Quantifier-Free Boolean Algebra with Presburger Arithmetic is NP-Complete

Viktor Kuncak

MIT Computer Science and AI Lab, Cambridge, MA

Abstract. Boolean Algebra with Presburger Arithmetic (BAPA) combines 1) Boolean algebras of sets of uninterpreted elements (BA) and 2) Presburger arithmetic operations (PA). BAPA can express the relationship between integer variables and cardinalities of unbounded finite sets and can be used to express verification conditions in verification of data structure consistency properties.
In this report I consider the Quantifier-Free fragment of Boolean Algebra with Presburger Arithmetic (QFBAPA). Previous algorithms for QFBAPA had non-deterministic exponential time complexity. In this report I show that QFBAPA is in NP, and is therefore NP-complete. My result yields an algorithm for checking satisfiability of QFBAPA formulas by converting them to polynomially sized formulas of quantifier-free Presburger arithmetic. I expect this algorithm to substantially extend the range of QFBAPA problems whose satisfiability can be checked in practice.

1 Introduction

This paper considers the satisfiability problem for a logic that allows reasoning about sets and their cardinalities. We call this logic quantifier-free Boolean Algebra with Presburger Arithmetic and denote it QFBAPA. Figure 1 shows the syntax of QFBAPA. The logic contains 1) arbitrary boolean algebra (BA) expressions denoting sets, 2) arbitrary quantifier-free Presburger arithmetic (PA, linear integer arithmetic) expressions, and 3) a cardinality operator for stating that the size of a set denoted by a set expression is equal to an integer denoted by a given PA expression. The constant MAXC denotes the size of the universal set, so $|\mathbf{1}| = \mathsf{MAXC}$. The expression $K \,\mathsf{dvd}\, T$ means that K divides integer T, whereas B^c denotes the complement of the set B.

QFBAPA satisfiability is clearly NP-hard, because QFBAPA has propositional operators on formulas. Moreover, QFBAPA contains Boolean algebra of sets that has its own propositional structure. The challenge is therefore to prove the membership in NP. The difficulty is that formulas such as $|A \setminus B \cup C| = 10000$ force the sizes of sets to be exponential in the length of the formula, leading to a doubly exponential number of interpretations of set variables.

Motivation for QFBAPA. Our motivation for QFBAPA is proving the validity of formulas arising from program verification [16]. The logic QFBAPA is a quantifier-free fragment of Boolean Algebra with Presburger Arithmetic (BAPA) which extends QFBAPA with arbitrary set and integer quantifiers. BAPA was implicitly used in [11, Section 8, Page 90] as an extension of set algebra that occurs in Feferman-Vaught construction. Subsequently, BAPA was found to be of interest in program verification [18, 19, 16, 33] and constraint databases [29]. In [18] we have shown that BAPA has the same complexity as PA, namely alternating doubly exponential time with a linear number of alternations, denoted $\mathsf{STA}(*, 2^{2^{n^{O(1)}}}, n)$ in [4], [15, Lecture 24].

BAPA has quantifier elimination property, which implies that QFBAPA formulas define the same class of relations on sets and integers as BAPA formulas, so they essentially have the same expressive power. In general, QFBAPA formulas may be exponentially larger than the equivalent quantified BAPA formulas with same free variables. However, it is often the case that the proof obligation (or other problem of interest) is already expressed in quantifier-free form. It is therefore interesting to consider the complexity of the satisfiability problem for QFBAPA.

Quantifier-free PA. Quantifier-free PA is in NP because it has a small model property implying that satisfiable formulas have solutions whose binary representation is polynomial. The small model property for quantifier-free PA follows from the small model property for conjunctions of atomic formulas, which in turn follows from bounds on solutions of integer linear programming problems [27]. In practice, quantifier-free PA formulas can be solved using implementations such as CVC Lite [3] and UCLID [21].

Previous algorithms for QFBAPA. Existing algorithms for QFBAPA [34, 29, 26] run in non-deterministic exponential time, because they explicitly introduce a variable for each Venn region. The same exponential explosion occurs in our previous algorithm α [17, 18, 19] that decides the entire BAPA.

