

On Rules and Parameter Free Systems in Bounded Arithmetic

Andres Cordón-Franco, Alejandro Fernández-Margarit,
and Francisco Felix Lara-Martín*

Facultad de Matemáticas. Universidad de Sevilla
C/ Tarfia, s/n, 41012 Sevilla, Spain
{acordon,afmargarit,fflara}@us.es

Abstract. We present model-theoretic techniques to obtain conservation results for first order bounded arithmetic theories, based on a hierarchical version of the well known notion of an existentially closed model.

Keywords: Bounded Arithmetic, conservation results, parameter free schemes.

1 Introduction

Bounded arithmetic theories are formal systems tailored to capture computational complexity classes. The foundational work in this area is [3], where S. Buss introduced the families of theories S_2^i and T_2^i ($i \geq 0$) and showed that they can be considered as formal counterparts of the Polynomial Time Hierarchy *PH*. Since then a variety of related systems have been introduced in order to deal with other complexity classes. Among the fundamental results on these systems two groups can be isolated: (a) characterizations of their computational strength, mainly, by determining their Σ_i^b -definable functions; and (b) relationship among different axiomatizations, especially, *conservation results*.

Here we present model-theoretic methods to obtain both kinds of results for restricted versions of Buss's theories S_2^i , T_2^i as well as for the Σ_i^b -replacement scheme $BB\Sigma_i^b$. Systems S_2^i and T_2^i are axiomatized over a certain base theory by *axiom schemes* expressing (respectively) the polynomial and the usual induction principles restricted to Σ_i^b -formulas. We shall weaken these theories in two ways: (1) by formalizing the corresponding induction or replacement principle as an *inference rule* instead of an axiom scheme, or (2) by restricting the induction schemes to *parameter free* formulas. In the first case we drop the axiom scheme and consider the closure of the base theory under first order logic and nested applications of the corresponding inference rule. In the second case we still deal with an axiom scheme but now it is restricted to formulas with no other free variables than the induction variable. While the effects of these restrictions have been extensively investigated for fragments of Peano Arithmetic, it is not the

* Partially supported by grants MTM2005-08658 of MEC and TIC-137 of Junta de Andalucía, Spain. Part of this work was done while the first author was visiting the Mathematical Institute of the Academy of Sciences of the Czech Republic in 2004.

case of Bounded Arithmetic. To our best knowledge, parameter free systems have only been systematically studied by S. Bloch in the second part of his thesis [2]. On the other hand, systems described by inference rules in the sense above seldom appear in an explicit manner in the literature. A recent exception is J. Johannsen and C. Pollett's work [7], where the authors study the Δ_1^b -bit-comprehension rule in connection with the complexity class TC^0 of functions computable by uniform threshold circuit families of polynomial size and constant depth. Moreover, both in [2] and in [7] the analysis of those systems has been carried out by means of proof-theoretic methods.

In this paper we shall develop a model-theoretic approach to the investigation of these restricted systems. To this end, the key ingredient is the notion of an $\exists\hat{\Pi}_i^b$ -maximal model, a hierarchical version of the well known notion of an existentially closed model. These models allow us to clarify the relationships between the considered theories and their restricted versions in a particularly simple way. Namely, if T denotes S_2^i , T_2^i or $BB\Sigma_i^b$ and T^R (resp. T^-) denotes its inference rule (resp. parameter free) version, then (see Thm. 1 and Prop. 3)

- every $\exists\hat{\Pi}_i^b$ -maximal model of T^R is a model of T , and
- every theory extending T^- is closed under the corresponding inference rule and, so, every $\exists\hat{\Pi}_i^b$ -maximal model of T^- is a model of T .

From these facts we shall derive our main results (see Theorems 3 and 4): (1) S_2^i , T_2^i and $BB\Sigma_i^b$ are $\forall\Sigma_i^b$ -conservative over their inference rule versions; and (2) S_2^i , T_2^i are $\exists\forall\Sigma_i^b$ -conservative over their parameter free versions. As far as we know, these results are new, and the $\exists\forall\Sigma_i^b$ -conservation results for parameter free systems improve previous $\forall\Sigma_i^b$ -conservation obtained in [2].

Finally, in Sect. 4 we apply the results obtained for Σ_1^b -replacement to the analysis of the Δ_1^b -bit-comprehension rule Δ_1^b -BCR. This rule was introduced in [7] to capture the complexity class TC^0 and is the final refinement of a series of theories introduced in [5,6,7] in the quest for natural theories for TC^0 . In [7] it is proved that TC^0 coincides with the Σ_1^b -definable functions of the system Δ_1^b -CR given by the closure under Δ_1^b -BCR of a certain base theory; and that the Σ_0^b -replacement scheme $BB\Sigma_0^b$ is $\forall\Sigma_1^b$ -conservative over Δ_1^b -CR. Here, we prove that TC^0 also coincides with the $\hat{\Sigma}_1^b$ -definable functions of the (apparently) weaker system $\hat{\Delta}_1^b$ -CR and reformulate this system in terms of Σ_i^b -replacement rule, obtaining as a corollary a new proof of the conservation result in [7]. Our analysis is of independent interest in view of the open problems on Δ_1^b -BCR posed in [7]; however, it also supports Johannsen-Pollett's claim on Δ_1^b -CR as a minimal natural theory for TC^0 and makes more transparent the close relationship between Δ_1^b -bit-comprehension and Σ_1^b -replacement.

