Existentially Closed Models and Conservation Results in Bounded Arithmetic

A. CORDÓN-FRANCO, A. FERNÁNDEZ-MARGARIT and
F. F. LARA-MARTÍN, *Facultad de Matemáticas. Universidad de Sevilla,*
C/ Tarfia, s/n, 41012 Sevilla (Spain).
E-mail: {acordon,afmargarit,fflara}@us.es

Abstract

We develop 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. We focus on the classical Buss' theories S_2^i and T_2^i and prove that they are $\forall \Sigma_i^b$ conservative over their inference rule counterparts, and $\exists \forall \Sigma_i^b$ conservative over their parameter-free versions. A similar analysis of the Σ_i^b-replacement scheme is also developed. The proof method is essentially the same for all the schemes we deal with and shows that these conservation results between schemes and inference rules do not depend on the specific combinatorial or arithmetical content of those schemes. We show that similar conservation results can be derived, in a very general setting, for every scheme enjoying some syntactical (or logical) properties common to both the induction and replacement schemes. Hence, previous conservation results for induction and replacement can be also obtained as corollaries of these more general results.

Keywords: Bounded arithmetic, existentially closed models, conservation results, parameter-free schemes.

1 Introduction

Existentially closed models (a.k.a. existentially complete structures) are model-theoretic analogous of existentially closed fields. Since the early work on model-completeness by Abraham Robinson in the 1950s these models have been thoroughly studied for a variety of theories (and/or classes of structures) including fields, rings, groups, modules, etc. The introduction of forcing in Model Theory by A. Robinson in the 1970s gave powerful tools (finite and infinite forcing, generic models, etc.) for the analysis of existentially closed models and a lot of interesting results were obtained with applications mainly in the so-called field of Model Theoretic Algebra.

Existentially closed models of fragments of first order Arithmetic were also studied in the early 1970s (see [6]) but mainly from a purely model-theoretic point of view. More recently, these structures have found some interesting applications in Arithmetic theories: alternative proofs of Gödel's second incompleteness theorem (see [1]), and (under the name of Herbrand saturated models) new proofs of conservation results between different subsystems of first and second order Arithmetic (see [2]). In this article we elaborate on the methods developed by Avigad in [2] and obtain conservation results for theories given by inference rules or parameter-free versions of the classical (and also polynomial or linear) Induction, Replacement and Comprehension principles considered in Bounded Arithmetic.

Bounded Arithmetic theories are formal systems tailored to capture computational complexity classes. The foundational work in this area is [4], where 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 consider model-theoretic methods to obtain both kinds of results for restricted versions of Buss' 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 case of Bounded Arithmetic. To our best knowledge, parameter-free systems have only been systematically studied by Bloch in the second part of his thesis [3]. However, Bloch only considered parameter-free *schemes* indirectly, focusing on conservation for parameter-free *inference rules*. On the other hand, systems described by inference rules over a Hilbert-style proof system in the sense above (which are not equivalent to the corresponding axiom schemes) seldom appear in an explicit manner in the literature. A recent exception is Johannsen and Pollett's work [9], 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.

Both in [3] and in [9] the analysis of those systems has been carried out by means of proof-theoretic methods. Instead, we adopt a model-theoretic approach to the investigation of these restricted systems. To this end, the key ingredient is the notion of an $\exists\hat{\Pi}_i^b$-*closed 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 Theorem 1 and Corollary 2)

- every $\exists\hat{\Pi}_i^b$-closed model of T^R is a model of T,
- every theory extending T^- is closed under the corresponding inference rule and
- every $\exists\hat{\Pi}_i^b$-closed 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 schemes improve previous $\forall\Sigma_i^b$-conservation that immediately follows from the work in [3].

The proof method is essentially the same for all the schemes we deal with and, more interestingly, it suggests that these conservation results between schemes and inference rules have little to do with the specific combinatorial or arithmetical content of the principles involved (induction or replacement). Instead, they rest on the logical/syntactical structure of the considered schemes. In Section 5 we show that similar conservation results can be derived, in a very general setting, for every scheme enjoying

some syntactical (or logical) properties common to both induction and replacement schemes. Hence, the conservation results for induction and replacement obtained in the previous sections can be considered as corollaries of these more general results.

Finally, in Section 6 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 [9] to capture the complexity class TC^0 and is the final refinement of a series of theories introduced in [7–9] in the quest for natural theories for TC^0. In [9] it is proved that TC^0 coincides with the class of 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 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 class of $\hat{\Sigma}_1^b$-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 [9]. Our analysis is of independent interest in view of the problems on Δ_1^b-CR posed in [9]; 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 article (we refer the reader to [4, 11] 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 $\overset{\bullet}{-}$; where $\lfloor\frac{x}{2}\rfloor$ is x divided by two 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 $x\overset{\bullet}{-}y$ is the subtraction function. 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, 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 non-logical symbols of $\mathcal{L}_2$. Following [9, 13], our base theory will be LIOpen $=$ *BASIC* $+\{LIND_\varphi :$ φ is open$\}$. As shown there, LIOpen allows for simple definitions of tuple- and sequence-encoding

functions. First, observe that the following $\mathcal{L}_2$-terms define useful functions in LIOpen:

$$2^{|x|} := 1\#x \qquad\qquad mod2(x) := x \overset{\bullet}{-} 2\cdot\lfloor\tfrac{x}{2}\rfloor$$
$$Bit(x,i) := mod2(MSP(x,i)) \qquad 2^{min(x,|y|)} := MSP(2^{|y|}, |y|\overset{\bullet}{-}x)$$
$$LSP(x,i) := x \overset{\bullet}{-} 2^{min(i,|x|)}\cdot MSP(x,i)$$

In words, $Bit(x,i)$ is the value of the bit in the 2^i position of the binary representation of x, and $LSP(x,|i|)$ returns the number consisting of the low $|i|$ bits of x. The code of a sequence $\{b_0, b_1, \ldots, b_{|s|}\}$ 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 [13] 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, $PIND$, $LIND$, $LLIND$ and let $\mathbf{E}R$ denote the corresponding inference rule. First, the theory $\mathbf{E}\Gamma$ is LIOpen$+\{\mathbf{E}_\varphi : \varphi \in \Gamma\}$. Second, fragment $[T, \Gamma\text{-}\mathbf{E}R]$ is the closure of T under first order logic and *unnested* applications of the $\mathbf{E}$-rule restricted to formulas in Γ, where T is an arbitrary $\mathcal{L}_2$-theory extending LIOpen. We define $[T, \Gamma\text{-}\mathbf{E}R]_k$ for every $k \in \omega$ by recursion on k:

$$[T, \Gamma\text{-}\mathbf{E}R]_0 = T, \quad \text{and} \quad [T, \Gamma\text{-}\mathbf{E}R]_{k+1} = [[T, \Gamma\text{-}\mathbf{E}R]_k, \Gamma\text{-}\mathbf{E}R]$$

Then $T + \Gamma\text{-}\mathbf{E}R = \bigcup_{k\in\omega}[T, \Gamma\text{-}\mathbf{E}R]_k$ denotes the closure of T under first order logic and nested applications of the $\mathbf{E}$-rule restricted to formulas in Γ.

Finally, $\mathbf{E}\Gamma^-$ is LIOpen$+\{\mathbf{E}_\varphi : \varphi(x) \in \Gamma^-\}$, where $\varphi(x) \in \Gamma^-$ means that x is the only free variable occurring in φ. In accordance with this notation, the parameter-free version of the $\mathbf{E}$-rule is denoted by $\Gamma^-\text{-}\mathbf{E}R$.

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' original theories S_2^i and T_2^i, which does not include the MSP and $\overset{\bullet}{-}$ symbols. In addition, Buss' 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_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 and *strict* Π_i^b. Namely, $\hat{\Sigma}_0^b$ or $\hat{\Pi}_0^b$ denotes the sharply bounded formulas; and, for each $i \geq 0$, $\hat{\Sigma}_{i+1}^b$ (resp. $\hat{\Pi}_{i+1}^b$) is the least class containing $\hat{\Pi}_i^b$ (resp. $\hat{\Sigma}_i^b$) and closed under conjunction, disjunction and bounded existential (resp. universal) quantification. Each Σ_i^b (resp. Π_i^b) formula is equivalent in the standard model to a $\hat{\Sigma}_i^b$ (resp. $\hat{\Pi}_i^b$) formula.

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

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 [9, 13] for details).

$BB\Gamma$ is LIOpen$+\{BB_\varphi : \varphi \in \Gamma\}$. Similarly, the inference rule versions $T+\Gamma$-BBR are defined. In [13] 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 Proposition 3.2 in [7], it is easy to show that $BB\hat{\Sigma}_{i+1}^b \equiv BB\hat{\Pi}_i^b$, and $T+\hat{\Sigma}_{i+1}^b$-BBR$\equiv T+\hat{\Pi}_i^b$-BBR.

