, which they call  $n\text{-WWKL}_0$ . They showed that  $2\text{-WWKL}_0$  is equivalent to a formalized version of the dominated convergence theorem in second-order arithmetic (DCT), and  $n\text{-WWKL}_0$  implies the existence of  $n$ -random reals ( $n\text{-RAN}$ ). In the same paper, Avigad et al. asked whether  $n\text{-RAN}$  is equivalent to  $n\text{-WWKL}_0$ . Slaman [unpublished] answered their question in the negative. Nevertheless, by relativizing an argument in Kučera [20, Lemma 3], one can show that, over the standard model  $\mathbb{N}$ , every  $n$ -random real computes some  $X \in [T]$ , whenever  $T$  is a  $\Delta_n^0$ -definable tree satisfying (1.1). Hence the use of classical computability-theoretic tools alone cannot answer the question of Avigad et al.

In this paper we improve on Slaman's result. We introduce a principle called  $2\text{-WWKL}_0(1/2)$  whose strength lies between those of  $2\text{-WWKL}_0$  and  $2\text{-RAN}$ . It states that if  $T$  is a  $\Delta_2^0$ -definable tree satisfying

$$(1.2) \quad \forall n (|T \cap 2^n| \geq 2^{n-1})$$

then there exists an  $X \in [T]$ . We prove that the first-order theory of  $2\text{-WWKL}_0(1/2)$  can be axiomatized by  $\text{I}\Sigma_1$  plus the principle  $\Sigma_2\text{-WPHP}$ , which is a variant of the finite pigeonhole principle strictly

weaker than  $B\Sigma_2$ . As  $2\text{-WWKL}_0$  implies  $B\Sigma_2^0$  [2, Theorem 3.7], we know  $2\text{-WWKL}_0(1/2)$  is strictly weaker than  $2\text{-WWKL}_0$ . From this we conclude that  $I\Sigma_1 + \Sigma_2\text{-WPHP}$  is an upper bound for the first-order theory of  $2\text{-RAN}$ . We also prove that  $\Sigma_2\text{-WPHP}$  is substantially different from the usual fragments of first-order arithmetic.

We organize this paper as follows. In Section 2, we introduce some basic notation and set up a few preliminary results about weak pigeonhole principles. In Section 3, we explore the notion of  $\Delta_2^0$  trees in the absence of  $B\Sigma_2^0$ . In particular, we define  $2\text{-WWKL}_0(1/2)$ , and verify that  $\text{RCA}_0 + 2\text{-WWKL}_0(1/2) \vdash \Sigma_2^0\text{-WPHP} \wedge 2\text{-RAN}$  there. In Section 4, we prove that  $2\text{-WWKL}_0(1/2)$  is  $\Pi_1^1$ -conservative over  $\text{RCA}_0 + \Sigma_2^0\text{-WPHP}$ . In Section 5, we prove that  $\Sigma_{n+1}\text{-WPHP}$  is strictly weaker than  $B\Sigma_{n+1}$ , but strictly stronger than the cardinality scheme for  $\Sigma_{n+1}$  formulas ( $C\Sigma_{n+1}$ ). We conclude in Section 6 with a discussion of the techniques developed in our study of pigeonhole principles, and a list of questions.

## 2. BASICS

Let us start with some notational conventions. The language  $L_1$  of first-order arithmetic has symbols  $0, 1, +, \cdot, <$  and a symbol for equality. By convention, if  $\mathfrak{M}, \mathfrak{N}, \mathfrak{N}_k, \dots$  are  $L_1$  structures, then their universes are always denoted by  $M, N, N_k, \dots$  respectively. The language  $L_2$  of second-order arithmetic has a first-order sort and a second-order sort, with a copy of  $L_1$  on the first-order sort, and a symbol  $\in$  relating a first-order object to a second-order object. We only consider  $L_2$  structures  $\mathfrak{M} = (M, \mathcal{S}, 0^{\mathfrak{M}}, 1^{\mathfrak{M}}, +^{\mathfrak{M}}, \cdot^{\mathfrak{M}}, <^{\mathfrak{M}}, \in^{\mathfrak{M}})$  where the second-order universe  $\mathcal{S}$  is a collection of subsets of the first-order universe  $M$ , and  $\in^{\mathfrak{M}} = \in$ . Since there is no risk of ambiguity in this paper, we abbreviate an  $L_2$  structure  $(M, \mathcal{S}, \dots)$  as  $(M, \mathcal{S})$ .

Recall that  $\text{PA}^-$  is a finite set of axioms saying that the model is the non-negative half of a discretely ordered commutative ring. The *induction axiom*  $I\varphi$  for a formula  $\varphi(x, \vec{Y})$ , where  $\vec{Y}$  is a finite tuple of first- or second-order parameters, is the sentence

$$\forall \vec{Y} (\varphi(0, \vec{Y}) \wedge \forall x (\varphi(x, \vec{Y}) \rightarrow \varphi(x+1, \vec{Y})) \rightarrow \forall x \varphi(x, \vec{Y})).$$

If  $\Gamma$  is a set of formulas, then  $I\Gamma$  denotes the *induction scheme* for  $\Gamma$ , i.e., the collection of all  $I\varphi$ 's in which  $\varphi \in \Gamma$ . The *bounding axiom*  $B\psi$  for a formula  $\psi(w, x, \vec{Y})$  states

$$\forall \vec{Y}, u (\forall w < u \exists x \psi(w, x, \vec{Y}) \rightarrow \exists v \forall w < u \exists x < v \psi(w, x, \vec{Y})).$$

If  $\Gamma$  is a set of formulas, then  $B\Gamma$  denotes the *bounding scheme* for  $\Gamma$ , i.e., the collection of all  $B\psi$ 's such that  $\psi \in \Gamma$ .

Fix a  $\Pi_2$  sentence  $\text{exp}$  which asserts the totality of exponentiation over  $I\Delta_0$ . One can expand a model  $\mathfrak{M} \models I\Delta_0 + \text{exp}$  with the function  $x \mapsto 2^x$  provided by  $\text{exp}$ . We sometimes identify  $\mathfrak{M}$  with this expansion. From Gaifman–Dimitracopoulos [12, Theorem 3.3], we know

$\mathfrak{M} \models \text{I}\Delta_0(\text{exp})$ , where  $\Delta_0(\text{exp})$  denotes the smallest collection of formulas that contains all the atomic formulas *in the expanded language*, and is closed under Boolean operations and bounded quantification. We may even allow the exponential function to appear in the bounding terms here [14, Lemma I.1.30].

We say a subset  $F \subseteq M$  is  $\mathfrak{M}$ -finite or *coded in  $\mathfrak{M}$*  if there is  $c \in M$  such that

$$F = \{x \in M : \mathfrak{M} \models \text{the } x\text{-th digit in the binary expansion of } c \text{ is } 1\}.$$

For example, the finite set  $F = \{0, 1, 3\}$  is  $\mathfrak{M}$ -finite in all models  $\mathfrak{M} \models \text{I}\Delta_0 + \text{exp}$ , with code  $c = 2^0 + 2^1 + 2^3 = 11$ . In a model  $\mathfrak{M} \models \text{I}\Delta_0 + \text{exp}$  many desirable properties of  $\mathfrak{M}$ -finite sets are available, for example, bounded  $\Delta_0$  comprehension [14, Theorem I.1.36], the pigeonhole principle for coded functions [14, Theorem I.1.41(2)], and the existence of cardinalities [14, Theorem I.1.41(1)]. Unless otherwise stated, we denote by  $|X|$  the cardinality of a set  $X$  in this sense. Recall that  $\text{I}\Delta_0$  is already enough to prove the usual properties of the Cantor pairing function  $(x, y) \mapsto \langle x, y \rangle$ . Similarly, many properties of the usual sequence-coding function  $(a_i : i < n) \mapsto \langle a_i : i < n \rangle$  can be proved in  $\text{I}\Delta_0 + \text{exp}$ . Note that both of these coding functions have  $\Delta_0$  graphs, and each component of the object being coded is bounded above by the code, provably in  $\text{I}\Delta_0$ .

The axiom system  $\text{RCA}_0$  consists of  $\text{PA}^-$ ,  $\text{I}\Sigma_1^0$  and

$$(2.1) \quad \forall \vec{Y} (\forall x (\varphi(x, \vec{Y}) \leftrightarrow \psi(x, \vec{Y})) \rightarrow \exists Z \forall x (x \in Z \leftrightarrow \varphi(x, \vec{Y})))$$

for each pair  $\varphi$  and  $\psi$  of  $\Sigma_1^0$  and  $\Pi_1^0$  formulas. From a computability-theoretic viewpoint, the scheme (2.1) tells us that if  $\vec{Y}$  is a tuple of parameters from a model  $(M, \mathcal{S}) \models \text{RCA}_0$ , and  $\bigoplus \vec{Y}$  pointwise computes a total, binary-valued function  $f: M \rightarrow 2$ , then  $f^{-1}(1) \in \mathcal{S}$ .

Since we will sometimes work with sets  $Z$  outside of  $\mathcal{S}$ , we must face a troublesome detail documented by Groszek and Slaman [13]: if  $\text{B}\Sigma_1^0$  does not hold relative to  $Z$ , then the ‘pointwise Turing reduction’ described in line (2.1) is not transitive, i.e., there are  $X$  and  $Y$  such that  $X$  is pointwise computable from  $Y$  and  $Y$  from  $Z$ , but  $X$  is not pointwise computable from  $Z$ —in symbols,  $X \in \Delta_1^{0,Y}$  and  $Y \in \Delta_1^{0,Z}$ , but  $X \notin \Delta_1^{0,Z}$ . For this reason we also introduce a stronger, ‘setwise’ notion of Turing reduction: we write  $X \leq_T Y$  if the sequence  $(X \upharpoonright n : n \in M)$ , viewed as a function from  $M$  to codes of finite sets, is pointwise computable from  $Y$ ; here  $X \upharpoonright n$  denotes the binary string representing the first  $n$  bits of  $X$ . In the language of [13], such an  $X$  is *strongly recursive in  $Y$* .

The classical Shoenfield Limit Lemma states that every total  $\Delta_2^0$  function  $F$  can be approximated by a recursive total function  $F_0$  in the sense that  $F = \lim_s F_0(\cdot, s)$ . A formalization of it will be used all over this

paper. We formulate the Limit Lemma in the following very general form because we want to exploit this extra generality in Section 5.

**Lemma 2.1** (Limit Lemma). *Let  $n \in \mathbb{N}$ . For every  $\theta(\bar{x}, y) \in \Sigma_{n+1}^0$ , there exist  $\eta(s, \bar{x}, y) \in \Sigma_n^0$  and  $\alpha(s, \bar{x}) \in \Pi_n^0$  such that  $\text{I}\Sigma_n^0$  proves*

$$\begin{aligned} & \forall \bar{x} \exists! y \theta(\bar{x}, y) \\ & \rightarrow \forall s, \bar{x} \exists! y \eta(s, \bar{x}, y) \\ & \wedge \forall \bar{x}, y (\theta(\bar{x}, y) \leftrightarrow \exists s \forall t \geq s \eta(t, \bar{x}, y)) \\ & \wedge \forall s, \bar{x} (\alpha(s, \bar{x}) \leftrightarrow \forall t \geq s \forall y, y' (\eta(s, \bar{x}, y) \wedge \eta(t, \bar{x}, y') \rightarrow y = y')). \end{aligned}$$

Here the formula  $\theta(\bar{x}, y)$  may contain undisplayed free variables, in which case we allow the same variables to appear free in  $\eta(s, \bar{x}, y)$  and  $\alpha(s, \bar{x})$ .

*Proof.* For the  $n \geq 1$  case, see Theorem I.3.2 in Hájek–Pudlák [14]. Note that  $\text{B}\Sigma_n^0$  is sufficient for this case. If  $n = 0$  and  $\theta = \exists u \theta_0$  where  $\theta_0 \in \Sigma_0^0$ , then we can define  $\eta(s, \bar{x}, y)$  and  $\alpha(s, \bar{x})$  to be

$$(y \leq s \wedge \exists u \leq s \theta_0(u, \bar{x}, y)) \vee (\forall y', u \leq s \neg \theta_0(u, \bar{x}, y') \wedge y = s)$$

and  $\exists y, u \leq s \theta_0(u, \bar{x}, y)$  respectively.  $\square$

To formulate our pigeonhole principles, let us borrow the Erdős–Rado arrow notation

$$\kappa \rightarrow (\lambda)_c^n,$$

which means that every  $c$ -coloring of the  $n$ -element subsets of  $\kappa$  admits a homogeneous subset of cardinality  $\lambda$ .

**Definition 2.2.** When  $\Gamma$  is a class of functions,

$$\Gamma: x \rightarrow (z)_y^1$$

means that if  $f: x \rightarrow y$  is in  $\Gamma$  then  $f$  is constant on a subset of  $x$  of cardinality  $z$ . Here every number  $x$  is identified with  $\{v : v < x\}$  as in set theory.

For example, in this notation, the usual  $\Sigma_{n+1}^0$  Pigeonhole Principle can be written as  $\forall x (\Sigma_{n+1}^0: x + 1 \rightarrow (2)_x^1)$ .

We first prove several first-order properties of these partition relations. The first-order results we present here can easily be relativized to the second-order setting.

As is well known, there is a level-by-level correspondence between the usual Pigeonhole Principle and the collection scheme.

**Theorem 2.3** (Dimitracopoulos and Paris [9]). *Over  $\text{I}\Delta_0 + \text{exp}$ ,*

$$\forall x (\Sigma_{n+1}^0: x + 1 \rightarrow (2)_x^1)$$

*is equivalent to  $\text{B}\Sigma_{n+1}$  for all  $n \in \mathbb{N}$ .*  $\square$

The next lemma, which can be viewed as a variant of Theorem 2.3, offers a number of  $B\Sigma_{n+1}$ -like characterizations of the property  $\Sigma_{n+1} : b \rightarrow (2)_a^1$ .

**Lemma 2.4.** *Fix  $n \in \mathbb{N}$ . Suppose that  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp}$  and  $a < b$  are elements of  $M$ . The following are equivalent:*

- (1)  $\mathfrak{M} \models \Sigma_{n+1} : b \rightarrow (2)_a^1$ .
- (2) *If  $\varphi$  is a  $\Pi_n$  formula and  $\mathfrak{M} \models (\forall x < b)(\exists y)\varphi(x, y)$ , then there exist an  $\mathfrak{M}$ -finite  $A \subseteq b$  of size  $|A| \geq a$  and an  $\ell \in M$  such that*

$$\mathfrak{M} \models (\forall x \in A)(\exists y < \ell)\varphi(x, y).$$

- (3) *As above, but  $\varphi$  is  $\Sigma_{n+1}$ .*
- (4) *For every  $\Delta_{n+1}$  function  $f : b \rightarrow M$  there exist an  $\mathfrak{M}$ -finite  $A \subseteq b$  of size  $|A| \geq a$  and  $\ell \in M$  such that  $f(x) < \ell$  for all  $x \in A$ .*
- (5) *As above, plus  $\{(x, f(x)) : x \in A\}$  is  $\mathfrak{M}$ -finite.*

*Proof.* (1  $\implies$  2). Define  $m : b \rightarrow M$  by setting

$$m(x) = \langle x, y \rangle \text{ whenever } y \text{ is least s.t. } \mathfrak{M} \models \varphi(x, y).$$

For each  $x < b$ , the set  $\{v < b : m(v) < m(x)\}$  is  $\mathfrak{M}$ -finite by  $\text{I}\Sigma_n + \text{exp}$ , and so it has a cardinality, say  $f(x)$ , from the point of view of  $\mathfrak{M}$ . This gives a function  $f : b \rightarrow M$  which is injective (since  $m$  is injective) and  $\Sigma_{n+1}$ . Hence  $\text{Im}(f) \not\subseteq a$  by (1). If  $x < b$  such that  $f(x) = |\{v < b : m(v) < m(x)\}| \geq a$ , then we can set  $A = \{v < b : m(v) < m(x)\}$  and  $\ell = m(x)$ .