The result of this paper. I have previously used a divide-and-conquer algorithm to show that it is not necessary to simultaneously generate all Venn region variables, proving that QFBAPA is in PSPACE [23, Section 3]. I here give a stronger result, which shows that QFBAPA is in NP. In the process, I identify a natural encoding of QFBAPA formulas into polynomially-sized quantifier-free PA formulas. I use a recent result [10] that if an element is in an integer cone generated by a set of vectors X, then it is also in an integer cone generated by a "small" subset of X. This result implies that a system of equations with bounded coefficients, if satisfiable, has a *sparse solution* with only polynomially many non-zero variables, even if the number of variables in the system is exponential. As a result, instead of using exponentially many Venn region cardinality variables to encode relationships between sets, we can use polynomially many "generic" variables along with polynomially many indices that determine which region each generic variable represents. In other words, every satisfiable QFBAPA formula has a witness of polynomial size, which indicates the values of integer variables in the original QFBAPA formula, lists the Venn regions that are non-empty, and indicates the cardinalities of these non-empty regions.

$$F ::= A \mid F_1 \wedge F_2 \mid F_1 \vee F_2 \mid \neg F$$
$$A ::= B_1 = B_2 \mid B_1 \subseteq B_2 \mid T_1 = T_2 \mid T_1 < T_2 \mid K \;\mathsf{dvd}\; T$$
$$B ::= x \mid \mathbf{0} \mid \mathbf{1} \mid B_1 \cup B_2 \mid B_1 \cap B_2 \mid B^c$$
$$T ::= k \mid K \mid \mathsf{MAXC} \mid T_1 + T_2 \mid K \cdot T \mid \ \mathsf{|}B\mathsf{|}$$
$$K ::= \ldots -2 \mid -1 \mid 0 \mid 1 \mid 2 \ldots$$

Fig. 1. Quantifier-Free Boolean Algebra with Presburger Arithmetic (QFBAPA)

2 Constructing Small Presburger Arithmetic Formulas

Given a QFBAPA formula, this section shows how to construct a polynomially larger quantifier-free PA formula. Section 3 then proves that this formula is equisatisfiable with the original one.

Consider an arbitrary QFBAPA formula in the syntax of Figure 1. To analyze the problem, we first separate PA and BA parts of the formula by replacing $b_1 = b_2$ with $b_1 \subseteq b_2 \wedge b_2 \subseteq b_1$, replacing $b_1 \subseteq b_2$ with $|b_1 \cap b_2^c| = 0$, and then introducing integer variables k_i for all cardinality expressions $|b_i|$ occurring in the formula. With a constant increase in size, we obtain an equisatisfiable QFBAPA formula of the form $G \wedge F$ where G is a quantifier-free PA formula and F is of the form

$$\bigwedge_{i=0}^{p} |b_i| = k_i \tag{1}$$

We assume $b_0 = \mathbf{1}$ and $k_0 = \mathsf{MAXC}$, i.e., the first constraint is $|\mathbf{1}| = \mathsf{MAXC}$.

Let $y_1, \ldots, y_e$ be the set variables in $b_1, \ldots, b_p$. If we view each Boolean algebra formula b_i as a propositional formula, then for $\beta = (p_1, \ldots, p_e)$ where $p_i \in \{0,1\}$ let $[\![b_i]\!]_\beta \in \{0,1\}$ denote the truth value of b_i under the propositional valuation assigning the truth value p_i to the variable y_i. Let further s_β denote the Venn region associated with β, given by $s_\beta = \cap_{j=1}^{e} y_j^{p_j}$ where $y_j^0 = y_j^c$ is set complement and $y_j^1 = y_j$. We then have $|b_i| = \sum_{\beta \models b_i} |s_\beta|$. For the sake of analysis, for each $\beta \in \{0,1\}^e$ introduce a non-negative integer variable l_β denoting $|s_\beta|$. Then (1) is equisatisfiable with the exponentially larger PA formula

$$\bigwedge_{i=0}^{p} \sum \left\{ l_\beta \mid \beta \in \{0,1\}^e \wedge [\![b_i]\!]_\beta = 1 \right\} = k_i \tag{2}$$

