INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Induction rules in bounded arithmetic

Emil Jeřábek

Preprint No. 47-2018

PRAHA 2018

Induction rules in bounded arithmetic

Emil Jeřábek*

Institute of Mathematics of the Czech Academy of Sciences
Žitná 25, 115 67 Praha 1, Czech Republic, email: jerabek@math.cas.cz

September 27, 2018

Abstract

We study variants of Buss's theories of bounded arithmetic axiomatized by induction schemes disallowing the use of parameters, and closely related induction inference rules. We put particular emphasis on $\hat{\Pi}^b_i$ induction schemes, which were so far neglected in the literature. We present inclusions and conservation results between the systems (including a witnessing theorem for T^i_2 and S^i_2 of a new form), results on numbers of instances of the axioms or rules, connections to reflection principles for quantified propositional calculi, and separations between the systems.

1 Introduction

Commonly studied theories of arithmetic, weak and strong alike, are typically axiomatized by variants of induction or other axiom schemes (comprehension, collection, ...) restricted to suitable classes of formulas, where these formulas may freely use *parameters*: arbitrary numbers or other objects manipulated by the theory that enter the induction formula by means of free variables, unrelated to the induction variable. This generally makes the theories robust in their formal properties, and intuitive to work with. Nevertheless, induction schemes without parameters proved fruitful to study in the context of strong subtheories of Peano arithmetic (Σ_n-induction), revealing a landscape of strange, and yet familiar systems: see e.g. Kaye, Paris, and Dimitracopoulos [29], Adamowicz and Bigorajska [1], Bigorajska [5], Beklemishev [3, 4], and Cordón-Franco and Lara-Martín [19].

On the one hand, the parameter-free induction schemes $I\Sigma^-_n$ and $I\Pi^-_n$ are close to the original schemes with parameters $I\Sigma_n$, as the theories are conservative over each other with respect to large classes of sentences (though the correspondence is a bit off, as $I\Pi^-_{n+1}$ is on the same level as $I\Sigma_n$ and $I\Sigma^-_n$). On the other hand, there are substantial differences: as already alluded to, the Π_n schemes without parameters become genuinely distinct from (and

*The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) / ERC grant agreement no. 339691. The Institute of Mathematics of the Czech Academy of Sciences is supported by RVO: 67985840.

weaker than) the matching Σ_n schemes, whereas $I\Sigma_n = I\Pi_n$; neither $I\Sigma_n^-$ nor $I\Pi_n^-$ are finitely axiomatizable, in contrast to $I\Sigma_n$.

The parameter-free schemes $I\Sigma_n^-$ and $I\Pi_n^-$ are intimately connected to induction *rules* $I\Sigma_n^R$ and $I\Pi_n^R$: here, instead of theories generated just by axioms on top of the usual rules of first-order logic, we consider a form of induction as an additional (Hilbert-style) rule of inference. It turns out $I\Sigma_n^-$ is the weakest theory all of whose extensions are closed under $I\Sigma_n^R$, and likewise for Π_n. An important role in the analysis of $I\Sigma_n^-$ and $I\Pi_n^-$ is played by *reflection principles* for fragments of arithmetic [3, 4]: while $I\Sigma_n$ is equivalent to a certain uniform (global) reflection principle, the theories $I\Sigma_n^-$ and $I\Pi_n^-$ can be characterized using relativized *local* reflection principles. There are also intricate connections relating the nesting of applications of rules and the number of instances of axioms. As an alternative to reflection principles, parameter-free induction schemes can be analysed using *local induction* [19].

In contrast to all these results, much less is known about parameter-free induction axioms and induction rules in the context of bounded arithmetic: the early work of Kaye [28] introduced the parameter-free subtheories IE_i^- of $I\Delta_0$, while the only investigation of parameter-free Buss's theories was done by Bloch [6], who studied proof-theoretically Σ_i^b parameter-free induction rules[1] in a sequent formalism, and Cordón-Franco, Fernandéz-Margarit, and Lara-Martín [18], whose main results concern conservativity of the theories S_2^i and T_2^i over the parameter-free and induction-rule versions of $\hat{\Sigma}_i^b$-*PIND* and $\hat{\Sigma}_i^b$-*IND*, and conservativity of $BB\Sigma_i^b$ over its rule version. They rely on model-theoretic methods exploiting variants of existentially closed models.

The purpose of this paper is to study parameter-free versions of Buss's theories in a more systematic way, filling in various gaps in our knowledge to obtain a more complete picture. Some highlights are as follows. We will investigate $\hat{\Pi}_i^b$ schemes and rules, which were so far entirely ignored in the literature, alongside their $\hat{\Sigma}_i^b$ counterparts; in particular, we will prove conservation results of T_2^i and S_2^i over $\hat{\Pi}_i^b$-$(P)IND^-$. We try to get as complete a description of the relationships among the systems in question as possible; to this end, we also include tentative separation results (conditional or relativized). While bounded arithmetic is too weak to prove the consistency of interesting first-order theories, it has a well-known connection to propositional proof systems; in accordance with this, we will present characterizations of our systems in terms of variants of reflection principles for fragments of the quantified propositional sequent calculus. We also include some results on the nesting of rules, namely conditions ensuring that closure under the induction rules collapses to unnested closure, and conservation results of n instances of parameter-free induction axioms over n applications of induction rules.

The paper is organized as follows. After some preliminary background in Section 2, we introduce in Section 3 the main axioms and rules that we are interested in, and we prove some of their elementary properties—primarily reductions between the rules (Theorem 3.5), but also a result on a collapse of $\hat{\Pi}_i^b$-$(P)IND^R$ to unnested applications (Theorem 3.7). We discuss various variants of the axioms and rules in Section 4, and we show them mostly equivalent to our main systems (Proposition 4.2).

[1]Warning: the proof of Theorem 27, which effectively claims that $\hat{\Sigma}_i^b$-$(P)IND^- \equiv \hat{\Sigma}_i^b$-$(P)IND^R$, is incorrect.

The most substantial technical part of the paper comes in Section 5, which is devoted to conservation results. We recall the conservation of T^i_2 and S^i_2 over $\hat{\Sigma}^b_i$-$(P)\mathit{IND}^R$ (Theorem 5.1) from [6, 18], and we set out to prove an analogous conservation result over $\hat{\Pi}^b_i$-$(P)\mathit{IND}^R$ (Theorem 5.9). A key part of the proof is a new witnessing theorem for $\forall\exists\forall\hat{\Sigma}^b_{i-1}$ consequences (and $\forall\exists\forall\hat{\Sigma}^b_i$ consequences) of T^i_2 and S^i_2, which may be of independent interest (Theorem 5.4 and Proposition 5.5). We obtain conservation results over Γ-$(P)\mathit{IND}^-$, summarized in Corollary 5.14, and a result on collapse of nesting of $\hat{\Sigma}^b_i$-$(P)\mathit{IND}^R$ (Theorem 5.10). We also prove more direct conservation results of $T + \Gamma$-$(P)\mathit{IND}^-$ over $T + \Gamma$-$(P)\mathit{IND}^R$ for arbitrary theories T (Theorem 5.20).

We discuss connections to propositional proof systems in Section 6, the main result being a characterization of Γ-$(P)\mathit{IND}^R$ and Γ-$(P)\mathit{IND}^-$ in terms of reflection principles for quantified propositional calculi (Theorem 6.5). Section 7 is devoted to separations between our systems: we present some conditional separations in Section 7.1, and unconditional relativized separations in Section 7.2. We conclude the paper with a few remarks in Section 8.

2 Notation and preliminaries

We assume the reader is familiar with the basics of bounded arithmetic. We will work in the framework of Buss's one-sorted theories S^i_2 and T^i_2, as presented e.g. in Buss [7], Hájek and Pudlák [20, Ch. V], or Krajíček [31]. It would not be too difficult to adapt our results to the setting of two-sorted theories V^i as in Cook and Nguyen [16], but we find the one-sorted setting simpler to use for the present purpose.

In order not to get bogged down in trivial technicalities, we will employ a robust base theory in a rich language in place of Buss's *BASIC*: let BTC^0 denote the basic first-order theory for TC^0, in a language L_{TC^0} with function symbols for all TC^0 functions so that BTC^0 is a universal theory. We are not very particular about its exact definition; for example, we may axiomatize it as the theory Δ^b_1-CR of Johannsen and Pollett [25] expanded with function symbols for all Σ^b_1-definable functions of the theory, or as the equivalent theory TTC^0 of Clote and Takeuti [13]. Note that BTC^0 is RSUV-isomorphic to the theory VTC^0 (or rather, $\overline{\mathit{VTC}^0}$) of Cook and Nguyen [16]. Unless stated otherwise, we will assume all first-order theories to be formulated in L_{TC^0} and to extend BTC^0.

If Γ is a (possibly empty) set of sentences, and φ a sentence, we write $\Gamma \vdash \varphi$ if φ is provable in the theory $\mathit{BTC}^0 + \Gamma$. We may omit outermost universal quantifiers when writing down Γ or φ, as is the customary fashion. We may also write $\Gamma \vdash \Delta$ for a set of sentences Δ, meaning $\Gamma \vdash \varphi$ for all $\varphi \in \Delta$. We stress that $\mathit{BTC}^0 + \Gamma$ is only closed under the standard deduction rules of first-order logic (i.e., it includes logically valid sentences, and it is closed under modus ponens); it is not supposed to be closed under the Δ^b_1-CR rule even if we define BTC^0 as in [25].

Let $\hat{\Sigma}^b_i$ and $\hat{\Pi}^b_i$ denote the classes of *strict* Σ^b_i and Π^b_i formulas in L_{TC^0}: that is, $\hat{\Sigma}^b_0 = \hat{\Pi}^b_0 = \Sigma^b_0 = \Pi^b_0$ is the class of sharply bounded formulas, and for $i > 0$, a $\hat{\Sigma}^b_i$ formula ($\hat{\Pi}^b_i$ formula) consists of i alternating (possibly empty) blocks of bounded quantifiers followed by a Σ^b_0 formula, where the first block is existential (universal, resp.). Equivalently, we could further restrict the blocks to a single quantifier apiece. Note that every Σ^b_0 formula is equivalent to an atomic formula in BTC^0. The class of all bounded formulas is denoted Σ^b_∞.

We will combine notations such as $\hat\Sigma^b_i$ and $\hat\Pi^b_i$ with symbolic prefixes denoting *unbounded* quantifiers: for example, $\forall\exists\hat\Sigma^b_i$ denotes the class of formulas (in most contexts, sentences) consisting of a block of universal quantifiers, followed by a block of existential quantifiers, followed by a $\hat\Sigma^b_i$ formula.

Let Γ be a class of sentences, and T a theory. The Γ-*fragment of* T is the theory axiomatized by $BTC^0 + \{\varphi \in \Gamma : T \vdash \varphi\}$. If S is another theory, T is Γ-*conservative over* S if the Γ-fragment of T is included in S.

Let Σ^*_1 denote the least class of formulas that includes bounded formulas, and is closed under existential and bounded universal quantifiers; Π^*_1 denotes the dual class. A model-theoretic characterization of these classes is that Π^*_1 formulas are preserved downwards in cuts, and Σ^*_1 formulas upwards.

Theorem 2.1 (Parikh) *Let T be a Π^*_1-axiomatized extension of BTC^0, and $\varphi \in \Sigma^*_1$. If $T \vdash \forall x\, \exists y\, \varphi(x, y)$, there exists a term t such that $T \vdash \forall x\, \exists y \le t(x)\, \varphi(x, y)$.* □

We will occasionally use that Σ^*_1-sentences true in the standard model of arithmetic $\mathbb{N}$ are provable in BTC^0.

Another fundamental tool for studying systems of bounded arithmetic is Buss's witnessing theorem. We are actually not interested in witnessing per se, but in the following consequence:

Theorem 2.2 (Buss) *For any $i \ge 0$, S^{i+1}_2 is a $\forall\hat\Sigma^b_{i+1}$-conservative extension of T^i_2.* □

We will in fact use it in an ostensibly stronger form:

Corollary 2.3 *For any $i \ge 0$ and $T \subseteq \forall\hat\Sigma^b_i$, $S^{i+1}_2 + T$ is $\forall\exists\hat\Sigma^b_{i+1}$-conservative over $T^i_2 + T$.*

Proof: Assume that $S^{i+1}_2 + \forall z\, \psi(z) \vdash \forall x\, \exists y\, \varphi(x, y)$, where $\psi \in \hat\Sigma^b_i$, and $\varphi \in \hat\Sigma^b_{i+1}$. Then S^{i+1}_2 proves $\forall x\, \exists y\, \bigl(\neg\psi(y) \vee \varphi(x, y)\bigr)$. By Parikh's theorem, we may bound the y quantifier by a term in x, which makes the statement (equivalent to) a $\forall\hat\Sigma^b_{i+1}$ sentence. Thus, it is provable in T^i_2 by Theorem 2.2, and this implies $T^i_2 + \forall z\, \psi(z) \vdash \forall x\, \exists y\, \varphi(x, y)$. □

Our basic objects of study will be *rules* rather than just axiom schemes. Here, a rule R is a set of pairs $\langle\Gamma, \varphi_0\rangle$, where φ_0 is a sentence, and $\Gamma = \{\varphi_1, \dots, \varphi_n\}$ is a finite set of sentences; each $\langle\Gamma, \varphi_0\rangle \in R$ is called an *instance* of R, and will be written more conspicuously as $\Gamma \,/\, \varphi_0$, or

$$\frac{\varphi_1 \quad \varphi_2 \quad \cdots \quad \varphi_n}{\varphi_0}. \tag{1}$$

The instance above is *n-ary*. We will identify axiom schemes with 0-ary rules. Again, we will often omit outermost universal quantifiers from the sentences φ_i when writing down rules like (1).

If T is a theory, and R a rule, then $T + R$ denotes the least theory T' (i.e., deductively closed set of sentences) which includes T, and which is *closed under* R, meaning that for any instance $\Gamma \,/\, \varphi$ of R, if $\Gamma \subseteq T'$, then $\varphi \in T'$.

A rule R is *weakly reducible* to a rule S if $T+R \subseteq T+S$ for all theories T, and R and S are *weakly equivalent* if they are weakly reducible to each other. Note that R is weakly reducible to S iff for any instance $\Gamma \,/\, \varphi$ of R, $\varphi \in BTC^0 + \Gamma + S$.

We may stratify this definition by counting the nesting depth of applications of the rules. Let $[T, R]$ denote the closure of T under *unnested* applications of R-instances, i.e., the theory axiomatized by

$$T \cup \{\varphi : \Gamma \,/\, \varphi \in R, T \vdash \Gamma\},$$

and we define $[T, R]_0 = T$, $[T, R]_{n+1} = [[T, R]_n, R]$ by induction on $n \in \omega$. Notice that $T + R = \bigcup_n [T, R]_n$. We say that R is *reducible* to S, written $R \leq S$, if $[T, R] \subseteq [T, S]$ for every theory T, and R and S are *equivalent*, written $R \equiv S$, if $R \leq S \leq R$. As above, we have that $R \leq S$ iff $\varphi \in [BTC^0 + \Gamma, S]$ for each instance $\Gamma \,/\, \varphi$ of R. See also Remark 3.6.

We remark that just like sets of axioms are represented uniquely up to equivalence by *theories*, rules can be represented up to weak equivalence by *finitary consequence relations,* extending the standard first-order consequence relation of BTC^0.

Aside from bounded arithmetic, we will also assume (especially in Section 6) familiarity with basic propositional proof complexity, and in particular with the quantified propositional sequent calculus G (see [31, 16]). The classes Σ^q_i and Π^q_i of quantified propositional formulas are defined as usual: $\Sigma^q_0 = \Pi^q_0$ consists of quantifier-free formulas; Σ^q_{i+1} and Π^q_{i+1} include $\Sigma^q_i \cup \Pi^q_i$, and are closed under $\wedge$ and $\vee$; Σ^q_{i+1} is closed under existential quantifiers, and Π^q_{i+1} under universal quantifiers; negations of Σ^q_{i+1} formulas are Π^q_{i+1}, and vice versa.

Following [16], we define G_i for $i > 0$ as G restricted so that all cut-formulas are Σ^q_i. When the sequent to be proved consists of Σ^q_i formulas, this is equivalent to the original definition as in [31]. Note that up to polynomial simulation, we could allow Π^q_i cut-formulas in G_i as well; on the other hand, we could restrict cut-formulas to *prenex* Σ^q_i formulas only [24]. Let G^*_i denote the tree-like version of G_i. For $i = 0$, we define G_0 as extended Frege, optionally considered as a proof system for prenex Σ^q_1 formulas (the system introduced as *ePK* in [16]).

If P is a quantified propositional proof system, and $j \geq 0$, then $\mathrm{RFN}_j(P)$ denotes the Σ^q_j-reflection principle for P. If $j = 0$, we take this to mean the $\hat{\Pi}^b_1$ reading of the principle: "for every proof of a quantifier-free formula A, and every evaluation of subformulas of A that respects the connectives, the value assigned to A is 1" (Π^q_0-RFN_P in the notation of [16, §X.2.3]). (This can make a difference, as BTC^0 does not necessarily prove that any given quantifier-free formula can be evaluated.) Note that for all proof systems we are going to consider, this form of RFN_0 is BTC^0-provably equivalent to consistency.

3 Main systems

We are ready to introduce the main axioms and rules that will be the topic of this paper. In the rest of this section, we will show their basic properties, most importantly reductions (inclusions) among the rules.

Definition 3.1 Let $\Gamma = \hat{\Sigma}^b_i$ or $\Gamma = \hat{\Pi}^b_i$, where $i \geq 0$. The *induction* and *polynomial induction* axiom schemes are defined as usual:

$$(\Gamma\text{-}\mathit{IND}) \qquad \varphi(0,y) \land \forall x\,(\varphi(x,y) \to \varphi(x+1,y)) \to \forall x\,\varphi(x,y),$$

$$(\Gamma\text{-}\mathit{PIND}) \qquad \varphi(0,y) \land \forall x\,(\varphi(\lfloor x/2 \rfloor,y) \to \varphi(x,y)) \to \forall x\,\varphi(x,y),$$

where $\varphi \in \Gamma$. The corresponding induction *rules* are

$$(\Gamma\text{-}\mathit{IND}^R) \qquad \frac{\varphi(0,y) \quad \varphi(x,y) \to \varphi(x+1,y)}{\varphi(x,y)},$$

$$(\Gamma\text{-}\mathit{PIND}^R) \qquad \frac{\varphi(0,y) \quad \varphi(\lfloor x/2 \rfloor,y) \to \varphi(x,y)}{\varphi(x,y)}.$$

The variable y is a *parameter* of these axioms and rules (we could equivalently allow a tuple of parameters, as this can be encoded by a single parameter using a pairing function). The corresponding *parameter-free* schemes, denoted by superscript $^-$, are obtained by omitting y, i.e., φ has no free variables besides x.

The familiar theories S^i_2 and T^i_2 are defined as $BTC^0 + \hat{\Sigma}^b_i$-*PIND* and $BTC^0 + \hat{\Sigma}^b_i$-*IND*, respectively.