(2  $\implies$  3). Write  $\varphi$  as  $(\exists z)\psi(x, y, z)$ , and apply (2) to  $\psi$ .

(3  $\implies$  4). Suppose  $f(x) = y$  is defined by the  $\Sigma_{n+1}$  formula  $\varphi(x, y)$  in  $\mathfrak{M}$ . Then  $M \models (\forall x < b)(\exists y)\varphi(x, y)$ , so (3) gives the required  $A$  and  $\ell$ .

(4  $\implies$  5). Suppose  $f(x) = y$  is defined by  $(\exists z)\psi(x, y, z)$ , where  $\psi$  is  $\Pi_n$ . Define

$$g(x) = \min\{\langle y, z \rangle \in M : \mathfrak{M} \models \psi(x, y, z)\}.$$

Then  $g$  is  $\Delta_{n+1}$ . Apply (4) to get  $A \subseteq b$  of size  $\geq a$  and  $\ell \in M$  such that  $x \in A$  implies  $g(x) < \ell$ . Now the set in (5) is equal to

$$\{(x, y) : x \in A \text{ and } \mathfrak{M} \models (\exists z < \ell)\psi(x, y, z)\},$$

which is  $\Pi_n$  and bounded, and is therefore  $\mathfrak{M}$ -finite by  $\text{I}\Sigma_n + \text{exp}$ .

(5  $\implies$  1). Suppose  $f : b \rightarrow a$  is a  $\Delta_{n+1}$  injection. Apply (5) to get an  $A \subseteq b$  of size  $\geq a$  on which the graph  $\{(x, y) : x \in A \text{ and } f(x) = y\}$  of  $f$  is  $\mathfrak{M}$ -finite. Through this and the injectivity of  $f$ , it follows that the image  $f(A)$  also has size  $\geq a$ , so that  $f(A)$  must equal the whole codomain  $a$ . But this is impossible, as  $f$  is an injection and  $A$  is a proper subset of the domain  $b$ . (As a closing sidenote: this is the only place where we use  $a < b$ .)  $\square$

Of particular interest to this paper is the pigeonhole principle for injections with domain twice as large as the codomain.

**Definition 2.5** ( $\Gamma$  Weak Pigeonhole Principle). Given a class  $\Gamma$  of functions,  $\Gamma$ -WPHP is the statement

$$\forall x \geq 1 (\Gamma : 2x \rightarrow (2)_x^1),$$

i.e., for no positive  $x$  is there an injection in  $\Gamma$  from  $2x$  to  $x$ .

It is natural to ask how the weak pigeonhole principle relates to other, similar statements about definable injections. We start by observing:

**Lemma 2.6.** *Let  $n \in \mathbb{N}$  and  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp}$ . Suppose we have a  $\Sigma_{n+1}$ -definable injection  $f: b \rightarrow a$  in  $\mathfrak{M}$ , where  $a < b$ . Then:*

- (1) *For each  $d \in M$  there is a  $\Sigma_{n+1}$  injection mapping  $db \rightarrow da$ .*
- (2) *For each  $c < a$  in  $M$ , either there is a  $\Sigma_{n+1}$  injection mapping  $b \rightarrow c$ , or there is one mapping  $(b - c) \rightarrow (a - c)$ .*
- (3) *For each nonzero, standard  $m \in \mathbb{N}$ , there is a  $\Sigma_{n+1}$  injection mapping  $\lceil b/m \rceil \rightarrow \lfloor a/m \rfloor$ .*

*Proof.* For the first part, define  $g: db \rightarrow da$  by  $g(qb + r) = qa + f(r)$  whenever  $0 \leq r < b$ . For the second part, assume that there is no  $\Sigma_{n+1}$  injection mapping  $b \rightarrow c$ , i.e., that  $\mathfrak{M} \models \Sigma_{n+1}: b \rightarrow (2)_c^1$ . Then by Lemma 2.4(5) there is an  $\mathfrak{M}$ -finite  $C \subseteq b$  of size  $c$  for which  $f(C)$  is  $\mathfrak{M}$ -finite and also of size  $c$ . So deleting  $C$  from the domain and  $f(C)$  from the codomain gives us an injection mapping a set of size  $b - c$  into one of size  $a - c$ . For the third part, by (2), either there is a  $\Sigma_{n+1}$  injection mapping  $b \rightarrow \lfloor a/m \rfloor$ , or there is one mapping  $(b - \lfloor a/m \rfloor) \rightarrow (a - \lfloor a/m \rfloor)$ . If it is the former, we are already done; if it is the latter, apply (2) again to get either an injection  $(b - \lfloor a/m \rfloor) \rightarrow \lfloor a/m \rfloor$ , or an injection  $(b - 2\lfloor a/m \rfloor) \rightarrow (a - 2\lfloor a/m \rfloor)$ . As the reader can readily check, continuing in this way for up to  $m$  steps, at some point we produce the required injection. (This is where we use that  $m$  is standard: nonstandardly many iterations of this sort would require stronger axioms in general.)  $\square$

This has interesting consequences when the three parts work in concert. For example, given an injection mapping  $b \rightarrow b/2$ , if we let  $c = b/4$ , then the lemma's second part can provide (in either outcome) an injection mapping  $(3/4)b \rightarrow b/4$ . Then the lemma's first and third parts together yield an injection  $b \rightarrow b/3$ —domain the same as we started with, but codomain markedly smaller. By carefully extending this reasoning, one can prove both parts of the following.

**Lemma 2.7.** *Suppose  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp}$  where  $n \in \mathbb{N}$ , and  $b \in M$ .*

- (1) *The following are equivalent.*
  - (a)  $\mathfrak{M} \models \Sigma_{n+1}: 2b \rightarrow (2)_b^1$ .
  - (b)  $\mathfrak{M} \models \Sigma_{n+1}: b \rightarrow (2)_{rb}^1$  for some  $r \in \mathbb{Q}$  strictly between 0 and 1.

- (c)  $\mathfrak{M} \models \Sigma_{n+1}: b \rightarrow (2)_{rb}^1$  for all  $r \in \mathbb{Q}$  strictly between 0 and 1.
- (2) If  $\mathfrak{M} \models \Sigma_{n+1}: 2b \rightarrow (2)_b^1$  then  $\mathfrak{M} \models \Sigma_{n+1}: 2a \rightarrow (2)_a^1$  for all  $a < b$ .  $\square$

### 3. THE WEAK WEAK KÖNIG LEMMA

We begin with a formalization of the principle  $2\text{-WWKL}_0(1/2)$ . In a model  $\mathfrak{M} = (M, \mathcal{S})$  of  $\text{RCA}_0$ , a (*binary*) *tree* is a set  $T$  of  $\mathfrak{M}$ -finite binary sequences such that every initial segment of an element of  $T$  is also an element of  $T$ . If  $T$  is a tree, then  $[T]$  denotes the collection of ( $\mathfrak{M}$ -)infinite binary sequences all of whose  $\mathfrak{M}$ -finite initial segments are in  $T$ . Here it is possible that  $T \notin \mathcal{S}$  or  $[T] \not\subseteq \mathcal{S}$ . When the tree  $T$  is  $L_2$ -definable, one can express  $[T] \neq \emptyset$  by an  $L_2$  formula. We denote this formula also by  $[T] \neq \emptyset$ , but beware that its truth in  $\mathfrak{M}$  means not only  $[T] \neq \emptyset$ , but also  $[T] \cap \mathcal{S} \neq \emptyset$ . If  $M = \mathbb{N}$  and  $T$  is a tree, then obviously  $T \cap 2^n$  is  $\mathfrak{M}$ -finite for each  $n \in M$ , and the Lebesgue measure of  $[T]$  is the limit of  $|T \cap 2^n|/2^n$  as  $n$  tends to infinity. However, when  $M \neq \mathbb{N}$ , the  $T \cap 2^n$ 's may not be  $\mathfrak{M}$ -finite. Even if all the  $T \cap 2^n$ 's are  $\mathfrak{M}$ -finite,  $\lim_n |T \cap 2^n|/2^n$  may not be as reasonable as in  $\mathbb{N}$ . So we have to impose some additional conditions on  $T$  when we want to talk about the measure of  $[T]$ . There are in fact several sets of such additional conditions, but as we shall see, the corresponding restrictions of  $2\text{-WWKL}_0$  are equivalent to one another.

**Definition 3.1.** Let  $T$  be a tree in a model  $\mathfrak{M} \models \text{RCA}_0$  and  $r \in \mathbb{Q}^{\mathfrak{M}}$ . We say that  $\mu([T]) \geq r$  if for each  $n \in M$ , there exists an  $\mathfrak{M}$ -finite  $S \subseteq T \cap 2^n$  such that  $|S| \geq r2^n$ .

Define  $2\text{-WWKL}(x)$  to be an  $L_2$  formula which expresses, over  $\text{RCA}_0$ , that  $x$  is a rational number (possibly nonstandard), and that

if  $T$  is a  $\Delta_2^0$  tree with  $\mu([T]) \geq x$  then  $[T] \neq \emptyset$ .

For a positive  $r \in \mathbb{Q}$ , let  $2\text{-WWKL}_0(r) = \text{RCA}_0 + 2\text{-WWKL}(r)$ .

If  $T$  is a  $\Delta_2^0$  tree then  $[T]$  is a  $\Pi_2^0$  class. So the principal axiom of  $2\text{-WWKL}_0$  is an instance of the axiom  $2\text{-POS}$ , which states that every  $\Pi_2^0$  class with positive measure is nonempty. Avigad et al. [2] proved that  $2\text{-WWKL}_0$  is equivalent over  $\text{RCA}_0$  to  $2\text{-POS}$ , and that  $2\text{-WWKL}_0$  implies  $2\text{-RAN}$ . We will establish parallel results here. Let us start with the definition of  $\Pi_2^0$  classes within  $\text{RCA}_0$ .

**Definition 3.2.** Fix  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0$ . A  $\Pi_2^0$  class in  $\mathfrak{M}$  is a subset  $\mathcal{A}$  of the power set of  $M$  (which we identify with the set  $2^M$  of all functions  $M \rightarrow 2$ ) that can be written in the form

$$\mathcal{A} = \bigcap_{i \in M} \bigcup_{j \in M} \mathcal{B}_{i,j},$$

where each  $\mathcal{B}_{i,j}$  is a basic open set (meaning  $\mathcal{B}_{i,j} = [\sigma_{i,j}] = \{X \in 2^M : \sigma_{i,j} \text{ is an initial segment of } X\}$  for some  $\mathfrak{M}$ -finite binary sequence  $\sigma_{i,j}$ ), and where the function mapping  $(i, j)$  to the code of  $\sigma_{i,j}$  is in  $\mathcal{S}$ . We say  $\mathcal{A}$  is *strictly presented* if  $i \leq i'$  implies  $\bigcup_{j < j'} \mathcal{B}_{i,j} \supseteq \bigcup_{j < j'} \mathcal{B}_{i',j}$  for every  $j' \in M$ . When  $\mathcal{A}$  is strictly presented,  $\delta \in \mathbb{Q}^{\mathfrak{M}}$ , and  $\mu(\bigcup_j \mathcal{B}_{i,j}) \geq \delta$  for all  $i \in M$ , we write  $\hat{\mu}(\mathcal{A}) \geq \delta$ . Every presentation of a  $\Pi_2^0$  class  $\mathcal{A}$  in  $\mathfrak{M}$  gives rise to an  $L_2$  formula, possibly with parameters from  $\mathfrak{M}$ , that expresses  $\mathcal{A} \neq \emptyset$ . We denote such a formula also by  $\mathcal{A} \neq \emptyset$ , so that

$$\mathfrak{M} \models \mathcal{A} \neq \emptyset \quad \Leftrightarrow \quad \mathcal{A} \cap \mathcal{S} \neq \emptyset.$$

It is routine to check that every  $\Pi_2^0$  class in a model of  $\text{RCA}_0$  (except perhaps the empty class) can be written in strictly presented form, and hence can have its measure compared with rationals in this way; let us stress, however, that without  $\text{B}\Sigma_2^0$ , this measure may depend on the choice of the ‘presentation’.

**Definition 3.3.** Let  $2\text{-POS}(x)$  be an  $L_2$  formula which expresses, over  $\text{RCA}_0$ , that  $x$  is a rational number (possibly nonstandard), and that

$$\text{if } \mathcal{A} \text{ is a } \Pi_2^0 \text{ class and } \hat{\mu}(\mathcal{A}) \geq x \text{ then } \mathcal{A} \neq \emptyset.$$

Recall that if  $T$  is a  $\Delta_2^0$  tree then  $[T]$  is a  $\Pi_2^0$  class. So we have two kinds of measure inequalities:  $\mu([T]) \geq r$  and  $\hat{\mu}([T]) \geq r$ . In Lemma 3.5 below we will show that  $\mu([T]) \geq r$  implies  $\hat{\mu}([T]) \geq r$  over  $\text{RCA}_0$ , and hence  $\text{RCA}_0 \vdash 2\text{-POS}(r) \rightarrow 2\text{-WWKL}_0(r)$  for all positive  $r \in \mathbb{Q}$ ; and later, in Theorem 3.14, we obtain a strengthening and a converse. But first, let us pause for a useful lemma about  $\Delta_2^0$  trees.

One way to state the  $n = 1$  case of the Limit Lemma 2.1 is: Given a  $\Delta_2^0$  set  $A$  of natural numbers, there is a sequence of sets  $\langle A_0, A_1, \dots \rangle$  which converges pointwise to  $A$  and is uniformly  $\Delta_1^0$ , i.e. the function  $F(x, s)$  which  $= 1$  if  $x \in A_s$  and  $= 0$  if  $x \notin A_s$ , is in the second-order part of the model. The following refinement of the Limit Lemma says that if, in addition,  $A$  is (the set of codes of strings in) a binary tree, then the approximating sets can also be trees (i.e. can be closed under initial segment).

**Lemma 3.4** ( $\text{RCA}_0$ ). *Every  $\Delta_2^0$  tree  $T \subseteq 2^{<M}$  is the limit of a uniformly  $\Delta_1^0$  sequence of trees  $\langle T_0, T_1, \dots \rangle$ .*

With access to  $\text{B}\Sigma_2^0$ , this lemma would be immediate. Since only  $\text{RCA}_0$  is available, however, we resort to a ‘tame cuts’-style construction.

*Proof.* Let  $\langle A_0, A_1, \dots \rangle$  be a uniformly  $\Delta_1^0$  sequence of sets approximating  $T$  as given by the Limit Lemma, and let

$$m(\sigma, s) = \min \{r \geq |\sigma| : \forall t (r \leq t \leq s \implies [\sigma \in A_t \iff \sigma \in A_s])\}.$$

In other words,  $m$  is a stage-by-stage approximation to the usual modulus function. For each  $s$ , define a set  $T_s$  of strings by:

$$\sigma \in T_s \text{ if } (\forall \tau \subseteq \sigma)(\exists \rho \supseteq \tau) [\rho \in A_s \text{ and } m(\rho, s) \leq m(\tau, s)].$$

It is immediate from this definition that each  $T_s$  is a tree. And since  $m(\rho, s)$  is by definition always  $\geq |\rho|$ , to determine whether a given  $\sigma$  is in  $T_s$ , we need only consider strings  $\rho$  and  $\tau$  of length  $\leq s$ ; hence the trees are uniformly  $\Delta_1^0$ . It remains only to verify that they converge pointwise to  $T$ . For a given  $\sigma \in 2^{<M}$ , there are two cases to consider: either  $\sigma$  is in  $T$ , or it is not.

If  $\sigma \in T$ : Let  $r \geq |\sigma|$  be least such that  $t \geq r$  implies  $\sigma \in A_t$ . Since  $T$  is downward closed, there is by  $\text{IS}_1$  an  $s \geq r$  such that for all  $\tau \subseteq \sigma$  either  $\tau \in A_s$  or  $m(\tau, s) > r$ . In particular, for any  $t \geq s$  and any  $\tau \subseteq \sigma$ , we have either  $\tau \in A_t$  or  $m(\tau, t) > r$ . Since  $r = m(\sigma, t)$ , this means  $\sigma$  is in  $T_t$  for all  $t \geq s$ .