2 Fragments of Bounded Arithmetic

In what follows we state some definitions and results on Bounded Arithmetic that will be used through this paper (see [3,8] for more information). The first order language of Bounded Arithmetic \mathcal{L}_2 comprises the usual language of Peano

Arithmetic $\{0, S, +, \cdot, \leq\}$ together with five new function symbols: $\lfloor \frac{x}{2} \rfloor$, $|x|$, $\#$, MSP and \bullet ; where $\lfloor \frac{x}{2} \rfloor$ is x divided by 2 rounded down, $|x|$ is the length of x in binary notation, $x\#y$ is $2^{|x| \cdot |y|}$, $MSP(x, i)$ is $\lfloor \frac{x}{2^i} \rfloor$, and $x \bullet y$ is the subtraction function. As usual, we also write $x + 1$ and $2^{|x|}$ for Sx and $1\#x$, respectively. Bounded formulas of \mathcal{L}_2 are classified in a hierarchy of sets Σ_i^b and Π_i^b by counting the alternations of bounded quantifiers ($\exists x \leq t, \forall x \leq t$), ignoring sharply bounded quantifiers ($\exists x \leq |t|, \forall x \leq |t|$).

The *induction* axiom for $\varphi(x)$, I_φ , is the formula

$$\varphi(0) \wedge \forall x (\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x \varphi(x)$$

The *length induction* axiom for $\varphi(x)$, $LIND_\varphi$, and the *double length induction* axiom for $\varphi(x)$, $LLIND_\varphi$, are obtained replacing the consequent of I_φ by $\forall x \varphi(|x|)$ and $\forall x \varphi(|x|)$, respectively. The *polynomial induction* axiom for $\varphi(x)$, $PIND_\varphi$, is the formula

$$\varphi(0) \wedge \forall x (\varphi(\lfloor \frac{x}{2} \rfloor) \rightarrow \varphi(x)) \rightarrow \forall x \varphi(x)$$

In all cases, $\varphi(x)$ may contain other free variables, which are called parameters. On a par with these induction axioms, we consider induction *inference rules*. The *induction rule* for $\varphi(x)$, IR , is

$$\frac{\varphi(0), \forall x (\varphi(x) \rightarrow \varphi(Sx))}{\forall x \varphi(x)}$$

Similarly, $PINDR$, $LINDR$ and $LLINDR$ are defined.

BASIC denotes a finite set of open (quantifier-free) axioms specifying the interpretations of the nonlogical symbols of \mathcal{L}_2 . Following [7,10], our base theory will be $LIOpen = BASIC + \{LIND_\varphi : \varphi \text{ is open}\}$. As shown there, $LIOpen$ allows for simple definitions of tuple-encoding and sequence-encoding functions. First, observe that there are \mathcal{L}_2 -terms $Bit(x, i)$ and $LSP(x, i)$ returning the value of the bit in the 2^i position of the binary representation of x , and the number consisting of the low i bits of x , respectively. The code of a sequence $\{b_0, b_1, \dots, b_{|s|}\}$ with all its elements less than or equal to some a is the number $w < 4(a\#2s)$ whose binary representation consists of a 1 followed by the binary representations of the elements b_i concatenated, each padded with zeroes to length $|a|$ (we shall write $bd(a, s)$ for the bounding term $4(a\#2s)$). Thus, the \mathcal{L}_2 -term $\beta_a(w, i) := MSP(LSP(w, Si \cdot |a|), i \cdot |a|)$ returns the i -th element of such a sequence. As for tuple-encoding, pairs are coded as $\langle x, y \rangle := (B+y) \cdot 2B + (B+x)$, where $B = 2^{|max(x,y)|}$. Then there is an open formula $ispair(u)$ defining the range of the function $\langle x, y \rangle$; and there are terms $(u)_0, (u)_1$ returning the left and right coordinates from a coded pair (see [10] for details). Interestingly, the encoding and decoding functions are all \mathcal{L}_2 -terms so can be used in an \mathcal{L}_2 -formula without altering its quantifier complexity.

The theories we shall deal with are defined as follows. Let Γ be a set of formulas and let \mathbf{E} denote one of the schemes: $I, PIND, LIND, LLIND$. First,

the theory $\mathbf{E}\Gamma$ is $\text{LIOpen} + \{\mathbf{E}_\varphi : \varphi \in \Gamma\}$. Second, the fragment $T + \Gamma\text{-ER}$ is the closure of T under first order logic and nested applications of the \mathbf{E} -rule restricted to formulas in Γ , where T is an arbitrary \mathcal{L}_2 -theory extending LIOpen . Finally, $\mathbf{E}\Gamma^-$ is $\text{LIOpen} + \{\mathbf{E}_\varphi : \varphi(x) \in \Gamma^-\}$, where $\varphi(x) \in \Gamma^-$ means that x is the only free variable occurring in φ .

With this terminology, the three classic families of Bounded Arithmetic theories T_2^i , S_2^i and R_2^i correspond to $I\Sigma_i^b$, $PIND\Sigma_i^b$ and $LLIND\Sigma_i^b$, respectively. Let us remark, however, that \mathcal{L}_2 differs from the language of Buss's original theories S_2^i and T_2^i , which does not include the MSP and \bullet symbols. In addition, Buss's theories are axiomatized over $BASIC$ instead of over LIOpen . But these facts are inessential for sufficiently strong theories since both additional functions are Σ_1^b -definable in Buss's S_2^1 , and this last theory implies LIOpen .

Bounded formulas of \mathcal{L}_2 are also classified in a hierarchy of sets *strict* Σ_i^b ($= \hat{\Sigma}_i^b$) and *strict* Π_i^b ($= \hat{\Pi}_i^b$), where no sharply bounded quantifier is allowed to precede a quantifier that is not sharply bounded. Each Σ_i^b (resp. Π_i^b) formula is equivalent to a $\hat{\Sigma}_i^b$ (resp. $\hat{\Pi}_i^b$) formula and the $\hat{\Pi}_{i-1}^b$ -replacement scheme $BB\hat{\Pi}_{i-1}^b$ is a natural theory which proves that equivalence. The *replacement* or *bounded collection* axiom for a formula $\varphi(x, y)$ and a term $t(x)$, BB_φ , is

$$\begin{aligned} \forall x \leq |s| \exists y \leq t(x) \varphi(x, y) \rightarrow \\ \exists w < bd(t^*(|s|), s) \forall x \leq |s| (\beta_{t^*(|s|)}(w, x) \leq t(x) \wedge \varphi(x, \beta_{t^*(|s|)})), \end{aligned}$$

where t^* denotes an \mathcal{L}_2 -term canonically associated to t so that, provably in LIOpen , t^* is monotonic and $t \leq t^*$ (see [7,10] for details).