Remark 3.2 The cases $i = 0$ of our schemes and rules are idiosyncratic in various ways: first, $\hat{\Sigma}^b_0 = \hat{\Pi}^b_0$; second, $\hat{\Sigma}^b_0$ is closed under neither bounded existential nor bounded universal quantifiers, which is going to break some constructions; and third, $\hat{\Sigma}^b_0$-*PIND* and their parameter-free and rule variants are already derivable in the base theory BTC^0 (that is, in our language, $S^0_2 = BTC^0$, whereas T^0_2 is essentially PV_1).

The standard theories with parameters T^i_2 and S^i_2 are axiomatizable by bounded formulas (i.e., $\forall\Sigma^b_\infty$ sentences), since the *IND* axiom as stated above is equivalent to

$$\forall z\,\bigl(\varphi(0,y) \land \forall x < z\,(\varphi(x,y) \to \varphi(x+1,y)) \to \varphi(z,y)\bigr),$$

and similarly for *PIND*. The proof of this equivalence uses z as a parameter, hence it is not obvious that this should hold for the parameter-free schemes as well. Nevertheless, the $\hat{\Pi}^b_i$-$(P)\mathit{IND}^-$ schemes do have, for $i > 0$, bounded axiomatizations (specifically, by $\forall\hat{\Sigma}^b_{i+1}$ sentences), similarly to the case with parameters: if $\varphi \in \hat{\Pi}^b_i$, then

$$(2) \qquad \forall x\,\bigl(\varphi(0) \land \forall y < x\,(\varphi(y) \to \varphi(y+1)) \to \varphi(x)\bigr)$$

is provable by induction on the $\hat{\Pi}^b_i$ formula $\psi(x) = \forall y \le x\,\varphi(y)$, as

$$\vdash \forall y < x\,(\varphi(y) \to \varphi(y+1)) \land \neg\varphi(x) \to \forall z\,(\psi(z) \to \psi(z+1)),$$

and similarly for *PIND*. This argument does not seem to work for $\hat{\Sigma}^b_i$-$(P)\mathit{IND}^-$, though.

A crucial property is that induction *rules* are equivalent to their parameter-free versions. The case of $\hat{\Sigma}^b_i$ was already proved in [18], but we include it for completeness anyway.

Lemma 3.3 *If* $\Gamma = \hat\Sigma^b_i$ *or* $\hat\Pi^b_i$ *for* $i \ge 0$, *then* $\Gamma\text{-}(P)IND^R \equiv \Gamma\text{-}(P)IND^{R-}$.

Proof: Let $\langle x,y\rangle$ be a TC^0 pairing function nondecreasing in x such that $\langle x,y\rangle \ge x+y$, provably in BTC^0. If $i \le j \le |x|$, let $x_{[i,j)}$ denote the number whose binary representation consists of the ith through $(j-1)$th binary digits of x, where the *most significant* digit has index 0; i.e., $x_{[i,j)} = \lfloor x/2^{|x|-j}\rfloor \bmod 2^{j-i}$.

An instance of $\hat\Sigma^b_i$-IND^R for a formula $\varphi(x,y)$ can be reduced to $\hat\Sigma^b_i$-IND^R for the formula $z = 0 \vee \varphi(z_{[m,|z|)}, z_{[1,m)})$, where $m = \lceil |z|/2\rceil$: we have either $z_{[m,|z|)} = 0$, or $|z| = |z-1|$, $z_{[m,|z|)} = (z-1)_{[m,|z|)} + 1$, and $z_{[1,m)} = (z-1)_{[1,m)}$.

Since $\hat\Pi^b_0 = \hat\Sigma^b_0$ and $BTC^0 \vdash \hat\Sigma^b_0$-$PIND$, we may assume $i > 0$ in the remaining cases.

For $\hat\Pi^b_i$-IND^R, let $\varphi(x,y) \in \hat\Pi^b_i$, and put $\psi(z) = \forall x,y \le z\,(\langle x,y\rangle \le z \to \varphi(x,y))$. Then

$$\begin{aligned} \varphi(0,y) &\vdash \psi(0),\\ \varphi(0,y), \varphi(x,y) \to \varphi(x+1,y) &\vdash \psi(z) \to \psi(z+1),\\ &\vdash \psi(\langle x,y\rangle) \to \varphi(x,y). \end{aligned}$$

For $\hat\Pi^b_i$-$PIND^R$, we may use $\psi(z) = \forall u \le |z|\,\varphi(z \bmod 2^u, \lfloor z/2^u\rfloor)$ in a similar fashion. In order to verify

$$\varphi(0,y), \varphi(\lfloor x/2\rfloor, y) \to \varphi(x,y) \vdash \psi(\lfloor z/2\rfloor) \to \psi(z),$$

assume $z > 0$, and let $u \le |z|$. Put $x = z \bmod 2^u$, $y = \lfloor z/2^u\rfloor$. If $u = 0$, we have $x = 0$, and $\varphi(0,y)$ holds by assumption. Otherwise put $z' = \lfloor z/2\rfloor$, $u' = u-1$, $x' = z' \bmod 2^{u'}$, and $y' = \lfloor z'/2^{u'}\rfloor$. We have $u' \le |z'|$, $x' = \lfloor x/2\rfloor$, and $y' = y$, hence $\varphi(\lfloor x/2\rfloor, y)$ by the induction hypothesis, which implies $\varphi(x,y)$ by assumption.

For $\hat\Sigma^b_i$-$PIND^R$, let $\varphi(x,y)$ be a $\hat\Sigma^b_i$ formula of the form $\exists u \le t(x,y)\,\theta(x,y,u)$ with $\theta \in \hat\Pi^b_{i-1}$. Fix a suitable sequence encoding with $(w)_i$ being the ith element of the sequence coded by w, and $b(z)$ a term such that every sequence w of length at most $|z|$, each of whose entries is bounded by $t(x,y)$ for some $x,y \le z$, satisfies $w \le b(z)$. Let $\psi(z)$ be the $\hat\Sigma^b_i$ formula

$$\exists w \le b(z)\,\forall i,j \le |z|\,\big(\langle i,j\rangle < |z| \to (w)_{\langle i,j\rangle} \le t(z_{[j,i+j)}, z_{[0,j)}) \wedge \theta(z_{[j,i+j)}, z_{[0,j)}, (w)_{\langle i,j\rangle})\big).$$

Again, the least obvious property to check is that assuming the premises of $\hat\Sigma^b_i$-$PIND^R$ for φ, we can derive $\psi(\lfloor z/2\rfloor) \to \psi(z)$. Let $z > 0$, $z' = \lfloor z/2\rfloor$, and assume that w' is a sequence of length $|z'|$ witnessing $\psi(z')$. We will construct a sequence w witnessing $\psi(z)$. If $\langle i,j\rangle < |z'| = |z|-1$, then $i+j < |z'|$, thus $z'_{[j,i+j)} = z_{[j,i+j)}$ and $z'_{[0,j)} = z_{[0,j)}$, and we may take $(w)_{\langle i,j\rangle} = (w')_{\langle i,j\rangle}$. If $\langle i,j\rangle = |z'|$, put $x = z_{[j,i+j)}$, $y = z_{[0,j)}$. Either $i = 0$, in which case $x = 0$ and $\varphi(0,y)$ holds, or $\langle i-1,j\rangle < \langle i,j\rangle$, $z'_{[0,j)} = y$, and $z'_{[j,j+i-1)} = \lfloor x/2\rfloor$. We have $\varphi(\lfloor x/2\rfloor, y)$ as witnessed by $(w)_{\langle i-1,j\rangle}$, hence $\varphi(x,y)$. Either way, we can extend w' to w so that $(w)_{\langle i,j\rangle}$ is a witness for $\varphi(x,y)$, and then w witnesses $\psi(z)$. □

Corollary 3.4 $BTC^0 + \Gamma\text{-}(P)IND^-$ *is the weakest theory all of whose extensions are closed under* $\Gamma\text{-}(P)IND^R$.

Proof: On the one hand, it is clear that any extension of $\Gamma\text{-}(P)IND^-$ derives $\Gamma\text{-}(P)IND^{R-}$, hence $\Gamma\text{-}(P)IND^R$ by Lemma 3.3. On the other hand, assume that all extensions of T are closed under Γ-$(P)IND^R$. Let $\varphi \to \psi$ be any instance of Γ-$(P)IND^-$ as in Definition 3.1 (here, φ and ψ are sentences). Then $\varphi \,/\, \psi$ is an instance of Γ-$(P)IND^R$, thus $T + \varphi \vdash \psi$ by assumption. The deduction theorem then gives $T \vdash \varphi \to \psi$. $\square$

The next result presents all reductions between our core rules that we know about; they are summarized in Fig. 3.1. We will argue in Section 7 that no other reductions are likely waiting to be discovered.

Theorem 3.5 *Let $i \geq 0$, and Γ be $\hat{\Sigma}^b_i$ or $\hat{\Pi}^b_i$.*

(i) Γ-$(P)IND^R \leq \Gamma$-$(P)IND^- \leq \Gamma$-$(P)IND$.

(ii) $\hat{\Sigma}^b_i$-$(P)IND \equiv \hat{\Pi}^b_i$-$(P)IND$.

(iii) $\hat{\Pi}^b_i$-$(P)IND^- \leq \hat{\Sigma}^b_i$-$(P)IND^-$, *and* $\hat{\Pi}^b_i$-$(P)IND^R \leq \hat{\Sigma}^b_i$-$(P)IND^R$.

(iv) Γ-$PIND \leq \Gamma$-IND, Γ-$PIND^- \leq \Gamma$-IND^-, *and* Γ-$PIND^R \leq \Gamma$-IND^R.

(v) $\hat{\Sigma}^b_i$-$IND \leq \hat{\Sigma}^b_{i+1}$-$PIND^R$. *(See also Corollary 5.12.)*

(vi) $\hat{\Sigma}^b_i$-$IND^- \leq \hat{\Pi}^b_{i+1}$-$PIND^-$, *and* $\hat{\Sigma}^b_i$-$IND^R \leq \hat{\Pi}^b_{i+1}$-$PIND^R$.

Proof: (i) is an immediate consequence of Lemma 3.3.

(ii) is well known: *IND* for $\varphi(x, y)$ follows from *IND* for $\neg\varphi(a \dot{-} x, y)$, and *PIND* for φ follows from *PIND* for $\neg\varphi(\lfloor a/2^{|x|}\rfloor, y)$, where a is an additional parameter.

(iii): We may assume $i > 0$. Consider an instance of $\hat{\Pi}^b_i$-IND^R for a formula $\varphi(x, y) = \forall z \leq t(x, y)\, \theta(x, y, z)$, where $\theta \in \hat{\Sigma}^b_{i-1}$, and let $\psi(x, y, a, z)$ be the $\hat{\Sigma}^b_i$ formula

$$\varphi(a \dot{-} x, y) \wedge z \leq t(a, y) \to \theta(a, y, z).$$

Then

$$\begin{aligned} &\vdash \psi(0, y, a, z), \\ \varphi(x, y) \to \varphi(x + 1, y) &\vdash \psi(x, y, a, z) \to \psi(x + 1, y, a, z), \\ \psi(x, y, x, z) &\vdash \varphi(0, y) \to \varphi(x, y), \end{aligned}$$

showing that φ-IND^R reduces to ψ-IND^R.

In order to show $\hat{\Pi}^b_i$-$IND^- \leq \hat{\Sigma}^b_i$-$IND^-$, assume further that $\varphi(x)$ is parameter-free. Then $BTC^0 + \hat{\Sigma}^b_i$-$IND^- + \varphi(0) \wedge \forall x\, (\varphi(x) \to \varphi(x + 1))$ proves $\varphi(x)$ as it is closed under $\hat{\Sigma}^b_i$-$IND^R \geq \hat{\Pi}^b_i$-IND^R by (i), hence $BTC^0 + \hat{\Sigma}^b_i$-IND^- proves φ-IND^- by the deduction theorem.

The cases of $PIND^R$ and $PIND^-$ are similar, using $\lfloor a/2^{|x|}\rfloor$ in place of $a \dot{-} x$, as in (ii).

(iv): We may assume $i > 0$, as $BTC^0 \vdash \hat{\Sigma}^b_0$-$PIND$. *PIND* for a $\hat{\Pi}^b_i$ formula $\varphi(x, y)$ follows from *IND* for the $\hat{\Pi}^b_i$ formula $\forall u \leq x\, \varphi(x, y)$, and likewise for $PIND^-$ or $PIND^R$. *PIND* for a $\hat{\Sigma}^b_i$ formula $\varphi(x, y)$ follows from *IND* for the formula $\varphi(\lfloor a/2^{|a| \dot{-} x}\rfloor, y)$ with an additional parameter a, and this also applies to $PIND^R$. The result for $PIND^-$ follows from the result for $PIND^R$ as in the proof of (iii).

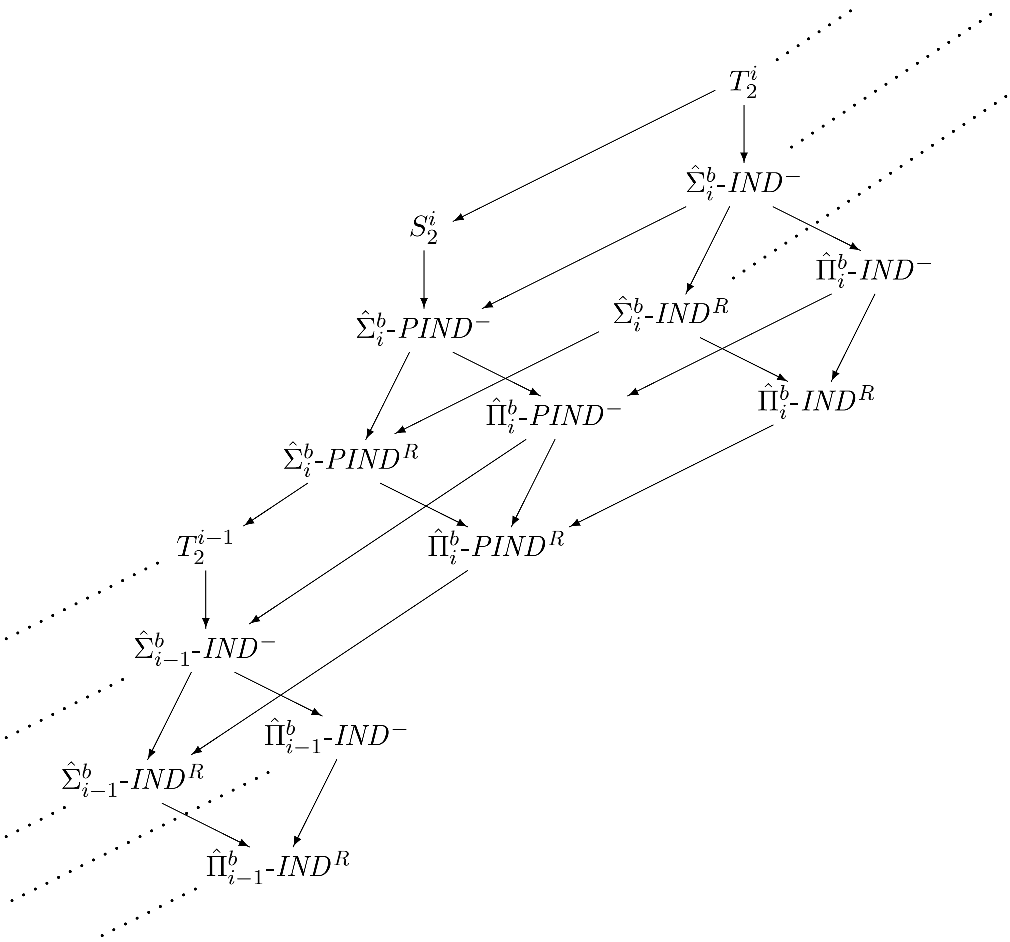

Figure 3.1: Reductions between the rules

(v): Let $\varphi(x,y) \in \hat{\Sigma}^b_i$, and let $\psi(x,y,a)$ be the $\hat{\Sigma}^b_{i+1}$ formula

$$\varphi(0,y) \wedge \neg\varphi(a,y) \to \exists u \le a, v \le \lceil a/2^{|x|} \rceil\, (u+v \le a \wedge \varphi(u,y) \wedge \neg\varphi(u+v,y)).$$

Then it is easy to check that BTC^0 proves

$$\psi(0,y,a),$$
$$\psi(\lfloor x/2 \rfloor, y, a) \to \psi(x,y,a),$$
$$\psi(a,y,a) \to \big(\varphi(0,y) \wedge \forall u < a\, (\varphi(u,y) \to \varphi(u+1,y)) \to \varphi(a,y)\big),$$

thus $[BTC^0, \hat{\Sigma}^b_{i+1}\text{-}PIND^R]$ derives the induction axiom for φ.

(vi): Let $\varphi(x,y) \in \hat{\Sigma}^b_i$, and let $\psi(x,y,z)$ be the $\hat{\Pi}^b_{i+1}$ formula

$$\forall x' \le z\, (\varphi(x',y) \wedge x + x' \le z \to \varphi(x+x',y)).$$

Then

$$\begin{aligned}
&\vdash \psi(0,y,z),\\
\varphi(x,y)\to\varphi(x+1,y) &\vdash \psi(1,y,z),\\
&\vdash \psi(x_0,y,z)\wedge\psi(x_1,y,z)\to\psi(x_0+x_1,y,z),\\
\psi(x,y,x) &\vdash \varphi(0,y)\to\varphi(x,y),
\end{aligned}$$

whence $\hat{\Sigma}^b_i$-$IND^R \le \hat{\Pi}^b_{i+1}$-$PIND^R$. The result for $\hat{\Sigma}^b_i$-IND^- follows as in (iii). □

Remark 3.6 Recall that we defined $[T,R]_n$ by counting the nesting depth of applications of R, which is in general necessary in order to make $[T,R]_n$ a deductively closed first-order theory. However, observe that unnested applications of $(P)IND^R$ for formulas $\varphi_0(x,\vec{y})$, ..., $\varphi_k(x,\vec{y})$ may be reduced to a single application of the same rule for the formula $\varphi(x,\vec{y}) = \bigwedge_{i\le k}\varphi_k(x,\vec{y})$. It follows that if Γ is closed under $\wedge$ (such as $\hat{\Sigma}^b_i$ or $\hat{\Pi}^b_i$), then $[T,\Gamma\text{-}(P)IND^R]_n$ coincides with the set of formulas provable using n *instances* of Γ-$(P)IND^R$; the same applies to $(P)IND^{R-}$.

Surprisingly, a simple argument shows that the closure of T under $\hat{\Pi}^b_i$-$(P)IND^R$ collapses to unnested applications of the rule (thus a single application is enough to prove any given consequence) under very mild assumptions on the complexity of the theory T. In particular, note that all traditional subsystems of S_2 such as S^i_2 are axiomatized by $\forall\Sigma^b_\infty \subseteq \Pi^*_1$ sentences.