3 On $\exists\hat{\Pi}_i^b$-closed 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. 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 $\mathfrak{A} \models \varphi(\vec{a}) \Longleftrightarrow \mathfrak{B} \models \varphi(\vec{a})$). The key ingredient for this construction is the notion of an $\exists\hat{\Pi}_i^b$-*closed model* for a theory T.

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

$$\mathfrak{A} \prec_{\hat{\Sigma}_i^b} \mathfrak{B} \quad \Longrightarrow \quad \mathfrak{A} \prec_{\exists\hat{\Pi}_i^b} \mathfrak{B}.$$

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 Avigad's [2] (our work is inspired by the methods in that paper). First of all, observe that $\exists\hat{\Pi}_i^b$-closed 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 $\exists\hat{\Pi}_i^b$-closed for T.

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

PROPOSITION 2
Let $\mathfrak{A}$ be $\exists\hat{\Pi}_i^b$-closed for T, $\vec{a} \in \mathfrak{A}$ and $\varphi(\vec{x}, \vec{v}) \in \hat{\Sigma}_i^b$ and let $\hat{\Pi}_i^b$-$Diag(\mathfrak{A})$ 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 $\hat{\Pi}_i^b$-$Diag(\mathfrak{A})$ satisfying $T+\theta(\vec{a}, \vec{b}) \vdash \forall\vec{x}\varphi(\vec{x}, \vec{a})$.

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

PROOF. (Replacement scheme): Assume that $\mathfrak{A}$ is an $\exists\hat{\Pi}_i^b$-closed model for $T+\hat{\Sigma}_i^b$-BBR and $\mathfrak{A} \models \forall x \leq |s| \exists y \leq t(x)\varphi(x, y, a)$, 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

$$(T + \hat{\Sigma}_i^b\text{-BBR}) + \theta(a, b) \vdash \forall x \leq |s| \exists y \leq t(x)\varphi(x, y, a).$$

So, $T + \hat{\Sigma}_i^b\text{-BBR} \vdash \theta(v, u) \to \forall x \leq |s| \exists y \leq t(x)\varphi(x, y, v)$. 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

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

Since $\mathfrak{A} \models \theta(a, b)$, $\mathfrak{A} \models \delta(x, y, a, b) \to \varphi(x, y, a)$ and hence the consequent of the bounded collection axiom for $\varphi(x, y, a)$ 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 $\exists\hat{\Pi}_i^b$-closed for $T + \hat{\Sigma}_i^b\text{-IR}$. To prove that $\mathfrak{A} \models I\hat{\Sigma}_i^b$, assume $\mathfrak{A} \models \varphi(0, a) \wedge \forall x(\varphi(x, a) \to \varphi(x+1, a))$, 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 re-expressed 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 $\exists\hat{\Pi}_i^b$-closed 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,

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

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 $\mathfrak{A}$. ∎

Combining Proposition 1 and Theorem 1, we 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

$(i \geq 1)$ Let **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 $T + \hat{\Sigma}_i^b\text{-ER} \nvdash \varphi$, 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 $\exists\hat{\Pi}_i^b$-closed 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. ∎

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 [13] stating that $BB\hat{\Sigma}_{i+1}^b$ implies S_2^i.

COROLLARY 1
1. LIOpen$+\hat{\Sigma}^b_{i+1}$-*BBR* implies S^i_2.
2. Both LIOpen$+\hat{\Sigma}^b_{i+1}$-*LINDR* and LIOpen$+\hat{\Sigma}^b_{i+1}$-*IR* imply T^i_2.
3. LIOpen$+\hat{\Sigma}^b_{i+1}$-*LLINDR* implies S^i_2.

PROOF. We only prove (2) for *LIND*, the remaining cases being similar. First, recall that $T^i_2 \equiv I\hat{\Sigma}^b_i$ and $LIND\hat{\Sigma}^b_{i+1} \equiv S^{i+1}_2$ contains $I\hat{\Sigma}^b_i$. It is easy to check that $I\hat{\Sigma}^b_i$ is $\forall\hat{\Sigma}^b_{i+1}$-axiomatizable and hence T^i_2 follows from the $\forall\hat{\Sigma}^b_{i+1}$-consequences of $LIND\hat{\Sigma}^b_{i+1}$. By Theorem 2 for $\mathbf{E}=LIND$ and $T=$LIOpen, $LIND\hat{\Sigma}^b_{i+1}$ is a $\forall\hat{\Sigma}^b_{i+1}$-conservative extension of LIOpen$+\hat{\Sigma}^b_{i+1}$-LINDR; so, the result follows. ∎

To extend previous conservation result to $\forall\Sigma^b_i$-formulas, we need the following lemma

LEMMA 1
($i \geq 1$) Let $\varphi(\vec{v}) \in \Sigma^b_i$. There exists $\hat{\varphi}(\vec{v}) \in \hat{\Sigma}^b_i$ such that:

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

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

PROOF. We proceed by induction on the the complexity of the Σ^b_i-formula $\varphi(\vec{v})$. If $\varphi(\vec{v}) \in \Pi^b_{i-1}$, there exists $\hat{\varphi}(\vec{v}) \in \hat{\Pi}^b_{i-1}$ provably equivalent in $BB\hat{\Pi}_{i-2}$ to $\varphi(\vec{v})$ (for $i=1$ it is trivial; for $i>1$ it follows from Pollett's work in [13]). If $\varphi(\vec{v})$ is obtained by conjunction, disjunction or bounded existential quantification, the induction step is immediate. Now consider the case where $\varphi(\vec{v})$ has the form $\forall x \leq |s(\vec{v})|\varphi_0(x, \vec{v})$. By the induction hypothesis, there is $\hat{\varphi}_0(x, \vec{v}) \in \hat{\Sigma}^b_i$ satisfying the conditions 1 and 2 in the statement of the lemma. Using a pairing function (available in LIOpen), we may assume that $\hat{\varphi}_0$ is $\exists y \leq t(x, \vec{v})\hat{\varphi}_1(x, y, \vec{v})$, where $\hat{\varphi}_1$ is $\hat{\Pi}^b_{i-1}$. The bounded collection axiom for $\hat{\varphi}_1$ says (for notational simplicity we omit the parameters in s, t, φ):

$$\forall x \leq |s|\exists y \leq t(x)\hat{\varphi}_1(x, y) \rightarrow$$
$$\exists w < bd(t^*(|s|), s)\forall x \leq |s|(\beta_{t^*(|s|)}(w, x) \leq t(x) \wedge \hat{\varphi}_1(x, \beta_{t^*(|s|)}(w, x))).$$

Observe that the converse of the implication above is trivially provable in LIOpen. Moreover,

$$BB\hat{\Pi}^b_{i-1} \vdash \varphi_0(x, \vec{v}) \leftrightarrow \exists y \leq t(x)\hat{\varphi}_1(x, y, \vec{v}); \text{ and}$$
$$BB\hat{\Pi}^b_{i-2} \vdash \exists y \leq t(x)\hat{\varphi}_1(x, y, \vec{v}) \rightarrow \varphi_0(x, \vec{v}).$$

Now define $\hat{\varphi}(\vec{v})$ to be the formula

$$\exists w < bd(t^*(|s|), s)\forall x \leq |s|(\beta_{t^*(|s|)}(w, x) \leq t(x) \wedge \hat{\varphi}_1(x, \beta_{t^*(|s|)}(w, x), \vec{v})).$$

Clearly, $\hat{\varphi}(\vec{v})$ is $\hat{\Sigma}^b_i$ and satisfies the two required conditions. ∎

LEMMA 2
Let Q denote a finite sequence of unbounded quantifiers of the form: $\exists\forall\ldots$ or $\forall\exists\ldots$. Suppose that T_1 is $Q\hat{\Sigma}^b_i$-conservative over T_2, T_1 implies $BB\hat{\Pi}^b_{i-1}$; and, for $i>1$, T_2 implies $BB\hat{\Pi}^b_{i-2}$. Then T_1 is $Q\Sigma^b_i$-conservative over T_2.

PROOF. Assume that T_1 proves $Q\vec{v}\varphi(\vec{v})$, where $\varphi(\vec{v})$ is Σ_i^b. By Lemma 1, there is $\hat{\varphi}(\vec{v}) \in \hat{\Sigma}_i^b$ satisfying (1) $T_1 \vdash \varphi(\vec{v}) \leftrightarrow \hat{\varphi}(\vec{v})$, and (2) $T_2 \vdash \hat{\varphi}(\vec{v}) \rightarrow \varphi(\vec{v})$. If follows by (1) that $T_1 \vdash Q\vec{v}\hat{\varphi}(\vec{v})$ and hence $T_2 \vdash Q\vec{v}\hat{\varphi}(\vec{v})$. By (2), $T_2 \vdash Q\vec{v}\varphi(\vec{v})$. ∎

THEOREM 3

($i \geq 1$) Let $\mathbf{E}$ denote one of the following schemes: *BB*, *I*, *PIND*, *LIND*. Then, $\mathrm{LIOpen} + \hat{\Sigma}_i^b\text{-}ER$ axiomatizes the $\forall\Sigma_i^b$-consequences of $\mathbf{E}\Sigma_i^b$.