$BB\Gamma$ is $\text{LIOpen} + \{BB_\varphi : \varphi \in \Gamma\}$. Similarly, the inference rule versions $T + \Gamma\text{-BBR}$ are defined. In [10] it is shown that every Σ_i^b formula is provably equivalent in $BB\hat{\Pi}_{i-1}^b$ to a $\hat{\Sigma}_i^b$ -formula, and that $PIND\hat{\Sigma}_i^b$ implies $BB\hat{\Pi}_{i-1}^b$ ($i \geq 1$). As a consequence, the author obtains the equivalences $T_2^i \equiv I\hat{\Sigma}_i^b$ and $S_2^i \equiv PIND\hat{\Sigma}_i^b \equiv LIND\hat{\Sigma}_i^b$. Finally, reasoning as in the proof of result 3.2 in [5], it is easy to show that $BB\hat{\Sigma}_{i+1}^b \equiv BB\hat{\Pi}_i^b$, and $T + \hat{\Sigma}_{i+1}^b\text{-BBR} \equiv T + \hat{\Pi}_i^b\text{-BBR}$.

3 On $\exists\hat{\Pi}_i^b$ -Maximal Models and Conservation Results

In this section we present our methods for proving conservation results. To illustrate these methods, we prove that S_2^i , T_2^i and $BB\Sigma_i^b$ are $\forall\Sigma_i^b$ -conservative over their inference rule versions; and we use these results to show that S_2^i and T_2^i are $\exists\forall\Sigma_i^b$ -conservative over their parameter free versions. The main idea involves a basic model-theoretic argument: we show that each (countable) model of the weak theory has a $\hat{\Sigma}_i^b$ -*elementary extension* to a model of the strong theory (\mathfrak{B} is a Γ -elementary extension of \mathfrak{A} , $\mathfrak{A} \prec_\Gamma \mathfrak{B}$, if $\mathfrak{A} \subseteq \mathfrak{B}$ and, for all $\varphi(\vec{x}) \in \Gamma$ and $\vec{a} \in \mathfrak{A}$, it holds that $\mathfrak{A} \models \varphi(\vec{a}) \iff \mathfrak{B} \models \varphi(\vec{a})$). The key ingredient for this construction is the notion of an $\exists\hat{\Pi}_i^b$ -*maximal model* for a theory T .

Definition 1. Let \mathfrak{A} be a model of T . We say that \mathfrak{A} is $\exists\hat{\Pi}_i^b$ -maximal for T if, for each $\mathfrak{B} \models T$, it holds that $\mathfrak{A} \prec_{\hat{\Sigma}_i^b} \mathfrak{B} \implies \mathfrak{A} \prec_{\exists\hat{\Pi}_i^b} \mathfrak{B}$.

This notion is a suitably modified version of the general concept of an existentially closed model. The use of similar notions to prove conservation results for arithmetic systems was presented in a general setting in J. Avigad's [1] (our work is inspired by the methods in that paper). First of all, observe that $\exists\hat{\Pi}_i^b$ -maximal models do exist. The proof is an easy modification of the standard iterative argument to construct existentially closed models.

Proposition 1. *Suppose T is $\forall\exists\hat{\Pi}_i^b$ -axiomatizable and \mathfrak{A} is a countable model of T . Then there is $\mathfrak{B} \models T$ such that $\mathfrak{A} \prec_{\hat{\Sigma}_i^b} \mathfrak{B}$ and \mathfrak{B} is $\exists\hat{\Pi}_i^b$ -maximal for T .*

Next, we prove the main property of these models of interest to us: each $\exists\hat{\Pi}_i^b$ -maximal model for $T + \hat{\Sigma}_i^b\text{-ER}$ also satisfies the corresponding scheme $\mathbf{E}\hat{\Sigma}_i^b$. We first need the following result (the proof is a standard compactness argument).

Proposition 2. *Let \mathfrak{A} be $\exists\hat{\Pi}_i^b$ -maximal for T , $\vec{a} \in \mathfrak{A}$ and $\varphi(\vec{x}, \vec{v}) \in \hat{\Sigma}_i^b$ and let $\hat{\Pi}_i^b\text{-Diag}(\mathfrak{A})$ denote the set of all the $\hat{\Pi}_i^b$ -formulas (with parameters in \mathfrak{A}) valid in \mathfrak{A} . The following conditions are equivalent.*

1. $\mathfrak{A} \models \forall \vec{x} \varphi(\vec{x}, \vec{a})$.
2. There is $\theta(\vec{a}, \vec{b})$ in $\hat{\Pi}_i^b\text{-Diag}(\mathfrak{A})$ satisfying $T + \theta(\vec{a}, \vec{b}) \vdash \forall \vec{x} \varphi(\vec{x}, \vec{a})$.