Theorem 3.7 *If T is Π^*_1-axiomatized, and $i>0$, then*

$$T+\hat{\Pi}^b_i\text{-}(P)IND^R = [T,\hat{\Pi}^b_i\text{-}(P)IND^R].$$

Proof: In view of Remark 3.6, it is enough to show that $[T,\hat{\Pi}^b_i\text{-}(P)IND^R]$ includes all formulas provable using two instances of $\hat{\Pi}^b_i$-$(P)IND^{R-}$: this implies $[T,\hat{\Pi}^b_i\text{-}(P)IND^R] = [T,\hat{\Pi}^b_i\text{-}(P)IND^R]_2$, i.e., $[T,\hat{\Pi}^b_i\text{-}(P)IND^R]$ is closed under $\hat{\Pi}^b_i$-$(P)IND^R$, and as such it equals $T+\hat{\Pi}^b_i$-$(P)IND^R$. So, let $\varphi,\psi\in\hat{\Pi}^b_i$ be formulas such that

$$\begin{aligned}
T &\vdash \varphi(0),\\
T &\vdash \varphi(y)\to\varphi(y+1),\\
T+\forall y\,\varphi(y) &\vdash \psi(0),\\
T+\forall y\,\varphi(y) &\vdash \psi(x)\to\psi(x+1).
\end{aligned}$$

(The case of *PIND* is completely analogous.) Since $\psi(0)$ is a bounded sentence, we may assume it is provable in T alone. By Parikh's theorem 2.1, there is a constant c such that

$$T\vdash \forall y\le 2^{|x|^c}\,\varphi(y)\to(\psi(x)\to\psi(x+1)).$$

Put

$$\chi(z)=\forall y\le z\,\varphi(y)\wedge\forall x\le z\,(2^{|x|^c}+x\le z\to\psi(x)).$$

Then T proves $\chi(z)\to\chi(z+1)$, and $\forall z\,\chi(z)$ implies $\forall x\,\psi(x)$. □

An analogous result for $\hat{\Sigma}^b_i$-$(P)IND^R$ only applies to theories T of bounded complexity (more in line with our expectations), and it seems to require a considerably more complicated proof, see Theorem 5.10.

4 Variants

Induction and polynomial induction axioms in bounded arithmetic have equivalent variants that differ in various details (see e.g. [31, §5.2]): we may consider the length-induction scheme, variants of minimization principles, or their dual "ordinal" induction axioms, and it is not a priori clear if such variants are still equivalent without parameters. The corresponding induction rules may be varied even more: e.g., the induction base case may be moved to the conclusion of the rule (cf. [3, §2])).

For completeness, we briefly discuss such variants in this section: fortunately, most of them turn out to be equivalent to some of the axioms and rules introduced in Section 3, except for a few pathological cases.

Definition 4.1 We consider the following schemes and rules, where Γ is a set of formulas, and φ is taken from Γ:

(Γ-*LIND*)	$\varphi(0,y) \wedge \forall x\,(\varphi(x,y) \to \varphi(x+1,y)) \to \forall x\,\varphi(\lvert x\rvert,y)$
(Γ-$IND_<$)	$\forall x\,(\forall x' < x\,\varphi(x',y) \to \varphi(x,y)) \to \forall x\,\varphi(x,y)$
(Γ-$LIND_<$)	$\forall x\,(\forall x' < x\,\varphi(x',y) \to \varphi(x,y)) \to \forall x\,\varphi(\lvert x\rvert,y)$
(Γ-$PIND_<$)	$\forall x\,(\forall x'\,(\lvert x'\rvert < \lvert x\rvert \to \varphi(x',y)) \to \varphi(x,y)) \to \forall x\,\varphi(x,y)$
(Γ-$PIND_\upharpoonright$)	$\forall x\,(\forall u \le \lvert x\rvert\,(u > 0 \to \varphi(\lfloor x/2^u\rfloor,y)) \to \varphi(x,y)) \to \forall x\,\varphi(x,y)$
(Γ-*MIN*)	$\exists x\,\varphi(x,y) \to \exists x\,(\varphi(x,y) \wedge \forall x' < x\,\neg\varphi(x',y))$
(Γ-*LMIN*)	$\exists x\,\varphi(x,y) \to \exists x\,(\varphi(x,y) \wedge \forall x'\,(\lvert x'\rvert < \lvert x\rvert \to \neg\varphi(x',y)))$
(Γ-$LIND^R$)	$\varphi(0,y), \varphi(x,y) \to \varphi(x+1,y) \;/\; \varphi(\lvert x\rvert,y)$
(Γ-IND^R_0)	$\varphi(x,y) \to \varphi(x+1,y) \;/\; \varphi(0,y) \to \varphi(x,y)$
(Γ-$PIND^R_0$)	$\varphi(\lfloor x/2\rfloor,y) \to \varphi(x,y) \;/\; \varphi(0,y) \to \varphi(x,y)$
(Γ-$LIND^R_0$)	$\varphi(x,y) \to \varphi(x+1,y) \;/\; \varphi(0,y) \to \varphi(\lvert x\rvert,y)$
(Γ-$IND^R_<$)	$\forall x' < x\,\varphi(x',y) \to \varphi(x,y) \;/\; \varphi(x,y)$
(Γ-$LIND^R_<$)	$\forall x' < x\,\varphi(x',y) \to \varphi(x,y) \;/\; \varphi(\lvert x\rvert,y)$
(Γ-$PIND^R_<$)	$\forall x'\,(\lvert x'\rvert < \lvert x\rvert \to \varphi(x',y)) \to \varphi(x,y) \;/\; \varphi(x,y)$
(Γ-$PIND^R_\upharpoonright$)	$\forall u \le \lvert x\rvert\,(u > 0 \to \varphi(\lfloor x/2^u\rfloor,y)) \to \varphi(x,y) \;/\; \varphi(x,y)$
(Γ-MIN^R)	$\exists x\,\varphi(x,y) \;/\; \exists x\,(\varphi(x,y) \wedge \forall x' < x\,\neg\varphi(x',y))$
(Γ-$LMIN^R$)	$\exists x\,\varphi(x,y) \;/\; \exists x\,(\varphi(x,y) \wedge \forall x'\,(\lvert x'\rvert < \lvert x\rvert \to \neg\varphi(x',y)))$

As before, the parameter-free versions of these rules are denoted by $^-$.

Proposition 4.2 *Let* $\Gamma = \hat\Sigma^b_i$ *or* $\hat\Pi^b_i$, *where* $i \ge 0$, *and* $\overline{\Gamma}$ *be its dual.*

$$\Gamma\text{-}(P)IND^{R-}_{(0)} \equiv \Gamma\text{-}(P)IND^R \tag{3}$$

$$\Gamma\text{-}LIND^{(R)}_{(<)} \equiv \Gamma\text{-}PIND^{(R)} \tag{4}$$

$$\hat\Sigma^b_i/\hat\Pi^b_{i+1}\text{-}(P)IND^{(R)(-)}_< \equiv \hat\Pi^b_{i+1}\text{-}(P)IND^{(R)(-)} \tag{5}$$

$$\Gamma\text{-}\mathit{PIND}^{(R)(-)}_{\restriction} \equiv \Gamma\text{-}\mathit{PIND}^{(R)(-)} \tag{6}$$

$$\Gamma\text{-}(P/L)\mathit{IND}^{R}_{0} \equiv \hat{\Sigma}^b_i\text{-}(P)\mathit{IND}^{R} \tag{7}$$

$$\Gamma\text{-}(L)\mathit{MIN}^{(-)} \equiv \overline{\Gamma}\text{-}(P)\mathit{IND}^{(-)}_{<} \tag{8}$$

$$\Gamma\text{-}(L)\mathit{MIN}^{R} \equiv \Gamma\text{-}(L)\mathit{MIN} \tag{9}$$

Proof (*sketch*):

(3): The position of $\varphi(0)$ is immaterial as it is a bounded sentence, and therefore provable or refutable in BTC^0. The rest was proved in Lemma 3.3.

(4): *PIND* for $\varphi(x, y)$ can be reduced to *LIND* for $\varphi(\lfloor z/2^{|z|-x}\rfloor, y)$, while *LIND* for $\varphi(x, y)$ can be reduced to *PIND* for $\varphi(|x|, y)$. In the case of $\mathit{LIND}_<$, we may use $\forall u \leq |x|\, \varphi(u, y)$; if $\Gamma = \hat{\Sigma}^b_i$ (where w.l.o.g. $i > 0$), we write $\varphi(x, y) = \exists z \leq t(x, y)\, \theta(x, y, z)$, and use *PIND* on $\exists w\, \forall u \leq |x|\, \theta(u, (w)_u, y)$ with a suitable bound on w.

(5): $(P)\mathit{IND}^{(R)(-)}_{<}$ for $\varphi(x, y)$ follows from $(P)\mathit{IND}^{(R)(-)}$ for $\forall z \leq x\, \varphi(z, y)$. On the other hand, let $\varphi(x) = \forall z < 2^{|x|^c}\, \theta(x, z)$ with $\theta \in \hat{\Sigma}^b_i$. Then the pairing function $\langle u, v\rangle := u2^{|u|^c} + v$ satisfies $\langle u, v\rangle < \langle u', v'\rangle$ or $|\langle u, v\rangle| < |\langle u', v'\rangle|$ as long as $u < u'$ or $|u| < |u'|$ (resp.), $v < 2^{|u|^c}$, and $v' < 2^{|u'|^c}$. Thus, defining $\psi(x)$ as $r(x) < 2^{|l(x)|^c} \to \theta(l(x), r(x))$, where $l(\langle u, v\rangle) = u$ and $r(\langle u, v\rangle) = v$, $\hat{\Pi}^b_{i+1}\text{-}(P)\mathit{IND}^{(R)-}$ for φ reduces to $\hat{\Sigma}^b_i\text{-}(P)\mathit{IND}^{(R)-}_{<}$ for ψ. The case with parameters is similar, but easier.

(6): $\mathit{PIND}_{\restriction}$ for $\varphi(x, y)$ reduces to *PIND* for $\forall u \leq |x|\, \varphi(\lfloor x/2^u\rfloor, y)$; in the case of $\Gamma = \hat{\Sigma}^b_i$, we swap the outermost quantifiers as in the proof of (4).

(7): $\hat{\Sigma}^b_i\text{-}(P/L)\mathit{IND}^{R}_{0}$ is equivalent to $\hat{\Pi}^b_i\text{-}(P/L)\mathit{IND}^{R}_{0}$ as in Theorem 3.5 (ii), and it is provable from $\hat{\Sigma}^b_i\text{-}(P/L)\mathit{IND}^{R}$ by replacing $\varphi(x, y) = \exists z \leq t(x, y)\, \theta(x, y, z)$ with $z \leq t(0, y) \wedge \theta(0, y, z) \to \varphi(x, y)$. (If $i = 0$, we just take $\theta = \varphi$.)

(8): $(L)\mathit{MIN}$ for $\varphi(x, y)$ amounts to $(P)\mathit{IND}_<$ for $\neg\varphi(x, y)$.

(9): Since $\hat{\Sigma}^b_{i+1}\text{-}(L)\mathit{MIN} \equiv \hat{\Pi}^b_i\text{-}(L)\mathit{MIN}$ by (5) and (8), it suffices to show $\hat{\Pi}^b_i\text{-}(L)\mathit{MIN} \leq \hat{\Pi}^b_i\text{-}(L)\mathit{MIN}^R$. Let $\varphi(x, y) \in \hat{\Pi}^b_i$. If $i = 0$, put $\theta = \varphi$, otherwise write $\varphi(x, y) = \forall v\, \theta(x, y, v)$, where $\theta \in \hat{\Sigma}^b_{i-1}$. Let $\psi(x, y, x_0)$ be the $\hat{\Pi}^b_i$ formula

$$\theta(x_0, y, x) \to \varphi(x, y).$$

Then BTC^0 proves $\exists x\, \psi(x, y, x_0)$: either $\neg\theta(x_0, y, x)$ for some x, or $\varphi(x_0, y)$ and we may take $x = x_0$. If $\exists x\, \varphi(x, y)$, fix x_0 such that $\varphi(x_0, y)$. Then a (length-)minimal x satisfying $\psi(x, y, x_0)$ is a (length-)minimal element satisfying $\varphi(x, y)$. □

Proposition 4.2 shows that each rule from Definition 4.1 is equivalent to one of the rules introduced in Definition 3.1, except for the following, which are too weak, and thus do not fit nicely in the main hierarchy:

- $\Gamma\text{-}\mathit{LIND}^{(R)-}_{(0/<)}$: Bounded formulas applied to lengths (without non-length parameters) are essentially sharply bounded, thus LIND^- (as well as all its variants) for bounded formulas whose bounding terms are polynomials is provable in BTC^0, and full $\Sigma^b_\infty\text{-}\mathit{LIND}^-$ is provable in $BTC^0 + \Omega_2$.

- Γ-$(L)MIN^{R-}$: The premises and conclusions of these rules are Σ^0_1 sentences, hence provable in BTC^0 if true. It follows that every Σ^0_1-sound theory, and every Π^*_1-axiomatized theory, is closed under these rules.

Other common variants of induction axioms include maximization schemes. In the presence of parameters, variants of maximization are easily seen to be equivalent to the corresponding variants of minimization. However, it is unclear how to sensibly formulate maximization axioms and rules without parameters: the problem is that unlike minimization, we need an upper bound for maximization, and if this is given by an extra variable, it can be abused to encode arbitrary parameters.

5 Conservation

In this section we investigate conservation results between induction schemes with and without parameters and induction rules. The main results state that for theories T of appropriate complexity, $T + T^i_2$ ($T + S^i_2$) is conservative over $T + \hat{\Sigma}^b_i$-$(P)IND^R$ and $T + \hat{\Pi}^b_i$-$(P)IND^R$ w.r.t. suitable classes of formulas. This will also imply certain conservativity of T^i_2 (S^i_2) over $\hat{\Sigma}^b_i$-$(P)IND^-$ and $\hat{\Pi}^b_i$-$(P)IND^-$.

We start with the easier, and already understood, case of $\hat{\Sigma}^b_i$ rules. The conservation result for $\hat{\Sigma}^b_i$-$(P)IND^R$ below, which also implies a conservation result for $\hat{\Sigma}^b_i$-$(P)IND^-$, was proved by Cordón-Franco, Fernández-Margarit, and Lara-Martin [18] by model-theoretic means. It generalizes the special case for $T \subseteq \forall\hat{\Sigma}^b_i$ shown proof-theoretically by Bloch [6]; an analogous result for IE^-_n was shown earlier by Kaye [28]. We include a proof-theoretic proof of the result for completeness.

Theorem 5.1 ([18]) *Let $i \geq 0$, and T be $\forall\exists\hat{\Sigma}^b_{i+1}$-axiomatized. Then the theory $T + S^i_2$ is $\forall\hat{\Sigma}^b_i$-conservative over $T + \hat{\Sigma}^b_i$-$PIND^R$, and $T + T^i_2$ is $\forall\hat{\Sigma}^b_i$-conservative over $T + \hat{\Sigma}^b_i$-IND^R.*

Proof: We may formulate $T + S^i_2$ in sequent calculus with quantifier-free initial sequents for axioms of BTC^0, bounded quantifier introduction rules, the $PIND$ rule

$$\frac{\Gamma, \varphi(\lfloor x/2 \rfloor) \Longrightarrow \varphi(x), \Delta}{\Gamma, \varphi(0) \Longrightarrow \varphi(t), \Delta}, \tag{10}$$

where $\varphi \in \hat{\Sigma}^b_i$ (possibly with parameters not shown) and x is not free in $\Gamma \cup \Delta$, and for every axiom of T of the form $\forall x\, \exists y\, \neg\theta(x, y)$ with $\theta \in \hat{\Sigma}^b_i$, the rule

$$\frac{\Gamma \Longrightarrow \theta(t, y), \Delta}{\Gamma \Longrightarrow \Delta},$$

where y is not free in Γ, Δ, or t. By the free-cut-elimination theorem, every $\hat{\Sigma}^b_i$ formula provable in $T + S^i_2$ has a sequent proof which only contains $\hat{\Sigma}^b_i$ formulas; in particular, the side formulas $\Gamma \cup \Delta$ in each instance of the $PIND$ rule are $\hat{\Sigma}^b_i$. Then we show by (meta-)induction on the length of the proof that all sequents in the proof (that is, their equivalent formulas) are provable in $T + \hat{\Sigma}^b_i$-$PIND^R$. The step for (10) goes as follows. First, we may replace each formula $\exists u \leq s\, \psi(u)$ in Γ with $v \leq s \wedge \psi(v)$, where v is a fresh variable. This turns all formulas in Γ into $\hat{\Pi}^b_{i-1}$ formulas, hence we may negate them and move them to the right-hand side. Taking disjunction of the side formulas on the right-hand side, we are left with a rule

$$\frac{\varphi(\lfloor x/2\rfloor) \to \varphi(x) \vee \psi}{\varphi(0) \to \varphi(t) \vee \psi},$$

where $\varphi, \psi \in \hat{\Sigma}^b_i$, and x is not free in ψ. This follows from an instance of $\hat{\Sigma}^b_i$-$PIND^R_0$ for the formula $\varphi(x) \vee \psi$, and it is reducible to $\hat{\Sigma}^b_i$-$PIND^R$ by Proposition 4.2 (7).

The argument for T^i_2 is similar. □

Parikh's theorem gives

Observation 5.2 *If T is Π^*_1-axiomatized, then the $\forall\hat{\Sigma}^b_i$- and $\forall\exists\hat{\Sigma}^b_i$-fragments of T are equivalent, for each $i > 0$.* □

Corollary 5.3 *Let $i > 0$, and T be $\forall\hat{\Sigma}^b_{i+1}$-axiomatized. Then $T + S^i_2$ is $\forall\exists\hat{\Sigma}^b_i$-conservative over $T + \hat{\Sigma}^b_i$-$PIND^R$, and $T + T^i_2$ is $\forall\exists\hat{\Sigma}^b_i$-conservative over $T + \hat{\Sigma}^b_i$-IND^R.* □

In order to obtain a similar conservation result for $\hat{\Pi}^b_i$-$(P)IND^R$ (Theorem 5.9), we will need a different method. Our starting point is the following witnessing theorem, somewhat reminiscent of the KPT theorem [34]. In the context of parameter-free schemes, it is related to a conservation result for the $L\Sigma^{-\infty}_n$ scheme (called $I\Pi^{-\infty}_n$ in Kaye [26]) proved by Kaye, Paris, and Dimitracopoulos [29, Thm 2.2].

Theorem 5.4 *Let $i > 0$, T be $\forall\exists\hat{\Sigma}^b_i$-axiomatized, and $\varphi(x) \in \exists\forall\hat{\Pi}^b_i$. If $T + T^i_2$ ($T + S^i_2$) proves $\forall x\, \varphi(x)$, then there are $k \in \mathbb{N}$ and $\hat{\Pi}^b_{i-1}$ formulas $\theta_1(x_0, x_1), \dots, \theta_k(x_0, \dots, x_k)$ such that*

$$T \vdash \varphi(x_0) \vee \exists y\, \theta_j(x_0, \dots, x_{j-1}, y), \qquad j = 1, \dots, k, \tag{11}$$

$$T \vdash \bigwedge_{j=1}^{k} \theta_j(x_0, \dots, x_j) \to \varphi(x_0) \vee \bigvee_{j,l=1}^{k} \big(x_l \prec x_j \wedge \theta_j(x_0, \dots, x_{j-1}, x_l)\big), \tag{12}$$

where $y \prec x$ denotes $y < x$ ($|y| < |x|$, respectively).