PROOF. The results follow from Theorem 2, Corollary 1 and Lemma 2 (to apply this lemma for $i > 1$, notice that S_2^{i-1} implies $BB\Pi_{i-2}^b$ by Buss' [4]). ∎

To close this section, observe that by Parikh's theorem we can replace $\forall\Sigma_i^b$-*consequences* by $\forall\exists\Sigma_i^b$-*consequences* in the theorem above.

4 Parameter-free systems of Bounded Arithmetic

In this section we use the results obtained in the previous one to show that S_2^i and T_2^i are $\exists\forall\Sigma_i^b$-conservative over their parameter free versions. 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 reduction of $\hat{\Sigma}_i^b$-IR and $\hat{\Sigma}_i^b$-PINDR to their parameter-free versions:

PROPOSITION 3

Let T be an extension of LIOpen. Then

1. $[T, \hat{\Sigma}_i^b\text{-}IR]$ and $[T, \hat{\Sigma}_i^{b,-}\text{-}IR]$ are equivalent.
2. ($i \geq 1$) $[T, \hat{\Sigma}_i^b\text{-}PINDR]$ and $[T, \hat{\Sigma}_i^{b,-}\text{-}PINDR]$ are equivalent.

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 $[T, \hat{\Sigma}_i^{b,-}\text{-}IR] \vdash \forall v \forall x \varphi(x,v)$. 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 $\hat{\Sigma}_i^{b,-}$-IR. 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 $T \vdash \varphi(x,v) \rightarrow \varphi(x+1,v)$.

So, it follows that $[T, \hat{\Sigma}_i^{b,-}\text{-}IR] \vdash \forall u \theta(u)$. To show $[T, \hat{\Sigma}_i^{b,-}\text{-}IR] \vdash \forall v \forall x \varphi(x,v)$, 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 rule: roughly speaking, if u codifies the pair (x, v) and $x > 0$, then $\lfloor \frac{u}{2} \rfloor$ must codify the pair $(\lfloor \frac{x}{2} \rfloor, v)$. In [3] Bloch proposed the following encoding function satisfying that property:

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

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 meeting some additional restrictions. Namely, by adapting Bloch's work in [3], we can show that LIOpen proves the following facts (we give only a sketch of the proofs, reasoning in a model of LIOpen):

(a) $|x| \le 2z \wedge |v| < z^2 \le |w| \to \exists! u([x, v, z] = u)$.

Indeed, given x, v, z and w, let $u = Concat(v + 2^{min(z^2, |w|)}, x + 2^{|x|})$. Then

$$z^2 \le |u| = z^2 + |x| \le z^2 + 2z < (z+1)^2$$

and, since $z^2 \le |u| \wedge z^2 \le |w| \to 2^{min(z^2, |u|)} = 2^{min(z^2, |w|)}$, we obtain $u = [x, v, z]$ and the uniqueness easily follows from this argument.

(b) $u > 0 \to \exists! x, v, z \le u([x, v, z] = u)$.

Let $\psi(z, u)$ be the open formula $z^2 \le |u|$. Then it holds that $\psi(0, u)$ and $\exists z \neg \psi(|z|, u)$ (since we have $|u + 2|^2 > |u|$); hence by LIOpen it holds that $\exists z(\psi(z, u) \wedge \neg \psi(z+1, u))$. So, it easily follows that $\forall u \exists! z(z^2 \le |u| < (z+1)^2)$.

Given $u > 0$, let z be such that $z^2 \le |u| < (z+1)^2$. Define

$$p = \lfloor \frac{u}{2^{min(z^2, |u|)}} \rfloor, \qquad x = LSP(u, |p|) \quad \text{and} \quad v = MSP(u, |p|) \bullet 2^{|MSP(u, |p|)| \bullet 1}$$

Then $[x, v, z] = u$. As for uniqueness, it is straightforward to check that z is unique and the uniqueness of x and v follows from the previous argument.

(c) $u = [x, v, z] \wedge x > 0 \to \lfloor \frac{u}{2} \rfloor = [\lfloor \frac{x}{2} \rfloor, v, z]$.

Let $x = 2q + r > 0$, where $r \le 1$. Then, $|q| = |x| \bullet 1$ and, since $z^2 < |u|$, it holds that

$$\begin{aligned} [\lfloor \tfrac{x}{2} \rfloor, v, z] &= Concat(v + 2^{min(z^2, |u|)}, q + 2^{|q|}) \\ &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|q + 2^{|q|}| \bullet 1} + (q + 2^{|q|}) \bullet 2^{|q + 2^{|q|}| \bullet 1} \\ &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|q|} + (q + 2^{|q|}) \bullet 2^{|q|} \\ &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|q|} + q \end{aligned}$$

and, as a consequence (recall that $|x| = |q| + 1 = |q + 2^{|q|}| = |x + 2^{|x|}| \bullet 1$)

$$\begin{aligned} 2 \cdot [\lfloor \tfrac{x}{2} \rfloor, v, z] + r &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|x|} + 2q + r \\ &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|x + 2^{|x|}| \bullet 1} + x \\ &= (v + 2^{min(z^2, |u|)}) \cdot 2^{|x + 2^{|x|}| \bullet 1} + (x + 2^{|x|}) \bullet 2^{|x + 2^{|x|}| \bullet 1} \\ &= Concat(v + 2^{min(z^2, |u|)}, x + 2^{|x|}) \\ &= u \end{aligned}$$

Hence, $\lfloor \frac{u}{2} \rfloor = [\lfloor \frac{x}{2} \rfloor, v, z]$.

Equipped with this encoding function, we can replace each application of the PIND rule for a $\hat{\Sigma}_i^b$-formula $\varphi(x,v)$ by one application of the PIND rule for the (parameter free) $\hat{\Sigma}_i^b$-formula $\theta(u) \equiv u > 0 \to \exists x, v, z \leq u([x,v,z] = u \wedge \varphi(x,v))$. ■

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

Observe that from Corollaries 1 and 2 it immediately follows that

COROLLARY 3
$T_2^{i+1,-}$ implies T_2^i, and $S_2^{i+1,-}$ implies S_2^i.

THEOREM 4
$(i \geq 1)$

1. T_2^i is $\exists\forall\Sigma_i^b$-conservative over $T_2^{i,-}$.
2. S_2^i is $\exists\forall\Sigma_i^b$-conservative over $S_2^{i,-}$.

PROOF. By Lemma 2 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 $\exists\hat{\Pi}_i^b$-closed model for T. By Corollary 2, 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. ■

REMARK 1
These results differ in several aspects from the ones obtained by Bloch in [3]. First of all, the proof-theoretic methods of Bloch deal with *general* Σ_i^b formulas while our model-theoretic methods are well suited to deal with *strict* Σ_i^b-formulas.

Secondly, Bloch focuses on *strongly* parameter free systems which in fact are given by parameter-free *inference rules*. For each Bounded Arithmetic theory, T, defined by an induction scheme (for instance T_2^i or S_2^i), he defines a parameter-free counterpart pfT by replacing the corresponding scheme by nested applications of a parameter-free rule. The considered rule is apparently stronger than the rule we have dealt with in this article. Namely, for $T = T_2^i$ Bloch's rule can be formulated in a Hilbert-style as follows:

$$\Sigma_i^{b,-}\text{-IR}_0: \quad \frac{\forall x(\varphi(x) \to \varphi(x+1))}{\varphi(0) \to \forall x \varphi(x)}$$

The resulting system pfT_2^i is apparently incomparable w.r.t. inclusion with our theory LIOpen$+$ $\hat{\Sigma}_i^{b,-}$-IR. However, using Proposition 3 it can be easily shown that the rules $\hat{\Sigma}_i^{b,-}$-IR$_0$, $\hat{\Sigma}_i^{b,-}$-IR, $\hat{\Sigma}_i^b$-IR and $\hat{\Sigma}_i^b$-IR$_0$ are equivalent over LIOpen and, by Theorem 3, axiomatize the $\forall\exists\Sigma_i^b$-consequences of T_2^i. Consequently, parameter-free $\hat{\Sigma}_i^b$ induction *rule* axiomatizes the $\forall\exists\Sigma_i^b$-consequences of T_2^i, whereas parameter-free $\hat{\Sigma}_i^b$ induction *scheme* axiomatizes its $\exists\forall\Sigma_i^b$-consequences. Similar remarks apply to *PIND* and S_2^i.

5 A general framework for conservativity between schemes and rules

In this section we prove that Theorems 1, 2 and 4 can be obtained as corollaries of very general conservation results between certain schemes and inference rules associated with them in a

natural way. The proof will make clear that these conservation results rest on the logical/syntactical structure of the schemes we deal with and that they can be derived with no use of the specific combinatorial or arithmetical content of the schemes.