Theorem 1. *Let \mathbf{E} denote one of the following schemes: $BB, I, PIND, LIND, LLIND$. If \mathfrak{A} is $\exists\hat{\Pi}_i^b$ -maximal for $T + \hat{\Sigma}_i^b\text{-ER}$, then $\mathfrak{A} \models \mathbf{E}\hat{\Sigma}_i^b$.*

Proof. (Collection scheme): Assume \mathfrak{A} is $\exists\hat{\Pi}_i^b$ -maximal for $T + \hat{\Sigma}_i^b\text{-BBR}$ and $\mathfrak{A} \models \forall x \leq |s| \exists y \leq t \varphi(x, y, a)$, where $\varphi(x, y, v) \in \hat{\Sigma}_i^b$, $a \in \mathfrak{A}$ and s, t are \mathcal{L}_2 -terms (for notational simplicity we omit the possible parameters in t, s). By Proposition 2 there are $b \in \mathfrak{A}$ and $\theta(v, u)$ in $\hat{\Pi}_i^b$ such that $\mathfrak{A} \models \theta(a, b)$, and $(T + \hat{\Sigma}_i^b\text{-BBR}) + \theta(a, b) \vdash \forall x \leq |s| \exists y \leq t \varphi(x, y, a)$. So, $T + \hat{\Sigma}_i^b\text{-BBR} \vdash \theta(v, u) \rightarrow \forall x \leq |s| \exists y \leq t \varphi(x, y, v)$. Define $\delta(x, y, v, u)$ to be $\neg\theta(v, u) \vee \varphi(x, y, v)$. Clearly, δ is $\hat{\Sigma}_i^b$ and $T + \hat{\Sigma}_i^b\text{-BBR}$ proves the antecedent of the bounded collection axiom for $\delta(x, y)$. Applying $\hat{\Sigma}_i^b\text{-BBR}$ in \mathfrak{A} and taking $v = a$ and $u = b$, we get

$$\mathfrak{A} \models \exists w < bd(t^*(|s|), s) \forall x \leq |s| (\beta_{t^*(|s|)}(w, x) \leq t \wedge \delta(x, \beta_{t^*(|s|)}(w, x), a, b))$$

Since $\mathfrak{A} \models \theta(a, b)$, $\mathfrak{A} \models \delta(x, y, a, b) \rightarrow \varphi(x, y, a)$ and hence the consequent of the bounded collection axiom for $\delta(x, y, a)$ is true in \mathfrak{A} .

(Induction schemes): We only write the proof for the usual induction scheme I , the remaining cases being analogous. Assume \mathfrak{A} is $\exists\hat{\Pi}_i^b$ -maximal for $T + \hat{\Sigma}_i^b\text{-IR}$. To prove that $\mathfrak{A} \models I\hat{\Sigma}_i^b$, assume $\mathfrak{A} \models \varphi(0, a) \wedge \forall x (\varphi(x, a) \rightarrow \varphi(x + 1, a))$, where $\varphi(x, v) \in \hat{\Sigma}_i^b$ and $a \in \mathfrak{A}$. We must show $\mathfrak{A} \models \forall x \varphi(x, a)$. Put $\varphi(x, v)$ as $\exists y \leq t(x, v) \varphi_0(x, y, v)$, where $\varphi_0(x, y, v) \in \hat{\Pi}_{i-1}^b$ and $t(x, v)$ is a term. By prenex operations, the antecedent of the induction axiom for φ can be reexpressed as

$$\forall x \forall y [\varphi(0, a) \wedge (\neg(y \leq t(x, a)) \vee \neg\varphi_0(x, y, a) \vee \varphi(x + 1, a))]$$

Let us denote by $\psi(x, y, a)$ the $\hat{\Sigma}_i^b$ -formula in brackets $[\]$ above. Since \mathfrak{A} is $\exists\hat{\Pi}_i^b$ -maximal for $T + \hat{\Sigma}_i^b\text{-IR}$ and $\mathfrak{A} \models \forall x, y \psi(x, y, a)$, by Proposition 2 it follows that

there are $b \in \mathfrak{A}$ and $\theta(v, u) \in \hat{\Pi}_i^b$ satisfying $\mathfrak{A} \models \theta(a, b)$, and $(T + \hat{\Sigma}_i^b\text{-IR}) + \theta(a, b) \vdash \forall x, y \psi(x, y, a)$. Hence,

$$T + \hat{\Sigma}_i^b\text{-IR} \vdash \theta(v, u) \rightarrow (\varphi(0, v) \wedge \forall x (\varphi(x, v) \rightarrow \varphi(x + 1, v)))$$

Now define $\delta(x, v, u)$ to be the $\hat{\Sigma}_i^b$ -formula $\neg\theta(v, u) \vee \varphi(x, v)$. Clearly, $T + \hat{\Sigma}_i^b\text{-IR}$ proves the antecedent of the induction axiom for $\delta(x, v, u)$. By applying $\hat{\Sigma}_i^b\text{-IR}$, we get $\mathfrak{A} \models \forall x, v, u \delta(x, v, u)$. In particular, $\mathfrak{A} \models \forall x (\neg\theta(a, b) \vee \varphi(x, a))$, and hence $\mathfrak{A} \models \forall x \varphi(x, a)$ since $\theta(a, b)$ is true in \mathfrak{A} . \square

Combining Proposition 1 and Theorem 1, we can derive our $\forall\Sigma_i^b$ -conservation results. The proof is in two steps. First, we prove this conservation result only for $\forall\hat{\Sigma}_i^b$ -formulas. Second, we show how to extend it to general $\forall\Sigma_i^b$ -formulas.

Theorem 2. *Let \mathbf{E} denote one of the following schemes: $BB, I, PIND, LIND, LLIND$ and let T be a $\forall\exists\hat{\Pi}_i^b$ -axiomatizable theory. Then $T + \mathbf{E}\hat{\Sigma}_i^b$ is $\forall\hat{\Sigma}_i^b$ -conservative over $T + \hat{\Sigma}_i^b\text{-ER}$.*

