Polynomially Closed Co-Clones

Victor Lagerkvist
Department of Computer and Information Science,
Linköpings Universitet, Sweden,
Email: victor.lagerkvist@liu.se

Magnus Wahlström
Department of Computer Science,
Royal Holloway, University of London, Great Britain,
Email: Magnus.Wahlstrom@rhul.ac.uk

***Abstract*—Two well-studied closure operators for relations are based on primitive positive (p.p.) definitions and quantifier free p.p. definitions. The latter do however have limited expressiveness and the corresponding lattice of strong partial clones is uncountable. We consider implementations allowing polynomially many existentially quantified variables and obtain a dichotomy for co-clones where such implementations are enough to implement any relation and prove (1) that all remaining co-clones contain relations requiring a superpolynomial amount of quantified variables and (2) that the strong partial clones corresponding to two of these co-clones are of infinite order whenever the set of invariant relations can be finitely generated.**

I. Introduction

A finite or infinite set of Boolean relations Γ is known as a *constraint language*. Given a constraint language Γ, a natural question to ask is which other relations R can be expressed by first order formulas over Γ, or, equivalently, which is the smallest set of relations closed under such definitions. In practice one often considers restricted first order formulas, and two common restrictions are *primitive positive (p.p.) definitions*, where one is allowed to use existential quantification, conjunction and equality, and *quantifier-free primitive positive definitions (q.p.p.)* where only conjunction and equality is allowed. In other words an n-ary relation R has a p.p. definition in Γ if $R(x_1, \ldots, x_n) \equiv \exists y_1, \ldots, y_m \,.\, R_1(\mathbf{x_1}) \wedge \ldots \wedge R_k(\mathbf{x_k})$, where each $R_i \in \Gamma \cup \{=\}$ and each $\mathbf{x_i}$ is a vector over $x_1, \ldots, x_n, y_1, \ldots, y_m$, and R has a q.p.p. definition in Γ if $R(x_1, \ldots, x_n) \equiv R_1(\mathbf{x_1}) \wedge \ldots \wedge R_k(\mathbf{x_k})$, where each $R_i \in \Gamma \cup \{=\}$ and each $\mathbf{x_i}$ is a vector over $x_1, \ldots, x_n$. In both cases $=$ denotes the equality relation $\{(0,0),(1,1)\}$. For a set of relations Γ we let $\langle\Gamma\rangle$ denote the smallest set of relations closed under p.p. definability over Γ and $\langle\Gamma\rangle_{\not\exists}$ denote the smallest set of relations closed under q.p.p. definability over Γ. Sets of the form $\langle\Gamma\rangle$ are known as *relational clones (or co-clones)* due to their relationship with clones and sets of the form $\langle\Gamma\rangle_{\not\exists}$ are known as *weak partial relational clones*[1] due to their relationship with strong partial clones. Throughout the article we deal exclusively with Boolean co-clones and hence often refer to these simply as co-clones. A set of functions is called a *clone* if it (1) is closed under composition of functions and (2) contains all projection functions of the form $e_i^n(x_1, \ldots, x_n) = x_i$. A set of (partial) functions F is called a *strong partial clone* if in addition to (1) and (2) it also contains all partial subfunctions, i.e. if $f \in \mathrm{F}$ then F also contains all partial functions g such that the domain of g is a subset of the domain of f and such that g agrees with f for all values for which it is defined. If F is a set of functions we let $[\mathrm{F}]$ ($[\mathrm{F}]_s$) denote the smallest (strong partial) clone containing F. In both cases the set F is said to be *the base* of $[\mathrm{F}]$ or $[\mathrm{F}]_s$. We make a similar definition for co-clones and weak partial co-clones. The *order* of a clone or co-clone is the cardinality of the smallest base. In particular we are often interested in whether this order is finite or infinite.

Clones can equivalently be described as sets of functions preserving a set of relations since any n-ary Boolean function f can be extended to work over an m-ary Boolean relation R as follows: $f(t_1, \ldots, t_n) = \big(f(t_1[1], \ldots, t_n[1]), \ldots, f(t_1[m], \ldots, t_n[m])\big)$ where $t_i[j]$ denotes the j-th element of the tuple $t_i \in R$. If R is closed under f we say that f *preserves* R or that f is a *polymorphism* of R. We define $\mathrm{Pol}(\Gamma)$ for a set of relations Γ to be the set of polymorphisms to Γ, and $\mathrm{Inv}(\mathrm{F})$ (abbreviated as IF) for a set of functions F to be the set of all relations preserved by F. The set $\mathrm{pPol}(\Gamma)$ of all partial polymorphisms of Γ is defined in the same manner but with the additional stipulation that a function is allowed to be undefined for some values. With these notions one can verify that for any set of relations Γ it holds that $\langle\Gamma\rangle = \mathrm{Inv}(\mathrm{Pol}(\Gamma))$ and $\langle\Gamma\rangle_{\not\exists} = \mathrm{Inv}(\mathrm{pPol}(\Gamma))$, and also that $[\mathrm{F}] = \mathrm{Pol}(\mathrm{Inv}(\mathrm{F}))$ and $[\mathrm{F}']_s = \mathrm{pPol}(\mathrm{Inv}(\mathrm{F}'))$ for a set of functions F and a set of partial functions F′. This yields the *Galois connections* between sets of relations and their preserving functions.

Theorem 1 ([4], [5], [9]): Let Γ and Δ be two sets of relations. Then $\langle\Gamma\rangle \subseteq \langle\Delta\rangle$ if and only if $\mathrm{Pol}(\Delta) \subseteq \mathrm{Pol}(\Gamma)$.

Theorem 2 ([4], [5], [15]): Let Γ and Δ be two sets of relations. Then $\langle\Gamma\rangle_{\not\exists} \subseteq \langle\Delta\rangle_{\not\exists}$ if and only if $\mathrm{pPol}(\Delta) \subseteq \mathrm{pPol}(\Gamma)$.