If  $\sigma \notin T$ : Let  $r \geq |\sigma|$  be least such that  $t \geq r$  implies  $\sigma \notin A_t$ . By  $\text{BS}_1$  (and using the fact that  $m(\tau, t)$  is always at least  $|\tau|$ ) there is an  $s \geq r$  such that for all  $\tau \supseteq \sigma$ , either  $\tau \notin A_s$  or  $m(\tau, s) > r$ . And in particular, for any  $t \geq s$  and any  $\tau \supseteq \sigma$ , we have either  $\tau \notin A_t$  or  $m(\tau, t) > r$ . Since  $r = m(\sigma, t)$ , this means  $\sigma$  is not in  $T_t$  for any  $t \geq s$ .  $\square$

We put this lemma straight to work.

**Lemma 3.5** ( $\text{RCA}_0$ ). *If  $r \in \mathbb{Q}$  is positive and  $T$  is a  $\Delta_2^0$  tree with  $\mu([T]) \geq r$  according to Definition 3.1, then  $[T]$  can be strictly presented as a  $\Pi_2^0$  class  $\mathcal{A}$  such that  $\hat{\mu}(\mathcal{A}) \geq r$  according to Definition 3.3.*

*Proof.* Fix  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0$  and  $T$  and  $r$  as in the statement. If  $T$  is bounded in height then  $[T]$  is empty, its measure is zero, and the proof is trivial, so assume that it is not. Let  $\langle T_0, T_1, \dots \rangle$  be a uniformly  $\Delta_1^0$  sequence of trees converging pointwise to  $T$ , as given by the previous Lemma. Further assume, by cutting the tops off if necessary, that each  $T_s$  is empty above level  $s$ . Our construction is as follows.

Whenever  $\sigma$  is in some  $T_s$ , select an unused  $j$  and set  $\mathcal{B}_{|\sigma|,j} = [\sigma] = \{X \in 2^M : \sigma \text{ is an initial segment of } X\}$ . Some care is needed in selecting the  $j$ 's so that the resulting  $\Pi_2^0$  class is strictly presented, but this is easily done: for instance,  $j = \langle f(\sigma), s \rangle$  will do, where  $f(\sigma) = 2^{|\sigma|} + \sum_{k < |\sigma|} \sigma(k) \cdot 2^k$ . For all other  $i, j$ , just set  $\mathcal{B}_{i,j}$  to equal some other  $\mathcal{B}_{i,j'}$  (which may involve waiting, if no other  $\mathcal{B}_{i,j'}$  has yet been defined).

It remains to verify that  $\mathcal{A} = \bigcap_i \bigcup_j \mathcal{B}_{i,j}$  equals  $[T]$ , and has measure  $\hat{\mu}(\mathcal{A}) \geq r$ . The former claim is a consequence of the starting assumption that each  $T_s$  be downward closed and of height  $\leq s$ . For the latter claim, it is enough to notice that if the tree's  $i$ -th level  $T \cap 2^i$  has exactly  $k$  elements, corresponding to a measure of  $k/2^i$ , then the union  $\bigcup_j \mathcal{B}_{i,j}$  contains at least  $k$ -many disjoint cylinders each of measure  $2^{-i}$ , totalling  $\geq k/2^i$ .  $\square$

**Corollary 3.6** ( $\text{RCA}_0$ ).  $2\text{-POS}(r)$  implies  $2\text{-WWKL}(r)$  for all positive  $r \in \mathbb{Q}$ .

*Proof.* Just apply Lemma 3.5.  $\square$

To get the other direction, we look at  $\Delta_2^0$  trees that behave regularly.

**Definition 3.7.** A regular  $\Delta_2^0$  tree in a model  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0$  is a  $\Delta_2^0$  tree  $T$  over  $\mathfrak{M}$  such that  $(T \cap 2^i : i < n)$  is  $\mathfrak{M}$ -finite for every  $n \in M$ .

Such a regularity property does not come for free in a nonstandard world: if  $\mathfrak{M} \not\models \text{B}\Sigma_2^0$ , then it is not hard to produce a  $\Delta_2^0$ -definable function  $F: M \rightarrow 2$  and  $n \in M$  such that  $(F(i) : i < n)$  is not  $\mathfrak{M}$ -finite. One can avoid this irregularity, provably in  $\text{RCA}_0$ , by passing on to a  $\Delta_2^0$ -definable non-decreasing cofinal sequence whose elements increase sufficiently rarely.

**Lemma 3.8.** In a model  $(M, \mathcal{S}) \models \text{RCA}_0$ , if  $F: M \rightarrow M$  is total  $\Delta_2^0$ , then there is a  $\Delta_2^0$  sequence  $(n_i)_{i \in M}$  such that  $i \leq j$  implies  $i \leq n_i \leq n_j$  and the function

$$j \mapsto \langle F(n_i) : i < j \rangle$$

is total  $\Delta_2^0$ .

*Proof.* Work in  $(M, \mathcal{S})$ . Let  $F = \lim_s F_0(\cdot, s)$  as in the Limit Lemma 2.1. Define the modulus function  $m: M \rightarrow M$  by

$$m(x) = \min\{s > x : \forall t \geq s F_0(x, s) = F_0(x, t)\}.$$

Then  $m$  is total and  $\Pi_1^0$  (i.e. its graph  $\{(x, s) : m(x) = s\}$  is  $\Pi_1^0$ ), as one can easily verify using  $\text{I}\Sigma_1^0$ . If we write  $m^k$  to mean  $m$  composed with itself  $k$  times, then the (possibly partial) function  $k \mapsto m^k(0)$  is strictly increasing and  $\Sigma_2^0$ . The domain of this function is clearly closed under successor. Therefore, since  $\text{I}\Sigma_1^0$  implies  $\{\langle k, n \rangle : m^k(0) = n \leq i\}$  is coded for every  $i \in M$ , we see that  $\forall i \exists k (m^k(0) > i)$ . Define the sequence  $(n_i)_{i \in M}$  by

$$n_i = \min\{m^k(0) : m^k(0) > i\}.$$

Notice if  $i, j, k, \ell \in M$  such that  $m^k(0) = n_i$  and  $m^\ell(0) = n_j$ , where  $i < j$ , then  $m(n_i) = m^{k+1}(0) \leq m^{\ell+1}(0) = m(n_j)$  and so  $F(n_i) = F_0(n_i, m(n_j))$ . A moment of thought then reveals  $j \mapsto \langle F(n_i) : i < j \rangle$  is total and  $\Delta_2^0$ .  $\square$

The following proposition is a strengthening of Avigad et al. [2, Proposition 3.4], which is in turn a formalization of Kurtz [21, p. 21, Lemma 2.2a], which is ultimately just an effective account of the regularity property of the Lebesgue measure on the Borel sets. The only difference between our proposition and that in [2] is that we require only  $\text{RCA}_0$ , rather than  $\text{RCA}_0 + \text{B}\Sigma_2^0$ ; and the only real difference between our proof and that in [2] is that by more carefully defining a

certain function  $F$  we are able to replace an appeal to  $\text{B}\Sigma_2^0$  with one to the lemma above.

**Proposition 3.9** ( $\text{RCA}_0$ ). *If  $\mathcal{A}$  is a  $\Pi_2^0$  class and  $\hat{\mu}(\mathcal{A}) \geq r > \delta > 0$ , then there exists a regular  $\Delta_2^0$  tree  $T$  such that  $[T] \subseteq \mathcal{A}$  and  $\mu([T]) \geq r - \delta$ .*

*Proof.* Work in a model  $(M, \mathcal{S}) \models \text{RCA}_0$ . Let

$$\mathcal{A} = \bigcap_i \bigcup_j \mathcal{B}_{i,j}$$

be a strictly presented  $\Pi_2^0$  class with  $\hat{\mu}(\mathcal{A}) \geq r$ , where  $\mathcal{B}_{i,j} = [\sigma_{i,j}]$  and the map  $(i, j) \mapsto \sigma_{i,j}$  is  $\Sigma_1^0$ . Moreover, we may assume that  $|\sigma_{i,j}| > i$ . For each  $i$ , let

$$F(i) = \min \left\{ k : \mu \left( \bigcup_{j < \ell} \mathcal{B}_{i',j} - \bigcup_{j < k} \mathcal{B}_{i',j} \right) < \frac{\delta}{2^{i+1}} \text{ for all } \ell > k \text{ and all } i' \leq i \right\}.$$

Then  $F$  is a  $\Delta_2^0$  total function by  $\text{I}\Sigma_1^0$ . Let  $(n_i)_{i \in M}$  be as in Lemma 3.8 for  $F$ . Intuitively, we will define  $T$  such that  $[T] = \bigcap_i \bigcup_{j < F(n_i)} \mathcal{B}_{i,j}$ . This will ensure  $[T] \subseteq \mathcal{A}$  and  $\mu([T]) \geq r - \delta$  if  $\hat{\mu}(\mathcal{A}) \geq r > \delta > 0$  by the definition of  $F$ .

Formally, we define a tree  $T$  as follows. For each binary sequence  $\tau$  of length  $\ell$ , put  $\tau \in T$  if and only if for each  $i \leq \ell$  there is some  $j < F(n_i)$  such that  $\sigma_{i,j}$  is comparable with  $\tau$ . It follows straight from the definition that  $T$  is a tree. Moreover  $T$  is a regular  $\Delta_2^0$  tree because  $T \cap 2^\ell$  can be computed uniformly from  $\langle F(n_i) : i \leq \ell \rangle$ . The rest is a simple exercise.  $\square$

From this lemma we can derive a partial reversal of Corollary 3.6.

**Proposition 3.10** ( $\text{RCA}_0$ ). *If  $q$  and  $r$  are elements of  $\mathbb{Q}$  and  $0 < q < r$  then  $2\text{-WWKL}(q)$  restricted to regular  $\Delta_2^0$  trees implies  $2\text{-POS}(r)$ .  $\square$*

To prove a full reversal of Corollary 3.6, i.e., the equivalence of  $2\text{-POS}(r)$  and  $2\text{-WWKL}(r)$  over  $\text{RCA}_0$ , we use the Weak Pigeonhole Principle. In the next lemma, let us establish our first connection between  $2\text{-WWKL}_0(1/2)$  and  $\Sigma_2^0\text{-WPHP}$ , by a proof similar to Avigad et al's [2] proof that  $2\text{-WWKL}_0$  implies  $\text{B}\Sigma_2^0$ , or the classical proof that  $\text{WKL}$  implies  $\text{B}\Sigma_1^0$  [11, Proposition 5].

**Lemma 3.11.**  *$2\text{-WWKL}_0(1/2)$  restricted to regular  $\Delta_2^0$  trees proves  $\Sigma_2^0\text{-WPHP}$ .*

*Proof.* Fix a model  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0 + \neg\Sigma_2^0\text{-WPHP}$ . Then there is a  $b \in M$  for which  $\mathfrak{M} \not\models \Sigma_2^0: 2b \rightarrow (2)_b^1$ . By Lemma 2.7(2), the set of such  $b$ 's is closed upwards in  $M$ ; so let us choose  $b$  to be a power of 2, say  $2b = 2^k$ . Identify the binary strings of length  $k$  uniquely with the numbers  $< 2b$ , e.g. by placing them in alphabetical order.

Use Lemma 2.4(2) to fix a  $\Pi_1^0$  formula  $\varphi$  such that  $\mathfrak{M} \models (\forall x < 2b)(\exists y)\varphi(x, y)$ , but for every  $\ell \in M$ , the set

$$\{x < 2b : \mathfrak{M} \models (\exists y < \ell)\varphi(x, y)\}$$

has size strictly less than  $b$ . Define a tree  $T$  by

$$T = 2^{<k} \cup \{\sigma \in 2^{\geq k} : \mathfrak{M} \models (\forall y < |\sigma|)\neg\varphi(\sigma \upharpoonright k, y)\}.$$

Then  $T$  is regular  $\Delta_2^0$  and has measure  $\geq 1/2$ , but has no infinite paths. So 2-WWKL $_0(1/2)$  fails.  $\square$

Our proof of the equivalence of 2-WWKL(1/2) and 2-POS(1/2) over  $\text{RCA}_0$  invokes a weak form of the Lebesgue Density Theorem, which will also be useful in Section 4 where we demonstrate the conservativity of 2-WWKL $_0(1/2)$  over  $\text{IS}_1 + \Sigma_2\text{-WPHP}$ .

**Definition 3.12.** If  $T$  is a tree and  $\sigma$  is a finite binary sequence, then let  $T_\sigma = \{\tau : \sigma\tau \in T\}$ .

Fix  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0$ . If  $T$  is a tree and  $\sigma$  is an  $\mathfrak{M}$ -finite binary sequence then  $T_\sigma$  is also a tree, and is computable in  $T$ . Moreover,  $T_\sigma$  is a regular  $\Delta_2^0$  tree if  $T$  is. If we work in a standard model—meaning  $M = \mathbb{N}$ —then by the Lebesgue Density Theorem every tree  $T$  with  $\mu([T]) > 0$  has nodes  $\sigma$  such that  $\mu([T_\sigma])$  is very close to 1. This fails in general for  $\Delta_2^0$  trees in the absence of  $\text{IS}_2^0$  (see [6]), but with  $\Sigma_2^0\text{-WPHP}$  it holds partially.

**Lemma 3.13** (Partial Lebesgue Density). *The following is provable in  $\text{RCA}_0 + \Sigma_2^0\text{-WPHP}$  for all nonzero  $n, m \in \mathbb{N}$ .*

*For any regular  $\Delta_2^0$  tree  $T$  of measure  $\geq 1/n$ , there is a string  $\sigma$  such that  $T_\sigma$  has measure  $\geq 1 - 1/m$ . Furthermore, given any  $b$ , we can ensure that  $|\sigma| \geq b$ .*

*Proof.* Fix a model  $(M, \mathcal{S}) \models \text{RCA}_0 + \Sigma_2^0\text{-WPHP}$  to work in. Let  $T$  be a regular  $\Delta_2^0$  tree of measure  $\geq 1/n$ . Since  $m$  and  $n$  are standard, we may choose the unique  $k$  such that

$$(3.1) \quad \frac{k}{2mn} \leq \mu([T]) < \frac{k+1}{2mn}.$$

(This is the only place where we use standardness.) Notice that  $k \geq 2m$ . Now fix a level  $\ell_0$  at which

$$\frac{k}{2mn} \leq \frac{|T \cap 2^{\ell_0}|}{2^{\ell_0}} < \frac{k+1}{2mn},$$

and assume towards a contradiction that  $\mu([T_\sigma]) < 1 - 1/m$  for all  $\sigma \in T \cap 2^{\ell_0}$ ; in other words, each  $\sigma \in T \cap 2^{\ell_0}$  has a level  $\ell$  at which  $|T_\sigma \cap 2^\ell|/2^\ell < 1 - 1/m$ . The function mapping  $\sigma$  to the least such  $\ell$  is

$\Delta_2^0$ , so by  $\Sigma_2^0$ -WPHP and Lemma 2.4(4), there is an  $\ell_1 \in M$  such that  $|T_\sigma \cap 2^{\ell_1}|/2^{\ell_1} < 1 - 1/m$  for more than half of these  $\sigma$ 's. Hence

$$\begin{aligned} \frac{|T \cap 2^{\ell_0+\ell_1}|}{2^{\ell_0+\ell_1}} &< \frac{1}{2^{\ell_0+\ell_1}} \left( \frac{|T \cap 2^{\ell_0}|}{2} \cdot 1 \cdot 2^{\ell_1} + \frac{|T \cap 2^{\ell_0}|}{2} \cdot \left(1 - \frac{1}{m}\right) \cdot 2^{\ell_1} \right) \\ &= \frac{1 + (1 - 1/m)}{2} \cdot \frac{|T \cap 2^{\ell_0}|}{2^{\ell_0}} \\ &< \frac{2m - 1}{2m} \cdot \frac{k + 1}{2mn}, \end{aligned}$$

which is  $< k/2mn$  since  $k \geq 2m$ . But this means  $\mu([T]) < k/2mn$ , contradicting line (3.1).