Proof. By contradiction, assume $T + \mathbf{E}\hat{\Sigma}_i^b \vdash \varphi$ but $T + \hat{\Sigma}_i^b\text{-ER} \not\vdash \varphi$, where $\varphi \in \forall\hat{\Sigma}_i^b$. Let \mathfrak{A} be a countable model of $(T + \hat{\Sigma}_i^b\text{-ER}) + \neg\varphi$. Since T is $\forall\exists\hat{\Pi}_i^b$ -axiomatizable, so is $T + \hat{\Sigma}_i^b\text{-ER}$ (for $\mathbf{E} = BB$, recall that $\hat{\Sigma}_i^b\text{-BBR}$ and $\hat{\Pi}_{i-1}^b\text{-BBR}$ are equivalent rules). By Proposition 1 there is $\mathfrak{B} \models T + \hat{\Sigma}_i^b\text{-ER}$ such that $\mathfrak{A} \prec_{\hat{\Sigma}_i^b} \mathfrak{B}$ and \mathfrak{B} is $\exists\hat{\Pi}_i^b$ -maximal for $T + \hat{\Sigma}_i^b\text{-ER}$. From Theorem 1 it follows that $\mathfrak{B} \models \mathbf{E}\hat{\Sigma}_i^b$. Hence, $\mathfrak{B} \models T + \mathbf{E}\hat{\Sigma}_i^b + \neg\varphi$, which is a contradiction. \square

Since T_2^i and S_2^i are $\forall\hat{\Sigma}_{i+1}^b$ -axiomatizable, a first application of Theorem 2 is the following strengthening of the well known facts that S_2^{i+1} implies T_2^i and R_2^{i+1} implies S_2^i , and of theorem 68 in [10] stating that $BB\hat{\Sigma}_{i+1}^b$ implies S_2^i .

Corollary 1

1. $LIOpen + \hat{\Sigma}_{i+1}^b\text{-LINDR}$ implies T_2^i .
2. Both $LIOpen + \hat{\Sigma}_{i+1}^b\text{-LLINDR}$ and $LIOpen + \hat{\Sigma}_{i+1}^b\text{-BBR}$ imply S_2^i .

To extend previous conservation result to $\forall\Sigma_i^b$ -formulas, we need the following lemma (the proof is by induction on the complexity of Σ_i^b -formulas).

Lemma 1. *($i \geq 1$) Let $\varphi(\vec{v}) \in \Sigma_i^b$. There exists $\hat{\varphi}(\vec{v}) \in \hat{\Sigma}_i^b$ such that:*

- (1) $BB\hat{\Pi}_{i-1}^b \vdash \varphi(\vec{v}) \leftrightarrow \hat{\varphi}(\vec{v})$, and
- (2) $BB\hat{\Pi}_{i-2}^b \vdash \hat{\varphi}(\vec{v}) \rightarrow \varphi(\vec{v})$.

(For $i = 1$, $BB\hat{\Pi}_{i-1}^b$ denotes $LIOpen$.)

Theorem 3. *Let \mathbf{E} denote one of the following schemes: $BB, I, PIND, LIND$. Then, $LIOpen + \hat{\Sigma}_i^b\text{-ER}$ axiomatizes the $\forall\Sigma_i^b$ -consequences of $\mathbf{E}\Sigma_i^b$.*

Proof. Assume $\mathbf{E}\Sigma_i^b \vdash \forall\vec{v} \varphi(\vec{v})$, where $\varphi \in \Sigma_i^b$. Let $\hat{\varphi}(\vec{v}) \in \hat{\Sigma}_i^b$ as in Lemma 1. Since S_2^i implies $BB\hat{\Pi}_{i-1}^b$ (see [3]), $\mathbf{E}\Sigma_i^b$ implies $BB\hat{\Pi}_{i-1}^b$ and $LIOpen + \hat{\Sigma}_i^b\text{-ER}$ implies $BB\hat{\Pi}_{i-2}^b$ by Corollary 1. Hence, $\mathbf{E}\Sigma_i^b \vdash \forall\vec{v} \hat{\varphi}(\vec{v})$ and $LIOpen + \hat{\Sigma}_i^b\text{-ER} \vdash \hat{\varphi}(\vec{v}) \rightarrow \varphi(\vec{v})$. So, this last theory proves $\forall\vec{v} \varphi(\vec{v})$ by Theorem 2. \square

In what follows we deal with parameter free versions of T_2^i and S_2^i . Notice that there are two natural candidates for their parameter free counterparts: restricting the axiom scheme to parameter free Σ_i^b -formulas, or to *strict* parameter free Σ_i^b -formulas. Since we are interested in conservation results over these theories, we choose the weakest ones to make the results stronger. That is, we fix $T_2^{i,-} \equiv I\hat{\Sigma}_i^{b,-}$ and $S_2^{i,-} \equiv PIND\hat{\Sigma}_i^{b,-}$. We derive the conservation theorems from our previous work on inference rules. The key observation is the following:

Proposition 3

1. If T implies $T_2^{i,-}$ then T is closed under $\hat{\Sigma}_i^b$ -IR.
2. ($i \geq 1$) If T implies $PIND\Sigma_1^{b,-} + S_2^{i,-}$ then T is closed under $\hat{\Sigma}_i^b$ -PINDR.

Proof. (1): Assume T proves $\varphi(0, v) \wedge \forall x (\varphi(x, v) \rightarrow \varphi(x + 1, v))$, where $\varphi(x, v)$ is $\hat{\Sigma}_i^b$. We must show $T \vdash \forall v \forall x \varphi(x, v)$. The idea is to codify the parameter v and the induction variable x in a single variable u using the pairing function and to apply $I\hat{\Sigma}_i^{b,-}$. To this end, define $\theta(u)$ to be the following $\hat{\Sigma}_i^b$ -formula:

$$(ispair(u) \wedge (u)_0 < (u)_1 \wedge ispair((u)_1)) \rightarrow \varphi((u)_0, (u)_{1,1})$$

Trivially, $T \vdash \theta(0)$ since $\neg ispair(0)$. Let us see that $T \vdash \forall u (\theta(u) \rightarrow \theta(u + 1))$. Reasoning in T , we assume $\theta(u)$ and $(ispair(u') \wedge (u')_0 < (u')_1 \wedge ispair((u')_1))$, where $u' = u + 1$. We must show $\varphi((u')_0, (u')_{1,1})$.