Proof: Let $\{\theta_j : j \geq 1\}$ be the list of all $\hat{\Pi}^b_{i-1}$ formulas $\theta(\vec{x}, y)$ such that

$$T \vdash \varphi(x_0) \vee \exists y\, \theta(\vec{x}, y),$$

enumerated in such a way that the free variables of θ_j are among $x_0, \dots, x_{j-1}, y$. Put

$$S = T + \neg\varphi(c_0) + \{\theta_j(c_0, \dots, c_j) : j \geq 1\} + \{c_l \prec c_j \to \neg\theta_j(c_0, \dots, c_{j-1}, c_l) : j, l \geq 1\},$$

where $C = \{c_j : j \in \omega\}$ is a set of fresh constants. If the conclusion of the theorem fails, S is consistent. Let U be a maximal set of $\forall\hat{\Sigma}^b_{i-1}(C)$ sentences consistent with S. Let us fix a model $M \vDash S + U$, and put $M_0 = \{c^M_j : j \in \omega\}$.

Claim 1 *Let* $\theta(x_0,\dots,x_n,y)$ *be a* $\hat\Pi^b_{i-1}$ *formula such that* $M\vDash\exists y\,\theta(c_0,\dots,c_n,y)$*.*

(*i*) *There are* $m\ge n$ *and* $\psi\in\forall\hat\Sigma^b_{i-1}$ *such that* $M\vDash\psi(c_0,\dots,c_m)$*, and*

$$T\vdash\psi(x_0,\dots,x_m)\to\varphi(x_0)\lor\exists y\,\theta(x_0,\dots,x_n,y).$$

(*ii*) *There exists* j *such that* $M\vDash\theta(c_0,\dots,c_n,c_j)$*, and* $M\vDash\neg\theta(c_0,\dots,c_n,c_l)$ *for all* l *such that* $c_l\prec c_j$*.*

Proof:

(i): If not, then $T+\mathrm{Th}_{\forall\hat\Sigma^b_{i-1}(C)}(M)+\neg\varphi(c_0)+\forall y\,\neg\theta(\vec c,y)$ is consistent. This theory includes $S+U$, but it also contains the $\forall\hat\Sigma^b_{i-1}(C)$ sentence $\forall y\,\neg\theta(\vec c,y)$ which is not in U (being false in M), contradicting the maximality of U.

(ii): Write ψ as $\forall y\,\xi(x_0,\dots,x_m,y)$ with $\xi\in\hat\Sigma^b_{i-1}$, and let $j>m$ be such that $\theta_j(\vec x,y)$ is equivalent to $\neg\xi(\vec x,y)\lor\theta(\vec x,y)$. Then $M\vDash\theta_j(c_0,\dots,c_j)$, which means $M\vDash\theta(c_0,\dots,c_n,c_j)$ as $M\vDash\xi(c_0,\dots,c_m,c_j)$. Likewise, $M\vDash c_l\prec c_j\to\neg\theta(c_0,\dots,c_n,c_l)$. □ (Claim 1)

By part (ii) of the claim, M_0 is a $\exists\hat\Pi^b_{i-1}$-elementary substructure of M. Since $S\subseteq\forall\exists\hat\Pi^b_{i-1}(C)$, we obtain $M_0\vDash S$, in particular $M_0\vDash T+\neg\forall x\,\varphi(x)$.

It remains to show $M_0\vDash T^i_2$ (S^i_2, resp.). If $\theta(\vec c,y)$ is a $\hat\Pi^b_{i-1}$ formula with parameters from M_0 such that $M_0\vDash\exists y\,\theta(\vec c,y)$, then using $M_0\preceq_{\exists\hat\Pi^b_{i-1}}M$ and the claim, there is j such that $M_0\vDash\theta(\vec c,c_j)$, and $M_0\vDash\neg\theta(\vec c,c_l)$ for all l such that $c_l\prec c_j$. Since all elements of M_0 are of the form c_l for some l, this in fact shows

$$M_0\vDash\theta(\vec c,c_j)\land\forall y\prec c_j\,\neg\theta(\vec c,y).$$

Thus, $M_0\vDash\hat\Pi^b_{i-1}$-$(L)MIN$, which is equivalent to $\hat\Sigma^b_i$-$(P)IND$. □

As an aside, an analogous argument shows the following property, whose special case with $\varphi\in\hat\Sigma^b_i$ may be employed to give a yet another alternative proof of Theorem 5.1:

Proposition 5.5 *Let* $i\ge0$*,* T *be* $\forall\exists\hat\Sigma^b_{i+1}$*-axiomatized, and* $\varphi(x)\in\exists\forall\hat\Pi^b_{i+1}$*. If* $T+T^i_2$ ($T+S^i_2$) *proves* $\forall x\,\varphi(x)$*, then there are* $k\in\mathbb N$ *and* $\hat\Pi^b_i$ *formulas* $\theta_1(x_0,x_1),\dots,\theta_k(x_0,\dots,x_k)$ *satisfying* (11) *and*

$$T\vdash\bigwedge_{j=1}^k\theta_j(x_0,\dots,x_j)\to\varphi(x_0)\lor\bigvee_{j=1}^k\bigl(x_j\ne0\land\theta_j(x_0,\dots,x_{j-1},P(x_j))\bigr),$$

where $P(x)$ *denotes* $x-1$ ($\lfloor x/2\rfloor$*, respectively*)*.*

Proof: We use the same proof as Theorem 5.4, with $i'=i+1$ in place of i, and with axioms

$$c_j=0\lor\neg\theta_j(c_0,\dots,c_{j-1},P(c_j))$$

in place of $c_l\prec c_j\to\neg\theta_j(c_0,\dots,c_{j-1},c_l)$ in S. By the same argument, M_0 is an $\exists\hat\Pi^b_{i'-1}$-elementary substructure of M (in particular, $M_0\vDash T+\neg\forall x\,\varphi(x)$), and $M_0\vDash\hat\Sigma^b_{i'-1}$-$(P)IND$. □

Remark 5.6 The conclusion of Theorem 5.4 (and, similarly, Proposition 5.5) implies that T proves

$$(13)\quad \Big[\bigwedge_{j=1}^{k} \forall x_1,\dots,x_{j-1}\,\exists y\,\theta_j(x_0,\dots,x_{j-1},y)$$
$$\to \exists x_1,\dots,x_k \bigwedge_{j=1}^{k}\big(\theta_j(x_0,\dots,x_j)\wedge \forall z\prec x_j\,\neg\theta_j(x_0,\dots,x_{j-1},z)\big)\Big] \to \varphi(x_0),$$

which means that $\varphi(x_0)$ follows over T from a form of k-times iterated $\hat\Pi^b_{i-1}$-minimization.

This k-dimensional minimization is, similarly to Kaye's $I\Pi_n^{-(k)}$, a form of induction over the ordinal ω^k, in contrast to the usual induction over ω; this is what makes $I\Pi_n^{-\infty}$ strictly stronger than $I\Pi_n^-$. However, we will see next that in our main case of interest, the $\exists y$ quantifiers above can be bounded by a term $t(x_0)$. In that case, the induction is really over the ordinal a^k for $a = t(x_0)$, which is finite, and as such should follow from ordinary induction. We will formalize this intuition below.

Lemma 5.7 *Let $i > 0$, $T \subseteq \forall\hat\Sigma^b_i$, and $\varphi(x) \in \exists\hat\Pi^b_i$. If $T + T^i_2$ ($T + S^i_2$) proves $\forall x\,\varphi(x)$, then there are $k \in \mathbb{N}$, $\hat\Pi^b_{i-1}$ formulas $\theta_1(x_0,x_1),\dots,\theta_k(x_0,\dots,x_k)$, and a term $t(x_0)$ such that*

$$(14)\qquad \vdash y \ge t(x_0) \to \theta_j(x_0,\dots,x_{j-1},y), \qquad j = 1,\dots,k,$$

$$(15)\qquad T \vdash \bigwedge_{j=1}^{k}\theta_j(x_0,\dots,x_j) \to \varphi(x_0) \vee \bigvee_{j=1}^{k}\exists z\prec x_j\,\theta_j(x_0,\dots,x_{j-1},z),$$

where $y \prec x$ denotes $y < x$ ($|y| < |x|$, respectively).

Proof: We modify the proof of Theorem 5.4 as follows. Let $\{\langle\theta_j,t_j\rangle : j \ge 1\}$ be an enumeration of pairs $\langle\theta,t\rangle$ where $t(x)$ is a term, and $\theta(\vec x,y)$ is a $\hat\Pi^b_{i-1}$ formula of the form $y \ge t(x_0)\vee\dots$. We define

$$S = T + \neg\varphi(c_0) + \{c_j \le t_j(c_0)\wedge\theta_j(c_0,\dots,c_j)\wedge\forall z\prec c_j\,\neg\theta_j(c_0,\dots,c_{j-1},z) : j \ge 1\},$$

and U, M, and M_0 as in Theorem 5.4. Since $S+U$ is Π^*_1-axiomatized, its validity is preserved downwards to cuts; thus, in view of the axioms $c_j \le t_j(c_0)$, we may assume that every element of M is bounded by a term in c_0.

In the proof of the Claim, there exists a term t such that $M \vDash \exists y \prec t(c_0)\,\theta(c_0,\dots,c_n,y)$, hence we may assume w.l.o.g. that θ has the form $y \prec t(x_0)\wedge\dots$. We change the definition of $\theta_j(\vec x,y)$ to $y \ge t(x_0)\vee\neg\xi(\vec x)\vee\theta(\vec x,y)$, with $t_j = t$. Then M satisfies $\theta_j(\vec c,c_j)$, and $\forall z \prec c_j\,\neg\theta_j(\vec c,z)$. Either $\theta(c_0,\dots,c_n,c_j)$, in which case we are done, or $c_j = t(c_0)$. But in the latter case, we have $\exists y\prec c_j\,\theta_j(\vec c,y)$, a contradiction.

The rest of the proof is as in Theorem 5.4. □

Lemma 5.8 *In Lemma 5.7, we may take $k = 1$. That is, under the assumptions of the lemma, there is a $\hat\Pi^b_{i-1}$ formula $\theta(x,y)$ and a term $t(x)$ such that*

$$\vdash y \ge t(x) \to \theta(x,y),$$
$$T \vdash \theta(x,y) \to \varphi(x)\vee\exists z\prec y\,\theta(x,z).$$

Proof: Let us first consider the case of *IND*. Let k, t, and $\theta_1, \dots, \theta_k$ be as in Lemma 5.7. We may assume w.l.o.g. that $t(x) = 2^{|x|^c} - 1$ for some constant $c \geq 1$. A k-tuple $\langle x_1, \dots, x_k \rangle$ where $x_1, \dots, x_k < 2^{|x|^c}$ may be represented by a number $y < 2^{k|x|^c}$ as

$$y = x_1 2^{(k-1)|x|^c} + x_2 2^{(k-2)|x|^c} + \dots + x_k. \tag{16}$$

With this encoding in mind, we define a $\hat{\Pi}^b_{i-1}$ formula $\theta(x, y)$ by

$$y \geq 2^{k|x|} \vee \bigwedge_{j=1}^{k} \theta_j \left(x, \left\lfloor \frac{y}{2^{(k-1)|x|^c}} \right\rfloor \bmod 2^{|x|^c}, \dots, \left\lfloor \frac{y}{2^{(k-j)|x|}} \right\rfloor \bmod 2^{|x|^c} \right).$$

Work in T, and assume for contradiction

$$\theta(x, y) \wedge \forall z < y \, \neg\theta(x, z) \wedge \neg\varphi(x).$$

Since $\theta(x, 2^{k|x|^c} - 1)$ by (14), we must have $y < 2^{k|x|^c}$. Write $x_0 = x$, and let $x_1, \dots, x_k < 2^{|x|^c}$ be as in (16). By (15), we have

$$\neg\theta_j(x_0, \dots, x_j) \vee \exists z < x_j \, \theta_j(x_0, \dots, x_{j-1}, z)$$

for some $j = 1, \dots, k$. However, $\neg\theta_j(x_0, \dots, x_j)$ is impossible because of $\theta(x, y)$, thus let us fix $z_j < x_j$ such that $\theta_j(x_0, \dots, x_{j-1}, z_j)$, and put

$$z = x_1 2^{(k-1)|x|^c} + \dots + x_{j-1} 2^{(k-j+1)|x|^c} + (z_j + 1) 2^{(k-j)|x|^c} - 1,$$

which represents the k-tuple $\langle x_1, \dots, x_{j-1}, z_j, 2^{|x|^c} - 1, \dots, 2^{|x|^c} - 1 \rangle$. We have $\theta_l(x_0, \dots, x_l)$ for $l < j$ as $\theta(x, y)$, $\theta_j(x_0, \dots, x_{j-1}, z_j)$ by the choice of z_j, and $\theta_l(x_0, \dots, x_{j-1}, z_j, 2^{|x|^c} - 1, \dots)$ for $l > j$ by (14), hence $\theta(x, z)$ and $z < y$, a contradiction.

In the case of *PIND*, we proceed similarly, except that we encode $\langle x_1, \dots, x_k \rangle$ by

$$2^{|x_1||x|^{(k-1)c} + |x_2||x|^{(k-2)c} + \dots + |x_k| + k|x|^c} + x_1 2^{(k-1)|x|^c} + x_2 2^{(k-2)|x|^c} + \dots + x_k,$$

and we define $\theta(x, y)$ to hold if $y \geq 2^{|x|^{kc} + k|x|^c}$, or if y is a valid encoding of $\langle x_1, \dots, x_k \rangle$ such that

$$\bigwedge_{j=1}^{k} \theta_j(x, x_1, \dots, x_j).$$

It is easy to see that if y encodes $\langle x_1, \dots, x_k \rangle$, and z encodes $\langle x_1, \dots, x_{j-1}, z_j, \dots, z_k \rangle$ with $|z_j| < |x_j|$, then $|z| < |y|$. Using this property, the same proof as above shows

$$T \vdash \theta(x, y) \to \varphi(x) \vee \exists z \left(|z| < |y| \wedge \theta(x, z) \right)$$

as required. □

Theorem 5.9 *If $i > 0$ and T is $\forall\hat{\Sigma}^b_i$-axiomatized, $T + S^{i+1}_2$ ($T + S^i_2$) is $\forall\hat{\Pi}^b_i$-conservative over* $[T, \hat{\Pi}^b_i\text{-}(P)IND^R]$.

Proof: $T + S_2^{i+1}$ is $\forall\hat\Sigma^b_{i+1}$-conservative over $T + T^i_2$ by Corollary 2.3, hence it suffices to deal with T^i_2 in place of S_2^{i+1}.

Assume that $T + T^i_2$ ($T + S^i_2$) proves $\forall x\,\varphi(x)$ with $\varphi \in \hat\Sigma^b_{i-1}$, and let θ and t be as in Lemma 5.8. Putting $\psi(x, y) = \varphi(x) \vee \neg\theta(x, y)$, we have

$$T \vdash \forall z \prec y\, \psi(x, z) \to \psi(x, y),$$

hence an application of $\hat\Sigma^b_{i-1}$-$(P)IND^R_<$, equivalent to $\hat\Pi^b_i$-$(P)IND^R$ by Proposition 4.2, yields $\psi(x, y)$. Since $\theta(x, y)$ holds for all sufficiently large y, this implies $\varphi(x)$. □

Using a similar strategy, we also obtain a $\hat\Sigma^b_i$ version of Theorem 3.7:

Theorem 5.10 *If $i > 0$ and T is $\forall\hat\Sigma^b_i$-axiomatized, $T + S_2^{i+1}$ ($T + S^i_2$) is $\forall\hat\Sigma^b_i$-conservative over $[T, \hat\Sigma^b_i$-$(P)IND^R]$. In particular, $T + \hat\Sigma^b_i$-$(P)IND^R = [T, \hat\Sigma^b_i$-$(P)IND^R]$.*

Proof: Assume $T + T^i_2$ ($T + S^i_2$) proves $\forall x\,\varphi(x)$ with $\varphi \in \hat\Sigma^b_i$, and let θ and t be as in Lemma 5.8. In the case of $\hat\Sigma^b_i$-*IND*, we put

$$\psi(x, w) = \varphi(x) \vee \exists y \le t(x)\, \big(w + y \le t(x) \wedge \theta(x, y)\big),$$

and observe

$$\begin{aligned} &\vdash \psi(x, 0),\\ T &\vdash \psi(x, w) \to \psi(x, w + 1),\\ &\vdash \psi(x, t(x) + 1) \to \varphi(x), \end{aligned}$$

thus $[T, \hat\Sigma^b_i$-$IND^R] \vdash \varphi(x)$. In the case of $\hat\Sigma^b_i$-*PIND*, we use

$$\psi(x, w) = \varphi(x) \vee \exists y \le t(x)\, \big(|w| + |y| \le |t(x)| \wedge \theta(x, y)\big)$$

in a similar way. □

As we will see in Corollary 6.7, Theorem 5.10 also holds for $i = 0$.

Corollary 5.11 *Let T be $\forall\hat\Sigma^b_i$-axiomatized.*

(i) $T + S^i_2$ is $\forall\exists\hat\Sigma^b_{i-1}$-conservative over $[T, \hat\Pi^b_i$-$PIND^R]$ for $i \ge 2$.

(ii) $T + S_2^{i+1}$ is $\forall\exists\hat\Sigma^b_{i+1}$-conservative over $T + T^i_2$, $\forall\exists\hat\Sigma^b_i$-conservative over $[T, \hat\Sigma^b_i$-$IND^R]$ for $i \ge 1$, and $\forall\exists\hat\Sigma^b_{i-1}$-conservative over $[T, \hat\Pi^b_i$-$IND^R]$ for $i \ge 2$.

Proof: By Observation 5.2, Theorems 5.9 and 5.10, and Corollary 2.3. □

We can draw a few conclusions from Theorems 5.1 and 5.9. First, some of our rules collapse over sufficiently simple base theories; this is analogous to the fact that $T + I\Pi^R_{n+1} = T + I\Sigma^R_n$ for $T \subseteq \Pi_{n+1}$ (Beklemishev [3]).