As for the ‘furthermore’ part of the lemma, simply note that when selecting  $\ell_0$ , we may select it to be larger than any given  $b \in M$ .  $\square$

**Theorem 3.14.** *The following statements are equivalent over  $\text{RCA}_0$  for all  $r \in (0, 1) \cap \mathbb{Q}$ .*

- (1) 2-POS(1/2).
- (2) 2-WWKL(1/2).
- (3) 2-WWKL(1/2) restricted to regular  $\Delta_2^0$  trees.
- (4) 2-POS( $r$ ).
- (5) 2-WWKL( $r$ ).
- (6) 2-WWKL( $r$ ) restricted to regular  $\Delta_2^0$  trees.

*Proof.* The implications (1)  $\Rightarrow$  (2) and (4)  $\Rightarrow$  (5) are special cases of Corollary 3.6. The implications (2)  $\Rightarrow$  (3) and (5)  $\Rightarrow$  (6) are trivial. We claim that 2-WWKL $_0(q)$  restricted to regular  $\Delta_2^0$  trees implies 2-POS( $r$ ) for all  $q, r \in (0, 1) \cap \mathbb{Q}$ . This suffices to entail the remaining implications (3)  $\Rightarrow$  (4) and (6)  $\Rightarrow$  (1).

Work over  $\text{RCA}_0$  plus 2-WWKL( $q$ ) restricted to regular  $\Delta_2^0$  trees. By Lemmas 3.11 and 3.13, we know 2-WWKL $_0(r/2)$  holds. So 2-POS( $r$ ) follows from Proposition 3.10.  $\square$

In view of the equivalences above, it does not matter whether we use Definition 3.1 or Definition 3.3 when we speak of measure of  $\Delta_2^0$  trees. For the sake of consistency, we will adopt the former in what follows.

Theorem 3.14 implies the following strengthening of a theorem by Avigad et al. [2, Proposition 3.6].

**Corollary 3.15.**  $\text{RCA}_0 \vdash 2\text{-WWKL}(1/2) \rightarrow 2\text{-RAN}$ .  $\square$

#### 4. CONSERVATIVITY

In this section, we will determine the first-order theory of 2-WWKL $_0(1/2)$ . In particular, we will see that it is finitely axiomatizable.

**Theorem 4.1.** *2-WWKL $_0(1/2)$  is  $\Pi_1^1$ -conservative over  $\text{RCA}_0 + \Sigma_2^0$ -WPHP. Hence  $\text{I}\Sigma_1 + \Sigma_2$ -WPHP axiomatizes the first-order theory of  $\text{RCA}_0 + 2\text{-WWKL}_0(1/2)$ .*

*Proof.* Conservativity will follow from the model expansion theorem 4.3 below in the usual manner. The remaining part is provided by Lemma 3.11.  $\square$

The relevant model expansion theorem will occupy us for the rest of this section.

**Definition 4.2.** A model  $(M, \mathcal{S}) \models \text{RCA}_0$  is *principal* if  $\mathcal{S} = \{X \subseteq M : X \leq_T Z\}$  for some  $Z \subseteq M$ . If  $\mathfrak{M} = (M, \mathcal{S})$  and  $X \subseteq M$ , then

$$\mathfrak{M}[X] = (M, \{Y \subseteq M : Y \leq_T Z \oplus X \text{ for some } Z \in \mathcal{S}\}).$$

We read  $\mathfrak{M}[X]$  as  $\mathfrak{M}$  *expanded by*  $X$ .

**Theorem 4.3.** *Let  $\mathfrak{M}$  be a countable principal model of  $\text{RCA}_0 + \Sigma_2^0$ -WPHP. If  $T$  is a regular  $\Delta_2^0(\mathfrak{M})$  tree of measure  $\geq 1/2 + \epsilon$ , where  $\epsilon \in \mathbb{Q}$  with  $0 < \epsilon < 1/2$ , then there exists  $X \in [T]$  such that  $\mathfrak{M}[X]$  is also a model of  $\text{RCA}_0 + \Sigma_2^0$ -WPHP.*

Our proof involves a relativized version of Jockusch–Soare forcing with the restriction that conditions must in a sense have large measure. The relativization will be carried out carefully so as to keep down the amount of induction used. In particular, we will start the forcing construction from a specially chosen condition, to be given by Lemma 4.6. Let us start by setting up some notation.

**Definition 4.4** ( $\text{RCA}_0^*$ ). Recall that  $\Phi_e^Z(x)$  denotes the  $e$ -th Turing functional, run with oracle  $Z$  on input  $x$ , and  $\Phi_{e,s}^Z(x)$  is the same but run for only  $s$ -many steps. Let  $\Phi_e^Z(\sigma; x)$  abbreviate  $\Phi_{e,|\sigma|}^{Z \oplus \sigma}(x)$ .

For example, with these notations, the Turing jump of  $Z \oplus X$  can be expressed in more than one way:

$$(Z \oplus X)' = \{e : \Phi_e^{Z \oplus X}(e) \downarrow\} = \{e : \exists \ell \Phi_e^Z(X \upharpoonright \ell; e) \downarrow\}.$$

Lemma 4.6 below is the same as Proposition 1.3 in Conidis–Slaman [7], except that the base theory is weakened from  $\text{RCA}_0 + \text{B}\Sigma_2^0$  to  $\text{RCA}_0$ . For this improvement, we carefully replace applications of  $\text{B}\Sigma_2^0$  with those of Lemma 4.5.

**Lemma 4.5** ( $\text{RCA}_0$ ). *Suppose  $F_0$  is a recursive function with two arguments, and there is a recursive function  $G$  such that for all  $x$ ,*

$$|\{s : F_0(x, s) \neq F_0(x, s+1)\}| < G(x).$$

*Then  $F(x) = \lim_s F_0(x, s)$  exists for all  $x$ , and the map  $n \mapsto \langle F(x) : x < n \rangle$  is total and  $\Delta_2^0$ .*

*Proof.* Define  $G^*(n) = G(0) + \dots + G(n-1)$ . Then  $G^*$  is total  $\Delta_1^0$ , and

$$|\{s : (\exists x < n) F_0(x, s) \neq F_0(x, s+1)\}| < G^*(n)$$

for all  $n$ . Since the set displayed above is coded, each  $n$  corresponds to some maximum value  $s_n$  of  $s$  (or, let us say, to zero should the

set be empty); moreover, the function mapping  $n$  to  $s_n$  is  $\Delta_2^0$ . From this we get a total  $\Delta_2^0$  function mapping each  $n$  to  $\langle F(x) : x < n \rangle = \langle F_0(x, s_n + 1) : x < n \rangle$ .  $\square$

Now we state and prove Lemma 4.6. The key idea of the proof is borrowed from the classical proof that every 2-random is generalized low.

**Lemma 4.6.** *Let  $\mathfrak{M} = (M, \mathcal{S}) \models \text{RCA}_0$  and  $Z \in \mathcal{S}$ . Then for every positive  $\epsilon \in \mathbb{Q}^{\mathfrak{M}}$  there are a non-decreasing  $\Delta_2^{0,Z}$  function  $H$  and a regular  $\Delta_2^{0,Z}$  tree  $\tilde{T}$  of measure  $\geq 1 - \epsilon$  such that for every  $X \in [\tilde{T}]$ ,*

$$(Z \oplus X)' = \{e : \Phi_e^Z(X \upharpoonright H(e); e) \downarrow\}$$

in  $\mathfrak{M}[X]$ .

*Proof.* Work in  $\mathfrak{M}$ . Fix a rational  $\epsilon > 0$ . For each pair  $(e, \ell)$ , define

$$p_{e,\ell} = \frac{|\{\sigma \in 2^\ell : \Phi_e^Z(\sigma; e) \uparrow\}|}{2^\ell}.$$

For any fixed  $e$ , we have  $1 \geq p_{e,0} \geq \dots \geq p_{e,\ell} \geq \dots \geq 0$ ; thus if we define a  $Z$ -recursive binary function  $F_0$  by  $F_0(e, 0) = 0$ , and

$$F_0(e, s + 1) = \begin{cases} s + 1, & \text{if } p_{e,F_0(e,s)} - p_{e,s+1} \geq \epsilon/2^{e+1}; \\ F_0(e, s), & \text{otherwise,} \end{cases}$$

there are at most  $\lfloor 2^{e+1}/\epsilon \rfloor$ -many  $s$ 's for which  $F_0(e, s) \neq F_0(e, s + 1)$ . Hence by Lemma 4.5 the pointwise limit  $F(\cdot) = \lim_s F_0(\cdot, s)$  exists, and the map  $n \mapsto \langle F(e) : e < n \rangle$  is total and  $\Delta_2^{0,Z}$ . From this we obtain a non-decreasing total function  $H : n \mapsto \max\{F(e) : e \leq n\}$ . As the reader can directly verify, the image of  $H$  is unbounded, and the map  $n \mapsto \langle H(e) : e < n \rangle$  is total  $\Delta_2^{0,Z}$ . Define a tree  $\tilde{T}$  level-by-level as follows: all the coded binary strings of length strictly less than  $H(0)$  are in  $\tilde{T}$ , and if  $H(n) \leq \ell < H(n + 1)$ , then the  $\ell$ -th level of  $\tilde{T}$  is

$$\tilde{T}_\ell = \{\sigma \in 2^\ell : \forall e < n \text{ (either } \Phi_e^Z(\sigma; e) \uparrow \text{ or } \Phi_e^Z(\sigma \upharpoonright F(e); e) \downarrow)\}.$$

It is straightforward to check that  $\tilde{T}$  is a regular  $\Delta_2^{0,Z}$  tree, and  $X \in [\tilde{T}]$  implies  $\forall e (\Phi_e^{Z \oplus X}(e) \uparrow \text{ or } \Phi_e^Z(X \upharpoonright H(e); e) \downarrow)$ .

As for the measure, consider the complement  $\tilde{T}_\ell^c$  of  $\tilde{T}_\ell$ . Whenever  $H(n) \leq \ell < H(n + 1)$ , we have  $F(e) \leq H(n) \leq \ell$  for all  $e < n$ , and so

$$\begin{aligned} \frac{|\tilde{T}_\ell^c|}{2^\ell} &= \frac{|\{\sigma \in 2^\ell : \exists e < n (\Phi_e^Z(\sigma; e) \downarrow \text{ and } \Phi_e^Z(\sigma \upharpoonright F(e); e) \uparrow)\}|}{2^\ell} \\ &\leq \sum_{e < n} \frac{|\{\sigma \in 2^\ell : \Phi_e^Z(\sigma; e) \downarrow \text{ and } \Phi_e^Z(\sigma \upharpoonright F(e); e) \uparrow\}|}{2^\ell} \\ &= \sum_{e < n} (p_{e,F(e)} - p_{e,\ell}) < \sum_{e < n} \frac{\epsilon}{2^{e+1}} < \epsilon. \end{aligned}$$

Hence  $\mu([\tilde{T}]) \geq 1 - \epsilon$ , as desired.  $\square$

We follow this up with a quick, technical lemma.

**Lemma 4.7.** *Fix a principal model  $\mathfrak{M} = (M, \{Y \subseteq M : Y \leq_T Z\}) \models \text{RCA}_0$ . Let  $H$  be a non-decreasing  $\Delta_2^{0,Z}$  function, and let  $X \subseteq M$  such that  $X \upharpoonright b$  is  $\mathfrak{M}$ -finite for every  $b \in M$ , and such that*

$$(Z \oplus X)' = \{e : \Phi_e^Z(X \upharpoonright H(e); e) \downarrow\}$$

*in  $\mathfrak{M}[X]$ . Then  $\mathfrak{M}[X] \models \text{RCA}_0$  and every  $\Sigma_2^{0,Z \oplus X}$  set is  $\Sigma_1^{0,Z' \oplus X}$  in  $\mathfrak{M}[X]$ .*

*Proof.* To get  $\text{RCA}_0$ , it is enough to show that  $\mathfrak{M}[X] \models \text{IS}_1^0$ , and for this it suffices to show that  $(Z \oplus X)' \upharpoonright b$  is  $\mathfrak{M}$ -finite for all  $b \in M$ . Since  $H$  is non-decreasing, in  $\mathfrak{M}[X]$  we have

$$(Z \oplus X)' \upharpoonright b = \{e < b : \Phi_e^Z(X \upharpoonright H(b); e) \downarrow\}.$$

This set, when evaluated in  $\mathfrak{M}$ , must be  $\mathfrak{M}$ -finite because  $X \upharpoonright H(b)$  is  $\mathfrak{M}$ -finite and  $\mathfrak{M} \models \text{IS}_0^0 + \text{exp}$ .

As  $\mathfrak{M}[X] \models \text{BS}_1^0 + \text{exp}$ , every  $\Sigma_2^{0,Z \oplus X}$  set is  $\Sigma_1^{0,(Z \oplus X)'}$  in  $\mathfrak{M}[X]$ . So for the Lemma's second claim, it is enough to show that  $(Z \oplus X)' \leq_T Z' \oplus X$ . Consider the definition of  $(Z \oplus X)' \upharpoonright b$  given in the line displayed above. Since  $H$  is  $\Delta_2^{0,Z}$ , we can compute this set from  $Z' \oplus X$  by first asking for the value of  $H(b)$ ; then asking for  $X \upharpoonright H(b)$  and  $Z \upharpoonright H(b)$ ; and finally checking whether each  $\Phi_e^{Z \upharpoonright H(b)}(X \upharpoonright H(b); e)$  halts or not.  $\square$

Now we are ready for the actual construction.

*Proof of Theorem 4.3.* Fix a principal countable model  $\mathfrak{M} = (M, \{Y \subseteq M : Y \leq_T Z\}) \models \text{RCA}_0 + \Sigma_2^0\text{-WPHP}$  and a rational  $\epsilon \in (0, 1/2)$ . Consider any regular  $\Delta_2^0(\mathfrak{M})$  tree  $T$  of measure  $\geq 1/2 + \epsilon$ . We will force a branch through  $T$  while preserving  $\text{RCA}_0 + \Sigma_2^0\text{-WPHP}$ .

Let  $\tilde{T}$  be the result of applying Lemma 4.6 to  $\mathfrak{M}$ ,  $Z$  and  $\epsilon$ . One can readily see that  $T^* := T \cap \tilde{T}$  is a regular  $\Delta_2^0(\mathfrak{M})$  tree of measure at least  $(1 - \epsilon) + (1/2 + \epsilon) - 1 = 1/2$  by the inclusion–exclusion principle. Forcing conditions are pairs  $(\sigma, S)$  where  $\sigma \in T^*$  and  $S$  is a regular  $\Delta_2^0(\mathfrak{M})$  subtree of  $T_\sigma^*$  with measure  $\geq 1/2$ . A condition  $(\hat{\sigma}, \hat{S})$  extends another condition  $(\sigma, S)$  if  $\hat{\sigma} = \sigma\tau$  and  $\hat{S} \subseteq S_\tau$  for some  $\mathfrak{M}$ -finite binary string  $\tau$ . Every sufficiently generic sequence  $(\sigma_i, S_i)_{i \in \mathbb{N}}$  of conditions gives rise to a subset  $G \subseteq M$  whose characteristic function is  $\bigcup_{i \in \mathbb{N}} \sigma_i$ .

**Claim 4.4.1.** *For every  $\ell \in M$  and every condition  $(\sigma, S)$ , there is a condition  $(\hat{\sigma}, \hat{S})$  extending  $(\sigma, S)$  in which the length of  $\hat{\sigma}$  is at least  $\ell$ . Hence if  $G$  is sufficiently generic then  $\mathfrak{M}[G] \models \text{RCA}_0$ .*

*Proof of Claim 4.4.1.* The first part is a direct consequence of the Partial Lebesgue Density Lemma 3.13. Thus every sufficiently generic set  $G \in [T^*] \subseteq [\tilde{T}]$ . So  $\mathfrak{M}[G] \models \text{RCA}_0$  by Lemma 4.6 and Lemma 4.7.