The Galois connection forms the fundamental principles of the algebraic approach to computational problems parameterized by constraint languages such as the *constraint satisfaction problem (CSP(Γ))*. In Jeavons [10] it is proved that the computational complexity of CSP(Γ) up to polynomial-time reductions for any finite constraint language Γ is determined by the set of polymorphisms of Γ. A similar classification in Jonsson et al. [11] shows that the set of partial polymorphisms of Γ preserves all languages Δ such that CSP(Δ) is solvable at least as fast as CSP(Γ), and therefore, in a sense, preserves the exact complexity of CSP(Γ). Hence for all problems where a Galois connection is applicable a complexity classification of the problem is tantamount to understanding the structure of the clone lattice. For the Boolean domain the lattice of clones is completely classified and known as *Post's lattice* due to Post's seminal work [14]. See Figure I for a visualization of the Boolean co-clone lattice. Unfortunately the corresponding lattice of strong partial clones

[1] We note that the term *weak system* has also been used in the literature [7].

is of uncountably infinite cardinality even for the Boolean domain [1]. Given the fact that the lattice of strong partial clones is extremely complicated it is reasonable to consider the expressive power of closure operators which lie between q.p.p. definitions and p.p. definitions. To find implementations of such intermediate complexity we restrict the number of existentially quantified variables occurring in the formula and are therefore interested in which n-ary relations can be implemented with $1, 2, \ldots, f(n)$ existentially quantified variables for some reasonably slowly growing function f. In the sequel we assume that f is a polynomial function. If $f(n)$ variables is sufficient to implement every n-ary relation R in a co-clone then we say that the co-clone is *polynomially closed*. If $f(n) \leq 2$ then the resulting set of definable relations over some language Γ closely corresponds to the *frozen partial co-clone* [13] of Γ, with the exception that variables are not required to be frozen to some constant domain value.

In Section III we give a complete classification of the polynomially closed co-clones. Our proofs are based on comparing the least expressive language in the co-clone with the most expressive language in the co-clone in order to obtain an upper bound of f. These languages are known as the weak base and plain base, respectively, and were introduced by Schnoor and Schnoor [16], and Creignou et al. [8]. We then proceed in Section IV by proving that the classification in Section III is indeed complete in the sense that all other co-clones either do not have a finite base (in which case the notion does not apply) or contain relations which for any finite base requires a superpolynomial amount of variables in any p.p. definition. We also prove that all strong partial clones C whose total component consists either only of projection functions, or a composition of projections and the negation function $f(x) = 1 - x$, are of infinite order whenever $\mathrm{Inv}(\mathrm{C})$ can be finitely generated. This suggests that describing the set of partial polymorphisms $\mathrm{pPol}(\Gamma)$ for any finite constraint languages Γ such that $\mathrm{CSP}(\Gamma)$ is NP-hard (e.g., k-SAT) may be a difficult task.

A related concept to polynomial closure is whether every relation in a co-clone can be implemented using only polynomially many constraints from the plain base. If this holds then we say that the co-clone has a *polynomial base*. Using results from Section IV we prove that none of the superpolynomially closed co-clones admit polynomial bases.

II. Preliminaries and notation

If Γ is a constraint language we let Γ^n be defined as $\{R \mid R \in \Gamma, \mathrm{ar}(R) \leq n\}$, where $\mathrm{ar}(R)$ is the arity of R. We typically represent relations and constraint languages by their defining Boolean formulas. We define the *weak base* and *plain base* of a co-clone IC to be the bases of the smallest and largest members of the set $\mathcal{I}(\mathrm{IC}) = \{\mathrm{IC}' \mid \mathrm{IC}' = \langle \mathrm{IC}' \rangle_{\nexists}$ and $\langle \mathrm{IC}' \rangle = \mathrm{IC}\}$. Following Schnoor and Schnoor [16] we typically refer to the set $\mathcal{I}(\mathrm{IC})$ as an *interval*. Plain bases for all Boolean co-clones are given in Creignou et al. [8]. Weak bases were first introduced in Schnoor and Schnoor [16] but were in many cases exponentially larger than the plain bases with respect to arity. Weak bases fulfilling additional minimality conditions was given in Lagerkvist [12] using relational descriptions. By construction the weak base of a co-clone can always be given as a single relation. See Table I for a complete list of weak and plain bases of the Boolean co-clones of finite order. As seen in the following two theorems the plain base can be regarded to be the most expressive language in a co-clone while the weak base is the least expressive.

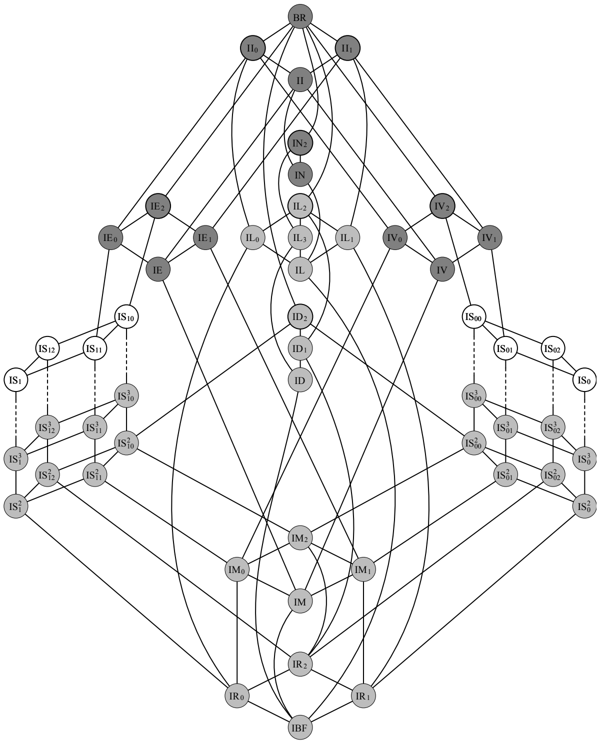

Fig. 1. The lattice of Boolean co-clones. The co-clones which are polynomially closed are coloured in grey. The co-clones which are not polynomially closed are coloured in dark grey. The remaining co-clones do not have a finite base.