Corollary 5.12 *If $i \ge 0$ and T is $\forall\hat\Sigma^b_i$-axiomatized, then $T + \hat\Pi^b_{i+1}$-$PIND^R = T + \hat\Sigma^b_i$-$IND^R$, and $T + \hat\Sigma^b_{i+1}$-$PIND^R = T + T^i_2$.*

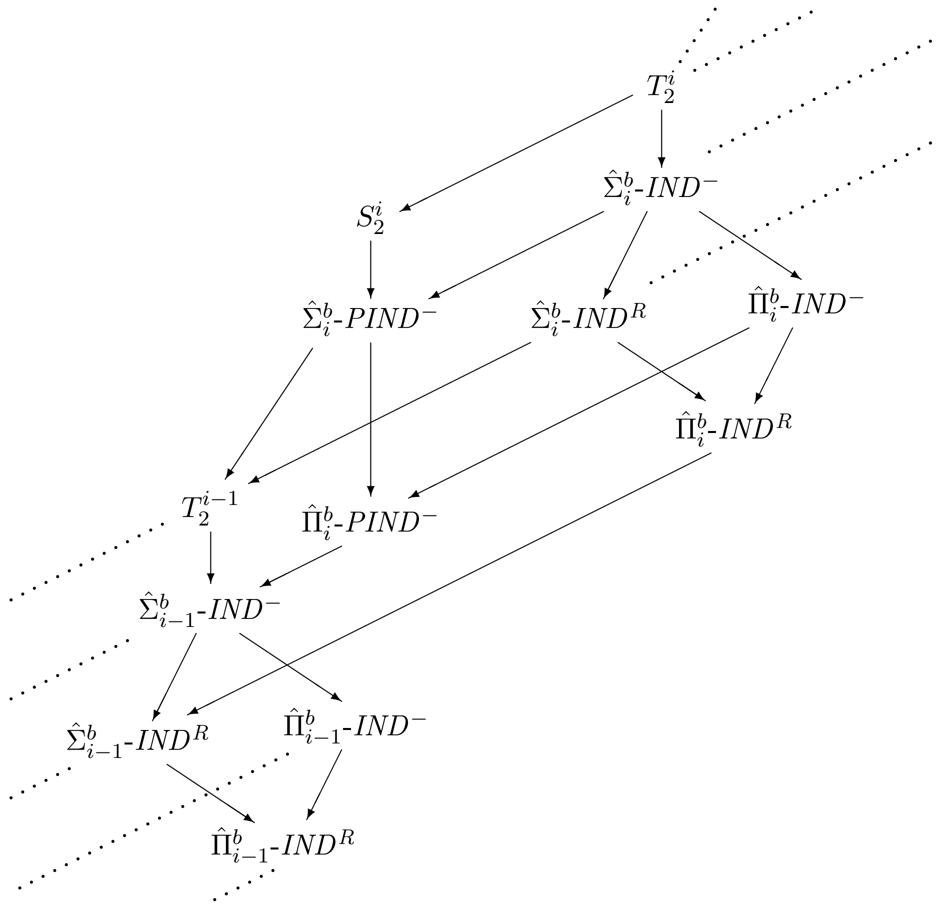

Figure 5.1: Inclusions between the theories

Proof: $T + \hat{\Pi}^b_{i+1}$-$PIND^R$ includes $T + \hat{\Sigma}^b_i$-IND^R by Theorem 3.5. On the other hand, $T + \hat{\Pi}^b_{i+1}$-$PIND^R \subseteq T + S_2^{i+1}$ is $\forall\hat{\Sigma}^b_i$-axiomatized, hence it is included in $T + \hat{\Sigma}^b_i$-IND^R by Theorem 5.10 if $i > 0$. For $i = 0$, $T + S_2^1$ is $\forall\hat{\Sigma}^b_1$-conservative over $T + T_2^0$ by Corollary 2.3, which is in turn $\forall\hat{\Sigma}^b_0$-conservative over $T + \hat{\Sigma}^b_0$-IND^R by Theorem 5.1.

Likewise, $T + T_2^i \subseteq T + \hat{\Sigma}^b_{i+1}$-$PIND^R \subseteq T + S_2^{i+1}$, and the $\forall\hat{\Sigma}^b_{i+1}$-fragment of $T + S_2^{i+1}$ is included in $T + T_2^i$ by Corollary 2.3. □

The inclusion diagram between theories axiomatized over BTC^0 by the rules from Definition 3.1, taking into account Corollary 5.12, is depicted in Figure 5.1. We will present evidence in Section 7 that no further inclusions hold.

Second, we obtain conservation results over parameter-free schemes from the corresponding results for rules and the deduction theorem. The following corollary summarizes conservativity of T_2^i or S_2^i over theories axiomatized over BTC^0 by parameter-free induction axioms or rules; since the conservations are generally for classes of sentences that include the complexity of the natural axiomatization of the theories in question, it provides their characterization as particular fragments of T_2^i or S_2^i.

Definition 5.13 If Γ is a set of *sentences*, then $\mathcal{B}(\Gamma)$ denotes the set of Boolean combinations of sentences from Γ, and $\mathcal{M}(\Gamma)$ monotone Boolean combinations of sentences from Γ.

Corollary 5.14 *Let $i \geq 0$.*

(i) $BTC^0 + \hat{\Sigma}^b_{i+1}$*-PIND*$^-$ *is the* $\mathcal{B}(\forall\hat{\Sigma}^b_{i+1})$*-fragment of* S^{i+1}_2*, and it is* $\exists\forall\hat{\Sigma}^b_{i+1}$*-conservative and* $\mathcal{M}(\exists\hat{\Pi}^b_{i+2}, \forall\exists\hat{\Sigma}^b_{i+1})$*-conservative under* S^{i+1}_2*.*

(ii) $BTC^0 + \hat{\Sigma}^b_{i+1}$*-PIND*$^R = T^i_2$ *is the* $\forall\hat{\Sigma}^b_{i+1}$*-fragment of* S^{i+1}_2*, and it is* $\forall\exists\hat{\Sigma}^b_{i+1}$*-conservative under* S^{i+1}_2*.*

(iii) $BTC^0 + \hat{\Pi}^b_{i+1}$*-PIND*$^-$ *is the* $\mathcal{M}(\exists\hat{\Pi}^b_{i+1}, \forall\hat{\Sigma}^b_i)$*-fragment of* S^{i+1}_2*, and if* $i > 0$*, it is* $\mathcal{M}(\exists\hat{\Pi}^b_{i+1}, \forall\exists\hat{\Sigma}^b_i)$*-conservative under* S^{i+1}_2*.*

(iv) $BTC^0 + \hat{\Sigma}^b_i$*-IND*$^-$ *is the* $\mathcal{B}(\forall\hat{\Sigma}^b_i)$*-fragment of* S^{i+1}_2 *or* T^i_2*, and it is* $\exists\forall\hat{\Sigma}^b_i$*-conservative under* T^i_2*. If* $i > 0$*, it is also* $\mathcal{M}(\exists\hat{\Pi}^b_{i+1}, \forall\exists\hat{\Sigma}^b_i)$*-conservative under* T^i_2*, and* $\mathcal{M}(\exists\hat{\Pi}^b_i, \forall\exists\hat{\Sigma}^b_i)$*-conservative under* S^{i+1}_2*.*

(v) $BTC^0 + \hat{\Sigma}^b_i$*-IND*$^R = BTC^0 + \hat{\Pi}^b_{i+1}$*-PIND*^{R} *is the* $\forall\hat{\Sigma}^b_i$*-fragment of* S^{i+1}_2 *or* T^i_2*, and if* $i > 0$*, it is* $\forall\exists\hat{\Sigma}^b_i$*-conservative under* S^{i+1}_2*.*

(vi) For $i > 0$*,* $BTC^0 + \hat{\Pi}^b_i$*-IND*$^-$ *is the* $\mathcal{M}(\exists\hat{\Pi}^b_i, \forall\hat{\Sigma}^b_{i-1})$*-fragment of* S^{i+1}_2 *or* T^i_2*. If* $i > 1$*, it is* $\mathcal{M}(\exists\hat{\Pi}^b_i, \forall\exists\hat{\Sigma}^b_{i-1})$*-conservative under* S^{i+1}_2*.*

(vii) For $i > 0$*,* $BTC^0 + \hat{\Pi}^b_i$*-IND*^{R} *is the* $\forall\hat{\Sigma}^b_{i-1}$*-fragment of* S^{i+1}_2 *or* T^i_2*, and if* $i > 1$*, it is* $\forall\exists\hat{\Sigma}^b_{i-1}$*-conservative under* S^{i+1}_2*.*

Proof: (i): On the one hand, each instance of $\hat{\Sigma}^b_{i+1}$-*PIND*$^-$ may be written as an implication between two $\forall\hat{\Sigma}^b_{i+1}$ sentences, and it is provable in S^{i+1}_2. On the other hand, if φ is a $\exists\forall\hat{\Sigma}^b_{i+1}$ sentence provable in S^{i+1}_2, then $BTC^0 + \neg\varphi + \hat{\Sigma}^b_{i+1}$-*PIND*$^- \supseteq BTC^0 + \neg\varphi + \hat{\Sigma}^b_{i+1}$-*PIND*^{R} is inconsistent by Theorem 5.1 and Lemma 3.3, thus $BTC^0 + \hat{\Sigma}^b_{i+1}$-*PIND*$^-$ proves φ. Likewise, an $\mathcal{M}(\exists\hat{\Pi}^b_{i+2}, \forall\exists\hat{\Sigma}^b_{i+1})$ sentence may be written as a conjunction of implications $\varphi \to \psi$, where $\varphi \in \forall\hat{\Sigma}^b_{i+2}$, and $\psi \in \forall\exists\hat{\Sigma}^b_{i+1}$. If $S^{i+1}_2 \vdash \varphi \to \psi$, then $BTC^0 + \varphi + \hat{\Sigma}^b_{i+1}$-*PIND*$^R \vdash \psi$ by Corollary 5.3, thus $BTC^0 + \hat{\Sigma}^b_{i+1}$-*PIND*$^- \vdash \varphi \to \psi$.

The other items are similar. □

Third, $\hat{\Sigma}^b_i$-induction schemes may be extended to variants of $\hat{\Delta}^b_{i+1}$-induction.

Proposition 5.15 *Let $i \geq 0$, and φ be a $\hat{\Pi}^b_{i+1}$ formula.*

(i) If φ *is provably equivalent to a* $\hat{\Sigma}^b_{i+1}$ *formula in* S^{i+1}_2*, then* $\hat{\Sigma}^b_i$*-IND*$^-$ *proves* φ*-IND*$^-$*, and* φ*-IND*^{R} *is weakly reducible to* $\hat{\Sigma}^b_i$*-IND*^{R}*.*

(ii) If φ *is provably equivalent to a* $\hat{\Sigma}^b_{i+1}$ *formula in* S^i_2*, then* $\hat{\Sigma}^b_i$*-PIND*$^-$ *proves* φ*-PIND*$^-$*, and* φ*-PIND*^{R} *is weakly reducible to* $\hat{\Sigma}^b_i$*-PIND*^{R}*.*

Proof:

(i): Let φ' be a $\hat{\Sigma}^b_{i+1}$ formula that S^{i+1}_2 proves equivalent to φ. First, recall that under the assumptions, φ-*IND* is provable in S^{i+1}_2: assuming $\forall x < a\,(\varphi(x,y) \to \varphi(x+1,y))$, we show $\forall x\,(x+z \le a \wedge \varphi'(x,y) \to \varphi(x+z,y))$ by $\hat{\Pi}^b_{i+1}$-*PIND* on z.

Now, the $\forall\hat{\Sigma}^b_i$ sentence $\forall x, y\,(\varphi'(x,y) \to \varphi(x,y))$ is provable in $[BTC^0, \hat{\Sigma}^b_i\text{-}IND^R]$ by Corollary 5.14 (v), and the $\forall\hat{\Sigma}^b_i$ / $\forall\hat{\Sigma}^b_i$ rule

$$\frac{\varphi(0,y) \quad \varphi'(x,y) \to \varphi(x+1,y)}{\varphi(x,y)}$$

is (strongly) reducible to $\hat{\Sigma}^b_i$-*IND*^{R} by Theorem 5.10, hence φ-*IND*^{R} is derivable from two instances of $\hat{\Sigma}^b_i$-*IND*^{R}. That $\hat{\Sigma}^b_i$-*IND*$^-$ proves φ-*IND* for φ parameter-free follows by the deduction theorem.

(ii) is analogous, using the fact that S^i_2 proves $\hat{\Delta}^b_{i+1}$-*PIND* [31, Cor. 8.2.7]. (For $i = 0$, if φ is $\hat{\Delta}^b_1$ in BTC^0, it is in fact Σ^b_0 in BTC^0, hence BTC^0 proves φ-*PIND*.) □

Remark 5.16 In contrast to Theorem 5.1, it is unclear whether the $\forall\hat{\Pi}^b_i$-conservativity of $T + T^i_2$ ($T + S^i_2$) over $T + \hat{\Pi}^b_i$-$(P)IND^R$ in Theorem 5.9 carries over to $\forall\exists\hat{\Sigma}^b_i$-axiomatized theories T, and whether T^i_2 (S^i_2) is $\exists\forall\hat{\Pi}^b_i$-conservative over $\hat{\Pi}^b_i$-$(P)IND^-$. (These two problems are in fact equivalent as a consequence of Theorem 5.20 below.)

Notice that the $\forall\hat{\Pi}^b_i$ consequences of $T + T^i_2$ ($T + S^i_2$) are axiomatized over T by the rule "from (13) infer $\forall x\,\varphi(x)$" for $\varphi \in \hat{\Sigma}^b_{i-1}$, and likewise, the $\exists\forall\hat{\Pi}^b_i$ consequences of T^i_2 (S^i_2) are axiomatized by the scheme

$$(17)\quad \bigwedge_{j=1}^{k} \forall x_1, \dots, x_{j-1}\,\exists y\,\theta_j(x_1, \dots, x_{j-1}, y)$$
$$\to \exists x_1, \dots, x_k \bigwedge_{j=1}^{k} \big(\theta_j(x_1, \dots, x_j) \wedge \forall z \prec x_j\,\neg\theta_j(x_1, \dots, x_{j-1}, z)\big)$$

for $k \in \mathbb{N}$ and $\theta_j \in \hat{\Pi}^b_{i-1}$. Thus, the question becomes whether $\hat{\Pi}^b_i$-$(P)IND^-$ proves (17). For $k = 1$, (17) is just $\hat{\Pi}^b_{i-1}$-$(L)MIN^-$, which is equivalent to $\hat{\Pi}^b_i$-$(P)IND^-$ by Proposition 4.2, hence another formulation is if the scheme (17) collapses to its case $k = 1$.

Question 5.17 *Let $i > 0$.*

(i) Is T^i_2 (S^i_2) $\exists\forall\hat{\Pi}^b_i$-conservative over $\hat{\Pi}^b_i$-$(P)IND^-$?

(ii) Is $T + T^i_2$ ($T + S^i_2$) $\forall\hat{\Pi}^b_i$-conservative over $T + \hat{\Pi}^b_i$-$(P)IND^R$ for every $\forall\exists\hat{\Sigma}^b_i$-axiomatized theory T?

Theorems 5.1 and 5.9 imply certain conservativity of $(P)IND^-$ over $(P)IND^R$. As we will see below, we can do better by a direct argument: the conservation results hold over base theories of arbitrary complexity, and they respect numbers of instances.

Kaye [27] gave a simple argument showing the conservativity of k instances of axioms of a particular form over k instances of the corresponding rule, with $I\Sigma^R_n$ as the main intended application. While he states the result more restrictively, his proof can be seen to give the following general statement.

Theorem 5.18 (Kaye [27]) *Let Γ and Δ be sets of sentences such that $\Gamma \vee \Delta \subseteq \Gamma$. Let $A^- = \{\alpha_j \to \beta_j : j < k\}$ be a set of k sentences satisfying $\alpha_j \in \Delta$, and A^R the set of corresponding rules $\alpha_j \vee \tau \ / \ \beta_j \vee \tau$ for $\tau \in \Gamma$. Then for any theory T, $T + A^-$ is Γ-conservative over $[T, A^R]_k$.* □

Theorem 5.18 implies a conservation result of $\hat{\Sigma}^b_i$-$(P)IND^-$ over $\hat{\Sigma}^b_i$-$(P)IND^R$ preserving numbers of instances, but it does not seem applicable to $\hat{\Pi}^b_i$-$(P)IND^R$, as the latter is not invariant under addition of $\hat{\Sigma}^b_i$ side-formulas. We remedy this defect using a modification of Kaye's argument that works under somewhat different assumptions, at the expense of employing more complicated rules (essentially, several rules from A^R working in parallel). The conservation results for $\hat{\Pi}^b_i$-$(P)IND^R$ we proved earlier then allow us to simulate these rules.

Lemma 5.19 *Let Γ and Δ be sets of sentences such that $\Gamma \vee \Delta \subseteq \Gamma$. Let $A^- = \{\alpha_j \to \beta_j : j < k\}$ be a set of k sentences satisfying $\beta_j \in \Delta$, and let $A^{R\|}$ denote the rules*

$$\frac{\bigvee_{j\in J} \alpha_j \vee \tau}{\bigvee_{j\in J} \beta_j \vee \tau}, \qquad \tau \in \Gamma, J \subseteq \{0, \dots, k-1\}.$$

Then for any theory T, $T + A^-$ is Γ-conservative over $[T, A^{R\|}]_k$.

Proof: Assume that

$$T \vdash \bigwedge_{j<k} (\alpha_j \to \beta_j) \to \varphi, \tag{18}$$

where $\varphi \in \Gamma$. We define the sentences

$$\tau_m = \varphi \vee \bigvee_{\substack{J\subseteq k \\ |J|=m}} \bigwedge_{j\in J} \beta_j,$$

$$\sigma_m = \varphi \vee \bigvee_{\substack{J\subseteq k \\ |J|=m}} \Big(\bigwedge_{j\in J} \beta_j \wedge \bigvee_{j\notin J} \alpha_j\Big)$$

for $m \le k$. Using (18), we can check easily

$$\vdash \tau_0,$$
$$\vdash \sigma_k \to \varphi,$$
$$T \vdash \tau_m \to \sigma_m,$$

it thus suffices to show $[\sigma_m, A^{R\|}] \vdash \tau_{m+1}$. Now, for every $I \subseteq k$ with $|I| = k - m$, we have

$$\sigma_m \vdash \varphi \vee \bigvee_{j\in I} \beta_j \vee \bigvee_{j\in I} \alpha_j$$

where $\varphi \vee \bigvee_{j \in I} \beta_j \in \Gamma$, hence

$$[\sigma_m, A^{R\|}] \vdash \varphi \vee \bigvee_{j \in I} \beta_j.$$

Since

$$\vdash \bigwedge_{\substack{I \subseteq k \\ |I| = k-m}} \Big(\varphi \vee \bigvee_{j \in I} \beta_j \Big) \to \tau_{m+1},$$

this gives $[\sigma_m, A^{R\|}] \vdash \tau_{m+1}$. □

Theorem 5.20 *Let $i \geq 0$, and $\Theta = \hat{\Sigma}^b_i$ or $\hat{\Pi}^b_i$. If T is an arbitrary extension of BTC^0, then $T + \Theta\text{-}(P)IND^-$ is $\forall\Theta$-conservative over $T + \Theta\text{-}(P)IND^R$.*

More precisely, all $\forall\Theta$ sentences provable from T and k instances of $\Theta\text{-}(P)IND^-$ are in $[T, \Theta\text{-}(P)IND^R]_k$.

Proof: We apply Lemma 5.19 with A^- being k instances of $\Theta\text{-}(P)IND^-$, and $\Gamma = \Delta = \forall\Theta$. The rules in $A^{R\|}$ have $\forall\hat{\Sigma}^b_i$ premises and $\forall\Theta$ conclusions, and they are clearly derivable in T^i_2 (S^i_2, resp.), hence each instance is reducible to an instance of $\Theta\text{-}(P)IND^R$ by Theorems 5.9 and 5.10. (For $i = 0$, we use Corollary 6.7 along with Theorem 5.1 instead.) □

Corollary 5.21 *If $\Theta\text{-}(P)IND^-$ is finitely axiomatizable, there is a constant k such that $T + \Theta\text{-}(P)IND^R = [T, \Theta\text{-}(P)IND^R]_k$ for every $T \supseteq BTC^0$.* □

6 Propositional proof systems

A fundamental tool for analysis of strong theories of arithmetic, especially in the context of induction rules and parameter-free schemes, are *reflection principles* for other theories of arithmetic (Beklemishev [3, 4]). This idea does not quite work for bounded arithmetic, which is too weak to prove even the consistency of the base theory Q. Instead, theories of bounded arithmetic may be studied using reflection principles for *propositional proof systems* by means of translation of bounded formulas to families of propositional formulas. Apart from the switch from first-order theories to propositional logic, there will be clear analogies between the form of our results and the classical case of strong systems.