Case 1: $(u')_0 = 0$. Then $\varphi(0, (u')_{1,1})$ since $T \vdash \forall v \varphi(0, v)$.

Case 2: $(u')_0 > 0$. Since $(u')_0 < (u')_1$, $max((u')_0 - 1, (u')_1) = (u')_1$ and hence by the definition of the pairing function u codifies the pair $\langle (u')_0 - 1, (u')_1 \rangle$ (that is, $(u)_0 = (u')_0 - 1$ and $(u)_1 = (u')_1$). Consequently, from $\theta(u)$ it follows $\varphi((u')_0 - 1, (u')_{1,1})$ and hence $\varphi((u')_0, (u')_{1,1})$ since $T \vdash \varphi(x, v) \rightarrow \varphi(x + 1, v)$.

From the induction axiom for $\theta(u)$ (available in T since it contains $I\hat{\Sigma}_i^{b,-}$) it follows that $T \vdash \forall u \theta(u)$. To show $T \vdash \forall v \forall x \varphi(x, v)$, observe that $\varphi(x, v)$ can be inferred from $\theta(\langle x, \langle x, v \rangle \rangle)$.

(2): The proof is similar to that of 1 but now we need to define a new tuple-encoding function *compatible* with the PIND axioms: roughly speaking, if u codifies the pair (x, \vec{v}) and $x > 0$, then $\lfloor \frac{u}{2} \rfloor$ must codify the pair $(\lfloor \frac{x}{2} \rfloor, \vec{v})$. In [2] Bloch proposed the following encoding function satisfying that property:

$$[x, v, z] = u \equiv \begin{cases} |v| < z^2 \leq |u| < (z + 1)^2 \wedge \\ u = Concat(v + 2^{\min(z^2, |u|)}, x + 2^{|x|}) \end{cases}$$

where $Concat(x, y) = x \cdot 2^{|y|^{\bullet 1}} + y \bullet 2^{|y|^{\bullet 1}}$. In words, we pad v to length z^2 and concatenate the result with x (notice that the *Concat* function operates on bit-strings rather than on binary numbers, that is, $Concat(1x, 1y) = 1xy$). Observe that the encoding function $[x, v, z]$ itself is not total, but it is total for all z sufficiently large. Namely, as shown in [2], $PIND\Sigma_1^{b,-}$ proves:

- (a) $|x| \leq 2z \wedge |v| < z^2 \rightarrow \exists! u ([x, v, z] = u)$,
- (b) $u > 0 \rightarrow \exists! x, v, z \leq u ([x, v, z] = u)$, and
- (c) $u = [x, v, z] \wedge x > 0 \rightarrow \lfloor \frac{u}{2} \rfloor = [\lfloor \frac{x}{2} \rfloor, v, z]$

Equipped with this encoding function, we can infer the PIND axiom for the $\hat{\Sigma}_i^b$ -formula $\varphi(x, v)$ from the PIND axiom for the (parameter free) $\hat{\Sigma}_i^b$ -formula $\theta(u) \equiv u > 0 \rightarrow \exists x, v, z \leq u ([x, v, z] = u \wedge \varphi(x, v))$. \square

Observe that from the previous result and Theorem 2 it immediately follows that $T_2^{i+1,-}$ implies T_2^i and that $PIND_{\Sigma_1^{b,-}} + S_2^{i+1,-}$ implies S_2^i .

Theorem 4

1. T_2^i is $\exists \forall \Sigma_i^b$ -conservative over $T_2^{i,-}$.
2. ($i \geq 1$) S_2^i is $\exists \forall \Sigma_i^b$ -conservative over $PIND_{\Sigma_1^{b,-}} + S_2^{i,-}$.

Proof. Using Lemma 1 as in Theorem 3, it suffices to show $\exists \forall \hat{\Sigma}_i^b$ -conservation. We only write the proof of 1. Assume φ is an $\exists \forall \hat{\Sigma}_i^b$ -sentence such that $T_2^i \vdash \varphi$ but $T_2^{i,-} \not\vdash \varphi$. Then $T = T_2^{i,-} + \neg \varphi$ is consistent and $\forall \exists \hat{\Pi}_i^b$ -axiomatizable. Let \mathfrak{A} be an $\exists \hat{\Pi}_i^b$ -maximal model for T . By Proposition 3, T is closed under $\hat{\Sigma}_i^b$ -IR. Hence, $\mathfrak{A} \models T + T_2^i$ by Theorem 1. So, $\mathfrak{A} \models T_2^i + \neg \varphi$, which is a contradiction. \square

As for parameter free $BB\Sigma_i^b$, we can prove that $BB\Sigma_i^b$ is $\exists \forall \Sigma_i^b$ -conservative over $UBB\hat{\Sigma}_i^b$ as in Theorem 4 (UBB_φ is obtained quantifying universally the parameters of $\varphi(x, y)$ in both the antecedent and the consequent of BB_φ).

4 On Replacement and Bit-Comprehension Rules

In this section we shall study an inference rule closely tied to Σ_1^b -replacement: Δ_1^b -bit-comprehension rule. This rule was defined in [7] as follows:

$$\Delta_1^b\text{-BCR} : \frac{\varphi(x) \leftrightarrow \psi(x)}{\exists y < 2^{|u|} \forall x < |u| (Bit(y, x) = 1 \leftrightarrow \varphi(x))}$$