Let us fix a finite first order language L. We are interested in conservation results between theories axiomatized by axiom schemes and theories described by inference rules associated with those schemes. The major axiom schemes in Bounded Arithmetic are the different forms of induction and replacement; so, in this section we shall deal with an arbitrary scheme enjoying some basic properties common to both principles. First of all we state the following general definition.

DEFINITION 2

A k-ary scheme $\mathbf{E}$ is a sentence of the language $L_P = L \cup \{P\}$, where P is a new k-ary predicate symbol. For each formula $\varphi(x_1, \ldots, x_k, \vec{v})$ of L with $\vec{v}$ not occurring in $\mathbf{E}$, we denote by $\mathbf{E}_{\varphi, x_1, \ldots, x_k}$ the formula obtained by substituting $\varphi(t_1, \ldots, t_k, \vec{v})$ for each atomic subformula of $\mathbf{E}$ of the form $P(t_1, \ldots, t_k)$, where $t_1, \ldots, t_k$ are terms. Variables $x_1, \ldots, x_k$ are called the *proper variables* of $\mathbf{E}_{\varphi, \vec{x}}$ and $\vec{v}$ are called *parameters*.

We are interested in schemes of the form $\mathbf{A} \to \mathbf{B}$. This is the common form of the usual formulations of induction and replacement principles and it allows us to attach an inference rule to each scheme $\mathbf{E}$ in a natural way. Observe that although $\mathbf{E}$ is a sentence of L_P, the formula $\mathbf{E}_{\varphi, \vec{x}}$ is not necessarily a sentence of L, since the possible parameters $\vec{v}$ of $\varphi(\vec{x}, \vec{v})$ are free variables of $\mathbf{E}_{\varphi, \vec{x}}$. So, if $\mathbf{E}$ is $\mathbf{A} \to \mathbf{B}$, then $\mathbf{E}_{\varphi, \vec{x}}(\vec{v})$ will have the form $\mathbf{A}_{\varphi, \vec{x}}(\vec{v}) \to \mathbf{B}_{\varphi, \vec{x}}(\vec{v})$. Thus, the inference rule associated with the scheme $\mathbf{E}$ is the rule $\mathbf{E}R$ defined by

$$\forall \vec{v} \mathbf{A}_{\varphi, \vec{x}}(\vec{v})$$
$$\forall \vec{v} \mathbf{B}_{\varphi, \vec{x}}(\vec{v})$$

Given a class of formulas of L, Γ and a theory T of language L, $[T, \Gamma\text{-}\mathbf{ER}]$ (resp. $T + \Gamma\text{-}\mathbf{ER}$) denotes the closure of T under first order logic and unnested (resp. nested) applications of $\mathbf{E}R$ restricted to formulas in Γ. If only parameter-free formulas are allowed then we obtain the theories $[T, \Gamma^-\text{-}\mathbf{ER}]$ and $T + \Gamma^-\text{-}\mathbf{ER}$.

In addition, for each scheme $\mathbf{E}$ of the form $\mathbf{A} \to \mathbf{B}$ we define three theories obtained by adding to T certain axiom scheme:

> $T + \mathbf{E}\Gamma$ is the theory axiomatized by T plus the universal closure of the formula $\mathbf{E}_{\varphi, \vec{x}}$ for every $\varphi \in \Gamma$.
>
> $T + \mathbf{UE}\Gamma$ is the theory axiomatized by T plus the sentence $\forall \vec{v} \mathbf{A}_{\varphi, \vec{x}}(\vec{v}) \to \forall \vec{v} \mathbf{B}_{\varphi, \vec{x}}(\vec{v})$ for every $\varphi \in \Gamma$.
>
> $T + \mathbf{E}\Gamma^-$ is the theory axiomatized by T plus the sentence $\mathbf{E}_{\varphi, \vec{x}}$ for every $\varphi(\vec{x}) \in \Gamma^-$.

In what follows we consider a fixed class of L-formulas Π containing all atomic formulas and closed under conjunction, disjunction, term substitution and subformulas. We denote by $\neg\Pi$ the set of formulas $\{\neg\varphi : \varphi \in \Pi\}$ and assume $\neg\Pi \subseteq \exists\Pi$.

Now we isolate the syntactic properties of induction and replacement schemes that we have used to derive Theorem 2.

DEFINITION 3

Let T be a theory and Γ a class of L-formulas. A scheme, $\mathbf{E} = \mathbf{A} \to \mathbf{B}$, is *T-monotonic* over Π and Γ if, for each formula $\theta(\vec{w}) \in \Pi$ and $\varphi(\vec{x}, \vec{v}) \in \Gamma$ such that $\vec{w}$ contains no variable in $\vec{x}$, it holds that:

1. Syntactical Conditions:
   (S1) $\theta(\vec{w}) \to \varphi(\vec{x}, \vec{v}) \in \Gamma$,

(S2) $\mathbf{A}_{\varphi,\vec{x}} \in \forall\neg\Pi$ and

(S3) $T + \Gamma\text{-}\mathbf{ER}$ is $\forall\exists\Pi$-axiomatizable.

2. Provability Conditions:

(P1) $T \vdash (\theta \to \mathbf{A}_{\varphi,\vec{x}}) \to \mathbf{A}_{\theta\to\varphi,\vec{x}}$, and

(P2) $T \vdash \mathbf{B}_{\theta\to\varphi,\vec{x}} \to (\theta \to \mathbf{B}_{\varphi,\vec{x}})$.

The main examples we have in mind are given by $\Pi = \hat{\Pi}_i^b$, $\Gamma = \hat{\Sigma}_i^b$ and one scheme $\mathbf{E} = $ I, PIND, LIND, LLIND or BB. It is easy to check that for every $\forall\exists\hat{\Pi}_i^b$-axiomatizable extension of *BASIC*, T, every induction scheme $\mathbf{E}$ is T-monotonic over $\hat{\Pi}_i^b$ and $\hat{\Sigma}_i^b$. In turn, for $i \geq 1$, *BB* is T-monotonic over $\hat{\Pi}_i^b$ and $\hat{\Sigma}_i^b$, since $T + \hat{\Sigma}_i^b\text{-BBR} \equiv T + \hat{\Pi}_{i-1}^b\text{-BBR}$.

Our emphasis on *strict* Σ_i^b-formulas here and in previous sections is motivated by the Syntactical Conditions. We restrict ourselves to $\hat{\Sigma}_i^b$-formulas to fulfil these conditions for induction schemes. As for general Σ_i^b-formulas we have that every induction scheme is *BASIC*-monotonic over $\mathcal{B}(\Sigma_i^b)$ and $\mathcal{B}(\Sigma_i^b)$, where $\mathcal{B}(\Sigma_i^b)$ denotes the class of boolean combinations of Σ_i^b-formulas.

Next definition introduces a straightforward generalization of the notion of an existentially closed model.

DEFINITION 4

Let $\mathfrak{A}$ be an L-structure. We say that $\mathfrak{A}$ is $\exists\Pi$-closed for T if $\mathfrak{A} \models T$ and for each $\mathfrak{B} \models T$,

$$\mathfrak{A} \prec_\Pi \mathfrak{B} \implies \mathfrak{A} \prec_{\exists\Pi} \mathfrak{B}.$$

The usual chain argument for constructing existentially closed models gives us an existence lemma. To this end, recall that Π contains all atomic formulas and it is closed under subformulas and term substitution; so, it can be showed by a typical induction argument that the union of a chain of Π-elementary structures is a Π-elementary extension of every structure in the chain.

LEMMA 3

Let T be a $\forall\exists\Pi$-axiomatizable consistent arithmetic theory. Then for each $\mathfrak{A} \models T$, there exists an $\exists\Pi$-closed model for T, $\mathfrak{B}$, such that $\mathfrak{A} \prec_\Pi \mathfrak{B}$.

Our basic device to prove conservation is next lemma, which is a general version of theorem 3.4 in [2].

LEMMA 4

Let T be a $\forall\exists\Pi$-axiomatizable theory and let T' be a theory such that every $\exists\Pi$-closed model for T is a model of T'. Then T' is $\forall\neg\Pi$-conservative over T.

PROOF. Let $\varphi \in \forall\neg\Pi$ be a sentence such that $T' \vdash \varphi$. If $T \not\vdash \varphi$ then there exists $\mathfrak{A} \models T + \neg\varphi$. Since $T + \neg\varphi$ is $\forall\exists\Pi$-axiomatizable, there exists an $\exists\Pi$-closed model for T, $\mathfrak{B}$, such that $\mathfrak{A} \prec_\Pi \mathfrak{B}$. Then $\mathfrak{B} \models T' + \neg\varphi$, a contradiction. ∎

The following result is a general version of Proposition 2 relating validity in an $\exists\Pi$-closed model for T and provability in the theory T.