Instead of this exponentially large formula where β ranges over all 2^e propositional assignments, we will check the satisfiability of a smaller formula

$$G \ \wedge \ \bigwedge_{i=0}^{p} \sum \left\{ l_\beta \mid \beta \in \{\beta_1, \ldots, \beta_N\} \wedge [\![b_i]\!]_\beta = 1 \right\} = k_i \tag{3}$$

where β ranges over a set of N assignments $\beta_1, \ldots, \beta_N$ for $\beta_i = (p_{i1}, \ldots, p_{ie})$ and p_{ij} are fresh free variables ranging over $\{0,1\}$. Let $d = p + 1$. We are interested in the best upper bound $N(d)$ on the number of non-zero Venn regions over all possible systems of equations. In the sequel we show that $N(d)$ is polynomial and therefore polynomial in the size of the original QFBAPA formula. This result will prove that QFBAPA is in NP and give an effective bound on how to construct a quantifier-free PA formula for checking the satisfiability of a given QFBAPA formula.

Some details on PA encoding of QFBAPA. We next provide some details on the encoding of the formula (3) in quantifier-free PA, to convince the reader that the resulting formula is indeed polynomially large as long as N is polynomial in d. Let $c_{ij} = [\![b_i]\!]_{\beta_j}$ for $1 \leq i \leq p$ and $1 \leq j \leq N$. Then we need to express in quantifier-free PA the sum $\sum_{j=1}^{N} c_{ij} l_{\beta_j} = k_i$. It suffices to show how to efficiently express sums with boolean variable (as opposed to constant) coefficients. We illustrate this encoding for our particular example. Introduce variables s_{ij} whose purpose is to store the value of the partial sum $s_{ij} = \sum_{k=1}^{j} c_{ik} l_{\beta_k}$. Introduce formula $s_{i0} = 0$ as well as

$$\begin{array}{l}(p \leftrightarrow [\![\lceil b_i \rceil]\!]_{\beta_j}) \wedge \\ (p \rightarrow s_{ij} = s_{i(j-1)} + l_{\beta_j}) \wedge \\ (\neg p \rightarrow s_{ij} = s_{i(j-1)})\end{array} \qquad (D_{ij})$$

where $[\![\lceil b_i \rceil]\!]_{\beta_j}$ denotes the propositional formula corresponding to b_i with propositional variables of β_j substituted for the corresponding sets. We therefore obtain dN polynomially sized expressions (D_{ij}), so if N is polynomial in d, the entire formula (3) is polynomial.

3 Upper Bound on the Number of Non-Zero Venn Regions

We next prove that the number of non-zero Venn regions can be assumed to be polynomial in d. Let $\mathbb{Z}$ denote the set of integers and $\mathbb{Z}_{\geq 0}$ denote the set of non-negative integers. We write $\sum X$ for $\sum\limits_{y \in X} y$.

Definition 1. *For $X \subseteq \mathbb{Z}^d$ a set of integer vectors, let*

$$\text{int\_cone}(X) = \{\lambda_1 x_1 + \ldots + \lambda_t x_t \mid t \geq 0 \wedge x_1, \ldots, x_t \in X \wedge \lambda_1, \ldots, \lambda_n \in \mathbb{Z}_{\geq 0}\}$$

denote the set of all non-negative integer linear combination of vectors from X.

To prove the bound on the number N of non-empty Venn regions from Section 2, we use a variation of the following result, established as Theorem 1(ii) in [10].

Fact 1 (Eisenbrand, Shmonina (2005)) *Let $X \subseteq \mathbb{Z}^d$ be a finite set of integer vectors and $M = \max\{(\max_{i=1}^{d} |x_j^i|) \mid (x_j^1, \ldots, x_j^d) \in X\}$ be the bound on the coordinates of vectors in X. If $b \in \text{int\_cone}(X)$, then there exists a subset $\tilde{X} \subseteq X$ such that $b \in \text{int\_cone}(\tilde{X})$ and $|\tilde{X}| \leq 2d \log(4dM)$.*

To apply Fact 1 to formula (2), let $X = \{x_\beta \mid \beta \in \{0,1\}^e\}$ where $x_\beta \in \{0,1\}^e$ is given by

$$x_\beta = (\llbracket b_0 \rrbracket_\beta, \llbracket b_1 \rrbracket_\beta, \ldots, \llbracket b_e \rrbracket_\beta).$$