Theorem 3: Let Γ be the plain base from Table I for some co-clone IC. Then $R \in \langle \Gamma^n \rangle_{\nexists}$ for any n-ary $R \in \mathrm{IC}$.

Proof: This follows from Creignou et al. [8] and the fact that the algorithm DescribePI [17] used there only works with k variables given a k-ary relation. ■

Theorem 4 ([16]): Let R_w be the weak base of some co-clone IC. Then $R_w \in \langle \Gamma \rangle_{\nexists}$ for any finite base Γ of IC.

We often use Eq as the equality relation $=$ and we let Γ_{SAT} denote the plain base from Table I of BR. By construction Γ^k_{SAT} is then the language corresponding to k-satisfiability. As is easily verified for every n-ary Boolean relation it holds that $R \in \langle \Gamma^n_{\mathrm{SAT}} \rangle_{\nexists}$. The full description of the relations involved is given in Lagerkvist [12].

III. Polynomially closed co-clones

In this section we formally introduce the notion of a polynomially closed co-clone. Intuitively the notion means that for any finite base a polynomial amount of variables is sufficient to p.p. implement any relation in the co-clone.

Definition 1: Let IC be a Boolean co-clone of finite order. We say that IC is *polynomially closed* if there exists a polynomial p such that for all finite bases Γ of IC and all n-ary $R \in \mathrm{IC}$ it holds that R can be p.p. defined in Γ with at most $p(n)$ existentially quantified variables.

TABLE I. WEAK AND PLAIN BASES OF ALL BOOLEAN CO-CLONES OF FINITE ORDER.

Co-clone	Weak base	Plain base
IBF	$\mathrm{Eq}(\mathrm{x}_1, \mathrm{x}_2)$	$\{\mathrm{Eq}(\mathrm{x}_1, \mathrm{x}_2)\}$
IR_0	$\mathrm{F}(c_0)$	$\{\mathrm{F}(c_0)\}$
IR_1	$\mathrm{T}(c_1)$	$\{\mathrm{T}(c_1)\}$
IR_2	$\mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{F}(c_0), \mathrm{T}(c_1)\}$
IM	$(x_1 \rightarrow x_2)$	$\{(x_1 \rightarrow x_2)\}$
IM_0	$(x_1 \rightarrow x_2) \wedge \mathrm{F}(c_0)$	$\{(x_1 \rightarrow x_2), \mathrm{F}(c_0)\}$
IM_1	$(x_1 \rightarrow x_2) \wedge \mathrm{T}(c_1)$	$\{(x_1 \rightarrow x_2), \mathrm{T}(c_1)\}$
IM_2	$(x_1 \rightarrow x_2) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \rightarrow x_2), \mathrm{F}(c_0), \mathrm{T}(c_1)\}$
$\mathrm{IS}_0^n, n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n)\}$
$\mathrm{IS}_{02}^n, n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), \mathrm{F}(c_0)\}$
$\mathrm{IS}_{01}^n, n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), (x_1 \rightarrow x_2)\}$
$\mathrm{IS}_{00}^n, n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), (x_1 \rightarrow x_2), \mathrm{F}(c_0)\}$
$\mathrm{IS}_1^n, n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n)\}$
$\mathrm{IS}_{12}^n, n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), \mathrm{T}(c_1)\}$
$\mathrm{IS}_{11}^n, n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), (x_1 \rightarrow x_2)\}$
$\mathrm{IS}_{10}^n, n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), (x_1 \rightarrow x_2), \mathrm{T}(c_1)\}$
ID	$(x_1 \oplus x_2 = 1)$	$\{(x_1 \oplus x_2 = 1)\}$
ID_1	$(x_1 \oplus x_2 = 1) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus x_2 = 1)\} \cup \{\mathrm{F}(c_0), \mathrm{T}(c_1)\}$
ID_2	$\mathrm{OR}^2_{2\neq}(x_1, x_2, x_3, x_4) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{F}(c_0), \mathrm{T}(c_1), (x_1 \vee x_2), (\neg x_1 \vee x_2), (\neg x_1 \vee \neg x_2)\}$
IL	$\mathrm{EVEN}^4(x_1, x_2, x_3, x_4)$	$\{(x_1 \oplus \ldots \oplus x_k = 0) \mid k\,\text{even}\}$
IL_0	$\mathrm{EVEN}^3(x_1, x_2, x_3) \wedge \mathrm{F}(c_0)$	$\{(x_1 \oplus \ldots \oplus x_k = 0) \mid k \in \mathbb{N}\}$
IL_1	$\mathrm{ODD}^3(x_1, x_2, x_3) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k \in \mathbb{N}, c = k \bmod 2\}$
IL_2	$\mathrm{EVEN}^3_{3\neq}(x_1, \ldots, x_6) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k \in \mathbb{N}, c \in \{0, 1\}\}$
IL_3	$\mathrm{EVEN}^4_{4\neq}(x_1, \ldots, x_8)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k\,\text{even}, c \in \{0, 1\}\}$
IV	$(\overline{x_1} \leftrightarrow \overline{x_2}\overline{x_3}) \wedge (\overline{x_2} \vee \overline{x_3} \rightarrow \overline{x_4})$	$\{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \geq 1\}$
IV_0	$(\overline{x_1} \leftrightarrow \overline{x_2}\overline{x_3}) \wedge \mathrm{F}(c_0)$	$\{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \in \mathbb{N}\}$
IV_1	$(\overline{x_1} \leftrightarrow \overline{x_2}\overline{x_3}) \wedge (\overline{x_2} \vee \overline{x_3} \rightarrow \overline{x_4}) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \geq 1\})$
IV_2	$(\overline{x_1} \leftrightarrow \overline{x_2}\overline{x_3}) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \in \mathbb{N}\})$
IE	$(x_1 \leftrightarrow x_2x_3) \wedge (x_2 \vee x_3 \rightarrow x_4)$	$\{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \geq 1\}$
IE_0	$(x_1 \leftrightarrow x_2x_3) \wedge (x_2 \vee x_3 \rightarrow x_4) \wedge \mathrm{F}(c_0)$	$\{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \in \mathbb{N}\}$
IE_1	$(x_1 \leftrightarrow x_2x_3) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \geq 1\})$
IE_2	$(x_1 \leftrightarrow x_2x_3) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \in \mathbb{N}\})$
IN	$\mathrm{EVEN}^4(x_1, x_2, x_3, x_4) \wedge x_1x_4 \leftrightarrow x_2x_3$	$\{\mathrm{Compl}_{m,n} \mid m, n \geq 1\}$
IN_2	$\mathrm{EVEN}^4_{4\neq}(x_1, \ldots, x_8) \wedge x_1x_4 \leftrightarrow x_2x_3$	$\{\mathrm{Compl}_{m,n} \mid m, n \in \mathbb{N}\}$
II	$(x_1 \leftrightarrow x_2x_3) \wedge (\overline{x_4} \leftrightarrow \overline{x_2}\overline{x_3})$	$\{(x_1 \vee \ldots x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m, n \geq 1\}$
II_0	$(\overline{x_1} \vee \overline{x_2}) \wedge (\overline{x_1}\overline{x_2} \leftrightarrow \overline{x_3}) \wedge \mathrm{F}(c_0)$	$\{(x_1 \vee \ldots x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m \in \mathbb{N}, n \geq 1\}$
II_1	$(x_1 \vee x_2) \wedge (x_1x_2 \leftrightarrow x_3) \wedge \mathrm{T}(c_1)$	$\{(x_1 \vee \ldots x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m \geq 1, n \in \mathbb{N}\}$
BR	$R^{1/3}_{3\neq}(x_1, \ldots, x_6) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \vee \ldots x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m, n \in \mathbb{N}\}$