$\square$  Claim 4.4.1

In view of Lemma 2.4, Lemma 4.7, and the claim above, it suffices to show that a sufficiently generic extension  $\mathfrak{M}[G]$  satisfies the following requirements for all Turing functionals  $\Phi$  and all  $b \in M$ :

$$\mathcal{R}_{\Phi,b}: (\exists a < b) \Phi^{Z' \oplus G}(a) \uparrow \text{ or } (\exists \ell) |\{a < b : \Phi^{Z'}(G \upharpoonright \ell; a) \downarrow\}| \geq \frac{b}{2}.$$

The next two claims demonstrate how one can satisfy these requirements and thus finish the proof. Fix a Turing functional  $\Phi$  and  $b \in M$ . Pick any condition  $(\sigma, S)$ . For each  $a < b$ , let

$$S_a = \{\tau \in S : \mathfrak{M} \models \Phi^{Z'}(\sigma\tau; a) \uparrow\}.$$

As  $S$  is a regular  $\Delta_2^0(\mathfrak{M})$  tree, so is  $S_a$ .

**Claim 4.4.2.** *If  $a \in M$  with  $a < b$  and  $\mu([S_a]) \geq 2^{-n}$  for some  $n \in \mathbb{N}$ , then  $(\sigma, S)$  has an extension forcing  $\Phi^{Z' \oplus X}(a) \uparrow$ .*

*Proof of Claim 4.4.2.* Immediate from the Partial Lebesgue Density Lemma 3.13.  $\square$  Claim 4.4.2

**Claim 4.4.3.** *Suppose that  $\mu([S_a]) < 2^{-n}$  for all  $a < b$  and all  $n \in \mathbb{N}$ . Then  $(\sigma, S)$  has an extension  $(\hat{\sigma}, \hat{S})$  such that*

$$|\{a < b : \mathfrak{M} \models \Phi^{Z'}(\hat{\sigma}; a) \downarrow\}| \geq \frac{b}{2}.$$

*Proof of Claim 4.4.3.* Work in  $\mathfrak{M}$ . For each  $\ell$ , define

$$q_\ell = \frac{|\{\tau \in S \cap 2^\ell : \Phi^{Z'}(\sigma\tau; a) \uparrow \text{ for more than half of all } a < b\}|}{2^\ell}.$$

If there is an  $\ell^*$  such that  $q_{\ell^*} < 1/4$ , then we can define  $(\hat{\sigma}, \hat{S}) = (\sigma\tau, S_\tau)$ , using a  $\tau \in S$  of length  $\geq \ell^*$  obtained by applying the Lebesgue Density Lemma 3.13 to the regular  $\Delta_2^{0,Z}$  tree

$$S \cap (2^{<\ell^*} \cup \{\tau \in 2^{\geq \ell^*} : \Phi^{Z'}(\sigma\tau; a) \downarrow \text{ for more than half of all } a < b\}),$$

which has measure  $\geq 1/4$ . So we only need to show that such an  $\ell^*$  exists.

Fix any value  $n \in \mathbb{N}$ . We know by the premise of this claim that for each  $a < b$  there is an  $\ell$  such that

$$\frac{|\{\tau \in S \cap 2^\ell : \Phi^{Z'}(\sigma\tau; a) \uparrow\}|}{2^\ell} < 2^{-n};$$

moreover, there exists a  $\Delta_2^{0,Z}$  function taking  $a$  to such an  $\ell$ . Note that if  $\ell$  satisfies the inequality above then so does every  $\ell' > \ell$ . Thus we may apply  $\Sigma_2^0$ -WPHP together with Lemmata 2.7(1) and 2.4(4) to obtain an  $\ell^*$  such that the inequality displayed above holds with  $\ell = \ell^*$  for more than  $(1 - 2^{-n})b$  many  $a < b$ .

Define three coded sets  $A, B, C$  as follows:

$$A = S \cap 2^{\ell^*},$$

$$B = \{\tau \in A : \Phi^{Z'}(\sigma\tau; a) \uparrow \text{ for more than half of all } a < b\},$$

$$C = \{(a, \tau) \in [0, b-1] \times A : \Phi^{Z'}(\sigma\tau; a) \downarrow\}.$$

On the one hand, by the choice of  $\ell^*$ ,

$$|C| > (1 - 2^{-n})b(|A| - 2^{\ell^* - n}).$$

On the other hand, if we count the elements  $(a, \tau) \in C$  with  $\tau \in B$  and those with  $\tau \in A - B$  separately, then we get

$$|C| < \frac{b}{2}|B| + b(|A| - |B|) = b|A| - \frac{b}{2}|B|.$$

Putting the two inequalities together, we deduce that

$$|B| < |A|2^{-n+1} + (1 - 2^{-n})2^{\ell^* - n+1} \leq 2^{\ell^*} (2^{-n+2} - 2^{-2n+1}).$$

In particular, when we use the value  $n = 4$ , this implies  $q_{\ell^*} = |B|/2^{\ell^*} < 1/4$ , as required.  $\square$  Claim 4.4.3

$\square$

## 5. STRENGTH

Here we compare  $\Sigma_{n+1}$ -WPHP with  $B\Sigma_{n+1}$  and the cardinality scheme for  $\Sigma_{n+1}$  formulas ( $C\Sigma_{n+1}$ ) introduced by Seetapun and Slaman [27]. For a set of formulas  $\Gamma$ , *the cardinality scheme for  $\Gamma$* , denoted by  $C\Gamma$ , asserts that every total injection defined by a formula in  $\Gamma$  has an unbounded range. (Note that this is different from what Kaye defines in his paper [18, Section 3.1].) In view of Theorem 2.3,

$$I\Delta_0 + \text{exp} \vdash (B\Sigma_{n+1} \rightarrow \Sigma_{n+1}\text{-WPHP}) \wedge (\Sigma_{n+1}\text{-WPHP} \rightarrow C\Sigma_{n+1}).$$

The main results of this section state that these arrows do not reverse for any  $n \in \mathbb{N}$ .

We will prove the independence of  $\Sigma_{n+1}$ -WPHP over  $I\Sigma_n + \text{exp} + C\Sigma_{n+1}$  using two drastically different model-theoretic constructions. In the first construction, we cofinally extend any countable  $\mathfrak{M} \models I\Sigma_n + \text{exp} + \neg\Sigma_{n+1}\text{-WPHP}$  to  $\mathfrak{N} \models I\Sigma_n + \text{exp} + \neg\Sigma_{n+1}\text{-WPHP} + C\Sigma_{n+1}$  by modifying a coded ultrapower construction that originated from Paris–Mills [23]. In the second construction, we build an end-extension chain  $(\mathfrak{M}_i : i \in \mathbb{N})$  of models of PA such that  $\bigcup_i \mathfrak{M}_i \models I\Sigma_n + \text{exp} + \neg\Sigma_{n+1}\text{-WPHP} + C\Sigma_{n+1}$ . The first construction has better control over which  $r$  we have  $\forall b (\Sigma_{n+1} : rb \rightarrow (2)_b^1)$  in the model, while the second construction naturally gives singular-like models. Both constructions will be revisited in Section 6.

For the independence of  $B\Sigma_{n+1}$  over  $I\Sigma_n + \text{exp} + \Sigma_{n+1}\text{-WPHP}$ , we modify the Paris–Mills coded ultrapower construction again to cofinally

extend any given countable  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp} + \neg\text{B}\Sigma_{n+1}$  to  $\mathfrak{N} \models \text{I}\Sigma_n + \text{exp} + \Sigma_{n+1}\text{-WPHP} + \neg\text{B}\Sigma_{n+1}$ .

As the reader may have already noticed, the Paris–Mills coded ultrapower construction is one of the key techniques we will use. A model of arithmetic  $\mathfrak{N}$  is a *cofinal extension* of another model  $\mathfrak{M}$ , denoted  $\mathfrak{M} \subseteq_{\text{cf}} \mathfrak{N}$ , if  $M \subseteq N$  and every  $b \in N$  is below some  $a \in M$ . We also write  $\mathfrak{M} \preceq_{\Sigma_n, \text{cf}} \mathfrak{N}$  if  $\mathfrak{M} \subseteq_{\text{cf}} \mathfrak{N}$  and  $\mathfrak{M} \preceq_{\Sigma_n} \mathfrak{N}$ , etc. The following theorem, which is extracted from the proofs in Paris–Mills [23], summarizes the major features of the construction; see also Paris [22, Theorem 11] and Enayat [10, Section 3].

**Theorem 5.1** (essentially Paris–Mills [23]). *Fix  $n \in \mathbb{N}$  and  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp}$ . Suppose  $b \in M$  and  $\mathcal{U}$  is an ultrafilter on the Boolean algebra of  $\mathfrak{M}$ -finite subsets of  $b$ . Let*

$$\mathfrak{N} = \mathfrak{M} \cap M^b / \mathcal{U}.$$

Then

- (1) *Łoś’s Theorem holds for every  $\Sigma_n$  or  $\Pi_n$  formula  $\theta$ , i.e., whenever  $[f_0], [f_1], \dots, [f_{k-1}] \in N$ ,*

$$\mathfrak{N} \models \theta([f_0], [f_1], \dots, [f_{k-1}])$$

*if and only if*

$$\{i \in M : i < b \text{ and } \mathfrak{M} \models \theta(f_0(i), \dots, f_{k-1}(i))\} \in \mathcal{U};$$

- (2)  $\mathfrak{M} \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N} \models \text{I}\Sigma_n + \text{exp}$ . □

It is not hard to see (from Paris’s application [22, Theorem 11] or from our applications below) that  $\Sigma_{n+1}$  elementarity here cannot be ‘improved’ to  $\Sigma_{n+2}$  elementarity in general.

**5.1. The cardinality scheme.** As alluded to above, we present two proofs of the following independence theorem.

**Theorem 5.2.**  $\text{I}\Sigma_n + \text{exp} + \text{C}\Sigma_{n+1} \not\prec \Sigma_{n+1}\text{-WPHP}$  for any  $n \in \mathbb{N}$ .

The ultrapower proof, which is the more elementary one here, comes first.

**Lemma 5.3.** *Fix  $n \in \mathbb{N}$ . Let  $\mathfrak{M}$  be a countable model of  $\text{I}\Sigma_n + \text{exp}$ , and  $a, b, e \in M$  with  $a \geq 2$  and  $e \notin \mathbb{N}$ . Suppose we have a  $\Sigma_{n+1}$  injection  $F: a^e b \rightarrow b$  in  $\mathfrak{M}$ . Then there exists a countable  $\mathfrak{N} \succ_{\Sigma_{n+1}, \text{cf}} \mathfrak{M}$  satisfying  $\text{I}\Sigma_n + \text{exp}$  in which  $[0, a]^{\mathfrak{M}} = [0, a]^{\mathfrak{N}}$  and  $F(c)^{\mathfrak{N}}$  is undefined at some  $c < a^e b$ .*

*Proof.* The extension  $\mathfrak{N}$  will be the ultrapower with respect to an ultrafilter  $\mathcal{U} \subset \mathcal{P}(a^e b)^{\mathfrak{M}} = \mathfrak{M} \cap \mathcal{P}(a^e b)$  constructed as follows.

First, apply Limit Lemma 2.1 to obtain a  $\Sigma_n$  function  $F_0$  approximating  $F$  over  $\text{I}\Sigma_n$ . This implies, in particular, that whenever  $\mathfrak{M} \preceq_{\Sigma_{n+1}}$

$\mathfrak{M}' \models \text{I}\Sigma_n$ , if  $F^{\mathfrak{M}'}$  is a total function  $[0, a^e b - 1]^{\mathfrak{M}'} \rightarrow M'$ , then  $F_0^{\mathfrak{M}'}$  is a total function  $[0, a^e b - 1]^{\mathfrak{M}'} \cdot M' \rightarrow M'$  and

$$\mathfrak{M}' \models \forall x < a^e b (F(x) = \lim_s F_0(x, s)).$$

For each  $s \in M$ , let

$$A_s = \{x < a^e b : \forall t > s (F_0(x, t) = F_0(x, s))\}^{\mathfrak{M}}.$$

Note that  $A_s \subseteq A_t$  whenever  $t > s$ . As  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp}$ , we know  $A_s \in M$  and  $\mathfrak{M} \models |A_s| \leq b$ . So  $(A_s : s \in M)$  generates a proper ideal  $\mathcal{A}$  on the Boolean algebra  $\mathcal{P}(a^e b)^{\mathfrak{M}}$ .

Then we construct  $\mathcal{U}$ . Let  $(h_k : k \in \mathbb{N})$  list all  $\mathfrak{M}$ -finite  $a^e b \rightarrow a$ . Let  $X_0 = [0, a^e b - 1]^{\mathfrak{M}}$ . If  $X_k \in M$  is defined and  $\mathfrak{M} \models |X_k| \geq a^{e-k} b$ , then define  $X_{k+1} = X_k \cap h_k^{-1}(i_k) \in M$  where  $i_k < a$  such that  $\mathfrak{M} \models |X_{k+1}| \geq a^{e-k-1} b$ . The result is a descending sequence  $(X_k : k \in \mathbb{N})$  in  $\mathcal{P}(a^e b)^{\mathfrak{M}}$  such that  $h_k(X_{k+1})$  is a singleton and  $\mathfrak{M} \models |X_k| > b$  for every  $k \in \mathbb{N}$ . Hence, the filter  $\mathcal{F}$  generated by  $(X_k : k \in \mathbb{N})$  is disjoint from  $\mathcal{A}$ . Let  $\mathcal{U}$  be any ultrafilter on  $\mathcal{P}(a^e b)^{\mathfrak{M}}$  that contains  $\mathcal{F}$  and is disjoint from  $\mathcal{A}$ .

Let

$$\mathfrak{N} = \mathfrak{M} \cap M^{a^e b} / \mathcal{U}.$$

For each  $[h] \in N$  below  $a$ , there exists  $i < a$  in  $M$  such that  $h^{-1}(i) \in \mathcal{U}$  by construction and by Łoś's Theorem for the Paris–Mills construction, i.e., Theorem 5.1(1). This implies, via Łoś's Theorem again, that  $[0, a]^{\mathfrak{M}} = [0, a]^{\mathfrak{N}}$ . Let  $c \in N$  represented by the identity function on  $a^e b$ . By Łoś's Theorem,  $c <^{\mathfrak{N}} a^e b$ . We claim that  $\lim_s F_0(c, s)$  is undefined in  $\mathfrak{N}$ . Suppose not. Then  $\mathfrak{N} \models \forall t > s (F_0(c, t) = F_0(c, s))$  for some  $s \in N$ . As  $\mathfrak{M} \subseteq_{\text{cf}} \mathfrak{N}$ , we may assume this  $s \in M$ . By Łoś's Theorem,

$$A_s = \{x < a^e b : \forall t > s (F_0(x, t) = F_0(x, s))\}^{\mathfrak{M}} \in \mathcal{U},$$

contradicting the condition that  $\mathcal{U}$  is disjoint from  $\mathcal{A}$ . Hence  $F(c)^{\mathfrak{N}}$  is undefined. The remaining properties of  $\mathfrak{N}$  follow from the Paris–Mills Theorem 5.1(2).  $\square$

Repeating the ultrapower construction above leads to the model extension theorem below.

**Theorem 5.4.** *Fix  $n \in \mathbb{N}$ . Let  $\mathfrak{M}$  be a countable model of  $\text{I}\Sigma_n + \text{exp}$  and  $a \in M$  with  $a \geq 2$ . Then there exists  $\mathfrak{N} \succ_{\Sigma_{n+1}, \text{cf}} \mathfrak{M}$  satisfying  $\text{I}\Sigma_n + \text{exp}$  such that  $[0, a]^{\mathfrak{M}} = [0, a]^{\mathfrak{N}}$  and  $\mathfrak{N} \models \Sigma_{n+1} : a^e b \rightarrow (2)_b^1$  for all  $e \in N - \mathbb{N}$  and  $b \in N$ .*

*Proof.* If  $\mathfrak{M}$  is standard, then there is nothing to do. So assume  $\mathfrak{M}$  is nonstandard. By repeated applications of Lemma 5.3, we can obtain a sequence  $(\mathfrak{N}_k : k \in \mathbb{N})$  such that for every  $k \in \mathbb{N}$ ,

- (1)  $\mathfrak{M} = \mathfrak{N}_0 \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N}_k \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N}_{k+1} \models \text{I}\Sigma_n + \text{exp}$ ;
- (2)  $[0, a]^{\mathfrak{M}} = [0, a]^{\mathfrak{N}_k}$ ; and

- (3) for each  $\Sigma_{n+1}$  injection  $F: a^e b \rightarrow b$  in  $\mathfrak{N}_k$  where  $b \in N_k$  and  $e \in N_k - \mathbb{N}$ , there exists  $\ell > k$  such that  $F^{\mathfrak{N}_\ell}$  is undefined at some  $c \in N_\ell$  below  $a^e b$ .