There are two main families of propositional translations of interest:

(i) A translation of bounded formulas to *quantified* propositional formulas, where number variables translate to sequences of propositional variables representing their bits, and bounded quantifiers translate to blocks of propositional quantifiers.

(ii) A translation of bounded formulas in a *relativized* language (i.e., with a new predicate $\alpha(x)$) to bounded-depth propositional formulas, where number variables are set to constants, atomic formulas involving α translate to propositional variables, and bounded quantifiers translate to large disjunctions and conjunctions.

Translation (i) goes back to Cook [14] who introduced it as a translation of the equational theory PV to EF; the extension to quantified propositional logic is due to Krajíček and Pudlák [33]. Under this translation, Buss's theories T^i_2 correspond to subsystems of the quantified propositional calculus G. See Krajíček [31] and Cook and Nguyen [16] for detailed treatments.

Translation (ii) was introduced by Paris and Wilkie [35] for $I\Delta_0(\alpha)$. Under this translation, relativized Buss's theories $T^i_2(\alpha)$ translate to quasipolynomial-size bounded-depth proofs. See [9, §3] for a thorough discussion of variants of the Paris–Wilkie translation[2].

The relationship between the two translations depends on the point of view. On the one hand, translation (ii) produces exponentially larger formulas than translation (i). On the other hand, if we identify Buss's theories with the two-sorted theories V^i using the $RSUV$-isomorphism, translation (ii) becomes essentially equivalent to a special case of translation (i) for sharply bounded formulas (this is how it appears in [16]).

In this paper, we are going to work with translation (i). For one thing, it is already well known that it leads to an exact correspondence of various subsystems of S_2 (with parameters) to reflection principles for subsystems of G, and the setup works smoothly enough so that it can be generalized to the theories we are interested in.

Perhaps more importantly, translation (ii) inherently needs relativized theories, and this is problematic in the context of parameter-free induction axioms. On the one hand, oracles are somewhat similar to parameters in that they provide black-box information shared by all parts of the induction axiom, and as such go against the idea of disallowing parameters; in some contexts, they may be used to sneak parameters back in. See Section 7.2 for more discussion. On the other hand, the Paris–Wilkie translation (ii) largely eliminates the distinction between induction axioms with and without parameters, as parameters (like all variables) are set to constants before the translation. This stands in contrast to translation (i), in which parameters explicitly manifest as tuples of propositional variables that appear both in premises and conclusions of translations of induction axioms, and thus their presence makes a difference.

In light of this discussion, for any formula $\varphi(\vec{x}) \in \Sigma^b_\infty$, let $\{[\![\varphi]\!]_n : n \in \omega\}$ denote a sequence of quantified propositional formulas obtained by a (i)-style translation of φ, where each first-order variable x_i translates to a vector of n propositional variables in $[\![\varphi]\!]_n$, representing an integer $< 2^n$. We do not want to get into the gory technical details of the translation; we can generally follow the definition of $\|\varphi\|^n_{q(n)}$ (for a suitably chosen bounding polynomial $q(n)$) from Krajíček [31, §9.2], or up to the $RSUV$ isomorphism, the definition of $\|\varphi(\vec{X})\|$ in [16, §VII.5]. In particular:

- bounded existential (universal) quantifiers translate to polynomial-size blocks of existential (universal, resp.) propositional quantifiers,
- sharply bounded existential (universal) quantifiers within $\hat{\Sigma}^b_0$ formulas translate to polynomial-size disjunctions (conjunctions, resp.), and
- propositional connectives translate to themselves.

[2]Their setup includes modular counting gates, but most of the results work also in the usual setup.

There is a bit of a problem in the definition of the translation for atomic formulas φ, which we would like to turn into Σ^q_0 (i.e., quantifier-free) formulas: the translation from [31] is not suitable as it translates atomic formulas to Σ^q_1 formulas (provably equivalent to Π^q_1 formulas in strong enough proof systems); the translation from [16] does translate atomic (and Σ^B_0) formulas to Σ^q_0 formulas—even of bounded depth—but it only works in a much less expressive language. It does not apply to our TC^0 language.

The solution is to construct, in a suitably canonical way depending on the exact definition of BTC^0, for each atomic formula φ a uniform sequence of TC^0 circuits that compute it, and expand them into (log-depth) propositional formulas $[\![\varphi]\!]_n$ by means of formulas computing majority. Something similar was done in [23] for a theory whose language includes NC^1 functions. Again, the details do not matter for us, as long as the translation is sufficiently well-behaved so that it can be operated by our theories and proof systems. We stress that the weakest proof system in which we will reason with the translations is extended Frege.

In this way, the translations of $\hat\Sigma^b_i$ formulas are Σ^q_i, and translations of $\hat\Pi^b_i$ formulas are Π^q_i, for any $i \ge 0$.

We recall the following characterization [16, X.2.23–24] (cf. [17]):

Theorem 6.1

- *(i) If $i \ge j > 0$, the $\forall\hat\Sigma^b_j$ consequences of S^i_2 are axiomatized by $BTC^0 + \mathrm{RFN}_j(G^*_i)$. If additionally $i > j$, they are also axiomatized by $BTC^0 + \mathrm{RFN}_j(G_{i-1})$.*
- *(ii) If $i > 0$, $S^i_2 = BTC^0 + \mathrm{RFN}_{i+1}(G^*_i)$.*
- *(iii) If $i \ge 0$, $T^i_2 = BTC^0 + \mathrm{RFN}_{i+1}(G^*_{i+1})$.* □

The main result of this section will be a characterization of parameter-free induction axioms and induction rules analogous to Theorem 6.1. It will involve the following proof systems:

Definition 6.2 Let $i \ge 0$. For any $\xi(x) \in \hat\Sigma^b_i$, we define the proof system $G_i + \xi$ as G_i with additional initial sequents of the form $\Longrightarrow [\![\xi]\!]_n(\vec{A})$, where $n \in \mathbb{N}$, and $A_0, \dots, A_{n-1}$ are quantifier-free formulas; $G^*_i + \xi$ is its tree-like version.

Proposition 6.3 *Let $i \ge 0$, $\xi \in \hat\Sigma^b_i$, and $\varphi \in \Sigma^b_\infty$.*

- *(i) If $i > 0$ and $S^i_2 + \forall x\, \xi(x) \vdash \forall x\, \varphi(x)$, then BTC^0 proves that the formulas $[\![\varphi]\!]_n$ have TC^0-constructible polynomial-size $(G^*_i + \xi)$-proofs.*
- *(ii) If $i > 0$ or $\varphi \in \hat\Sigma^b_1$, and $T^i_2 + \forall x\, \xi(x) \vdash \forall x\, \varphi(x)$, then BTC^0 proves that the formulas $[\![\varphi]\!]_n$ have TC^0-constructible polynomial-size $(G_i + \xi)$-proofs.*

Proof: For $i > 0$, the standard proofs of these results without ξ as in [16, VII.5.2, X.1.21] proceed as follows. We formulate S^i_2 (T^i_2) in a sequent calculus with bounded quantifier introduction rules, and an appropriate induction rule. By the free-cut-elimination theorem, each bounded consequence of the theory has a proof that only contains bounded formulas such that all cut-formulas are $\hat{\Sigma}^b_i$. Then we translate the proof to propositional logic line by line, supplying short subderivations for each step. This argument works in our situation just the same: if we enhance the first-order calculus with substitution instances of $\xi \in \hat{\Sigma}^b_i$ as additional axioms, the free-cut-elimination theorem again makes all cuts $\hat{\Sigma}^b_i$, and then the same translation as before produces a valid $G^{(*)}_i$ proof except for instances of ξ, which translate to the additional axioms of $G^{(*)}_i + \xi$. The case $i = 0$ needs a different argument (either direct as in [16, X.1.23], or by simulation of G^*_1 [16, VII.4.16]), but again it works in the presence of additional quantifier-free axioms. □

Lemma 6.4 *Let $i \geq 0$, and $\xi \in \hat{\Sigma}^b_i$.*

(i) $T^i_2 + \forall x\, \xi(x)$ proves $\mathrm{RFN}_{\max\{i,1\}}(G_i + \xi)$.

(ii) If $i > 0$, $S^i_2 + \forall x\, \xi(x)$ proves $\mathrm{RFN}_{i+1}(G^*_i + \xi)$.

(iii) If $i = 0$, $[\forall x\, \xi(x), \Sigma^b_0\text{-}IND^R]$ proves $\mathrm{RFN}_0(G_0 + \xi)$.

Proof: (i): The implication $\forall x\, \xi(x) \to \mathrm{RFN}_i(G_i + \xi)$ is $\forall\exists\hat{\Sigma}^b_{i+1}$, hence it is enough to prove it in S^{i+1}_2, which is straightforward for $i > 0$: a $(G_i + \xi)$-proof of a Σ^q_i formula contains only Σ^q_i formulas, hence we may show by $\hat{\Pi}^b_{i+1}$-*LIND* on the length of the proof that every sequent in the proof is valid. For $i = 0$, we may e.g. show that the given assignment can be extended to satisfy all extension axioms in the proof using $\hat{\Sigma}^b_1$-*LIND*, and then show that all lines of the proof are true under this assignment by $\hat{\Delta}^b_1$-*LIND*. This shows that the target Σ^q_1 formula has a true witness, and therefore is itself true.

(ii): We may get rid of each axiom $\Longrightarrow [\![\xi]\!]_n(A_0, \dots, A_{n-1})$ in a $(G^*_i + \xi)$-proof by adding the Σ^q_{i+1} sentence $\exists x_0, \dots, x_{n-1}\, \neg[\![\xi]\!]_n(x_0, \dots, x_{n-1})$ to the succedent of every sequent in the proof. It follows using Theorem 6.1 that the original end-sequent or one of the new formulas is true under any given assignment, however, the latter contradicts $\forall x\, \xi(x)$.

(iii): It suffices to prove the consistency of $G_0 + \xi$, i.e., $EF + \xi$. By introducing extension variables for all subformulas used in the proof and other standard manipulations, BTC^0 knows that if there is an $(EF + \xi)$-proof of $\bot$, there is one where all formulas have bounded size (in particular, we can evaluate them on any given assignment in TC^0), and the only variables that occur in the proof are extension variables. Let $\pi(z)$ be a Σ^b_0 formula stating that z is a proof of this form. Let

$$q_{m-1} \leftrightarrow A_{m-1},\ q_{m-2} \leftrightarrow A_{m-2}(q_{m-1}),\ \dots,\ q_0 \leftrightarrow A_0(q_1, \dots, q_{m-1})$$

be the list of all extension axioms used in z. Writing u_i for the ith bit of u, let $\theta(u, z)$ be the formula

$$\pi(z) \to u < 2^m \wedge \forall i < m \left[\forall j < m \left(j > i \to u_j = A_j(u_{j+1}, \dots, u_{m-1})\right) \wedge u_i = 1 \right. \\ \left. \to A_i(u_{i+1}, \dots, u_{m-1}) = 1\right].$$

Notice that assuming $\pi(z)$, we can extract m (which is a length) and A_i from z by a TC^0 function, hence we can write $\theta(u, z)$ as a Σ^b_0 formula. Clearly, BTC^0 proves $\theta(0, z)$, and

$\pi(z) \to \neg\theta(2^m, z)$, that is, $\forall u\, \theta(u, z) \to \neg\pi(z)$, which in view of the preceding discussion means that

$$\vdash \forall u, z\, \theta(u, z) \to \mathrm{RFN}_0(G_0 + \xi).$$

It thus suffices to verify

$$\forall x\, \xi(x) \vdash \theta(u, z) \to \theta(u + 1, z).$$

Assume for contradiction that $\theta(u, z) \wedge \neg\theta(v, z)$, where $v = u + 1$. We must have $\pi(z)$ and $u < 2^m$. Let $i_0 \le m$ be the least index of a 0-bit of u, so that $u_j = v_j$ for $j > i_0$; $u_{i_0} = 0$, $v_{i_0} = 1$; and $u_j = 1$, $v_j = 0$ for $j < i_0$. If $v = 2^m$, we can show $A_i(x_{i+1}, \dots, x_{m-1}) = 1$ by reverse induction on $i < m$ (i.e., Σ^b_0-*LIND*, available in BTC^0). If $v < 2^m$, let $i < m$ be a witness that $\neg\theta(v, z)$, i.e.,

$$\forall j < m\, \big(j > i \to v_j = A_j(v_{j+1}, \dots, v_{m-1})\big) \wedge v_i = 1 \wedge A_i(v_{i+1}, \dots, v_{m-1}) = 0.$$

Since $v_i = 1$, this makes $i \ge i_0$. On the other hand, we cannot have $i > i_0$, as then the same would hold for u in place of v, contradicting $\theta(u, z)$. Thus, $i = i_0$. This implies

$$\forall j < m\, \big(j > i_0 \to u_j = A_j(u_{j+1}, \dots, u_{m-1})\big) \wedge u_{i_0} = 0 = A_{i_0}(u_{i_0+1}, \dots, u_{m-1}).$$

Using $\theta(u, z)$ and $u_j = 1$ for $j < i_0$, we then prove $A_j(x_{j+1}, \dots, u_{m-1}) = 1$ for $j < i_0$ by reverse induction on j (Σ^b_0-*LIND* again), hence in either case,

$$u_j = A_j(u_{j+1}, \dots, u_{m-1})$$

for all $j < m$. In other words, the bits of u taken as an assignment to the q_j variables satisfy all the extension axioms. Using Σ^b_0-*LIND* once more, we show that the assignment in fact satisfies *all* formulas in the proof: the induction steps for Frege rules follow from the fact that the rules are sound, and the $[\![\xi]\!]$ axioms are true because we assume $\forall x\, \xi(x)$. However, the last formula of the proof, $\bot$, is false, which is a contradiction. $\Box$

Theorem 6.5 *Let $i \ge 0$.*

(i) $\hat\Sigma^b_i$-*IND*^{R} *is equivalent to the rule*

$$\frac{\xi(x)}{\mathrm{RFN}_i(G_i + \xi)}, \qquad \xi \in \hat\Sigma^b_i.$$

(ii) $\hat\Sigma^b_i$-*IND*$^-$ *is equivalent to the scheme*

$$\forall x\, \xi(x) \to \mathrm{RFN}_i(G_i + \xi), \qquad \xi \in \hat\Sigma^b_i.$$

(iii) For $i > 0$, $\hat\Pi^b_i$-*IND*^{R} *is equivalent to the rule*

$$\frac{\xi(x)}{\mathrm{RFN}_{i-1}(G_i + \xi)}, \qquad \xi \in \hat\Sigma^b_i.$$

(iv) For $i > 0$, $\hat{\Pi}^b_i$-IND$^-$ is equivalent to the scheme

$$\forall x\, \xi(x) \to \mathrm{RFN}_{i-1}(G_i + \xi), \qquad \xi \in \hat{\Sigma}^b_i.$$

If $i > 0$, analogous equivalences hold with PIND in place of IND, and G^*_i in place of G_i.

Proof: (ii) and (iv) follow from (i) and (iii) and the deduction theorem.

(i): On the one hand, $[\forall x\, \xi(x), \hat{\Sigma}^b_i\text{-}IND^R] \vdash \mathrm{RFN}_i(G_i+\xi)$ by Lemma 6.4 and Theorem 5.10. On the other hand, let $\forall x\, \xi(x)$ be a $\forall\hat{\Sigma}^b_i$ sentence equivalent to $\varphi(0) \wedge \forall x\, (\varphi(x) \to \varphi(x+1))$, where $\varphi \in \hat{\Sigma}^b_i$. Then $T^i_2 + \forall x\, \xi(x)$ proves $\forall x\, \varphi(x)$, hence by Proposition 6.3, the formulas $[\![\varphi]\!]_n$ have short $(G_i+\xi)$-proofs, provably in BTC^0. Consequently, $BTC^0 + \mathrm{RFN}_i(G_i+\xi)$ proves that $[\![\varphi]\!]_n$ are tautologies for every length n, which implies $\forall x\, \varphi(x)$ by reasoning in BTC^0. (Note for $i = 0$ or the $\hat{\Pi}^b_1$ cases that even the $\hat{\Pi}^b_1$-definition of validity of $[\![\varphi]\!]_n$ ensures $\forall x < 2^n\, \varphi(x)$ for $\varphi \in \Sigma^b_0$: BTC^0 can construct the evaluation of $[\![\varphi]\!]_n$ and its subformulas under a given assignment using a TC^0 function, even though it may not prove that propositional formulas can be evaluated in general.)

(iii) is similar to (i), and the arguments for *PIND* and G^*_i are analogous. □

Corollary 6.6 *If $i \geq 0$, and T is a finitely $\forall\hat{\Sigma}^b_i$-axiomatized extension of BTC^0, then the theories $T + \hat{\Sigma}^b_i$-(P)IND^R and $T + \hat{\Pi}^b_i$-(P)IND^R are finitely axiomatizable.*

Specifically, if $T = BTC^0 + \forall x\, \xi(x)$ with $\xi \in \hat{\Sigma}^b_i$, then

$$T + \hat{\Sigma}^b_i\text{-}IND^R = BTC^0 + \mathrm{RFN}_i(G_i + \xi),$$

and for $i > 0$,

$$\begin{aligned} T + \hat{\Sigma}^b_i\text{-}PIND^R &= BTC^0 + \mathrm{RFN}_i(G^*_i + \xi), \\ T + \hat{\Pi}^b_i\text{-}IND^R &= T + \mathrm{RFN}_{i-1}(G_i + \xi), \\ T + \hat{\Pi}^b_i\text{-}PIND^R &= T + \mathrm{RFN}_{i-1}(G^*_i + \xi). \end{aligned}$$

Proof: The inclusions $\supseteq$ are special cases of Theorem 6.5. On the other hand, $T + \hat{\Sigma}^b_i$-*IND*^{R} is $\forall\hat{\Sigma}^b_i$-axiomatized, and if $T + \hat{\Sigma}^b_i\text{-}IND^R \subseteq T^i_2 + \forall x\, \xi(x)$ proves a $\hat{\Sigma}^b_i$ formula $\varphi(x)$, then $BTC^0 + \mathrm{RFN}_i(G_i + \xi)$ proves $\varphi(x)$ by the argument in the proof of Theorem 6.5. The other cases are similar, except that the arguments work just for $\varphi \in \hat{\Sigma}^b_{i-1}$ if we have only RFN_{i-1}. This is fine as $T + \hat{\Pi}^b_i$-(P)*IND*^{R} is $\forall\hat{\Sigma}^b_{i-1}$-axiomatized over T. □

Using this characterization, we can extend Theorem 5.10 to the case $i = 0$:

Corollary 6.7 *If $T \subseteq \forall\Sigma^b_0$, $T + \Sigma^b_0\text{-}IND^R = [T, \Sigma^b_0\text{-}IND^R]$.*

Proof: W.l.o.g., T is finitely axiomatizable, hence we may write $T = BTC^0 + \forall x\, \xi(x)$ with $\xi \in \Sigma^b_0$. Then $T + \Sigma^b_0\text{-}IND^R = BTC^0 + \mathrm{RFN}_0(G_0 + \xi) \subseteq [T, \Sigma^b_0\text{-}IND^R]$ by Corollary 6.6 and Lemma 6.4. □

A direct proof of Corollary 6.7 is also possible, but it is not particularly illuminating.