For co-clones without a finite base this notion is not relevant. If a co-clone is not polynomially closed then we say that it is *superpolynomially closed*. In order to prove that a co-clone is polynomially closed it is sufficient to prove that there exists some polynomial p such that the weak base of the co-clone can implement any n-ary relation with $p(n)$ variables. This also implies that Definition 1 can be rephrased as the seemingly weaker condition of for every finite base finding a polynomial which limits the number of quantifiers. Similarly we say that a co-clone IC has a *polynomial base* if there exists a polynomial p such that every n-ary $R \in \mathrm{IC}$ has a q.p.p. implementation Γ^n, where Γ is the plain base from Table I, with at most $p(n)$ constraints. Obviously this trivially holds for all co-clones with a finite plain base. Polynomial bases and polynomially closed co-clones are related by the following lemma which states that a polynomial base for a co-clone implies polynomial closure under some additional conditions.

Lemma 1: Let IC be a co-clone with plain base Γ and weak base R_w from Table I such that (1) R_w can p.p. implement Γ^n with $p(n)$ variables for some polynomial p and (2) IC has a polynomial base for some polynomial g. Then IC is polynomially closed.

Proof: Let $R \in \mathrm{IC}$ be an n-ary relation. By Theorem 3 and the original assumption it follows that Γ^n can q.p.p. implement R using at most $g(n)$ constraints. Let ϕ denote the q.p.p. implementation of R in Γ^n. For every constraint C_i in ϕ we then replace C_i with its p.p. implementation in $\{R_w, =\}$. Let the resulting formula be ϕ'. Since ϕ had $g(n)$ constraints and each constraint in ϕ' introduced at most $p(n)$ new existentially quantified variables, the total number of variables in ϕ' is $g(n) \cdot p(n)$, clearly polynomial with respect to n. Hence IC is polynomially closed. ■

The rest of the article is devoted to proving a complete dichotomy theorem between polynomially closed and superpolynomially closed co-clones. The proof consists of two cases depending on whether the co-clone in question has a plain base of finite or infinite cardinality.

Lemma 2: If IC has a finite plain base then IC is polynomially closed.

Proof: Assume that IC has a plain base Γ of finite cardinality and let R_w denote a weak base of IC. Since Γ is finite there exists a polynomial p such that R_w can p.p. implement Γ^n for every $n \geq 1$ with $p(n)$ variables. To see this simply take the number of existentially quantified variables of the relation requiring the largest number of quantified variables in the p.p. definition in Γ. Such a relation must exist since Γ is finite. The result then follows from Lemma 1 since IC has a polynomial base whenever the plain base Γ is finite. ■

Lemma 2 is however not applicable for IL, IL_0, IL_1, IL_3 and IL_2 since their plain bases are infinite. It is however easy to prove that all these co-clones admit polynomial bases since the included relations can be viewed as linear equations over the field GF(2).

Lemma 3: $\mathrm{IL}, \mathrm{IL}_0, \mathrm{IL}_1, \mathrm{IL}_3$ and IL_2 have polynomial bases.