LEMMA 5

Let $\mathfrak{A} \models T$ be an $\exists\Pi$-closed model for T and $\varphi(\vec{x}) \in \forall\neg\Pi$, $\vec{a} \in \mathfrak{A}$ such that $\mathfrak{A} \models \varphi(\vec{a})$. Then there exist $\vec{c} \in \mathfrak{A}$ and $\theta(\vec{x},\vec{z}) \in \Pi$ such that

$$\mathfrak{A} \models \theta(\vec{a},\vec{c}) \quad \text{and} \quad T \vdash \theta(\vec{x},\vec{z}) \to \varphi(\vec{x}).$$

PROOF. Let $D_\Pi(\mathfrak{A})$ denote the Π-diagram of $\mathfrak{A}$. Since $\mathfrak{A} \models \varphi(\vec{a})$ and $\mathfrak{A}$ is $\exists\Pi$-closed, $T + D_\Pi(\mathfrak{A}) + \neg\varphi(\vec{a})$ is inconsistent. Therefore, $T + D_\Pi(\mathfrak{A}) \vdash \varphi(\vec{a})$ and, since Π is closed under conjunctions, there exist $\vec{c} \in \mathfrak{A}$ and $\theta(\vec{x}, \vec{z}) \in \Pi$ such that $\theta(\vec{a}, \vec{c}) \in D_\Pi(\mathfrak{A})$ and $T \vdash \theta(\vec{x}, \vec{z}) \to \varphi(\vec{x})$. ∎

Now we are ready to prove our general conservation theorems.

THEOREM 5
Let T be a $\forall\exists\Pi$-axiomatizable theory and let $\mathbf{E} = \mathbf{A} \to \mathbf{B}$ be a T-monotonic scheme over Π and Γ. Then $T + \mathbf{E}\Gamma$ is $\forall\neg\Pi$-conservative over $T + \Gamma\text{-}\mathbf{ER}$.

PROOF. By condition (S3), $T + \Gamma\text{-}\mathbf{ER}$ is $\forall\exists\Pi$-axiomatizable; so, by Lemma 4, it is enough to prove that every $\exists\Pi$-closed model for $T + \Gamma\text{-}\mathbf{ER}$ is a model of $T + \mathbf{E}\Gamma$.

Let $\varphi(\vec{x}, \vec{v}) \in \Gamma$ and $\vec{a} \in \mathfrak{A}$ such that $\mathfrak{A} \models \mathbf{A}_{\varphi, \vec{x}}(\vec{a})$. By condition (S2), $\mathbf{A}_{\varphi, \vec{x}} \in \forall\neg\Pi$ and, therefore, by Lemma 5, there exist $\theta(\vec{v}, \vec{z}) \in \Pi$ and $\vec{c} \in \mathfrak{A}$ such that $\mathfrak{A} \models \theta(\vec{a}, \vec{c})$ and

$$T + \Gamma\text{-}\mathbf{ER} \vdash \theta(\vec{v}, \vec{z}) \to \mathbf{A}_{\varphi, \vec{x}}(\vec{v})$$

By condition (P1), it holds that

$$T + \Gamma\text{-}\mathbf{ER} \vdash \mathbf{A}_{\theta \to \varphi, \vec{x}}$$

Now, by condition (S1), $\theta(\vec{v}, \vec{z}) \to \varphi(\vec{x}, \vec{z}) \in \Gamma$; so, applying $\Gamma\text{-}\mathbf{ER}$, we get $T + \Gamma\text{-}\mathbf{ER} \vdash \mathbf{B}_{\theta \to \varphi, \vec{x}}$. Finally, by condition (P2), it follows that

$$T + \Gamma\text{-}\mathbf{ER} \vdash \theta(\vec{v}, \vec{z}) \to \mathbf{B}_{\varphi, \vec{x}}(\vec{v})$$

Therefore, $\mathfrak{A} \models \mathbf{B}_{\varphi, \vec{x}}(\vec{a})$, since $\mathfrak{A} \models T + \Gamma\text{-}\mathbf{ER}$ and $\mathfrak{A} \models \theta(\vec{a}, \vec{c})$. ∎

Now, Theorem 2 can be derived as a corollary of Theorem 5, since for every $\forall\exists\hat{\Pi}_i^b$-axiomatizable theory, T, the schemes $I, PIND, LIND, LLIND, BB$ are T-monotonic over $\hat{\Pi}_i^b$ and $\hat{\Sigma}_i^b$, $(i \geq 1)$.

THEOREM 6
Let T and $\mathbf{E}$ be as in previous theorem. Then

1. If $\mathbf{UE}\Gamma$ is $\forall\exists\Pi$-axiomatizable then $T + \mathbf{E}\Gamma$ is $\exists\forall\neg\Pi$-conservative over $T + U\mathbf{E}\Gamma$.
2. If $\mathbf{UE}\Gamma$ is $\exists\forall\neg\Pi$-axiomatizable and every $\forall\exists\Pi$-axiomatizable extension of $T + \mathbf{E}\Gamma^-$ is closed under $\Gamma\text{-}\mathbf{ER}$, then

$$T + \mathbf{E}\Gamma^- \equiv T + \mathbf{UE}\Gamma.$$

PROOF. Let us assume that $\mathbf{UE}\Gamma$ is $\forall\exists\Pi$-axiomatizable. Then part (1) follows from Theorem 5 applied to $T' = T + \mathbf{UE}\Gamma$ and $\mathbf{E}$, since $T' + \mathbf{E}\Gamma$ is equivalent to $T + \mathbf{E}\Gamma$ and T' is, obviously, closed under $\Gamma\text{-}\mathbf{ER}$.

On the other hand, it is obvious that $T + \mathbf{UE}\Gamma$ extends $T + \mathbf{E}\Gamma^-$; so, in order to get part (2) it suffices to prove the reverse implication. Let us assume that $\mathbf{UE}\Gamma$ is $\exists\forall\neg\Pi$-axiomatizable and let $\psi \in \exists\forall\neg\Pi$ be an axiom of $\mathbf{UE}\Gamma$. We reason by contradiction. If $T + \mathbf{E}\Gamma^- \not\vdash \psi$ then $T'' = (T + \mathbf{E}\Gamma^-) + \neg\psi$ is a $\forall\exists\Pi$-axiomatizable consistent extension of $T + \mathbf{E}\Gamma^-$. By hypothesis, T'' is closed under $\Gamma\text{-}\mathbf{ER}$, thus by Theorem 5, $T'' + \mathbf{E}\Gamma$ is $\forall\exists\Pi$-conservative over T''. In particular, $T'' + \mathbf{E}\Gamma$ is consistent, and this provides the required contradiction, since $T'' + \mathbf{E}\Gamma$ extends $\mathbf{UE}\Gamma$ and, thus, $T'' + \mathbf{E}\Gamma$ proves both ψ and $\neg\psi$. ∎

Let us illustrate the content of Theorem 6 by considering the case of $\hat{\Sigma}_i^b$-replacement scheme that we have omitted in Section 4 on parameter-free schemes. Recall that for $i \geq 1$, BB is $BASIC$-monotonic over $\hat{\Pi}_i^b$ and $\hat{\Sigma}_i^b$. Therefore, since $UBB\hat{\Sigma}_i^b \equiv UBB\hat{\Pi}_{i-1}^b$ and this theory is $\forall\exists\hat{\Pi}_i^b$-axiomatizable, by part 1. of Theorem 6 and Lemma 2, it follows that:

COROLLARY 4

$(i \geq 1)$ $BB\hat{\Sigma}_i^b$ is $\exists\forall\Sigma_i^b$-conservative over $UBB\hat{\Sigma}_i^b$.

Note that the scheme UBB_φ is obtained quantifying universally the parameters of $\varphi(x,y)$ in both the antecedent and the consequent of BB_φ. The corresponding parameter-free system $BB\hat{\Sigma}_i^{b,-}$, is not known to be equivalent to $UBB\hat{\Sigma}_i^b$.

For $\hat{\Sigma}_i^b$-induction the situation is different. First of all, note that the corresponding schemes $UI\hat{\Sigma}_i^b$ and $UPIND\hat{\Sigma}_i^b$ are $\exists\forall\hat{\Sigma}_i^b$-axiomatizable. Secondly, by Proposition 3, every extension of $T_2^{i,-}$ (resp. $S_2^{i,-}$) is closed under $\hat{\Sigma}_i^b$-IR (resp. $\hat{\Sigma}_i^b$-PINDR). Hence, by the second part of Theorem 6, it holds that

COROLLARY 5

$(i \geq 1)$ $UI\hat{\Sigma}_i^b$ and $UPIND\hat{\Sigma}_i^b$ are equivalent to $T_2^{i,-}$ and $S_2^{i,-}$, respectively.