Remark 6.8 We could extend the definition of $G_i + \xi$ to $\xi \in \hat{\Sigma}^b_{i+1}$ as follows: write $\xi(x) = \exists y < 2^{|x|^c} \neg\theta(x, y)$ with $\theta \in \hat{\Sigma}^b_i$, and let $G_i + \xi$ denote G_i augmented by the rule

$$\frac{\Gamma \Longrightarrow \Delta, [\![\theta]\!]_{n,n^c}(A_0, \dots, A_{n-1}, x_0, \dots, x_{n^c-1})}{\Gamma \Longrightarrow \Delta},$$

where A_j are quantifier-free, and x_j are not free in Γ, Δ, or $A_{j'}$; likewise for $G^*_i + \xi$. (This is easily seen to be p-equivalent to the original definition if $\xi \in \hat{\Sigma}^b_i$.) Proposition 6.3 continues to hold in this setting, and the proof of Lemma 6.4 gives $S^{i+1}_2 + \forall x\, \xi(x) \vdash \mathrm{RFN}_i(G_i + \xi)$. Since this extension does not seem to yield new insights about parameter-free induction schemes or rules, we skip the details.

7 Separations

We have seen in the previous sections many results relating subsystems of bounded arithmetic with and without parameters, but in order for these results to be useful, it would be nice to know that the systems do not collapse: what if the parameter-free induction schemes are actually equivalent to the usual schemes with parameters, so that e.g. $T^i_2 = \hat{\Sigma}^b_i$-IND^-? This would make the investigation of IND^- rather pointless. Likewise, since we spent so much effort on $\hat{\Pi}^b_i$ schemes and rules, we would like to know that they are genuinely distinct from the corresponding $\hat{\Sigma}^b_i$ rules.

In general, we are interested if there are any reductions between our schemes and rules that do not follow from Theorem 3.5 (as depicted in Figure 3.1), and furthermore if there are any inclusions between the theories generated by our rules over the base theory that do not follow from Theorem 3.5 and Corollary 5.12 (as depicted in Figure 5.1).

Checking all the cases naively would be a gargantuan task: we have 10 rules at each level of the hierarchy, and we need to consider reductions spanning three levels: e.g., S^i_2 is supposed not to be included in $BTC^0 + \hat{\Pi}^b_{i+1}$-$IND^-$, which is two levels higher up, being $\forall\hat{\Sigma}^b_i$-conservative under S^{i+2}_2. However, we do not actually have to consider all possible pairs, as there is a lot of redundancy: for example, we do not need to check separately that $BTC^0 + \hat{\Sigma}^b_i$-$IND^- \nvdash T^i_2$, because $T^i_2 \supseteq S^i_2$, $\hat{\Sigma}^b_i$-$IND^- \subseteq \hat{\Pi}^b_{i+1}$-$IND^-$, and we want to make sure that $BTC^0 + \hat{\Pi}^b_{i+1}$-$IND^- \nvdash S^i_2$ anyway. Let us put our job into a more formal setting:

Definition 7.1 A *basis of non-inequalities* of a poset $\langle P, \leq \rangle$ is a set $B \subseteq P^2$ such that

(i) $a \nleq b$ for any $\langle a, b \rangle \in B$, and

(ii) for each $a, b \in P$ such that $a \nleq b$, there is $\langle a', b' \rangle \in B$ such that $a' \leq a$ and $b \leq b'$.

A *critical pair* of P is $\langle a, b \rangle \in P$ such that $a \nleq b$, but $a' \leq b$ for all $a' < a$, and $a \leq b'$ for all $b' > b$. Observe that any basis of non-inequalities of P has to include all critical pairs.

Let $\langle P_R, \leq_R \rangle$ denote the poset with formal elements representing BTC^0 and the axioms and rules $\hat{\Sigma}^b_i$-IND, $\hat{\Sigma}^b_i$-IND^-, $\hat{\Sigma}^b_i$-IND^R, $\hat{\Pi}^b_{i+1}$-IND^-, $\hat{\Pi}^b_{i+1}$-IND^R, $\hat{\Sigma}^b_{i+1}$-$PIND$, $\hat{\Sigma}^b_{i+1}$-$PIND^-$,

$\hat{\Sigma}^b_{i+1}$-$PIND^R$, $\hat{\Pi}^b_{i+1}$-$PIND^-$, and $\hat{\Pi}^b_{i+1}$-$PIND^R$ for $i \geq 0$, and with $\leq_R$ being the transitive reflexive closure of the relation given by Theorem 3.5. (BTC^0 is a least element of P_R.)

Let $\langle P_T, \leq_T \rangle$ be the quotient of $\langle P_R, \leq_R \rangle$ identifying $\hat{\Sigma}^b_{i+1}$-$PIND^R$ with $\hat{\Sigma}^b_i$-IND, and $\hat{\Pi}^b_{i+1}$-$PIND^R$ with $\hat{\Sigma}^b_i$-IND^R, for each $i \geq 0$.

Beware that neither P_R nor P_T is a lattice.

Lemma 7.2 *Let $\langle P, \leq \rangle$ be a poset in which all strictly increasing infinite sequences are upwards cofinal, and all strictly decreasing infinite sequences are downwards cofinal*[3]*. Then the set of critical pairs is a basis of non-inequalities of P.*

Proof: The assumptions may be restated such that for each $u \in P$, $<$ is well-founded on $\{x \in P : x \nleq u\}$, and converse well-founded on $\{x \in P : u \nleq x\}$. Thus, given $a \nleq b$, we can find a minimal $a' \leq a$ such that $a' \nleq b$, and then a maximal $b' \geq b$ such that $a' \nleq b'$. Then $\langle a', b' \rangle$ is a critical pair. □

The critical pairs of P_R and P_T can be determined by a somewhat tedious, but straightforward computation, chasing the diagrams in Figures 3.1 and 5.1. We see that P_R and P_T have common critical pairs

$$\langle \hat{\Sigma}^b_i\text{-}PIND, \hat{\Pi}^b_{i+1}\text{-}IND^- \rangle, \qquad \langle \hat{\Sigma}^b_0\text{-}IND, \hat{\Pi}^b_1\text{-}IND^- \rangle,$$
$$\langle \hat{\Pi}^b_i\text{-}PIND^-, \hat{\Pi}^b_{i+1}\text{-}IND^R \rangle, \qquad \langle \hat{\Sigma}^b_0\text{-}IND^-, \hat{\Pi}^b_1\text{-}IND^R \rangle,$$
$$\langle \hat{\Pi}^b_i\text{-}IND^R, \hat{\Sigma}^b_i\text{-}PIND \rangle, \qquad \langle \hat{\Sigma}^b_0\text{-}IND^R, BTC^0 \rangle$$

for $i \geq 1$. Moreover, P_R has critical pairs

$$\langle \hat{\Pi}^b_{i+1}\text{-}PIND^R, \hat{\Sigma}^b_i\text{-}IND \rangle$$

for $i \geq 0$, but we can disregard these: $\hat{\Pi}^b_{i+1}$-$PIND^R \leq T^i_2$ implies $T^i_2 \vdash \hat{\Pi}^b_{i+1}$-$PIND^-$ using the deduction theorem, hence also $BTC^0 + \hat{\Pi}^b_{i+2}$-$IND^R \vdash \hat{\Pi}^b_{i+1}$-$PIND^-$, which is an instance of another critical pair. Thus, we obtain:

Proposition 7.3 *If there is a reduction between the rules from Definition 3.1 which does not follow from Theorem 3.5, or an additional inclusion between the first-order theories they generate over BTC^0 not warranted by Corollary 5.12, it implies one of the following:*

(19) $$S^i_2 \vdash BTC^0 + \hat{\Pi}^b_i\text{-}IND^R \text{ for some } i \geq 0,$$

(20) $$\hat{\Pi}^b_{i+1}\text{-}IND^- \vdash S^i_2 \text{ for some } i > 0,$$

(20′) $$\hat{\Pi}^b_1\text{-}IND^- \vdash T^0_2,$$

(21) $$BTC^0 + \hat{\Pi}^b_{i+1}\text{-}IND^R \vdash \hat{\Pi}^b_i\text{-}PIND^- \text{ for some } i > 0, \text{ or}$$

(21′) $$BTC^0 + \hat{\Pi}^b_1\text{-}IND^R \vdash \Sigma^b_0\text{-}IND^-.$$

(Recall that in our setup, $S^0_2 = BTC^0$.) □

[3]In fact, weaker assumptions suffice: it is enough if $\mathbb{Q}$, $\omega \sqcup 1$, and $\omega^* \sqcup 1$ do not embed in P, where $\sqcup$ denotes disjoint union of posets.

The remaining goal is to convince ourselves that (19)–(21$'$) are likely false, or at least suspect. We are not very picky, and do not attempt to devise sophisticated separation arguments optimized for the particular theories; rather, we are content with any evidence that we did not overlook something in Theorem 3.5. We will present run-of-the-mill separations of two kinds, as commonly done for systems of bounded arithmetic: separations conditional on plausible complexity-theoretic assumptions, and unconditional separations of relativized versions of our theories.

7.1 Unrelativized separations

The state of our knowledge does not allow us to disprove even $BTC^0 = S_2$ unconditionally—this would require a major breakthrough. We thus cannot disprove (19)–(21$'$) either. What we can do instead is to show that they imply other statements (from computational and proof complexity) that are more commonly recognized as implausible.

Theorem 7.4 *If $S^i_2 \vdash BTC^0 + \hat\Pi^b_i$-$IND^R$, then T^i_2 is $\forall\hat\Sigma^b_{\max\{i-1,0\}}$-conservative over S^i_2 (and thus over T^{i-1}_2 for $i > 0$). Consequently:*

(i) If $i = 0$, TC^0-Frege p-simulates EF.

(ii) If $i > 0$, G^*_i and G_{i-1} p-simulate G_i with respect to Σ^q_{i-1} sequents.

(iii) If $i > 1$, the game induction principle GI_i (Skelley and Thapen [36]) is reducible to GI_{i-1}.

Proof: The conservativity of T^i_2 over S^i_2 is a consequence of the characterization of $BTC^0 + \hat\Pi^b_i$-IND^R from Corollary 5.14. Then (i) and (ii) follow by a standard argument: T^i_2, hence S^i_2 and T^{i-1}_2 by assumption, proves $\mathrm{RFN}_{i-1}(G_i)$. Thus, BTC^0 proves that the tautologies $[\![\mathrm{RFN}_{i-1}(G_i)]\!]_n$ have TC^0-constructible proofs in G^*_i and G_{i-1}, which in turn implies that these two proof systems p-simulate G_i-proofs of Σ^q_{i-1} sequents. Similarly, (iii) follows from the fact that GI_i is complete for the class of NP-search problems provably total in T^i_2. □

Recall that $\mathrm{FP}^{\Sigma^P_i[O(g(n)),\mathrm{wit}]}$ denotes the class of total search problems computable by a polynomial function that makes $O(g(n))$ queries to a witnessing Σ^P_i oracle, meaning that for any positive answer, the oracle also has to produce a witness to the outermost existential quantifier. For any $i > 0$, the $\hat\Sigma^b_{i+1}$-definable search problems provably total in S^i_2 comprise exactly $\mathrm{FP}^{\Sigma^P_i[O(\log n),\mathrm{wit}]}$, and the Σ^b_{i+1}-definable search problems provably total in T^{i-1}_2 comprise exactly $\mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$ (see e.g. [16, Thm. VIII.7.17]; the original results are due to Krajíček, Pudlák, and Takeuti [34] and Krajíček [30]).

Theorem 7.5

(i) If $\hat\Pi^b_1$-$IND^- \vdash T^0_2$, then $\mathrm{P} = \mathrm{TC}^0$.

(ii) If $\hat\Pi^b_{i+1}$-$IND^- \vdash S^i_2$ for some $i > 0$, then $\mathrm{FP}^{\Sigma^P_i[O(\log n),\mathrm{wit}]} = \mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$, and $\mathrm{PH} = \mathcal{B}(\Sigma^P_{i+1})$.

Proof: First, observe that $\hat{\Pi}^b_{i+1}$-IND^- follows from the set of all true $\forall\hat{\Sigma}^b_i$ sentences: it is axiomatized by sentences of the form $\varphi \to \psi$, where $\varphi \in \forall\Sigma^b_\infty$, and $\psi \in \forall\hat{\Sigma}^b_i$. If φ is false, $\neg\varphi$ (and a fortiori $\varphi \to \psi$) is provable in BTC^0, being a true Σ^0_1 sentence. Otherwise, ψ is true, hence included in $\mathrm{Th}_{\forall\hat{\Sigma}^b_i}(\mathbb{N})$.

(i): Every poly-time function f has a provably total $\hat{\Sigma}^b_1$-definition in T^0_2, hence by assumption, in $\mathrm{Th}_{\forall\Sigma^b_0}(\mathbb{N})$, i.e., in the set of true universal sentences of L_{TC^0}. By Herbrand's theorem (and closure under definitions by cases), f is definable by an L_{TC^0}-term, i.e., it is a TC^0-function. In particular, every poly-time predicate is computable in TC^0.

(ii): Every $\mathrm{FP}^{\Sigma^P_i[O(\log n),\mathrm{wit}]}$ search problem has a $\hat{\Sigma}^b_{i+1}$-definition provably total in S^i_2, hence by assumption, in $\mathrm{Th}_{\forall\hat{\Sigma}^b_i}(\mathbb{N})$. We claim that, just like for T^{i-1}_2, the provably total $\hat{\Sigma}^b_{i+1}$-definable search problems of $\mathrm{Th}_{\forall\hat{\Sigma}^b_i}(\mathbb{N})$ are in $\mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$: if

$$\forall u\, \psi(u) \vdash \forall x\, \exists y\, \varphi(x,y),$$

where $\psi \in \hat{\Sigma}^b_i$, $\varphi \in \hat{\Sigma}^b_{i+1}$, and $\mathbb{N} \vDash \forall u\, \psi(u)$, we have

$$T^{i-1}_2 \vdash \forall x\, \exists y\, \bigl(\neg\psi(y) \lor \varphi(x,y)\bigr)$$

(the T^{i-1}_2 is not really doing anything for us here). We may bound the y using Parikh's theorem, and then by the above-mentioned characterization of $\forall\hat{\Sigma}^b_{i+1}$ consequences of T^{i-1}_2, we obtain

$$T^{i-1}_2 \vdash \forall x\, \bigl(\neg\psi(f(x)) \lor \varphi(x, f(x))\bigr)$$

for some search problem $f \in \mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$, $\hat{\Sigma}^b_{i+1}$-definable in T^{i-1}_2; but the first disjunct cannot happen in the real world:

$$\mathbb{N} \vDash \forall x\, \varphi(x, f(x)).$$

Thus, $\mathrm{FP}^{\Sigma^P_i[O(\log n),\mathrm{wit}]} = \mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$. This implies $\mathrm{P}^{\Sigma^P_i[O(\log n)]} = \mathrm{P}^{\Sigma^P_i[O(1)]} = \mathcal{B}(\Sigma^P_i)$, as predicates (i.e., $\{0,1\}$-valued functions) in $\mathrm{FP}^{\Sigma^P_i[O(1),\mathrm{wit}]}$ are in $\mathrm{P}^{\Sigma^P_i[O(1)]}$ (cf. [31, 6.3.4–5]). This in turn implies the collapse of PH to $\mathcal{B}(\Sigma^P_{i+1})$ by Chang and Kadin [10]. □

Remark 7.6 The second point of Theorem 7.5 is a variant of the well-known result that $T^{i-1}_2 = S^i_2$ implies the collapse of PH, originally proved in [34], and subsequently improved in [8, 38, 15, 22]. The current state of the art is that $T^{i-1}_2 = S^i_2$ implies $T^{i-1}_2 \vdash \mathrm{PH} = \mathcal{B}(\Sigma^P_i)$ [22, Cor. 4.7], which is a one whole level deeper collapse than in Theorem 7.5.

While we did not attempt to check the details, it is not implausible that these improvements also work in the presence of additional true $\forall\hat{\Sigma}^b_i$ axioms; if correct, this would strengthen the conclusion of Theorem 7.5 (ii) to $\mathrm{PH} = \mathcal{B}(\Sigma^P_i)$.

Question 7.7 *Can we disprove* (21) *or* (21′) *under a credible hypothesis?*

7.2 Relativized separations

Rather than relying on unproven hypotheses, we may want to look at unconditional separations of relativized theories. All theories we work with may be relativized in the standard way: we include a new predicate symbol $\alpha(x)$ in the language, and extend all schemes to allow the use of α along with other atomic formulas, but do not include any axioms to fix its particular values.

Relativization is commonly employed in bounded arithmetic to obtain separation results, exploiting the fact that we can unconditionally separate various complexity classes in the relativized setting. The usefulness of this technique of course hinges on our belief that for the classes in question (e.g., levels of the polynomial hierarchy), noninclusions between their relativized versions truly reflect properties of the original unrelativized classes. (Relativized bounded arithmetic is also useful in connection with bounded-depth propositional proof systems, as the Paris–Wilkie translation only makes sense for relativized theories.)

Relativization of parameter-free schemes may seem somewhat more dubious than in the case of usual theories of bounded arithmetic, as it goes against the spirit of parameter removal: similar to parameters, the oracle provides access to additional black-box information that is shared by antecedents and succedents of induction axioms. This worry is for the most part unsubstantiated, as there is a crucial difference in that the oracle is arbitrary but *fixed*, whereas parameters of a scheme are universally quantified, and as such represent all numbers in the domain even in the context of a single statement. Nevertheless, we will see that the idea that an oracle can simulate parameters works out in certain situations, and some of our relativized separation results rely on it.

Perhaps the best way to argue that relativized separations are useful is that they show unprovability of inclusions or reductions between rules by means of the techniques we employed elsewhere in this paper, as all positive results we proved earlier *do* relativize. This is easy to observe[4] for the results in Sections 3–5. For Section 6, we may relativize the proof systems by expanding the propositional language with a new unbounded fan-in connective representing α, and then everything works out.

Theorem 7.8 *$\hat{\Pi}^b_1(\alpha)$-IND$^-$ $\nvdash T^0_2(\alpha)$, and $\hat{\Pi}^b_{i+1}(\alpha)$-IND$^-$ $\nvdash S^i_2(\alpha)$ for $i > 0$.*

Proof: If we fix an oracle $A \subseteq \mathbb{N}$, then $\hat{\Pi}^b_{i+1}(\alpha)$-*IND*$^-$ follows from the set of all $\forall\hat{\Sigma}^b_i(\alpha)$ sentences true in $\langle \mathbb{N}, A\rangle$. The same argument as in the proof of Theorem 7.5 then shows that if $\hat{\Pi}^b_{i+1}(\alpha)$-*IND*$^- \vdash S^i_2(\alpha)$, then the relativized polynomial hierarchy PH^A collapses. However, it is well known that we can find A such that this does not happen [37, 21].

Similarly, $\hat{\Pi}^b_1(\alpha)$-*IND*$^- \vdash T^0_2(\alpha)$ implies $\mathrm{P}^A = (\mathrm{TC}^0)^A$ for every $A \subseteq \mathbb{N}$. The proper notion of relativized TC^0 corresponding to $\forall\hat{\Sigma}^b_1(\alpha)$-witnessing of universal extensions of BTC^0 is explained in Aehlig, Cook, and Nguyen [2], where they also exhibit an oracle separating NL^A (hence $(\mathrm{TC}^0)^A$) from P^A. □

[4]The one possible exception is that we used a couple of times the fact that every bounded sentence is provable or refutable in the base theory. This is not literally true in the relativized setting, but it may be replaced by the weaker property that every bounded sentence is equivalent to a Boolean combination of sentences of the form $\alpha(k)$ for standard constants k.