Proof: We only consider IL_2 since the other cases follow through similar arguments. Every n-ary relation $R \in \mathrm{IL}_2$ can according to Theorem 3 be expressed by a Γ^n formula ϕ with m constraints, where Γ is the plain base in Table I. Thus every constraint C_i in ϕ is of the form $(x_{i_1} \oplus \ldots \oplus x_{i_n}) = c_i$, where $c_i \in \{0, 1\}$. Create an $m \times (n+1)$-matrix M such that each entry $r_{i,j}$, $1 \leq j \leq n$, is equal to 1 if the variable x_j is included in the constraint C_i, and 0 otherwise. The entry $r_{i,n+1}$ is equal to the constant c_i in C_i. Then it is not hard to verify that if the row r_{i+1} is linearly dependent on $r_1, \ldots, r_i$ then $C_1, \ldots C_i$ entails C_{i+1} in any satisfying assignment. Hence we only need to keep the rows that are linearly independent which gives the bound $\min(n+1, m)$ on the number of constraints. ■

Lemma 4: IL, IL_0, IL_1, IL_3 and IL_2 are polynomially closed.

Proof: We only present the proof of IL_2 since the other co-clones follow through entirely analogous arguments. Let Γ and R_w be the plain and weak base of IL_2 from Table I, respectively. Since IL_2 has a polynomial base by Lemma 3 all we need to prove is that R_w can p.p. define Γ^n with poly(n) variables. We first and most crucially show that Γ^n can implement Γ^{n+1} with only one extra variable, for every $n \geq 3$, with the implementation $(x_1 \oplus \ldots \oplus x_{n+1} = c) \equiv \exists x.(x_1 \oplus \ldots \oplus x_{n-1} \oplus x = c) \wedge (x_n \oplus x_{n+1} \oplus x = 0)$. In addition to one quantified variable this requires one extra Γ^3-constraint. Hence if $3 \leq n \leq n'$ then Γ^n can implement every relation in $\Gamma^{n'}$ with $O(n'-n)$ variables and $n'-n$ additional Γ^3-constraints. By this it first follows that Γ^3 can p.p. define any relation in Γ^n with at most $n-3$ variables and $n-2$ constraints. The weak base R_w can then p.p. implement Γ^3 with a fixed number of variables since the arity of each relation is bounded, for example we have that $(x_1 \oplus x_2 \oplus x_3 = 0) \equiv \exists y_1, y_2, y_3, c_0, c_1 . R_w(x_1, x_2, x_3, y_1, y_2, y_3, c_0, c_1)$ and $(x_1 \oplus x_2 \oplus x_3 = 1) \equiv \exists\, y_1, y_2, y_3, c_0, c_1 \,.\, R_w(y_1, y_2, y_3, x_1, x_2, x_3, c_0, c_1)$. Put together this implies that R_w can p.p. define any Γ^n with $O(n)$ existentially quantified variables, and by Lemma 1 that IL_2 is polynomially closed. ■

Theorem 5: If $\mathrm{IC} \subseteq \mathrm{IX}$ for some $\mathrm{IX} \in \{\mathrm{IL}_2, \mathrm{ID}_2\} \cup \{\mathrm{IS}^n_{00}, \mathrm{IS}^n_{10} \mid n \geq 2\}$ then IC is polynomially closed.

IV. Superpolynomially closed co-clones

From Theorem 5 we now have a classification of the polynomially closed co-clones. In this section we strengthen this further and prove that the classification is also complete in the sense that all other co-clones of finite order are superpolynomially closed. The heart of the proof is based on counting the number of n-ary relations in the co-clone. If this number is sufficiently large we can prove that there exists relations which for any finite base cannot be expressed with a polynomial amount of existentially quantified variables.

Lemma 5: Let IC be a co-clone of finite order. If IC is polynomially closed, then the number of n-ary relations in IC is at most $2^{p(n)}$ for some polynomial p.

Proof: Let Γ be a finite base of IC and let R be the relation with the highest arity k in Γ. We make a few observations before the proof: first, $\langle\Gamma\rangle_{\nexists} \subseteq \langle\Gamma^k_{\mathrm{SAT}}\rangle_{\nexists}$; second, if some $R' \notin \langle\Gamma^k_{\mathrm{SAT}}\rangle_{\nexists}$ then $R' \notin \Gamma$. This also implies that if Γ can p.p. define some n-ary relation R with $p(n)$ existentially quantified then the same is true for Γ^k_{SAT}. By contraposition this also implies that if Γ^k_{SAT} cannot p.p. define some n-ary relation R with $p(n)$ variables then neither can Γ. Now let $n > 0$. The number of k-SAT formulas over n variables is bounded by $2^{2^k n^k}$, which is also the number of q.p.p. implementations with Γ^k_{SAT} over n variables since $\mathrm{Eq} \in \langle\Gamma^{\mathrm{k}}_{\mathrm{SAT}}\rangle_{\nexists}$. Hence the maximum amount of relations q.p.p. definable with Γ^k_{SAT} is at most $2^{2^k \cdot n^k}$. Since IC is polynomially closed we are allowed to introduce at most $p(n)$ existentially quantified variables to implement any n-ary relation, hence the number of definable relations is at most $2^{2^k p(n)^k}$. ■

Since the number of n-ary Boolean relations is 2^{2^n} it immediately follows that BR is superpolynomially closed. To handle the other cases in the co-clone lattice where it is not directly obvious how to count the number of relations we provide mappings from BR to prove the requisite lower bound.

We say that a relation R is *downward closed* (respectively upward closed) if for every tuple $t \in R$ it holds that $t' \in R$ whenever $t' \leq t$ (respectively $t' \geq t$), where $t' \leq t$ is applied componentwise.

Lemma 6: If R is downward closed (respectively upward closed) then $R \in \mathrm{IS}_1$ (respectively IS_0).

Proof: First note that S_1 can be generated by the 1-separating Boolean function $f(x, y) = x \wedge \neg y$ [6]. Assume that R is an n-ary downward closed relation. We prove that it is then closed under f from which it follows that $R \in \mathrm{IS}_1$. Let $t_1 = (x_1, \ldots, x_n), t_2 = (y_1, \ldots, y_n) \in R$. Then the tuple $(f(x_1, y_1), \ldots, f(x_n, y_n)) \in R$ since for every i we have that $f(x_i, y_i) = x_i \wedge \neg y_i \leq x_i$. The proof for upward closed relations is similar. ■

For the following proof we slightly abuse notation and let IS^n_1 (and respectively for IS^n_0) denote the finite set $\{R \mid R \in \mathrm{IS}_1, \mathrm{ar}(R) \leq n\}$, and not as in Table I, the co-clone generated by NAND^n.