where $\varphi(x) \in \Sigma_1^b$ and $\psi(x) \in \Pi_1^b$. In [7], it is proved that $BB\Sigma_0^b$ (denoted there by C_0^2) is a $\forall \Sigma_1^b$ -conservative extension of Δ_1^b -CR (the theory $LIOpen + \Delta_1^b$ -BCR). So, in view of Theorem 3, it is natural to investigate the relationship between Δ_1^b -BCR and $\hat{\Sigma}_1^b$ -BBR. In this section, we consider the apparently weaker rule for *strict* formulas $\hat{\Delta}_1^b$ -BCR and show that $LIOpen + \hat{\Delta}_1^b$ -BCR (denoted in what follows by $\hat{\Delta}_1^b$ -CR) is equivalent to $LIOpen + \hat{\Sigma}_1^b$ -BBR, see Theorem 5. In fact, over $LIOpen$, the four rules Σ_1^b -BBR, $\hat{\Sigma}_1^b$ -BBR, Δ_1^b -BCR and $\hat{\Delta}_1^b$ -BCR are equivalent and, by Theorem 3, provide axiomatizations of the $\forall \Sigma_1^b$ -consequences of $BB\Sigma_1^b$. Moreover, in [9], answering a question posed in [7], it is shown that Δ_1^b -CR is finitely axiomatizable; so, a finite number of *nested* applications of any of the rules above axiomatizes the $\forall \Sigma_1^b$ -consequences of $BB\Sigma_1^b$. However,

Problem 1. *Is $LIOpen + \hat{\Sigma}_1^b$ -BBR equivalent to $[LIOpen; \hat{\Sigma}_1^b$ -BBR], the closure of $LIOpen$ under first order logic and unnested applications of $\hat{\Sigma}_1^b$ -BBR?*

Our work suggests a positive answer to Problem 1 since this is the case for the analogous problem for collection rule in the usual language of Peano Arithmetic.

Now we prove that $LIOpen + \hat{\Sigma}_1^b$ -BBR and $\hat{\Delta}_1^b$ -CR are equivalent. Our work also provides a new proof of Johannsen-Pollett's conservation theorem.

Firstly, observe that it can be easily shown that $\text{LIOpen} + \hat{\Sigma}_1^b\text{-BBR}$ is closed under $\hat{\Delta}_1^b\text{-BCR}$. On the other hand, since C_2^0 coincides with $BB\Sigma_0^b$, by Theorem 2, C_2^0 is $\forall\hat{\Sigma}_1^b$ conservative over $\text{LIOpen} + \hat{\Sigma}_1^b\text{-BBR}$. So, in order to simultaneously get the equivalence of $\text{LIOpen} + \hat{\Sigma}_1^b\text{-BBR}$ and $\hat{\Delta}_1^b\text{-CR}$, and Johanssen–Pollett’s theorem, it suffices to prove that $\hat{\Delta}_1^b\text{-CR}$ is closed under $\hat{\Sigma}_1^b\text{-BBR}$. Next two lemmas are the key ingredients of the proof. The first one provides a weak form of replacement available in $\hat{\Delta}_1^b\text{-CR}$ (the proof is straightforward and we omit it). The second one is a selection (or witnessing) principle for $\hat{\Delta}_1^b\text{-CR}$.

Lemma 2. *Let $\varphi(x, y) \in \hat{\Sigma}_1^b$ such that $\hat{\Delta}_1^b\text{-CR} \vdash \forall x \leq |s| \exists! y \leq t \varphi(x, y)$. Then*

$$\hat{\Delta}_1^b\text{-CR} \vdash \exists w < bd(t^*(|s|), s) \forall x \leq |s| (\beta_{t^*(|s|)}(w, x) \leq t \wedge \varphi(x, \beta_{t^*(|s|)}(w, x)))$$

Lemma 3. *Let $\varphi(x, y) \in \hat{\Sigma}_1^b$ such that $\hat{\Delta}_1^b\text{-CR} \vdash \forall x \exists y \leq t \varphi(x, y)$, then there exists $\psi(x, y) \in \hat{\Sigma}_1^b$ such that $\hat{\Delta}_1^b\text{-CR}$ proves*

$$(1) \quad \forall x \exists! y \leq t \psi(x, y), \quad \text{and} \quad (2) \quad \forall x \forall y (\psi(x, y) \rightarrow \varphi(x, y))$$

Proof. (Sketch) The proof we present here leans upon an analysis of the class of $\hat{\Sigma}_1^b$ -definable functions $\hat{\Delta}_1^b\text{-CR}$. We refine corollary 1 in [7] and prove that TC^0 is the class of $\hat{\Sigma}_1^b$ -definable functions of $\hat{\Delta}_1^b\text{-CR}$. The basic result is a machine-independent characterization of TC^0 given by Clote and Takeuti in [4]:

Let BF be the set of basic functions $\{o, s_0, s_1, \#, \times, |\cdot| \} \cup \{II_i^n : 1 \leq i \leq n\}$, where $o(x) = 0$, $s_0(s) = 2x$, $s_1(x) = 2x + 1$, $|x| = \lceil \log_2(x + 1) \rceil$, $x\#y = 2^{|x| \cdot |y|}$, \times denotes the usual product and $II_i^n(x_1, \dots, x_n) = x_i$.

Given $g : \omega^n \rightarrow \omega$ and $h_0, h_1 : \omega^{n+1} \rightarrow \{0, 1\}$, a function f is defined by *concatenation recursion on notation* (CRN) from g, h_0 and h_1 if

$$\begin{aligned} f(0, \vec{x}) &= g(\vec{x}) \\ f(2n, \vec{x}) &= 2 \cdot f(n, \vec{x}) + h_0(n, \vec{x}), \quad \text{provided } n \neq 0 \\ f(2n + 1, \vec{x}) &= 2 \cdot f(n, \vec{x}) + h_1(n, \vec{x}) \end{aligned}$$

Clote and Takeuti proved that TC^0 is the smallest class of functions containing BF and closed under composition and CRN. In order to show that every function in TC^0 is $\hat{\Sigma}_1^b$ -definable in $\hat{\Delta}_1^b\text{-CR}$, we show a stronger technical result:

For each function $f \in TC^0$ there exist a formula $\psi(\vec{x}, y, z_1, \dots, z_n) \in \Sigma_0^b$ and terms $t(x), t_1(x, y), t_2(x, y, z_1), \dots, t_n(x, y, z_1, \dots, z_{n-1})$ such that the $\hat{\Sigma}_1^b$ -formula $\exists z_1 \leq t_1 \dots \exists z_n \leq t_n \psi(x, y, \vec{z})$ defines f in the standard model and

$$\text{LIOpen} + \hat{\Delta}_1^b\text{-BCR} \vdash \forall \vec{x} \exists! y \leq t \exists! z_1 \leq t_1 \dots \exists! z_n \leq t_n \psi(x, y, \vec{z})$$

The proof proceeds by induction, using Clote–Takeuti’s characterization of TC^0 . The claim obviously holds for the basic functions and for f defined by composition from functions verifying the claim. So it suffices to prove the result for functions defined by CRN and this can be done as in theorem 4 in [7].