6   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 [9] as follows:

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

where $\varphi(x) \in \Sigma_1^b$ and $\psi \in \Pi_1^b$. We shall consider the (apparently) weaker rule for *strict* formulas $\hat{\Delta}_1^b$-BCR and show that, over LIOpen, it is equivalent to $\hat{\Sigma}_1^b$-replacement rule. In fact, the four rules Σ_1^b-BBR, $\hat{\Sigma}_1^b$-BBR, Δ_1^b-BCR and $\hat{\Delta}_1^b$-BCR are equivalent over LIOpen and therefore, by Theorem 3, axiomatize the class of $\forall\Sigma_1^b$-consequences of $BB\Sigma_1^b$.

In [7], Σ_1^b-replacement and Δ_1^b-bit-comprehension schemes were considered in connection with the class of functions computable by uniform threshold circuit families of polynomial size and constant depth, TC^0. The main result of [7] states that TC^0 is the class of functions Σ_1^b-definable in the system $\bar{R}_2^0$. This theory was proposed as a fairly natural alternative to the theory TTC^0 introduced in [5] in order to characterize the class TC^0. A more elegant axiomatization of $\bar{R}_2^0$ is presented in [8] where it is proved that $\bar{R}_2^0$ is equivalent to $BB\Sigma_0^b$, denoted there by C_2^0. Finally, in [9], it is proved that C_2^0 is a $\forall\Sigma_1^b$-conservative extension of LIOpen$+\Delta_1^b$-BCR (this theory is denoted in [9] by Δ_1^b-CR). So the class of Σ_1^b-definable functions of this last theory is TC^0 and the authors argue that it is the weakest natural theory with this property. Here, we prove that TC^0 coincides with the class of $\hat{\Sigma}_1^b$-definable functions of LIOpen$+\hat{\Delta}_1^b$-BCR and show how $\hat{\Sigma}_1^b$-replacement rule can be used together with this fact to prove the conservation result between C_2^0 and Δ_1^b-CR obtained by Johannsen and Pollett. Our analysis is of independent interest; nevertheless, it also supports Johannsen–Pollett's claim on Δ_1^b-BCR 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. As a by-product our approach also provides a strict version of the KPT-witnessing result for Δ_1^b-BR obtained in [9].

Now we shall prove that LIOpen+$\hat{\Sigma}_1^b$-BBR and LIOpen+$\hat{\Delta}_1^b$-BCR are equivalent, see Theorem 8. In what follows LIOpen+Δ_1^b-BCR and LIOpen+$\hat{\Delta}_1^b$-BCR will be denoted by Δ_1^b-CR and $\hat{\Delta}_1^b$-CR, respectively. The proof will occupy the most part of this section and it is strongly related to our proof of Johannsen–Pollett's conservation theorem. First of all let us observe the following fact:

LEMMA 6

Let T be an extension of LIOpen. Then $[T, \hat{\Sigma}_1^b$-BBR$]$ extends $[T, \hat{\Delta}_1^b$-BCR$]$.

PROOF. Let us assume that $T \vdash \forall x(\varphi(x) \leftrightarrow \psi(x))$, where $\varphi(x) \in \hat{\Sigma}_1^b$, $\psi(x) \in \hat{\Pi}_1^b$. Let $\theta(x, y)$ be the $\hat{\Sigma}_1^b$ formula $(\varphi(x) \wedge y = 1) \vee (\neg \psi(x) \wedge y = 0)$. Then $T \vdash \forall x \le |u| \exists y \le 1 \, \theta(x, y)$ and, by $\hat{\Sigma}_1^b$-BBR,

$$[T, \hat{\Sigma}_1^b\text{-BBR}] \vdash \exists w < bd(t^*, u) \forall x \le |u| (\beta_{t^*}(w, x) \le t \wedge \theta(x, \beta_{t^*}(w, x)).$$

Let w be such that $\forall x \le |u| (\beta_{t^*}(w, x) \le t \wedge \theta(x, \beta_{t^*}(w, x)))$ and let $v = LSP(w, |u|)$. Then, $[T, \hat{\Sigma}_1^b\text{-BBR}] \vdash \forall x < |u| (\varphi(x) \leftrightarrow Bit(v, x) = 1)$, as required. ∎

Next three lemmas will allow us to obtain a partial converse of Lemma 6, revealing a close relationship between both rules.

LEMMA 7

Let T be an extension of LIOpen. Then

1. $[T, \hat{\Delta}_1^b$-BCR$]$ extends $[T, \hat{\Delta}_1^b$-LINDR$]$.
2. (Extensionality) LINDΣ_0^b (hence, also $[T, \hat{\Delta}_1^b$-BCR$]$) proves

$$(|v| = |w| \wedge \forall x < |v| Bit(v, x) = Bit(w, x)) \to v = w.$$

PROOF. The proof is straightforward, see lemma 5 in [9]. ∎

Next lemma provides a weak form of replacement available in $\hat{\Delta}_1^b$-CR.

LEMMA 8

Let T denote an extension of LIOpen. Let $\varphi(x, y) \in \hat{\Sigma}_1^b$ and let t and s be terms such that $T \vdash \forall x \le |s| \exists! y \le t(x) \varphi(x, y)$. Then

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

PROOF. First of all let us observe that

(•) LIOpen$\vdash |a| \le |s| \cdot |t^*(|s|)| \to \forall i \le |a| \exists! q \le |s| \exists! r < |t^*(|s|)|(i = q|t^*(|s|)| + r)$

Fix $a = bd(t^*(|s|), s)$ and $\theta(i) \equiv i = |a| - 1 \vee (i < |a| - 1 \wedge \delta(i))$, where $\delta(i)$ is

$$\exists r < |t^*(|s|)| \exists q \le |s| \exists y \le t(i)(i = q|t^*(|s|)| + r \wedge \varphi(q, y) \wedge Bit(y, r) = 1)$$

Then $\theta(i)$ is $\hat{\Delta}_1^b$ in T, since, by (•) and the functional character of φ, it is equivalent in T to the $\hat{\Pi}_1^b$-formula: $i = |a| - 1 \vee (i < |a| - 1 \wedge \sigma(i))$, where $\sigma(i)$ is

$$\forall r < |t^*(|s|)| \forall q \le |s| \forall y \le t(i)(i = q|t^*(|s|)| + r \wedge \varphi(q, y) \to Bit(y, r) = 1)$$

By $\hat{\Delta}_1^b$-BCR, $\exists w < 2^{|a|} \forall i < |a| (Bit(w, i) = 1 \leftrightarrow \theta(i))$. Then, by construction, it holds that $\forall x \le |s| (\beta_{t^*(|s|)}(w, x) \le t \wedge \varphi(x, \beta_{t^*(|s|)}(w, x)))$, as required. ∎

DEFINITION 5
We say that a theory T has $\hat{\Sigma}_1^b$-selection if for every formula $\varphi(x,y) \in \hat{\Sigma}_1^b$ such that $T \vdash \forall x \exists y \leq t(x)\varphi(x,y)$, there exists $\psi(x,y) \in \hat{\Sigma}_1^b$ such that T proves

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

LEMMA 9
Let T be an extension of LIOpen such that T has $\hat{\Sigma}_1^b$-selection. Then the theories $[T, \hat{\Delta}_1^b\text{-BCR}]$ and $[T, \hat{\Sigma}_1^b\text{-BBR}]$ are equivalent.

PROOF. By Lemma 6 it is enough to prove that $[T, \hat{\Delta}_1^b\text{-BCR}]$ extends $[T, \hat{\Sigma}_1^b\text{-BBR}]$.

Let $\varphi(x,y) \in \hat{\Sigma}_1^b$ and t,s be terms such that $T \vdash \forall x \leq |s| \exists y \leq t(x)\varphi(x,y)$. Define $\theta(x,y) \in \hat{\Sigma}_1^b$ to be the formula

$$(x > |s| \wedge y = 0) \vee (x \leq |s| \wedge \varphi(x,y)).$$

Then $T \vdash \forall x \exists y \leq t(x)\theta(x,y)$ and, since T has $\hat{\Sigma}_1^b$-selection, there is $\psi(x,y) \in \hat{\Sigma}_1^b$ such that T proves

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

By (1) and Lemma 8, it holds that

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

Hence, $[T, \hat{\Delta}_1^b\text{-BCR}]$ also proves

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

since, by (2) and the definition of $\theta(x,y)$, we have $T \vdash x \leq |s| \wedge \psi(x,y) \rightarrow \varphi(x,y)$. ∎

Since C_2^0 coincides with $BB\Sigma_0^b$, by Theorem 2, C_2^0 is $\forall \hat{\Sigma}_1^b$ conservative over LIOpen $+ \hat{\Sigma}_1^b$-BBR. So, in order to get the equivalence between LIOpen $+ \hat{\Sigma}_1^b$-BBR and $\hat{\Delta}_1^b$-CR (and, as a consequence, Johannsen–Pollett's theorem) it suffices to prove that $\hat{\Delta}_1^b$-CR is closed under $\hat{\Sigma}_1^b$-BBR. By the preceeding lemmas in order to prove that fact it is enough to show that $\hat{\Delta}_1^b$-CR has $\hat{\Sigma}_1^b$-selection. Hence, we need a $\hat{\Sigma}_1^b$ witnessing theorem for this last theory. The proof we present here will require an analysis of the class of $\hat{\Sigma}_1^b$-definable functions. A key ingredient in such analysis is a machine-independent characterization of the complexity class TC^0 given by Clote and Takeuti in [5]:

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

DEFINITION 6
Let $g : \omega^n \rightarrow \omega$ and $h_0, h_1 : \omega^{n+1} \rightarrow \omega$ such that $h_0(n,\vec{x}), h_1(n,\vec{x}) \leq 1$. A function f is defined by *concatenation recursion on notation* (CRN) from g, h_0 and h_1 if

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

THEOREM 7

(Clote–Takeuti) TC^0 is the smallest class of functions containing BF and closed under composition and CRN.