Lemma 7: For every n there exists an injective map from BR^n to IS^{2n}_1 (respectively IS^{2n}_0).

Proof: Let R be an n-ary Boolean relation. We define the $2n$-ary Boolean relation R' such that:

- $(x_1, y_1, \ldots, x_n, y_n) \in R'$ if $x_i + y_i \leq 1$ for all $i \in [n]$ with $x_i = y_i = 0$ for at least one $i \in [n]$, and
- if $x_i + y_i = 1$ for all $i \in [n]$ then $(x_1, y_1, \ldots, x_n, y_n) \in R'$ if and only if $(x_1, \ldots, x_n) \in R$.

We first prove that $R' \in \mathrm{IS}_1$ by showing that it is downward closed. Let $t = (x_1, y_1, \ldots, x_n, y_n) \in R'$ and let $t' = (x'_1, y'_1, \ldots, x'_n, y'_n)$ be any tuple such that $t' < t$. Then we first see that the property $x'_i + y'_i \leq 1$ still holds. Second assume that $(x_1, \ldots, x_n) \in R$. This means that for all i it holds that $x_i + y_i = 1$. Since $t' < t$ there is at least one i such that $x_i + y_i = 0$ which means that $t' \in R'$.

Injectivity follows from the fact that any two relations results in two distinct relations in IS_1. The proof for IS_0 is similar but using $x_i + y_i \geq 1$ and upward closure instead. ■

In the previous proof we map every n-ary relation to a $2n$-ary relation. The number of n-ary relations in IS_1 and IS_0 is therefore $2^{2^{\Theta(n)}}$ since it is at least $2^{2^{\frac{n}{2}}}$ if n is even and

$2^{2^{\frac{n-1}{2}}}$ if n is odd (it is easy to create an $(n+1)$-ary relation from an n-ary relation by adding a new variable not occurring elsewhere). The mapping from IS_0 to IV is similar but based on a translation from a positive clause $(x_1 \vee \ldots \vee x_n)$ to a dual Horn clause of the form $(x_1 \vee \ldots \vee x_n \vee \neg y)$.

Lemma 8: For every n there exists an injective map from IS_0^n (respectively IS_1^n) to IV^{n+1} (respectively IE^{n+1}).

Proof: Let R be an n-ary relation in IS_0. We define the $(n+1)$-ary relation R' such that:

- R' contain all tuples $(x_1, \ldots, x_n, y)$ where $y = 0$, and
- if $y = 1$ then $(x_1, \ldots, x_n, y) \in R'$ if and only if $(x_1, \ldots, x_n) \in R$.

Due to the second condition the mapping is injective. The difficulty lies in proving that R' is indeed included in IV. The clone V can be generated by the binary or function $x_1 \vee x_2$ and the constant functions $c_0(x) = 0$ and $c_1(x) = 1$ [6]. Since $c_1 \in \text{S}_0$ all we have to do is to verify that R' is closed under c_0 and disjunction. That R' is closed under c_0 follows from the first condition which ensures that R' will be 0-valid. Let $(x_1, \ldots, x_n, x)$ and $(y_1, \ldots, y_n, y)$ be any two tuples in R' and let $(z_1, \ldots, z_n, z) = (x_1 \vee y_1, \ldots, x_n \vee y_n, x \vee y)$. If $z = 0$ then $(z_1, \ldots, z_n, z) \in R'$ so instead assume that $z = 1$. This means that $x + y \geq 1$ and hence that $(x_1, \ldots, x_n) \in R$ or $(y_1, \ldots, y_n) \in R$. Then for every i we have both that $z_i \geq x_i$ and $z_i \geq y_i$, which means that $(z_1, \ldots, z_n)$ is larger than both $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, and since at least one of these tuples are included in R it follows that $(z_1, \ldots, z_n) \in R$ and $(z_1, \ldots, z_n, z) \in R'$ since R is closed upwards.

The mapping from IS_1 to IE is defined dually. ■

Theorem 6: If $\text{IC} \supseteq \text{IX}$ for some $\text{IX} \in \{\text{IV}, \text{IE}, \text{IN}\}$ then IC is superpolynomially closed.

Proof: Due to the mappings from BR in Lemma 7 and 8 we get the bound $2^{2^{\Theta(n)}}$ for IV and IE, and hence also for IE_0, IE_1, IE_2, IV_0 IV_1, IV_2, II, II_0 and II_1. The remaining two cases IN and IN_2 are handled by the bound for IN which is $2^{2^{\Theta(n)}}$ since there are $2^{2^{n-1}}$ n-ary relations closed under complement and $2^{2^{n-1}-1}$ n-ary 0- and 1-valid relations which are closed under complement. ■

There is a close connection between polynomially closed co-clones and a certain class of algebras [3]. If Γ is a constraint language the algebra $\mathbf{A}_\Gamma = (\{0,1\}, \text{Pol}(\Gamma))$ is said to have *few subpowers* if the base 2 logarithm of the cardinality of the set of all subuniverses of $\mathbf{A}_\Gamma^n$, where $\mathbf{A}_\Gamma^n$ is the n-ary direct power of $\mathbf{A}_\Gamma$, can be bounded by some polynomial. This furthermore holds if and only if the number of n-ary relations in $\langle \Gamma \rangle$ is bounded by $2^{p(n)}$ for some polynomial p, if and only if $\text{Pol}(\Gamma)$ contains a certain *edge* polymorphism [3]. These facts together with Lemma 5 can be used as a less direct proof of Theorem 6.