Then  $\mathfrak{N} = \bigcup_{k \in \mathbb{N}} \mathfrak{N}_k$  satisfies the requirements, as the reader can readily verify.  $\square$

*First Proof of Theorem 5.2.* Groszek and Slaman [13, Proposition 3.1] produced a countable model  $\mathfrak{M} \models \text{IS}_1$  with  $a \in M$  such that some  $\Delta_2^{\mathfrak{M}}$  injection  $M \rightarrow [0, a]$ . Relativization gives a countable model  $\mathfrak{M} \models \text{IS}_n + \text{exp}$  with  $a \in M$  such that some  $\Delta_{n+1}^{\mathfrak{M}}$  injection  $M \rightarrow [0, a]$ . In particular, there exists a  $\Sigma_{n+1}^{\mathfrak{M}}$  injection  $F: 2a \rightarrow a$ . Apply Theorem 5.4 to this  $\mathfrak{M}$  and this  $a$  to get  $\mathfrak{N}$ . Note that  $[0, a]^{\mathfrak{M}} = [0, a]^{\mathfrak{N}}$  implies  $[0, 2a]^{\mathfrak{M}} = [0, 2a]^{\mathfrak{N}}$ . Therefore, by the elementarity between the models,  $F^{\mathfrak{N}}$  is a  $\Sigma_{n+1}$  injection  $2a \rightarrow a$  in  $\mathfrak{N}$ . So  $\mathfrak{N} \models \text{IS}_n + \text{exp} + \text{CS}_{n+1} + \neg \Sigma_{n+1}\text{-WPHP}$ .  $\square$

The second proof of Theorem 5.2 originates from a construction devised by Theodore A. Slaman in around 2011; see Haken [15, Chapter 3]. What allows us to improve on Slaman's construction is the following recent theorem from Blanck [3, Theorem 5] and Hamkins [16, Theorem 18]. Here  $\Pi_n\text{-Tr}$  denotes the set of all (standard and non-standard)  $\Pi_n$  sentences that are declared true by the usual satisfaction predicate for  $\Pi_n$  formulas. An extension of a model of arithmetic is an *end extension* if all new elements are above all old elements. End extensions are indicated by a subscript  $e$ .

**Theorem 5.5** (Blanck and Hamkins, independently). *Let  $n \in \mathbb{N}$  and  $T \supseteq \text{PA}$  be a recursively axiomatized theory in a language extending the language of first-order arithmetic. Then there exists a  $\Sigma_{n+1}$  formula  $\theta(x)$  such that*

- (1)  $\text{PA} \vdash \exists s \forall x (x \in s \leftrightarrow \theta(x))$ ;
- (2)  $\text{PA} \vdash \text{Con}(T + \Pi_n\text{-Tr}) \leftrightarrow \forall x \neg \theta(x)$ ; and
- (3) for every  $\mathfrak{M} \models T$  and every  $s \in M$ , if  $\mathfrak{M} \models \forall x (\theta(x) \rightarrow x \in s)$ , then  $\mathfrak{M}$  has a  $\Sigma_n$ -elementary end extension  $\mathfrak{K} \models T + \forall x (\theta(x) \leftrightarrow x \in s)$ .  $\square$

The next theorem illustrates how one can use the Blanck–Hamkins theorem to define any specific countable set in a singular-like end extension. Recall that a linearly ordered structure  $\mathfrak{M}$  is  $\kappa$ -like, where  $\kappa$  is a cardinal, if  $\mathfrak{M}$  has cardinality  $\kappa$ , but every proper initial segment of  $\mathfrak{M}$  has cardinality strictly less than  $\kappa$ .

**Theorem 5.6.** *Fix  $n \in \mathbb{N}$ , a recursive theory  $T_0 \supseteq \text{PA}$ , and a cardinal  $\kappa$  of countable cofinality. Let  $\mathfrak{M} \models T_0$  of cardinality strictly less than  $\kappa$ . For every countable  $A \subseteq M$ , there is a  $\kappa$ -like  $\Sigma_n$ -elementary end extension  $\mathfrak{K}$  of  $\mathfrak{M}$  satisfying  $\Sigma_{n+3}\text{-Th}(T_0)$  in which  $A \in \Sigma_{n+1}\text{-Def}(\mathfrak{K})$ .*

*Proof.* Without loss of generality, assume  $\mathfrak{M}$  is nonstandard. Fix a strictly increasing sequence of cardinals  $(\kappa_j)_{j \in \mathbb{N}}$  whose supremum is  $\kappa$  and whose first element  $\kappa_0$  is strictly bigger than the cardinality of  $\mathfrak{M}$ . Use Craig's Trick to find a recursive sequence  $(\pi_k(v))_{k \in \mathbb{N}}$  of  $\Pi_{n+2}$  formulas such that  $\{\exists v \pi_k(v) : k \in \mathbb{N}\}$  axiomatizes  $\Sigma_{n+3}\text{-Th}(T_0)$ . Then use  $\text{IS}_{n+3}$  to get  $c \in M$  which makes  $\mathfrak{M} \models \pi_k((c)_k)$  for all  $k \in \mathbb{N}$ . Let  $T = T_0 + \{\pi_k((c)_k) : k \in \mathbb{N}\}$ , where  $c$  is a fresh constant symbol. Notice  $(\mathfrak{M}, c) \models T$ . We will build a sequence

$$\mathfrak{M} = \mathfrak{M}_0 \preceq_{\Sigma_n, e} \mathfrak{M}_1 \preceq_{\Sigma_n, e} \mathfrak{M}_2 \preceq_{\Sigma_n, e} \cdots$$

such that each  $(\mathfrak{M}_{j+1}, c)$  is a  $\kappa_j$ -like model of  $T$ . This automatically ensures  $\mathfrak{K} = \bigcup \{\mathfrak{M}_j : j \in \mathbb{N}\}$  is a  $\kappa$ -like  $\Sigma_n$ -elementary end extension of  $\mathfrak{M}$  satisfying  $\Sigma_{n+3}\text{-Th}(T_0)$ .

Let  $\theta(x)$  be a  $\Sigma_{n+1}$  formula satisfying conditions (1)–(3) in Theorem 5.5. Use condition (1) there to find  $b \in M$  such that  $\mathfrak{M} \models \forall x (\theta(x) \rightarrow x < b)$ . Fix an enumeration  $(a_j)_{j \in \mathbb{N}}$  of  $A$ .

Now, given any model  $(\mathfrak{M}_j, c) \models T + \forall x \geq b (\theta(x) \leftrightarrow \bigvee_{i < j} x = b + a_i)$  of cardinality less than  $\kappa_j$ , we can apply Theorem 5.5(3) to obtain  $\mathfrak{M}_{j+1} \succ_{\Sigma_n, e} \mathfrak{M}_j$  satisfying  $T + \forall x \geq b (\theta(x) \leftrightarrow \bigvee_{i < j+1} x = b + a_i)$ . Moreover, in view of the Löwenheim–Skolem Theorem and the Mac Dowell–Specker Theorem [17, Theorem 8.6], this  $\mathfrak{M}_{j+1}$  can be chosen to be  $\kappa_j$ -like. This ensures

$$A = \{a \in K : \mathfrak{K} \models \theta(b + a)\} \in \Sigma_{n+1}\text{-Def}(\mathfrak{K})$$

at the end.  $\square$

*Second Proof of Theorem 5.2.* Let  $\mathfrak{M}$  be a countable nonstandard model of PA. Fix any nonstandard  $a \in M$  and any bijection  $f: 2a \rightarrow a$ . Apply Theorem 5.6 to  $\kappa = \beth_\omega$  and  $A = \{\langle x, f(x) \rangle : x < 2a\}$  with  $T_0 = \text{PA}$ .  $\square$

**5.2. The collection scheme.** To show this independence, we will start with a countable model of  $\text{IS}_n + \text{exp} + \neg \text{B}\Sigma_{n+1}$ , then repeatedly apply a suitable version of the Paris–Mills coded ultrapower construction to achieve  $\Sigma_{n+1}$ -WPHP in a cofinal extension while preserving  $\text{IS}_n + \text{exp} + \neg \text{B}\Sigma_{n+1}$ .

First, let us see how to preserve the failure of  $\text{B}\Sigma_{n+1}$  in an extension. In view of Slaman [29], between models of  $\text{I}\Delta_0 + \text{exp}$ , this is equivalent to preserving some proper  $\Delta_{n+1}$ -definable cut.

**Definition 5.7.** If  $\mathfrak{N}$  is a linearly ordered structure and  $X \subseteq N$ , then

$$\sup_{\mathfrak{N}} X = \{x \in N : \exists y \in X \ x \leq^{\mathfrak{N}} y\}$$

and

$$\inf_{\mathfrak{N}} X = \{x \in N : \forall y \in X \ x <^{\mathfrak{N}} y\}.$$

**Lemma 5.8** (Keita Yokoyama). *Fix  $\mathfrak{M} \models \text{PA}^-$  and  $n \in \mathbb{N}$ . If  $I$  is a  $\Delta_{n+1}$ -definable proper cut of  $\mathfrak{M}$ , and  $\mathfrak{N} \succ_{\Sigma_{n+1}} \mathfrak{M}$  in which  $\sup_{\mathfrak{N}} I = \inf_{\mathfrak{N}}(M - I)$ , then  $J := \sup_{\mathfrak{N}} I$  is a  $\Delta_{n+1}$ -definable proper cut of  $\mathfrak{N}$ .*

*Proof.* The following piece of notation will be convenient here: if  $\mathfrak{K}$  is an  $L_1$  structure and  $\theta(v)$  is an  $L_1$  formula, possibly with parameters from  $K$ , then we write  $\theta(\mathfrak{K})$  for

$$\{v \in K : \mathfrak{K} \models \theta(v)\}.$$

Suppose  $I = \varphi(\mathfrak{M}) = \psi(\mathfrak{M})$ , where  $\varphi(v) \in \Sigma_{n+1}$  and  $\psi(v) \in \Pi_{n+1}$ , both of which may involve parameters from  $M$ . Define

$$\begin{aligned} \varphi'(v) &= \exists w \geq v \varphi(w), \quad \text{and} \\ \psi'(v) &= \forall u \leq v \psi(u). \end{aligned}$$

Since  $I$  is a cut of  $\mathfrak{M}$ , we know  $\varphi'(\mathfrak{M}) = \psi'(\mathfrak{M}) = I$ . Notice  $\varphi'(v)$  is  $\Sigma_{n+1}$  and  $\psi'(v)$  is  $\Pi_{n+1}$ . So it suffices to show that  $J = \varphi'(\mathfrak{N}) = \psi'(\mathfrak{N})$ .

The two directions are symmetric. So we only show one of them here. Take  $c \in N - J$ . Recall  $J = \sup_{\mathfrak{N}} I = \inf_{\mathfrak{N}}(M - I)$ . So we get  $d \in M - I$  such that  $d \leq c$ . Then, since  $d \notin I = \varphi'(\mathfrak{M}) = \psi'(\mathfrak{M})$ ,

$$\mathfrak{M} \models \forall w \geq d \neg \varphi(w) \wedge \exists u \leq d \neg \psi(u).$$

By  $\Sigma_{n+1}$  elementarity, the same formula is true in  $\mathfrak{N}$ . Thus, as  $d \leq c$ ,

$$\mathfrak{N} \models \forall w \geq c \neg \varphi(w) \wedge \exists u \leq c \neg \psi(u). \quad \square$$

Our ultrapower construction below is designed to correct a failure of  $\Sigma_{n+1}$ -WPHP. That it naturally produces an extension satisfying the hypotheses of Lemma 5.8 is rather remarkable, at least at first sight.

**Lemma 5.9.** *Fix  $n \in \mathbb{N}$ . Let  $\mathfrak{M}$  be a countable model of  $\text{I}\Sigma_n + \text{exp}$  and  $(I_k : k \in \mathbb{N})$  be a countable family of cuts of  $M$ . Suppose  $a \in M$  and  $F$  is a  $\Sigma_{n+1}^{\mathfrak{M}}$  injection  $2a \rightarrow a$ . Then there exists a countable  $\mathfrak{N} \succ_{\Sigma_{n+1}, \text{cf}} \mathfrak{M}$  satisfying  $\text{I}\Sigma_n + \text{exp}$  in which  $\sup_{\mathfrak{N}} I_k = \inf_{\mathfrak{N}}(M - I_k)$  for all  $k \in \mathbb{N}$ , and  $F(c)^{\mathfrak{N}}$  is undefined for some  $c < 2a$ .*

*Proof.* We will construct an ultrafilter  $\mathcal{U}$  on  $\mathcal{P}(2a)^{\mathfrak{M}}$ , i.e., the power set of  $2a$  computed in  $\mathfrak{M}$ , so that the coded ultrapower

$$\mathfrak{N} := \mathfrak{M} \cap M^{2a} / \mathcal{U}$$

has the required properties.

First, apply Limit Lemma 2.1 to obtain a  $\Sigma_n$  function  $F_0$  approximating  $F$  over  $\text{I}\Sigma_n$  as in the proof of Lemma 5.3. For each  $s \in M$ , let

$$A_s = \{i < 2a : \forall t > s (F_0(i, t) = F_0(i, s))\}^{\mathfrak{M}}.$$

As in the proof of Lemma 5.3, each  $A_s$  is  $\mathfrak{M}$ -finite and  $\mathfrak{M} \models |A_s| < a$ . Let  $\mathcal{A}$  be the ideal generated by  $(A_s : s \in M)$  in  $\mathcal{P}(2a)^{\mathfrak{M}}$ . Say an  $\mathfrak{M}$ -finite set  $X$  is *large* if and only if  $\mathfrak{M} \models \forall s (|X \cap A_s| < |X|/2)$ , and *very large* if and only if  $\mathfrak{M} \models \forall s (|X \cap A_s| < |X|/4)$ .

Then we construct a descending sequence  $(X_\ell : \ell \in \mathbb{N})$  of large  $\mathfrak{M}$ -finite sets starting from  $X_0 = [0, 2a - 1]^{\mathfrak{M}}$ , which is clearly large. We employ a forcing-style construction to ensure, for each  $I_k$  and each  $\mathfrak{M}$ -finite  $h: 2a \rightarrow M$ , the existence of  $y \in M$  and  $\ell \in \mathbb{N}$  such that either

- $y \in I_k$  and  $X_\ell \subseteq \{i < 2a : h(i) \leq y\}^{\mathfrak{M}}$ ; or
- $y \notin I_k$  and  $X_\ell \subseteq \{i < 2a : h(i) > y\}^{\mathfrak{M}}$ .

The two claims below can be viewed as density properties of an appropriate forcing notion.