It is not difficult to verify that, if $f \in TC^0$ is defined by CRN from g, h_0 and h_1 then the proof of the previous technical result provide $\hat{\Sigma}_1^b$ -formulas defining the functions involved and such that the relations stated by the recursion equations of CRN can be proved in $\hat{\Delta}_1^b\text{-CR}$. Bearing this fact in mind, we can introduce a universally axiomatized and conservative extension of $\hat{\Delta}_1^b\text{-CR}$, denoted by

CRNA. This universal theory can be defined in such a way that the functions in TC^0 are defined by terms of CRNA. In this way we can prove that every function $\hat{\Sigma}_1^b$ -definable in $\hat{\Delta}_1^b$ -CR is in TC^0 by a typical application of Herbrand's theorem. The whole argument is very similar to the Herbrand analyses of S_2^i developed by W. Sieg in [11].

Finally, we derive Lemma 3 from Herbrand's theorem applied to CRNA. \square

Theorem 5. *The theories $\hat{\Delta}_1^b$ -CR and $LIOpen + \Sigma_1^b$ -BBR are equivalent and axiomatize the class of the $\forall\Sigma_1^b$ -consequences of C_2^0 .*

Proof. Observe that C_2^0 extends $LIOpen + \Sigma_1^b$ -BBR, which in turn extends $\hat{\Delta}_1^b$ -CR; so, since $LIOpen + \Sigma_1^b$ -BBR is $\forall\Sigma_1^b$ -axiomatized, it suffices to prove that C_2^0 is $\forall\Sigma_1^b$ -conservative over $\hat{\Delta}_1^b$ -CR. Finally, by Theorem 3 it is enough to show that $\hat{\Delta}_1^b$ -CR is closed under $\hat{\Sigma}_1^b$ -BBR. Let us work in $\hat{\Delta}_1^b$ -CR.

Let $\varphi(x, y) \in \hat{\Sigma}_1^b$ and t, s be terms such that $\forall x \leq |s| \exists y \leq t \varphi(x, y)$. Define $\theta(x, y) \in \hat{\Sigma}_1^b$ to be $(x > |s| \wedge y = 0) \vee (x \leq |s| \wedge \varphi(x, y))$. Then $\forall x \exists y \leq t \theta(x, y)$ and, by Lemma 3, there is $\psi(x, y) \in \hat{\Sigma}_1^b$ such that (1) $\forall x \exists! y \leq t \psi(x, y)$, and (2) $\forall x \forall y (\psi(x, y) \rightarrow \theta(x, y))$. By (1) and Lemma 2, it holds that

$$\exists w < bd(t^*(|s|), s) \forall x \leq |s| (\beta_{t^*(|s|)}(w, x) \leq t \wedge \psi(x, \beta_{t^*(|s|)}(w, x)))$$

Hence, $\exists w < bd(t^*(|s|), s) \forall x \leq |s| (\beta_{t^*(|s|)}(w, x) \leq t \wedge \varphi(x, \beta_{t^*(|s|)}(w, x)))$, since, by (2) and the definition of θ , we have $x \leq |s| \wedge \psi(x, y) \rightarrow \varphi(x, y)$. \square

Corollary 2. *(Johannsen–Pollett) C_2^0 is $\forall\Sigma_1^b$ -conservative over Δ_1^b -CR.*

References

1. Avigad, J.: Saturated models of universal theories. *Annals of Pure and Applied Logic* 118, 219–234 (2002)
2. Bloch, S.: *Divide and Conquer in Parallel Complexity and Proof Theory*, Ph. D. Thesis. University of California, San Diego (1992)
3. Buss, S.: *Bounded Arithmetic*. Bibliopolis, Napoli (1986)
4. Clote, P., Takeuti, G.: First order bounded arithmetic and small boolean circuit complexity classes. In: Clote, P., Rémel, J (eds.) *Feasible Mathematics II*, pp. 154–218. Birkhäuser, Boston (1995)
5. Johannsen, J.: A Bounded Arithmetic Theory for Constant Depth Threshold Circuits. In: Gödel'96. *Lecture Notes in Logic*, vol. 6, pp. 224–234. Springer, Heidelberg (1996)
6. Johannsen, J., Pollett, C.: On Proofs About Threshold Circuits and Counting Hierarchies. In: *Proc. 13th IEEE Symposium on Logic in Computer Science* (1998)
7. Johannsen, J., Pollett, C.: On the Δ_1^b -Bit-Comprehension Rule. *Logic Colloquium'98. Lecture Notes in Logic*, ASL 13, 262–279 (2000)
8. Krajíček, J.: *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Cambridge University Press, Cambridge (1995)
9. Nguyen, P., Cook, S.: Theories for TC^0 and other small complexity classes. *Logical Methods in Computer Science* 2, 1–40 (2006)
10. Pollett, C.: Structure and Definability in General Bounded Arithmetic Theories. *Annals of Pure and Applied Logic* 100, 189–245 (1999)
11. Sieg, W.: Herbrand Analyses. *Archive for Mathematical Logic* 30, 409–441 (1991)