We now show that the strong partial clones corresponding to $\mathcal{I}(\text{BR})$ and $\mathcal{I}(\text{IN}_2)$ have a highly complex structure and are of infinite order whenever the set of invariant relations can be finitely generated. We need the following construction of a universal hash family, due to Alon et al. [2].

Theorem 7 (Section 4 of [2]): For any k and n, there is a family H of $2^{O(k)} \log n$ functions $h_i : \{1, \ldots, n\} \mapsto \{1, \ldots, k\}$ such that for every $S \subset \{1, \ldots, n\}$ of size k there is a function in H that is injective on S.

Note that the bound $O(k)$ has no hidden dependency on n. Hence, if k is a constant, then $2^{O(k)} \log n \in O(\log(n))$.

Lemma 9: Let Γ be a constraint language with $\langle \Gamma \rangle \in \{\text{BR}, \text{IN}_2\}$. If $\text{pPol}(\Gamma)$ has finite order, then Γ can p.p. implement all n-ary relations $R \in \langle \Gamma \rangle$ with at most $O(n)$ existentially quantified variables.

Proof: Let $R(x_1, \ldots, x_n)$ be an n-ary relation in IN_2 or BR, and let $m \leq 2^n$ be the number of tuples in R. Let S be a finite basis of $\text{pPol}(\Gamma)$, let r be the largest arity of any function in S, and let H be the r-universal hash family from $[m]$ to $[r]$ of Theorem 7. We will create a relation R' on variable set $\{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_p\}$ where $p = 2^r |H|$, such that $R(x_1, \ldots, x_n) \equiv \exists y_1, \ldots, y_p . R'(x_1, \ldots, x_n, y_1, \ldots, y_p)$, and such that R' is closed under $\text{pPol}(\Gamma)$. Since $|H| = 2^{O(r)} \log m = 2^{O(r)} O(n)$, and since r is a constant, we have that the number of existentially quantified variables introduced is $2^{O(r)} O(n) = O(n)$. The relation R' is defined as follows. Write R down in matrix form, with each column representing a variable and each row representing a tuple of R. Let M be this matrix. For each hash function $h_i \in H$ and each mapping $g_j : [r] \to \{0,1\}$, we add a new column $y_{i,j}$ to M, which in row t takes value $g_j(h_i(t))$, $t \in [m]$. In other words, for each hash function $h_i : [m] \to [r]$ we get a block of 2^r variables $y_{i,j}$, corresponding to composing h_i with each of the 2^r possible maps g_j. We split into cases depending on $\langle \Gamma \rangle$.

If $\langle \Gamma \rangle = \text{BR}$, then we use M as described. Let R' be the relation whose tuple set is enumerated by M. We claim that for every q-tuple $T = (t_1, \ldots, t_q)$ of distinct rows from M, with $q \leq r$, and for every $(b_1, \ldots, b_q) \in \{0,1\}^q$, there is at least one column $y_{i,j}$ from M such that $(t_1[y_{i,j}], \ldots, t_q[y_{i,j}]) = (b_1, \ldots, b_q)$. Let $P = (p_1, \ldots, p_q) \in [m]^q$ be the row indices of T, i.e., $t_i = M[p_i, \cdot]$ for each $i \in [q]$. Since H is a universal hash family, there is some $h \in H$ which is injective on P, hence (by the extra mappings g_ℓ) the columns $y_{i,j}$, restricted to the rows P, enumerate all 2^q vectors. Thus there is a value j such that the claim holds.

If $\langle \Gamma \rangle = \text{IN}_2$, we proceed as follows. First, we add one further column z to M, making it the all-zero column. Next we close M under complement by adding the complement of every row of M to M. Let R' be the relation corresponding to the new matrix M, and let $m' = 2m$ be the number or rows. Note that for a tuple $t \in R'$, $t(z) = 0$ if t is one of the "original" rows, while $t(z) = 1$ if t was added in the closure step. Let $T = (t_1, \ldots, t_q)$ where $t_1, \ldots, t_q$ are distinct rows from M, with $q \leq r$, and let $P = (p_1, \ldots, p_q) \in [m']^q$ be the row indices of T, i.e., $t_i = M[p_i, \cdot]$ for each $i \in [q]$. Define a new tuple $(c_1, \ldots, c_q)$ where $c_i = b_i \oplus t_i(z)$, and let $y_{i,j}$ be the column which takes values $(c_1, \ldots, c_q)$ on the original tuples corresponding to T. This column exists by the proof of the previous paragraph. Then clearly, $y_{i,j}$ takes values $(b_1, \ldots, b_q)$ evaluated on the rows P.

We find that in both cases, restricted to any q-tuple of distinct rows, the variables $y_{i,j}$ enumerate all 2^q bit vectors. It follows that R' is closed under $\text{pPol}(\Gamma)$: Consider an application $f(t_1, \ldots, t_q)$ of some $f \in \text{pPol}(\Gamma)$. We may assume that all $t_1, \dots, t_q$ come from different rows, as otherwise the application of f is equivalent to the application of some q'-ary partial polymorphism f' on distinct rows, where q' is the number of distinct rows represented in $(t_1, \dots, t_q)$. If f is partial (i.e., has at least one undefined value), then by the above there is some column $y_{i,j}$ which blocks its application; otherwise we must have $\langle \Gamma \rangle = \mathrm{IN}_2$ and f must be a combination of projection and complement, and R' is already closed under both. Thus Γ can q.p.p. implement R', and we are done. ■

Theorem 8: Let $\mathrm{IC} \in \{\mathrm{BR}, \mathrm{IN}_2\}$. Then, if Γ is a finite base of IC, $\mathrm{pPol}(\Gamma)$ is of infinite order.