PROPOSITION 4

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

1. $\forall \vec{x} \exists! y \le t \exists z_1 \le t_1 \ldots \exists z_n \le t_n \psi(x,y,\vec{z})$
2. $\forall \vec{x}, y (\exists \vec{z} \le \vec{t} \psi(x,y,\vec{z}) \to \exists! \vec{z} \le \vec{t} \psi(x,y,\vec{z}))$

PROOF. The proof proceeds by induction. The result obviously holds for the basic functions and it is straightforward to check that it holds for f defined by composition from functions verifying conditions 1 and 2. So, it suffices to prove the result for functions defined by CRN. We adapt the proof of theorem 4 in [9].

Let us assume that f is defined by CRN from $g(\vec{x})$, $h_0(u,x)$ and $h_1(u,x)$ and that g, h_0 and h_1 verify the claim. Let us define

$$h(a,x) = \sum_{i=0}^{|a|} \mathrm{cond}(Bit(a, |a| \overset{\bullet}{-} i), h_0(i,x), h_1(i,x)) \cdot 2^i$$

where $\mathrm{cond}(x,y,z) = y$, if $x = 0$, and z otherwise. Then $f(a,x) = g(x) \cdot 2^{|h(a,x)|} + h(a,x)$. So it is enough to prove that h can be defined by a $\hat{\Sigma}_1^b$ formula verifying the two conditions of the proposition.

Let us consider $\theta(u,x,y,a,z_1,\ldots,z_m) \in \Sigma_0^b$ and $t_1,\ldots,t_m$ terms such that the $\hat{\Sigma}_1^b$-formula $\exists z_1 \le t_1 \ldots \exists z_m \le t_m \theta(u,x,y,a,z_1,\ldots,z_m)$ defines in the standard model the function $k(u,x,a) = \mathrm{cond}(Bit(a, |a| \overset{\bullet}{-} i), h_0(i,x), h_1(i,x))$ and $\hat{\Delta}_1^b$-CR proves

$$\forall u,x,a \exists! y \le 1 \exists \vec{z} \le \vec{t} \theta(u,x,y,a,\vec{z}) \quad \wedge$$
$$\forall u,x,a,y (\exists \vec{z} \le \vec{t} \theta(u,x,y,a,\vec{z}) \to \exists! \vec{z} \le \vec{t} \theta(u,x,y,a,\vec{z}))$$

Using a pairing function and the monotonic terms t_j^* we can assume that $m = 1$. Let us work in a model of $\hat{\Delta}_1^b$-CR. Then, we have $\exists z_1 \le t_1 \theta(u,x,1,a,z_1) \in \hat{\Delta}_1^b$ and, by $\hat{\Delta}_1^b$-BCR,

$$\exists w \le 2^{|a|} \forall u \le |a| (Bit(w,u) = 1 \leftrightarrow \exists z_1 \le t_1 \theta(u,x,1,a,z_1))$$

(and w is unique by Extensionality). Let $\psi(u,x,a,z_1,w) \in \Sigma_0^b$ be the formula

$$(Bit(w,u) = 1 \wedge \theta(u,x,1,a,z_1)) \vee (Bit(w,u) = 0 \wedge \theta(u,x,0,a,z_1)).$$

Since z_1 is unique, the formula $\forall u \le |a| (Bit(w,u) = 1 \leftrightarrow \exists z_1 \le t_1 \theta(u,x,1,a,z_1))$ is equivalent to $\forall u \le |a| \exists z_1 \le t_1 \psi(u,x,a,z_1,w)$. Hence, by Lemma 8 and uniqueness of z_1, $\forall u \le |a| \exists z_1 \le t_1 \psi(u,x,a,z_1,w)$ is equivalent to $\exists w_1 \le \mathrm{Bd} \varphi(a,x,w_1,w)$, where $\varphi(a,x,w_1,w) \in \Sigma_0^b$ is

$$\forall u \le |a| (|w_1| = |2a| \cdot |t_1^*| + 1 \wedge \beta_{t_1^*(|a|)}(w_1,u) \le t_1 \wedge \psi(u,x,a,\beta_{t_1^*(|a|)}(w_1,x),w))$$

and Bd stands for the term $bd(t_1^*(|a|), a)$. Then, by Extensionality, it holds that

$$\forall a \forall x \exists! w \le 2^{|a|} \exists! w_1 \le \mathrm{Bd}(|w| = |a| \wedge \varphi(a,x,w_1,w)).$$

Finally, the $\hat{\Sigma}_1^b$ formula $\exists w_1 \le \mathrm{Bd}(|w| = |a| \wedge \varphi(a,x,w_1,w))$ defines h in the standard model. ∎

REMARK 2

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 previous proposition provides $\hat{\Sigma}_1^b$-formulas defining the functions involved and such that the relations stated by the recursion equations of CRN can be proved in $\hat{\Delta}_1^b$-CR.

COROLLARY 6

Every function in TC^0 is $\hat{\Sigma}_1^b$-definable in $\hat{\Delta}_1^b$-CR.

In order to get $\hat{\Sigma}_1^b$-selection we introduce a universally axiomatized extension of LIOpen$+$ $\hat{\Delta}_1^b$-BCR, denoted by CRNA. The proof we present here is a typical application of Herbrand's theorem and it is very similar to the Herbrand analyses of S_2^i developed by Sieg in [14].

LEMMA 10

$\hat{\Delta}_1^b$-CR has $\hat{\Sigma}_1^b$-selection.

PROOF. (Sketch) Let CRNA be the following theory:

- Language: $\mathcal{L}_{cr} = \bigcup_{i \in \omega} L_i$, where
  - $L_0 = \mathcal{L}_2$ plus a function symbol B_f for each basic function, $f \in BF$.
  - $L_{j+1} = L_j$ plus a function symbol, $\mathbf{f}_t$ for each term of L_j and a function symbol $\mathbf{f}_{t_0, t_1, t_2}$ for each terms $t_0(\vec{x})$, $t_1(y, \vec{x})$, $t_2(y, \vec{x})$ of L_j.
- Axioms: (in (4) $sg(x)$ is the term $1 \overset{\bullet}{-}(1 \overset{\bullet}{-}x)$)

      (1) BASIC
      (2) $B_{s_0}(x) = 2x$, $B_{s_1}(x) = 2x+1$, $B_{\Pi_i^n}(x_1, \ldots, x_n) = x_i$, $B_o(x) = 0$.
      (3) $\mathbf{f}_t(\vec{x}) = t(\vec{x})$.
      (4) $\mathbf{f}_{t_0, t_1, t_2}(\vec{x}, 0) = t_0(\vec{x})$
          $y \neq 0 \rightarrow \mathbf{f}_{t_0, t_1, t_2}(2y, \vec{x}) = 2 \cdot \mathbf{f}_{t_0, t_1, t_2}(y, \vec{x}) + sg(t_1(y, \vec{x}))$
          $\mathbf{f}_{t_0, t_1, t_2}(2y+1, \vec{x}) = 2 \cdot \mathbf{f}_{t_0, t_1, t_2}(y, \vec{x}) + sg(t_2(y, \vec{x}))$.
      (5) Induction: the induction scheme, *LIND*, for *open* formulas of $\mathcal{L}_{cr}$.

The proof of the following result is not hard in view of the techniques in [5].

CLAIM 1

1. In CRNA the open formulas are closed under sharply bounded quantification.
2. In CRNA, every formula Σ_0^b is equivalent to an open formula.
3. CRNA supports definition by cases.
4. CRNA is universally axiomatizable.

Since CRNA is a universal theory, a standard argument involving part (2) of Claim 1 and Herbrand's theorem proves that

CLAIM 2

In CRNA, every $\hat{\Delta}_1^b$ formula (of $\mathcal{L}_2$) is equivalent to an open formula.

Using Claim 2, Remark 2 and the formal version of CRN available in CRNA by the axiom group (4), we obtain the following

CLAIM 3

CRNA is a (conservative) extension of $\hat{\Delta}_1^b$-CR.