Theorem 7.9 *$BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$IND^R \nvdash \hat{\Pi}^b_i(\alpha)$-$PIND^-$ for $i > 0$, and $BTC^0(\alpha) + \hat{\Pi}^b_1(\alpha)$-$IND^R \nvdash \Sigma^b_0(\alpha)$-$IND^-$.*

Proof: Assume for contradiction that $BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$IND^R \vdash \hat{\Pi}^b_i(\alpha)$-$PIND^-$, where $i > 0$. We will argue that parameters of the *PIND* scheme can be encoded into the oracle.

Given a term $t(x)$, let us fix a proof π of *PIND* for the parameter-free $\hat{\Pi}^b_i(\alpha)$ formula

$$\forall x_1 \le t(x)\, \exists x_1 \le t(x) \cdots Q x_i \le t(x)\, \alpha(\langle x, x_1, \dots, x_i \rangle) \tag{22}$$

in $BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$IND^R$, and let $\varphi(x, y)$ be a $\hat{\Pi}^b_i(\alpha)$ formula of the form

$$\forall x_1 \le t(x)\, \exists x_1 \le t(x) \cdots Q x_i \le t(x)\, \theta(x, y, x_1, \dots, x_i), \tag{23}$$

where $\theta \in \hat{\Sigma}^b_0(\alpha)$. We may assume without loss of generality that y does not occur in π. If we substitute $\theta\big((z)_0, y, (z)_1, \dots, (z)_i\big)$ for $\alpha(z)$ everywhere in the proof, the result is still a valid $BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$IND^R$ proof as IND^R allows parameters, hence the theory proves *PIND* for $\varphi(x, y)$.

This is not yet a general instance of $\hat{\Pi}^b_i(\alpha)$-*PIND*, as all quantifiers in φ have to be bounded by a term in the induction variable. However, this restriction is immaterial: if $\varphi(x, y) \in \hat{\Pi}^b_i(\alpha)$ is arbitrary, *PIND* for φ follows from *PIND* for the formula $|x| < |y| \lor \varphi(\lfloor x/2^{|y|} \rfloor, y)$, which may be equivalently rewritten so that all quantifiers are bounded by a term in x alone.

Thus, $BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$IND^R \vdash S^i_2(\alpha)$, but this contradicts Theorem 7.8.

Likewise, $BTC^0(\alpha) + \hat{\Pi}^b_1(\alpha)$-$IND^R \vdash \Sigma^b_0(\alpha)$-$IND^-$ would imply $BTC^0(\alpha) + \hat{\Pi}^b_1(\alpha)$-$IND^R \vdash T^0_2(\alpha)$. □

We do not have an unconditional disproof of (19) in its full generality, but several partial results that come close:

Theorem 7.10 *Let $i \ge 0$.*

(i) If $i > 0$, $S^i_2(\alpha) \nvdash BTC^0(\alpha) + \hat{\Sigma}^b_i(\alpha)$-$IND^R = BTC^0(\alpha) + \hat{\Pi}^b_{i+1}(\alpha)$-$PIND^R$.

(ii) $S^2_2(\alpha) \nvdash BTC^0(\alpha) + \hat{\Pi}^b_2(\alpha)$-$IND^R$.

(iii) $S^i_2(\alpha) \nvdash \hat{\Pi}^b_i(\alpha)$-$IND^-$.

(iv) $\hat{\Pi}^b_i(\alpha)$-$IND^R \nleq S^i_2(\alpha)$.

Proof: (i): In view of Corollary 5.14, the claim is equivalent to the fact that $T^i_2(\alpha)$ is not $\forall\hat{\Sigma}^b_i(\alpha)$-conservative over $S^i_2(\alpha)$ due to Buss and Krajíček [32].

(ii): This amounts to the $\forall\hat{\Sigma}^b_1(\alpha)$-non-conservativity of $T^2_2(\alpha)$ over $S^2_2(\alpha)$, proved by Chiari and Krajíček [11] (see also [12]).

(iii): Assume that $S^i_2(\alpha) \vdash \hat{\Pi}^b_i(\alpha)$-$IND^-$; we will argue that $S^i_2(\alpha) \vdash \hat{\Pi}^b_i(\alpha)$-*IND*, contradicting $S^i_2(\alpha) \ne T^i_2(\alpha)$. As in the proof of Theorem 7.9, if $\varphi(x, y)$ is a formula of the form (23), we construct a proof of φ-*IND* in $S^i_2(\alpha)$ by taking a proof (not containing y) of *IND* for the formula (22), and substituting $\theta\big((z)_0, y, (z)_1, \dots, (z)_i\big)$ for $\alpha(z)$. If $\varphi(x, y)$ is an arbitrary $\hat{\Pi}^b_i$ formula, then φ-*IND* (with x being the induction variable, and y a parameter) follows from *IND* for the formula $x < y \lor \varphi(x - y, y)$, which is equivalent to a formula of the form (23).

(iv) follows from (iii) using the deduction theorem. □

Remark 7.11 By inspection of critical pairs of P_R and P_T, the net effect of Theorems 7.8, 7.9, and 7.10 is that in the relativized setting:

- all valid reductions between the rules from Definition 3.1 follow from Theorem 3.5;
- all valid inclusions between theories generated by these rules follow from Theorem 3.5 and Corollary 5.12, except possibly

  (24) $$BTC^0(\alpha) \vdash BTC^0(\alpha) + \hat{\Sigma}^b_0(\alpha)\text{-}IND^R,$$

  or

  (25) $$T \vdash BTC^0(\alpha) + \hat{\Pi}^b_i(\alpha)\text{-}IND^R$$

  for some $i \geq 1$, $i \neq 2$, and a theory T between $BTC^0(\alpha) + \hat{\Sigma}^b_{i-1}(\alpha)\text{-}IND^R$ and $S^i_2(\alpha)$ (apart from the two indicated, these are $\hat{\Sigma}^b_{i-1}(\alpha)\text{-}IND^-$, $T^{i-1}_2(\alpha)$, $\hat{\Pi}^b_i\text{-}PIND^-$, and $\hat{\Sigma}^b_i\text{-}PIND^-$).

Note that for any given i, (25) holds either for all the theories T, or for none of them; that is, the following are equivalent:

(i) $S^i_2(\alpha) \vdash BTC^0(\alpha) + \hat{\Pi}^b_i(\alpha)\text{-}IND^R$,

(ii) $BTC^0(\alpha) + \hat{\Sigma}^b_{i-1}(\alpha)\text{-}IND^R = BTC^0(\alpha) + \hat{\Pi}^b_i(\alpha)\text{-}IND^R$,

(iii) $T^i_2(\alpha)$ is $\forall\hat{\Sigma}^b_{i-1}(\alpha)$-conservative over $S^i_2(\alpha)$ (or equivalently, over $T^{i-1}_2(\alpha)$).

Likewise, (24) is equivalent to the $\forall\hat{\Sigma}^b_0(\alpha)$-conservativity of $T^0_2(\alpha)$ over $BTC^0(\alpha)$.

Even though it is commonly believed that $T^i_2(\alpha)$ is not $\forall\Sigma^b_0(\alpha)$-conservative over $S^i_2(\alpha)$ for any $i \geq 0$, it is a major open problem to improve the above-quoted results of [32, 11] even just by one level, thus (24) and (25) are open.

In this connection, we mention a possibly interesting consequence of Theorem 7.10 (iv):

Corollary 7.12 *For any $i \geq 1$, there is a $\forall\hat{\Sigma}^b_i(\alpha)$ sentence φ such that $T^i_2(\alpha) + \varphi$ is not $\forall\hat{\Sigma}^b_{i-1}(\alpha)$-conservative over $S^i_2(\alpha) + \varphi$.* □

8 Conclusion

We have undertaken a comprehensive investigation of parameter-free and inference-rule variants of the $\hat{\Sigma}^b_i$ and $\hat{\Pi}^b_i$ induction and polynomial induction axioms. We found which rules and axioms reduce to other rules, and which do not. We have seen conservation results among the systems; in particular, each of our theories can be characterized as the Γ-fragment of some S^i_2 for a suitable class of sentences Γ. We also found equivalent expressions for our axioms and rules in terms of reflection principles for axiomatic extensions of the quantified propositional calculi G_i, and we proved a few other results, in particular concerning nesting depth of rules.

In some respects, the properties of our systems resemble the situation of strong theories of arithmetic $I\Sigma^-_n$ and $I\Pi^-_n$: the $\hat{\Pi}^b_i$ schemes and rules are weaker than their $\hat{\Sigma}^b_i$ counterparts, there are conservation results connecting the systems to the usual theories S^i_2, the parameter-free schemes do not seem to be finitely axiomatizable, and our systems correspond to reflection principles and rules (albeit of different nature) of similar overall shape as for the strong systems.

On the other hand, there are also notable differences. Most importantly, the hierarchies fit together in different ways: $I\Pi^-_{n+1}$ is equiconsistent with (and $\mathcal{B}(\Sigma_{n+1})$-conservative over) $I\Sigma^-_n$ and $I\Sigma_n$, whereas in our case, $\hat{\Pi}^b_i$-$(P)IND^-$ is $\mathcal{M}(\exists\hat{\Pi}^b_i, \forall\hat{\Pi}^b_i)$-conservative under $\hat{\Sigma}^b_i$-$(P)IND^-$ and $\hat{\Sigma}^b_i$-$(P)IND$. On a related note, the systems $I\Pi^-_{n+1}$ and $I\Sigma_n$ on the same level of the hierarchy are incomparable, and their join $I\Pi^-_{n+1} + I\Sigma_n$ has strictly stronger consistency strength—it proves the consistency of $I\Sigma_n$ (cf. [4]); no such phenomenon is possible in our setup, as all the systems on each level of our hierarchy are included in the largest one among them, namely S^i_2.

There are some problems that we left open, specifically if T^i_2 is $\exists\forall\hat{\Pi}^b_i$-conservative over $\hat{\Pi}^b_i$-IND^-, and similarly for S^i_2 and $\hat{\Pi}^b_i$-$PIND^-$ (Question 5.17). It would be also desirable to prove unrelativized separation of $\hat{\Pi}^b_i$-$PIND^-$ from $BTC^0 + \hat{\Pi}^b_{i+1}IND^R$ (Question 7.7) under plausible assumptions.

We tried our best to conduct an in-depth examination of parameter-free and inference-rule versions of the *IND* and *PIND* schemes, that also applies, by the results of Section 4, to their common variants like *LIND* and minimization schemes. However, we left out other schemes of interest in bounded arithmetic: in particular, the choice (aka replacement or bounded collection) scheme *BB* (which was studied in [18]), and analogues of *LIND* with induction up to bounds given by more general classes of terms (including *LLIND*, etc.). Related to *BB*, we might be interested in variants of $(P)IND$ and other schemes for the non-strict Σ^b_i and Π^b_i formula classes: it is well known that with parameters, the strict and non-strict $(P)IND$ schemes are equivalent—both define the familiar theories S^i_2 and T^i_2. It is however likely that the situation will get more complicated without parameters. We also left out various combinations of our base systems such as $S^i_2 + \hat{\Pi}^b_i$-$IND^- + \hat{\Sigma}^b_i$-IND^R.

The reason we decided not to discuss any of these potentially interesting topics is sheer complexity: we have 10 systems per each level of the hierarchy as is, which already leads to a complex network of relations among them. If we added more schemes and rules to the mix, the number of combinations would multiply, rendering the global picture unmanageable. That is to say, there are certainly many aspects of these systems that are worth further investigation, but we deem them out of scope of this paper.

Acknowledgements

I would like to thank Andrés Cordón-Franco and Félix Lara-Martín for stimulating discussions of the topic, and for drawing my attention to [27].

References

[1] Zofia Adamowicz and Teresa Bigorajska, *Functions provably total in $I^-\Sigma_1$*, Fundamenta Mathematicae 132 (1989), pp. 189–194.

[2] Klaus Aehlig, Stephen Cook, and Phuong Nguyen, *Relativizing small complexity classes and their theories*, Computational Complexity 25 (2016), no. 1, pp. 177–215.

[3] Lev D. Beklemishev, *Induction rules, reflection principles, and provably recursive functions*, Annals of Pure and Applied Logic 85 (1997), no. 3, pp. 193–242.

[4] ———, *Parameter free induction and provably total computable functions*, Theoretical Computer Science 224 (1999), pp. 13–33.

[5] Teresa Bigorajska, *On Σ_1-definable functions provably total in $I\Pi_1^-$*, Mathematical Logic Quarterly 41 (1995), pp. 135–137.

[6] Stephen Austin Bloch, *Divide and conquer in parallel complexity and proof theory*, Ph.D. thesis, University of California, San Diego, 1992.

[7] Samuel R. Buss, *Bounded arithmetic*, Bibliopolis, Naples, 1986, revision of 1985 Princeton University Ph.D. thesis.

[8] ———, *Relating the bounded arithmetic and polynomial time hierarchies*, Annals of Pure and Applied Logic 75 (1995), no. 1–2, pp. 67–77.

[9] Samuel R. Buss, Leszek A. Kołodziejczyk, and Konrad Zdanowski, *Collapsing modular counting in bounded arithmetic and constant depth propositional proofs*, Transactions of the American Mathematical Society 367 (2015), no. 11, pp. 7517–7563.

[10] Richard Chang and Jim Kadin, *The Boolean hierarchy and the polynomial hierarchy: a closer connection*, SIAM Journal on Computing 25 (1996), no. 2, pp. 340–354.

[11] Mario Chiari and Jan Krajíček, *Witnessing functions in bounded arithmetic and search problems*, Journal of Symbolic Logic 63 (1998), no. 3, pp. 1095–1115.

[12] ———, *Lifting independence results in bounded arithmetic*, Archive for Mathematical Logic 38 (1999), no. 2, pp. 123–138.

[13] Peter Clote and Gaisi Takeuti, *First order bounded arithmetic and small boolean circuit complexity classes*, in: Feasible Mathematics II (P. Clote and J. B. Remmel, eds.), Progress in Computer Science and Applied Logic vol. 13, Birkhäuser, 1995, pp. 154–218.

[14] Stephen A. Cook, *Feasibly constructive proofs and the propositional calculus*, in: Proceedings of the 7th Annual ACM Symposium on Theory of Computing, 1975, pp. 83–97.

[15] Stephen A. Cook and Jan Krajíček, *Consequences of the provability of* $\mathbf{NP} \subseteq \mathbf{P/poly}$, Journal of Symbolic Logic 72 (2007), no. 4, pp. 1353–1371.

[16] Stephen A. Cook and Phuong Nguyen, *Logical foundations of proof complexity*, Perspectives in Logic, Cambridge University Press, New York, 2010.

[17] ———, *Corrections for [16]*, `http://www.cs.toronto.edu/~sacook/homepage/corrections.pdf`, 2013.

[18] Andrés Cordón-Franco, Alejandro Fernández-Margarit, and Francisco Félix Lara-Martín, *Existentially closed models and conservation results in bounded arithmetic*, Journal of Logic and Computation 19 (2009), no. 1, pp. 123–143.

[19] Andrés Cordón-Franco and Francisco Félix Lara-Martín, *Local induction and provably total computable functions*, Annals of Pure and Applied Logic 165 (2014), no. 9, pp. 1429–1444.

[20] Petr Hájek and Pavel Pudlák, *Metamathematics of first-order arithmetic*, Perspectives in Mathematical Logic, Springer, 1993, second edition 1998.

[21] Johan Håstad, *Almost optimal lower bounds for small depth circuits*, in: Randomness and Computation (S. Micali, ed.), Advances in Computing Research: A Research Annual vol. 5, JAI Press, 1989, pp. 143–170.

[22] Emil Jeřábek, *Approximate counting by hashing in bounded arithmetic*, Journal of Symbolic Logic 74 (2009), no. 3, pp. 829–860.

[23] ———, *On theories of bounded arithmetic for* NC^1, Annals of Pure and Applied Logic 162 (2011), no. 4, pp. 322–340.

[24] Emil Jeřábek and Phuong Nguyen, *Simulating non-prenex cuts in quantified propositional calculus*, Mathematical Logic Quarterly 57 (2011), no. 5, pp. 524–532.

[25] Jan Johannsen and Chris Pollett, *On the* Δ^b_1*-bit-comprehension rule*, in: Logic Colloquium '98: Proceedings of the 1998 ASL European Summer Meeting held in Prague, Czech Republic (S. R. Buss, P. Hájek, and P. Pudlák, eds.), ASL, 2000, pp. 262–280.

[26] Richard Kaye, *Parameter free induction in arithmetic*, in: Proceedings of the 5th Easter Conference on Model Theory, Sektion Mathematik der Humboldt-Universität zu Berlin, 1987, pp. 70–81, Seminarbericht Nr. 93.

[27] ———, *Axiomatizations and quantifier complexity*, in: Proceedings of the 6th Easter Conference on Model Theory, Sektion Mathematik der Humboldt-Universität zu Berlin, 1988, pp. 65–84, Seminarbericht Nr. 98.

[28] ———, *Diophantine induction*, Annals of Pure and Applied Logic 46 (1990), no. 1, pp. 1–40.

[29] Richard Kaye, Jeff Paris, and Costas Dimitracopoulos, *On parameter free induction schemas*, Journal of Symbolic Logic 53 (1988), no. 4, pp. 1082–1097.

[30] Jan Krajíček, *Fragments of bounded arithmetic and bounded query classes*, Transactions of the American Mathematical Society 338 (1993), no. 2, pp. 587–598.

[31] ——————, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications vol. 60, Cambridge University Press, 1995.

[32] Jan Krajíček and Samuel R. Buss, *An application of boolean complexity to separation problems in bounded arithmetic*, Proceedings of the London Mathematical Society 69 (1994), no. 3, pp. 1–21.

[33] Jan Krajíček and Pavel Pudlák, *Quantified propositional calculi and fragments of bounded arithmetic*, Zeitschrift für mathematische Logik und Grundlagen der Mathematik 36 (1990), no. 1, pp. 29–46.

[34] Jan Krajíček, Pavel Pudlák, and Gaisi Takeuti, *Bounded arithmetic and the polynomial hierarchy*, Annals of Pure and Applied Logic 52 (1991), no. 1–2, pp. 143–153.

[35] Jeff B. Paris and Alex J. Wilkie, *Counting problems in bounded arithmetic*, in: Methods in Mathematical Logic (C. A. Di Prisco, ed.), Lecture Notes in Mathematics vol. 1130, Springer, 1985, pp. 317–340.

[36] Alan Skelley and Neil Thapen, *The provably total search problems of bounded arithmetic*, Proceedings of the London Mathematical Society 103 (2011), no. 1, pp. 106–138.

[37] Andrew C.-C. Yao, *Separating the polynomial-time hierarchy by oracles*, in: Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science (R. E. Tarjan, ed.), 1985, pp. 1–10.

[38] Domenico Zambella, *Notes on polynomially bounded arithmetic*, Journal of Symbolic Logic 61 (1996), no. 3, pp. 942–966.