Proof: First assume that $\mathrm{pPol}(\Gamma)$ can be finitely generated. By Lemma 9 we then have that Γ can p.p. implement all n-ary relations in IC with $O(n)$ existentially quantified variables which contradicts Theorem 6. ■

Note that the proof does not work for the other superpolynomially closed co-clones since in general there is no guarantee that the extended relation has the same set of total polymorphisms as the original relation. However, the co-clones BR and IN_2 are interesting for practical considerations since these are the only cases when the CSP problem is NP-hard. Theorem 8 then says that describing the set of partial polymorphisms (which strongly correlates to worst-case running times [11]) even for very simple languages such as $R_{1/3} = \{(0,0,1), (0,1,0), (1,0,0)\}$ is complicated since these are always of infinite order.

V. Polynomial bases for co-clones

A further strengthening of Theorem 6 involves proving that all superpolynomially closed co-clones lack polynomial bases. This can be shown by a simple counting argument, using the bounds from the previous section on the number of n-ary relations in each of these co-clones.

Theorem 9: If $\mathrm{IC} \supseteq \mathrm{IX}$ for some $\mathrm{IX} \in \{\mathrm{IV}, \mathrm{IE}, \mathrm{IN}\}$ then IC does not admit a polynomial base.

Proof: We show the theorem with a counting argument, using the results of Section IV. We make two quick observations: first, recall from Section IV that every co-clone IC as specified above contains $2^{2^{\Theta(n)}}$ n-ary relations; second, if Γ is the plain base for IC from Table I, then $|\Gamma^n| = p(n)$ for some polynomial $p(n)$ (the proof will however go through with any plain base satisfying $|\Gamma^n| \leq 2^{p(n)}$). For each $R \in \Gamma^n$, there are at most n^n different possible constraints one can form with R; thus the number of different possible constraints overall is bounded by $|\Gamma^n| \cdot n^n$. The number of possible formulas with at most $c(n)$ constraints is then bounded by $(|\Gamma^n| \cdot n^n)^{c(n)} \leq 2^{q(n)}$ for a polynomial $q(n)$. The theorem follows. ■

Thus a co-clone of finite order has a polynomial base if and only if it is polynomially closed. In conjunction the results of Sections III, IV and V therefore imply the following corollary.

Corollary 1: Let $\langle \Gamma \rangle$ be a Boolean co-clone of finite order. Then the following statements are equivalent.

- $\langle \Gamma \rangle$ is polynomially closed.
- $\langle \Gamma \rangle$ has a polynomial base.
- The algebra $(\{0,1\}, \mathrm{Pol}(\Gamma))$ has few subpowers.
- There exists a polynomial p such that the number of n-ary relations in $\langle \Gamma \rangle$ is not larger than $2^{p(n)}$.

Acknowledgements

The authors are grateful toward Peter Jonsson, Karsten Schölzel, Bruno Zanuttini, and the anonymous referees, for helpful comments and suggestions.

References

[1] V. B. Alekseev and A. A. Voronenko. On some closed classes in partial two-valued logic. *Discrete Math. Appl.*, 4(5):401–419, 1994.

[2] N. Alon, R. Yuster, and U. Zwick. Color-coding. *J. ACM*, 42(4):844–856, July 1995.

[3] J. Berman, P. Marković, R. Mckenzie, M. Valeriote, and R. Willard. Varieties with few subalgebras of powers. *Trans. Amer. Math. Soc.*, 362(3):1445 – 1473, 2006.

[4] V. G. Bodnarchuk, L. A. Kaluzhnin, V. N. Kotov, and B. A. Romov. Galois theory for Post algebras. I. *Cybernetics*, 5:243–252, 1969.

[5] V. G. Bodnarchuk, L. A. Kaluzhnin, V. N. Kotov, and B. A. Romov. Galois theory for Post algebras. II. *Cybernetics*, 5:531–539, 1969.

[6] E. Böhler, N. Creignou, S. Reith, and H. Vollmer. Playing with Boolean blocks, part I: Post's lattice with applications to complexity theory. *ACM SIGACT-Newsletter*, 34(4):38–52, 2003.

[7] F. Börner, L. Haddad, and R. Poschel. A note on minimal partial clones. In *Proc. ISMVL-1991*, pages 262–267, 1991.

[8] N. Creignou, P. Kolaitis, and B. Zanuttini. Structure identification of Boolean relations and plain bases for co-clones. *J. Comput. Syst. Sci.*, 74(7):1103–1115, November 2008.

[9] D. Geiger. Closed systems of functions and predicates. *Pac. J. Math.*, 27(1):95–100, 1968.

[10] P. Jeavons. On the algebraic structure of combinatorial problems. *Theor. Comput. Sci.*, 200:185–204, 1998.

[11] P. Jonsson, V. Lagerkvist, G. Nordh, and B. Zanuttini. Complexity of SAT problems, clone theory and the exponential time hypothesis. In *Proc. SODA-2013*, pages 1264–1277, 2013.

[12] V. Lagerkvist. Weak bases of Boolean co-clones. *CoRR*, abs/1310.3674, 2013.

[13] G. Nordh and B. Zanuttini. Frozen Boolean partial co-clones. In *Proc. ISMVL-2009*, pages 120 –125, 2009.

[14] E. Post. The two-valued iterative systems of mathematical logic. *Ann. of Math. Stud.*, 5:1–122, 1941.

[15] B.A. Romov. The algebras of partial functions and their invariants. *Cybernetics*, 17(2):157–167, 1981.

[16] H. Schnoor and I. Schnoor. Partial polymorphisms and constraint satisfaction problems. In N. Creignou, P. G. Kolaitis, and H. Vollmer, editors, *Complexity of Constraints*, volume 5250 of *Lect. Notes. Comput. Sc.*, pages 229–254. Springer Berlin Heidelberg, 2008.

[17] B. Zanuttini and J.J. Hébrard. A unified framework for structure identification. *Inform. Process. Lett.*, 81(6):335 – 339, 2002.