**Claim 5.9.1.** *Every large  $\mathfrak{M}$ -finite set has a very large subset.*

*Proof of Claim 5.9.1.* Fix a large  $X \in M$  and work in  $\mathfrak{M}$ . Let

$$j = \max\{i < 8 : \exists s(|X \cap A_s| \geq i|X|/8)\},$$

and let  $s$  be such that  $|X \cap A_s| \geq j|X|/8$ . Let  $Y = X - A_s$ . Then, for each  $t$ , the maximality of  $j$  implies  $|X \cap A_t| < (j+1)|X|/8$  and hence, if  $t \geq s$ , then

$$|Y \cap A_t| = |X \cap A_t| - |X \cap A_s| < (j+1)\frac{|X|}{8} - j\frac{|X|}{8} = \frac{|X|}{8} < \frac{|Y|}{4}$$

by the largeness of  $X$ . □ Claim 5.9.1

**Claim 5.9.2.** *For each  $k \in \mathbb{N}$ , each  $\mathfrak{M}$ -finite  $h: 2a \rightarrow M$  and each very large  $\mathfrak{M}$ -finite  $Y$ , there exist a large  $\mathfrak{M}$ -finite  $Z \subseteq Y$  and  $y \in M$  such that either  $y \in I_k$  and  $Z \subseteq \{i < 2a : h(i) \leq y\}^{\mathfrak{M}}$ , or  $y \notin I_k$  and  $Z \subseteq \{i < 2a : h(i) > y\}^{\mathfrak{M}}$ .*

*Proof of Claim 5.9.2.* Work in  $\mathfrak{M}$ . Since  $h \in M$ , the range of  $h$  is bounded. Define

$$z = \min\{x : |\{i \in Y : h(i) \leq x\}| \geq |Y|/2\}.$$

If  $z \in I_k$ , then let  $y = z$  and  $Z = \{i \in Y : h(i) \leq y\}$ . Clearly  $|Z| \geq |Y|/2$  in this case. If  $z \notin I_k$ , then let  $y = z - 1 \in M - I_k$  and  $Z = \{i \in Y : h(i) > y\}$ . The minimality of  $z$  ensures  $|Z| \geq |Y|/2$  in this case too. As  $Y$  is very large,

$$|Z \cap A_s| < |Y|/4 \leq |Z|/2$$

for all  $s \in M$ . So  $Z$  and  $y$  are as desired. □ Claim 5.9.2

By the countability of  $M$  and also of  $(I_k)$ , we can inductively apply the above claims to obtain the  $X_\ell$ 's we want. As the reader can readily see, the largeness of the  $X_\ell$ 's implies that the filter  $\mathcal{F}$  generated by  $(X_\ell)$  is disjoint from  $\mathcal{A}$ . Let  $\mathcal{U}$  be an ultrafilter on  $\mathcal{P}(2a)^{\mathfrak{M}}$  which contains  $\mathcal{F}$  and is disjoint from  $\mathcal{A}$ , and let

$$\mathfrak{N} = \mathfrak{M} \cap M^{2a}/\mathcal{U}.$$

**Claim 5.9.3.** *For each  $k \in \mathbb{N}$ ,  $\sup_{\mathfrak{N}} I_k = \inf_{\mathfrak{N}}(M - I_k)$ .*

*Proof of Claim 5.9.3.* Let  $k \in \mathbb{N}$  and  $[h] \in N$ . There are  $X \in \mathcal{U}$  and  $y \in M$  such that either  $h(i) \leq y \in I_k$  for all  $i \in X$ , or  $h(i) > y > I_k$  for all  $i \in X$ . So by Łoś's Theorem, either  $[h] \in \sup_{\mathfrak{M}} I_k$  or  $[h] \notin \inf_{\mathfrak{M}}(M - I_k)$ .  $\square$  Claim 5.9.3

As in the proof of Lemma 5.3, we know  $\mathfrak{M} \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N} \models \text{I}\Sigma_n + \text{exp}$  and  $F^{\mathfrak{M}}$  is undefined at the element represented by the identity function on  $[0, 2a - 1]^{\mathfrak{M}}$ . So  $\mathfrak{N}$  satisfies all the properties required by the lemma.  $\square$

With Lemmata 5.8 and 5.9 at hand, separating  $\Sigma_{n+1}$ -WPHP and  $\text{B}\Sigma_{n+1}$  is only a matter of routine iteration.

**Theorem 5.10.** *Fix  $n \in \mathbb{N}$ . Let  $\mathfrak{M}$  be a countable model of  $\text{I}\Sigma_n + \text{exp}$  and  $(I_k : k \in \mathbb{N})$  be a countable family of cuts of  $\mathfrak{M}$ . Then there exists  $\mathfrak{N} \succ_{\Sigma_{n+1}, \text{cf}} \mathfrak{M}$  satisfying  $\text{I}\Sigma_n + \text{exp} + \Sigma_{n+1}$ -WPHP such that for all  $k \in \mathbb{N}$ ,*

$$\sup_{\mathfrak{N}} I_k = \inf_{\mathfrak{N}}(M - I_k).$$

Hence  $\text{I}\Sigma_n + \text{exp} + \Sigma_{n+1}$ -WPHP  $\not\equiv \text{B}\Sigma_{n+1}$ .

*Proof.* By repeated applications of Lemma 5.9, obtain a sequence  $(\mathfrak{N}_\ell : \ell \in \mathbb{N})$  such that for all  $k, \ell \in \mathbb{N}$ ,

- (1)  $\mathfrak{M} = \mathfrak{N}_0 \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N}_\ell \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N}_{\ell+1} \models \text{I}\Sigma_n + \text{exp}$ ;
- (2)  $\sup_{\mathfrak{N}_\ell} I_k = \inf_{\mathfrak{N}_\ell}(M - I_k)$ ;
- (3) for each  $\Sigma_{n+1}$  injection  $F: 2a \rightarrow a$  in  $\mathfrak{N}_\ell$  where  $a \in N_\ell$ , there exists  $m > \ell$  such that  $F^{\mathfrak{N}_m}$  is undefined at some  $c \in N_m$  below  $2a$ .

Let  $\mathfrak{N} = \bigcup_{\ell \in \mathbb{N}} \mathfrak{N}_\ell$ . Then  $\mathfrak{M} \preceq_{\Sigma_{n+1}, \text{cf}} \mathfrak{N}$  and  $\sup_{\mathfrak{N}} I_k = \inf_{\mathfrak{N}}(M - I_k)$  for all  $k \in \mathbb{N}$ . Since  $(\mathfrak{N}_\ell)$  is a  $\Sigma_{n+1}$ -elementary chain, the union  $\mathfrak{N}$  satisfies  $\bigcap_{\ell \in \mathbb{N}} \Pi_{n+3}\text{-Th}(\mathfrak{N}_\ell) \supseteq \text{I}\Sigma_n + \text{exp}$ .

Suppose  $a \in N$  and  $F: 2a \rightarrow a$  is a  $\Sigma_{n+1}^{\mathfrak{N}}$  injection. Pick a large enough  $\ell$  such that  $N_\ell$  contains  $a$  and all the parameters in the definition of  $F$ . Then  $F^{\mathfrak{N}_\ell}$  is a  $\Sigma_{n+1}^{\mathfrak{N}_\ell}$  injection  $[0, 2a - 1]^{\mathfrak{N}_\ell} \rightarrow [0, a - 1]^{\mathfrak{N}_\ell}$  since  $\mathfrak{N}_\ell \preceq_{\Sigma_{n+1}} \mathfrak{N}$ . So (3) gives  $m > \ell$  and  $c \in N_m$  below  $2a$  such that  $F^{\mathfrak{N}_m}$  is undefined at  $c$ . As  $\mathfrak{N}_m \preceq_{\Sigma_{n+1}} \mathfrak{N}$ ,  $F^{\mathfrak{N}}$  is undefined at  $c$  as well, contradicting the assumption on  $F$ . This shows that  $\mathfrak{N} \models \Sigma_{n+1}$ -WPHP.

If the  $\mathfrak{M}$  above does not satisfy  $\text{B}\Sigma_{n+1}$ , then we can choose  $I_0$  to be a proper  $\Delta_{n+1}^{\mathfrak{M}}$  cut, which ensures  $\mathfrak{N} \not\models \text{B}\Sigma_{n+1}$  in view of Lemma 5.8. The last part of the theorem follows.  $\square$

*Remark 5.11.* It is apparent that the assumption  $\text{exp}$  can be weakened in the cofinal extension constructions in this section: we only need a theory in which we can count the elements of  $\Delta_0$ -definable sets somehow. For example, Theorem 5.1 remains true even without  $\text{exp}$  if we replace ‘ $\mathfrak{M}$ -finite’ by ‘bounded  $\Delta_0(\Sigma_n)$ -definable’ everywhere.

## 6. MORE ABOUT PIGEONHOLE PRINCIPLES

As we saw in Sections 3 and 4, the principle  $\Sigma_n$ -WPHP arises naturally when one studies the Weak Weak König Lemma. Clearly one can obtain a hierarchy of weaker pigeonhole principles by similarly changing the domains of the functions involved:

$$\begin{array}{lll}
\Sigma_n\text{-PHP}(\text{num} + 1, \text{num}) & := & \forall a \quad \Sigma_n: a + 1 \rightarrow (2)_a^1; \\
\Sigma_n\text{-PHP}(2\text{num}, \text{num}) & := & \forall a \geq 1 \quad \Sigma_n: 2a \rightarrow (2)_a^1; \\
\Sigma_n\text{-PHP}(\text{num}^2, \text{num}) & := & \forall a \geq 2 \quad \Sigma_n: a^2 \rightarrow (2)_a^1; \\
\Sigma_n\text{-PHP}(2^{\text{num}}, \text{num}) & := & \forall a \quad \Sigma_n: 2^a \rightarrow (2)_a^1; \\
\vdots & & \vdots \\
\Sigma_n\text{-PHP}(H(\text{num}), \text{num}) & := & \forall a \quad \Sigma_n: H(a) \rightarrow (2)_a^1; \\
\vdots & & \vdots \\
\Sigma_n\text{-PHP}(<\infty, \text{num}) & := & \forall a \exists b \quad \Sigma_n: b \rightarrow (2)_a^1; \\
\Sigma_n\text{-PHP}(\infty, \text{num}) & := & \forall a \quad \Sigma_n: \infty \rightarrow (2)_a^1.
\end{array}$$

Here  $\Sigma_n\text{-PHP}(\text{num} + 1, \text{num})$  is simply the usual  $\Sigma_n\text{-PHP}$ ; the principle  $\Sigma_n\text{-PHP}(2\text{num}, \text{num})$  is what we have called  $\Sigma_n\text{-WPHP}$ ; and  $\Sigma_n\text{-PHP}(\infty, \text{num}) = \text{C}\Sigma_n$ . In general, any unary function  $H$  that is provably total in the base theory gives rise to a pigeonhole principle  $\Sigma_n\text{-PHP}(H(\text{num}), \text{num})$ . Kaye [18, Section 3.2] refers to  $\{\Sigma_k\text{-PHP}(<\infty, \text{num}) : k \in \mathbb{N}\}$  as a *generalized pigeonhole principle*.

We saw several separation results for this hierarchy over  $\text{I}\Delta_0 + \text{exp}$  for *positive*  $n \in \mathbb{N}$ . On the one hand, Theorem 5.10 separates  $\Sigma_n\text{-PHP}(\text{num} + 1, \text{num})$  from  $\Sigma_n\text{-PHP}(2\text{num}, \text{num})$ . On the other hand, both constructions in Section 5.1 can separate  $\Sigma_n\text{-PHP}(2\text{num}, \text{num})$ ,  $\Sigma_n\text{-PHP}(\text{num}^2, \text{num})$ ,  $\dots$  from  $\Sigma_n\text{-PHP}(<\infty, \text{num})$ . In fact, one can squeeze a little more out of the second construction. The following improves Theorem 13 in Haken [15].

**Theorem 6.1.** *Let  $n \in \mathbb{N}$ . For any set of  $\Sigma_{n+3}$  sentences  $S$  that is consistent with PA and any  $S$ -provably total unary function  $H$  with a  $\Sigma_{n+1}$ -definable graph,*

$$S + \{\Sigma_k\text{-PHP}(<\infty, \text{num}) : k \in \mathbb{N}\} \not\vdash \Sigma_{n+1}\text{-PHP}(H(\text{num}), \text{num}).$$

*Proof.* Using a universal  $\Sigma_{n+1}$  predicate, one can finitely axiomatize  $\Sigma_{n+1}\text{-PHP}(H(\text{num}), \text{num})$  over  $\text{I}\Delta_0 + \text{exp} + S$ , and  $\text{I}\Delta_0 + \text{exp}$  itself is well known to be finitely axiomatizable [12, §6]. So we may assume  $S$  is finite without loss of generality. Then run our second proof of Theorem 5.2, changing  $T_0$  to  $\text{PA} + S$  and  $2a$  to  $H(a)$ .  $\square$

Clearly one can strengthen  $\{\Sigma_k\text{-PHP}(<\infty, \text{num}) : k \in \mathbb{N}\}$  in Theorem 6.1 to any theory satisfied in all  $\sqsupset_\omega$ -like models of  $\text{I}\Delta_0 + \text{exp}$ . As shown by Kaye [18, Theorem 3.20], such a theory cannot be too strong, in the sense that it is always weaker than

$$\text{IB} + \text{exp} := \text{I}\Delta_0 + \text{exp} + \{\text{I}\Sigma_k \rightarrow \text{B}\Sigma_k : k \in \mathbb{N}\},$$

which is partially conservative over all the usual fragments of Peano arithmetic [18, Theorem 4.1]. In fact, from our proof of Theorem 6.1, one sees this implication is strict. The strictness of this implication, which answers a question in Kaye [18, Problem 4.3], was first shown by Theodore A. Slaman in around 2011 using a similar method; see Haken [15, Section 3.3].

When  $n = 0$ , the situation is somewhat different: as shown by Paris–Wilkie–Woods [24, Corollary 2] and Thapen [30, Lemma 2.1], there is a way to construct a counterexample to  $\Sigma_0\text{-PHP}(<\infty, \text{num})$  from a counterexample to  $\Sigma_0\text{-PHP}(2\text{num}, \text{num})$  in  $\text{I}\Delta_0 + \Omega_1$ , where  $\Omega_1$  denotes an axiom asserting the totality of  $x \mapsto x^{\log x}$  over  $\text{I}\Delta_0$ . This construction does not work at higher levels of the arithmetic hierarchy because apparently one cannot iterate a  $\Sigma_n$ -definable function without increasing the complexity of the defining formula when  $n \geq 1$  and  $\text{B}\Sigma_n$  is absent. Using a diagonal argument, Paris–Wilkie–Woods [24, Theorem 1] showed  $\text{I}\Delta_0 + \Omega_1 \vdash \Sigma_0\text{-PHP}(\text{num}^2, \text{num})$ . So  $\text{I}\Delta_0 + \Omega_1 \vdash \Sigma_0\text{-PHP}(2\text{num}, \text{num})$  too. The question whether  $\text{I}\Delta_0 + \Omega_1 \vdash \Sigma_0\text{-PHP}(\text{num} + 1, \text{num})$ , first raised by Macintyre, is a fundamental open question in bounded arithmetic [1, Problem B(c)].

As observed by Dimitracopoulos and Paris [9, Remarks on page 79], there is some connection between the  $\Sigma_0$  and the  $\Sigma_1$  level: one can deduce from the Paris–Wilkie–Woods theorem in the previous paragraph that  $\text{B}\Sigma_1 + \Omega_1 \vdash \Sigma_1\text{-PHP}(2\text{num}, \text{num})$ , but the question whether  $\text{B}\Sigma_1 + \Omega_1 \vdash \Sigma_1\text{-PHP}(\text{num} + 1, \text{num})$  is open because it is equivalent to Macintyre’s question.

Although  $\Sigma_0\text{-PHP}(2\text{num}, \text{num})$  is known to be strictly weaker than  $\Sigma_0\text{-PHP}(\text{num} + 1, \text{num})$  in the relativized setting [19, 26], we do not yet have an unrelativized separation to date. In this context, the coded ultrapower construction that we used to prove our unrelativized separation at higher levels of the arithmetic hierarchy (i.e., Theorem 5.10) may provide useful information.

Our coded ultrapower constructions in Section 5 is of independent model-theoretic interest. Surprisingly little is known about non-elementary cofinal extensions of models of arithmetic. For instance, all such constructions known so far make a new collection axiom true in the extension. Our construction, on the contrary, can preserve all failures of collection at the appropriate level.

**Question 6.2.** *Given  $n \in \mathbb{N}$ , can one find a model of  $\text{B}\Sigma_{n+1}$  with a cofinal extension satisfying  $\text{I}\Delta_0$  but not  $\text{B}\Sigma_{n+1}$ ?*

In some sense, one can use Lemma 2.7(2) and Theorem 5.4 to characterize  $\Sigma_{n+1}\text{-PHP}(2\text{num}, \text{num})$ .