Now the result can be proved as follows: We may assume that $\varphi(x,y) \in \hat{\Delta}_1^b$ has the form $\exists z \le t_0 \varphi_0(x,y,z)$, where $\varphi_0(x,y,z) \in \Sigma_0^b$. By Claim 3, CRNA is an (conservative) extension of $\hat{\Delta}_1^b$-CR and, by Claim 1, there exists an open formula $\theta(x,y,z)$ such that

$$\text{CRNA} \vdash (y \le t \wedge z \le t_0 \wedge \varphi_0(x,y,z)) \leftrightarrow \theta(x,y,z)$$

Hence, $\text{CRNA} \vdash \forall x \exists y \exists z \theta(x,y,z)$ and, by Herbrand's theorem, there exist terms $s_1(x), \ldots, s_k(x)$ and $t_1(x), \ldots, t_k(x)$ such that

$$\text{CRNA} \vdash \theta(x, s_1(x), t_1(x)) \vee \cdots \vee \theta(x, s_k(x), t_k(x)).$$

Since CRNA admits definitions by cases, there exist terms $s(x)$ and $t'(x)$ such that $\text{CRNA} \vdash \theta(x, s(x), t'(x))$. By Proposition 4 and Remark 2 there exists an extension by $(\hat{\Sigma}_1^b)$ definitions of $\hat{\Delta}_1^b$-CR to the language $\mathcal{L}_{cr}$, which extends CRNA. As a consequence, by Claim 3, there exists $\psi(x,y) \in \hat{\Delta}_1^b$ in CRNA such that

$$\text{CRNA} \vdash \psi(x,y) \leftrightarrow s(x) = y \wedge \theta(x,y,t'(x)).$$

Then CRNA proves (1) $\forall x \exists! y \le t \, \psi(x,y)$, and (2) $\forall x \forall y (\psi(x,y) \to \varphi(x,y))$. But CRNA is a conservative extension of $\hat{\Delta}_1^b$-CR so it also proves (1) and (2). ∎

Now we can derive the main result of this section.

THEOREM 8
The theories $\hat{\Delta}_1^b$-CR, Δ_1^b-CR, LIOpen $+ \hat{\Sigma}_1^b$-BBR and LIOpen $+ \Sigma_1^b$-BBR are equivalent and axiomatize the class of all $\forall \Sigma_1^b$-consequences of C_2^0.

PROOF. Since C_2^0 is $\forall \Sigma_1^b$ axiomatizable and it is equivalent to $BB\hat{\Sigma}_1^b$ the conservation result follows from Theorem 3, modulo the equivalence of the four theories.

In order to obtain the equivalence of the four theories it is enough to prove that LIOpen $+ \hat{\Delta}_1^b$-BCR and LIOpen $+ \Sigma_1^b$-BBR are equivalent.

By Lemma 6, LIOpen $+ \Sigma_1^b$-BBR extends $\hat{\Delta}_1^b$-CR. On the other hand, since $\hat{\Delta}_1^b$-CR has $\hat{\Sigma}_1^b$-selection, by Lemma 9 the theory $[\hat{\Delta}_1^b$-CR, $\hat{\Sigma}_1^b$-BBR] is equivalent to $[\hat{\Delta}_1^b$-CR, $\hat{\Delta}_1^b$-BCR] $(= \hat{\Delta}_1^b$-CR). So, $\hat{\Delta}_1^b$-CR is closed under $\hat{\Sigma}_1^b$-BBR and extends LIOpen $+ \hat{\Sigma}_1^b$-BBR. ∎

COROLLARY 7
(Johannsen–Pollett) C_2^0 is a $\forall \Sigma_1^b$-conservative extension of Δ_1^b-CR.

The proof of Lemma 10 allows us to complete Corollary 6 deriving a characterization of the $\hat{\Sigma}_1^b$-definable functions in $\hat{\Delta}_1^b$-CR.

THEOREM 9
TC^0 is the class of the $\hat{\Sigma}_1^b$-definable functions of $\hat{\Delta}_1^b$-CR.

PROOF. By Corollary 6 every function in TC^0 is $\hat{\Sigma}_1^b$-definable in $\hat{\Delta}_1^b$-CR. The converse is easily obtained by an application of Herbrand's Theorem:

By Claim 2 of the proof of Lemma 10, if a function f is $\hat{\Sigma}_1^b$-definable in $\hat{\Delta}_1^b$-CR then it can be defined in CRNA by an open formula. Hence, by Herbrand's theorem, f can be defined in CRNA by a term of $\mathcal{L}_2$. But, by Clote–Takeuti's characterization stated in Theorem 7, each term of CRNA defines in the standard model a function in TC^0; so, $f \in TC^0$. ∎

Let us note in passing that the analysis of $\hat{\Delta}_1^b$-CR we have developed provides a strict version of the Johannsen–Pollett KPT-witnessing theorem for Δ_1^b-CR. The proof relies on Herbrand's theorem and is similar to the first proof given by Krajíček, Pudlák and Takeuti for their own witnessing theorem in [10].

COROLLARY 8
Let $\varphi(x,y,z) \in \hat{\Delta}_1^b$ in Δ_1^b-CR such that Δ_1^b-CR $\vdash \forall x \exists y \forall z \varphi(x,y,z)$. Then there exist $k \in \omega$ and functions $f_1, \ldots, f_k \in TC^0$, which are $\hat{\Sigma}_1^b$-definable in Δ_1^b-CR and such that Δ_1^b-CR proves

$$\varphi(x, f_1(x), z_1) \vee \varphi(x, f_2(x, z_1), z_2) \vee \cdots \vee \varphi(x, f_k(x, z_1, \ldots, z_{k-1}), z_k).$$

We conclude with an open problem. In [12], answering a question posed in [9], 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 $BB\Sigma_1^b$. In fact, the work in [12] suggests that Δ_1^b-CR is equivalent to [LIOpen, Δ_1^b-BCR]. However,

PROBLEM 1
Is LIOpen $+ \hat{\Sigma}_1^b$-BBR equivalent to [LIOpen, $\hat{\Sigma}_1^b$-BBR]?

In view of our results, a positive answer to Problem 1 seems to be likely, since this is the case for the analogous problem for collection rule in the usual language of Peano Arithmetic. On the other hand, the equivalence between $[T, \hat{\Delta}_1^b$-BCR] and $[T, \hat{\Sigma}_1^b$-BBR] has been proved only for theories T having $\hat{\Sigma}_1^b$-selection. So, it could be possible to answer Problem 1 in the positive, although the corresponding question for $\hat{\Delta}_1^b$-BCR could have a negative answer.

Acknowledgements

References

[1] Z. Adamowicz and T. Bigorajska. Existentially closed structures and Gödel's second incomplenteness theorem. *The Journal of Symbolic Logic* **66**, 349–356, 2001.

[2] J. Avigad. Saturated models of universal theories. *Annals of Pure and Applied Logic*, **118**, 219–234, 2002.

[3] S. Bloch. *Divide and Conquer in Parallel Complexity and Proof Theory*. PhD Thesis, University of California, San Diego, 1992.

[4] S. Buss. Bounded Arithmetic. Bibliopolis, Napoli, 1986.

[5] P. Clote and G.Takeuti. First order bounded arithmetic and small boolean circuit complexity classes. In *Feasible Mathematics II*, P. Clote and J. Remmel, eds, pp. 154–218. Birkhäuser, Boston, 1995.

[6] J. Hirschfeld and W. A. Wheeler. *Forcing, Arithmetic, Division Rings*. Lecture Notes in Mathematics, vol. 454, Springer Verlag, Berlin, Heidelberg, 1975.

[7] J. Johannsen. *A Bounded Arithmetic Theory for Constant Depth Threshold Circuits*. Gödel'96. Lecture Notes in Logic, vol. 6, pp. 224–234, Springer, Berlin, Heidelberg, 1996.

[8] J. Johannsen and C. Pollett. On proofs about threshold circuits and counting hierarchies. In *Proceedings of the 13th IEEE Symposium on Logic in Computer Science*, pp. 444–453, IEEE Computer Society, Los Alamitos, CA, USA, 1998.

[9] J. Johannsen and C. Pollett. In *On the Δ_1^b–bit–comprehension rule. Logic Colloquium'98*. Lecture Notes in Logic, **13**, 262–279. ASL/A.K. Peters, Natick, Massachusetts, 2000.

[10] J. Krajíček, P. Pudlák and Takeuti, G. Bounded arithmetic and the polynomial herarchy. *Annals of Pure and Applied Logic*, **52**, 143–153, 1991.

[11] J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Cambridge University Press, Cambridge, New York, Melbourne, 1995.

[12] P. Nguyen and S. Cook. Theories for TC^0 and other small complexity classes. *Logical Methods in Computer Science*, **2**, 1–40, 2006.

[13] C. Pollett. Structure and definability in general bounded arithmetic theories. *Annals of Pure and Applied Logic*, **100**, 189–245, 1999.

[14] W. Sieg. Herbrand analyses. *Archive for Mathematical Logic*, **30**, 409–441, 1991.