**Proposition 6.3.** *Let  $n \in \mathbb{N}$  and  $H$  be a provably total unary function in  $\text{I}\Sigma_n + \text{exp}$  with a  $\Sigma_{n+1}$ -definable graph. If  $\text{I}\Sigma_n + \text{exp} + \Sigma_{n+1}\text{-PHP}(<\infty, \text{num})$*

proves

$$\forall x H(x) > x \quad \text{and} \quad \forall r \forall w \exists x \geq w H(x) \geq rx,$$

then it cannot prove

$$\forall a (\exists x \geq a (\Sigma_{n+1}: H(x) \rightarrow (2)_x^1) \rightarrow (\Sigma_{n+1}: H(a) \rightarrow (2)_a^1)).$$

*Proof.* Use Theorem 6.1 to find a countable  $\mathfrak{M} \models \text{I}\Sigma_n + \text{exp} + \neg \Sigma_{n+1}\text{-PHP}(H(\text{num}), \text{num})$ . Let  $a \in M$  such that  $\mathfrak{M} \not\models \Sigma_{n+1}: H(a) \rightarrow (2)_a^1$ . Fix  $e \in M - \mathbb{N}$ . Set  $r = \max\{a, H(a)\}$ . Apply Theorem 5.4 to find  $\mathfrak{N} \succ_{\Sigma_{n+1}} \mathfrak{M}$  satisfying  $\text{I}\Sigma_n + \text{exp} + \forall b (\Sigma_{n+1}: r^e b \rightarrow (2)_b^1)$  such that  $[0, r]^{\mathfrak{M}} = [0, r]^{\mathfrak{N}}$ . Notice  $\mathfrak{N} \not\models \Sigma_{n+1}: H(a) \rightarrow (2)_a^1$  as a result. Hence if the first conjunct in the hypothesis of the proposition is true, but the conclusion is not, then  $\mathfrak{N} \models \forall x \geq a \neg(\Sigma_{n+1}: H(x) \rightarrow (2)_x^1)$ , and so  $\mathfrak{N} \models \forall x \geq a H(x) < H(a)^e x$ .  $\square$

Let us modify our second proof of Theorem 5.2 to show a similar characterization for  $\Sigma_{n+1}\text{-PHP}(\text{num}^2, \text{num})$  in terms of what we call  $\Sigma_{n+1}$ -cardinalities of numbers.

**Definition 6.4.** Let  $n \in \mathbb{N}$ . If  $\mathfrak{M} \models \text{I}\Delta_0$  and  $a \in M$ , then

$$\Sigma_n\text{-Card}^{\mathfrak{M}}(a) = \{b \in M : \text{in } \mathfrak{M} \text{ some } \Sigma_n\text{-definable injection } b \rightarrow a\}.$$

Clearly, the  $\Sigma_n$ -Card of a number is closed downwards and always contains the number itself. In view of the usual set-theoretic convention, it is probably more appropriate to define  $\Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  to be

$$\{b \in M : \text{in } \mathfrak{M} \text{ some } \Sigma_n\text{-definable injection } b + 1 \rightarrow a\}.$$

We choose to adopt a slightly different definition because (1) it actually does not make any difference in the cases we are interested in, and (2) it makes the next proof neater.

**Proposition 6.5.** Fix  $n \in \mathbb{N}$ . Let  $\mathfrak{M} \models \text{I}\Delta_0$  and  $a \in M$ .

- (1)  $a + 1 \in \Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  if and only if  $\Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  is closed under  $x \mapsto x + 1$ .
- (2)  $2a \in \Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  if and only if  $\Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  is closed under  $x \mapsto 2x$ .
- (3)  $a^2 \in \Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  if and only if  $\Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  is closed under  $x \mapsto x^2$ .

*Proof.* The right-to-left directions are obvious. So let us concentrate on the left-to-right directions. Fix  $b \in \Sigma_n\text{-Card}^{\mathfrak{M}}(a)$  and a  $\Sigma_n$ -definable injection  $F: b \rightarrow a$ .

- (1) Define  $F_1: b + 1 \rightarrow a + 1$  by setting, for each  $x < b + 1$ ,

$$F_1(x) = \begin{cases} F(x) + 1, & \text{if } x < b; \\ 0, & \text{if } x = b. \end{cases}$$

- (2) Define  $F_2: 2b \rightarrow 2a$  by setting, for each  $i < 2$  and  $v < b$ ,

$$F_2(ib + v) = ia + F(v).$$

(3) Define  $F_3: b^2 \rightarrow a^2$  by setting, for all  $u, v < b$ ,

$$F_3(ub + v) = F(u) \cdot a + F(v).$$

Composing  $F_j$  with a witness to the left-hand-side condition gives the injection we want.  $\square$

Although one may not expect that this list of equivalences goes on forever, one may expect at least an analogous equivalence for  $x \mapsto 2^x$ . Nevertheless, this extrapolated equivalence is not true, as one can deduce from the following theorem by Paris and Mills [23, Theorem 2].

**Theorem 6.6** (Paris–Mills). *Let  $\mathfrak{M}_0$  be a countable model of PA and  $I$  be a cut of  $\mathfrak{M}_0$  closed under multiplication. Then  $\mathfrak{M}_0$  has an elementary extension  $\mathfrak{M}$  in which  $\sup_{\mathfrak{M}} I = I$  and every interval  $[0, b]^{\mathfrak{M}}$  where  $b \in M - I$  is uncountable.*  $\square$

Although Corollaries 6.7 and 6.8 are formulated in terms of PA, it is not hard to see that they remain true when PA is replaced by any recursively axiomatized consistent extension of PA.

**Corollary 6.7.** *Fix  $n \in \mathbb{N}$  and a countable  $\mathfrak{M}_0 \models \text{PA}$ . Let  $I$  be a cut of  $\mathfrak{M}_0$  closed under multiplication and  $a \in I - \mathbb{N}$ . Then  $\mathfrak{M}_0$  has a  $\Sigma_{n+1}$ -elementary extension  $\mathfrak{K} \models \Sigma_{n+3}\text{-Th}(\text{PA})$  in which*

$$\Sigma_{n+1}\text{-Card}^{\mathfrak{K}}(a) = \Sigma_{n+2}\text{-Card}^{\mathfrak{K}}(a) = \dots = I.$$

*Proof.* Let  $\mathfrak{M}$  be an extension of  $\mathfrak{M}_0$  given by Theorem 6.6. By the Löwenheim–Skolem Theorem, this  $\mathfrak{M}$  can be chosen to have cardinality  $\aleph_1$ . Fix any bijection  $f: I \rightarrow a$ . Apply Theorem 5.6 to  $\kappa = \beth_\omega$  and  $A = \{\langle x, f(x) \rangle : x \in I\}$  with  $T_0 = \text{PA}$ .  $\square$

**Corollary 6.8.** *Let  $n \in \mathbb{N}$  and  $H$  be a provably total unary function in PA with a  $\Sigma_{n+1}$ -definable graph. If  $\text{PA} \vdash \forall w \exists x \geq w H(x) \geq x^k$  for all  $k \in \mathbb{N}$ , then there exist  $\mathfrak{K} \models \Sigma_{n+3}\text{-Th}(\text{PA})$  and  $a \in K$  such that  $H(a) \in \Sigma_{n+1}\text{-Card}^{\mathfrak{K}}(a)$  but  $\Sigma_{n+1}\text{-Card}^{\mathfrak{K}}(a)$  is not closed under  $H$ .*

*Proof.* Take any nonstandard element  $a$  in a countable model  $\mathfrak{M}_0 \models \text{PA}$ . By the hypothesis, we know  $\mathfrak{M}_0 \models \exists x \geq \max\{a, H(a)\} H(x) \geq x^k$  for all  $k \in \mathbb{N}$ . So overspill gives  $b \geq \max\{a, H(a)\}$  in  $\mathfrak{M}_0$  such that  $H(b) \geq b^k$  for all  $k \in \mathbb{N}$ . Then apply Corollary 6.7 to  $I = \sup_{\mathfrak{M}_0} \{b^k : k \in \mathbb{N}\}$  to obtain the model  $\mathfrak{K}$  we want.  $\square$

Let us conclude with two general questions on the strength of weak pigeonhole principles. Recall that the question whether  $\text{I}\Delta_0 + \neg \text{exp} + \neg \text{B}\Sigma_1$  is consistent is widely open [1, Question 29].

**Question 6.9.** *Does  $\text{I}\Delta_0 + \text{exp} + \Sigma_{n+1}\text{-PHP}(2\text{num}, \text{num})$  prove  $\text{I}\Sigma_n$  or  $\text{B}\Sigma_n$  for any  $n \geq 1$ ?*

**Question 6.10.** *Is  $\text{I}\Delta_0 + \neg \text{exp} + \neg \Sigma_1\text{-PHP}(\infty, \text{num})$  consistent?*

## ACKNOWLEDGEMENTS

We thank Keita Yokoyama for numerous fruitful discussions which led to a simplification of the proof of Theorem 5.10 and to the conception of Proposition 6.5. We thank Leszek Kołodziejczyk for introducing to us the references related to the  $\Delta_0$  pigeonhole principle, and for suggesting various improvements to the first draft of this paper. We thank Ali Enayat and Rasmus Blanck for bringing to our attention several pieces of work in the literature [3, 10, 16, 23] that are relevant to this research. We also thank the anonymous referee for many helpful suggestions of improvements.

## REFERENCES

- [1] *Open problems*, Arithmetic, proof theory, and computational complexity (Prague, 1991), Oxford Logic Guides, vol. 23, Oxford Univ. Press, New York, 1993, Compiled by Peter Clote and Jan Krajíček, pp. 1–19. MR 1236455
- [2] Jeremy Avigad, Edward T. Dean, and Jason Rute, *Algorithmic randomness, reverse mathematics, and the dominated convergence theorem*, Ann. Pure Appl. Logic **163** (2012), no. 12, 1854–1864. MR 2964874
- [3] Rasmus Blanck, *Hierarchical incompleteness results for arithmetically definable fragments of arithmetic*, arxiv:1803.01762v1 [math.LO], March 2018.
- [4] C. T. Chong, Theodore A. Slaman, and Yue Yang, *The metamathematics of stable Ramsey’s theorem for pairs*, J. Amer. Math. Soc. **27** (2014), no. 3, 863–892. MR 3194495
- [5] ———, *The inductive strength of Ramsey’s Theorem for Pairs*, Adv. Math. **308** (2017), 121–141. MR 3600057
- [6] C.T. Chong, W. Li, W. Wang, and Y. Yang, *On the computability of perfect subsets of sets with positive measure*, Proceedings of the American Mathematical Society **147** (2019), no. 9, 4021–4028.
- [7] Chris J. Conidis and Theodore A. Slaman, *Random reals, the rainbow Ramsey theorem, and arithmetic conservation*, J. Symbolic Logic **78** (2013), no. 1, 195–206. MR 3087070
- [8] Jared Corduan, Marcia J. Groszek, and Joseph R. Mileti, *Reverse mathematics and Ramsey’s property for trees*, J. Symbolic Logic **75** (2010), no. 3, 945–954. MR 2723776
- [9] C. Dimitracopoulos and J. Paris, *The pigeonhole principle and fragments of arithmetic*, Z. Math. Logik Grundlag. Math. **32** (1986), no. 1, 73–80. MR 832863
- [10] Ali Enayat, *Automorphism of models of bounded arithmetic*, Fund. Math. **192** (2006), no. 1, 37–65. MR 2283628
- [11] Fernando Ferreira, *A feasible theory for analysis*, J. Symbolic Logic **59** (1994), no. 3, 1001–1011. MR 1295985
- [12] Haim Gaifman and Constantine Dimitracopoulos, *Fragments of Peano’s arithmetic and the MRDP theorem*, Logic and algorithmic (Zurich, 1980), Monograph. Enseign. Math., vol. 30, Univ. Genève, Geneva, 1982, pp. 187–206. MR 648303
- [13] Marcia J. Groszek and Theodore A. Slaman, *On Turing reducibility*, Preprint, 1994.
- [14] Petr Hájek and Pavel Pudlák, *Metamathematics of first-order arithmetic*, Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1993. MR 1219738

- [15] Ian Robert Haken, *Randomizing Reals and the First-Order Consequences of Randoms*, Ph.D. thesis, University of California, Berkeley, 2014, p. 54. MR 3295326
- [16] Joel David Hamkins, *The modal logic of arithmetic potentialism and the universal algorithm*, arXiv:1801.04599v3 [math.LO], May 2018.
- [17] Richard Kaye, *Models of Peano arithmetic*, Oxford Logic Guides, vol. 15, The Clarendon Press Oxford University Press, New York, 1991, Oxford Science Publications.
- [18] ———, *The theory of  $\kappa$ -like models of arithmetic*, Notre Dame J. Formal Logic **36** (1995), no. 4, 547–559, Special Issue: Models of arithmetic. MR 1368466
- [19] Jan Krajíček, Pavel Pudlák, and Alan Woods, *An exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle*, Random Structures Algorithms **7** (1995), no. 1, 15–39. MR 1346282
- [20] Antonín Kučera, *Measure,  $\Pi_1^0$ -classes and complete extensions of PA*, Recursion theory week (Oberwolfach, 1984), Lecture Notes in Math., vol. 1141, Springer, Berlin, 1985, pp. 245–259. MR 87e:03102
- [21] Stuart Kurtz, *Randomness and genericity in the degrees of unsolvability*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1981.
- [22] J. B. Paris, *Some conservation results for fragments of arithmetic*, Model theory and arithmetic (Paris, 1979–1980), Lecture Notes in Math., vol. 890, Springer, Berlin-New York, 1981, pp. 251–262. MR 645006
- [23] J. B. Paris and George Harlow Mills, *Closure properties of countable nonstandard integers*, Fund. Math. **103** (1979), no. 3, 205–215. MR 547554
- [24] J. B. Paris, A. J. Wilkie, and A. R. Woods, *Provability of the pigeonhole principle and the existence of infinitely many primes*, J. Symbolic Logic **53** (1988), no. 4, 1235–1244. MR 973114
- [25] Ludovic Patey and Keita Yokoyama, *The proof-theoretic strength of Ramsey’s theorem for pairs and two colors*, Adv. Math. **330** (2018), 1034–1070. MR 3787563
- [26] Toniann Pitassi, Paul Beame, and Russell Impagliazzo, *Exponential lower bounds for the pigeonhole principle*, Comput. Complexity **3** (1993), no. 2, 97–140. MR 1233662
- [27] David Seetapun and Theodore A. Slaman, *On the strength of Ramsey’s theorem*, Notre Dame J. Formal Logic **36** (1995), no. 4, 570–582, Special Issue: Models of arithmetic.
- [28] Stephen G. Simpson, *Subsystems of Second Order Arithmetic*, second ed., Perspectives in Logic, Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, 2009. MR 2517689 (2010e:03073)
- [29] Theodore A. Slaman,  *$\Sigma_n$ -bounding and  $\Delta_n$ -induction*, Proc. Amer. Math. Soc. **132** (2004), no. 8, 2449–2456. MR 2052424
- [30] Neil Thapen, *A model-theoretic characterization of the weak pigeonhole principle*, Ann. Pure Appl. Logic **118** (2002), no. 1-2, 175–195. MR 1934122
- [31] Xiaokang Yu and Stephen G. Simpson, *Measure theory and weak König’s lemma*, Arch. Math. Logic **30** (1990), no. 3, 171–180. MR MR1080236 (91i:03112)

INSTITUTE FOR MATHEMATICAL SCIENCES, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 118402

*Email address:* `belanger@nus.edu.sg`

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 119076

*Email address:* `chongct@nus.edu.sg`

INSTITUTE OF LOGIC AND COGNITION AND DEPARTMENT OF PHILOSOPHY, SUN YAT-SEN UNIVERSITY, GUANGZHOU, CHINA

*Email address:* `wwang.cn@gmail.com`

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 119076

*Email address:* `matwong@nus.edu.sg`

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 119076

*Email address:* `matyangy@nus.edu.sg`