Fact 1 implies is that if $(k_0, k_1, \ldots, k_p) \in \text{int\_cone}(X)$ where k_i are as in formula (2), then $(k_0, k_1, \ldots, k_p) \in \text{int\_cone}(\tilde{X})$ where $|\tilde{X}| = 2d \log(4d)$ (note that $M = 1$ because x_β are $\{0,1\}$-vectors). The subset $\tilde{X}$ corresponds to selecting a polynomial subset of N Venn region cardinality variables l_β and assuming that the remaining ones are zero. This implies that formulas (2) and (3) are equisatisfiable.

A direct application of Fact 1 yields $N = 2d \log(4d)$ bound, which is sufficient to prove that QFBAPA is in NP. However, because this bound is not tight, in the sequel we prove results that slightly strengthen the bound and provide additional insight into the problem.

4 Properties of Nonredundant Integer Cone Generators

Definition 2. *Let X be a set of integer vectors. We say that X is a* nonredundant integer cone generator *for b, and write $NICG(X, b)$, if $b \in \text{int\_cone}(X)$, and for every $y \in X$, $b \notin \text{int\_cone}(X \setminus \{y\})$.*

Lemma 1 says that if $\text{NICG}(X, b)$ for some b, then the sums of vectors $\sum Y$ for $Y \subseteq X$ are uniquely generated elements of $\text{int\_cone}(X)$.

Lemma 1. *Suppose $NICG(X, b)$. If $\lambda_1, \lambda_2 : X \to \mathbb{Z}_{\geq 0}$ are non-negative integer coefficients for vectors in X such that*

$$\sum_{x \in X} \lambda_1(x) x = \sum_{x \in X} \lambda_2(x) x \tag{4}$$

and $\lambda_1(x) \in \{0,1\}$ for all $x \in X$, then $\lambda_2 = \lambda_1$.

Proof. Suppose $\text{NICG}(X, b)$, $\lambda_1, \lambda_2 : X \to \mathbb{Z}_{\geq 0}$ are such that (4) holds and $\lambda_1(x) \in \{0,1\}$ for all $x \in X$, but $\lambda_2 \neq \lambda_1$. If there are vectors x on the left-hand side of (4) that also appear on the right-hand side, we can cancel them. We obtain an equality of the form (4) for distinct λ_1', λ_2' with the additional property that $\lambda_1'(x) = 1$ implies $\lambda_2'(x) = 0$. Moreover, not all $\lambda_1'(x)$ are equal to zero. By $b \in \text{int\_cone}(X)$, let $\lambda : X \to \mathbb{Z}_{\geq 0}$ be such that $b = \sum_{x \in X} \lambda(x) x$. Let x_0 be such that $\lambda_1'(x_0) = \min\{\lambda(x) \mid \lambda_1'(x) = 1\}$. By construction, $\lambda_1'(x_0) = 1$ and $\lambda_2'(x_0) = 0$. We then have, with x in sums ranging over X:

$$\begin{aligned}
b &= \sum_{\lambda_1'(x)=1} \lambda(x) x + \sum_{\lambda_1'(x)=0} \lambda(x) x \\
&= \sum_{\lambda_1'(x)=1} (\lambda(x) - \lambda(x_0)) x + \lambda(x_0) \sum_{\lambda_1'(x)=1} x + \sum_{\lambda_1'(x)=0} \lambda(x) x \\
&= \sum_{\lambda_1'(x)=1} (\lambda(x) - \lambda(x_0)) x + \lambda(x_0) \sum \lambda_2'(x) x + \sum_{\lambda_1'(x)=0} \lambda(x) x
\end{aligned}$$

In the last sum, the coefficient next to x_0 is zero in all three terms. We conclude $b \in \text{int\_cone}(X \setminus \{x_0\})$, contradicting $\text{NICG}(X, b)$.

We write NICG(X) as a shorthand for NICG$(X, \sum X)$. Theorem 1 gives several equivalent characterizations of NICG(X).

Theorem 1. *Let $X \subseteq \{0,1\}^d$. The following statements are equivalent:*

1) *there exists a vector $b \in \mathbb{Z}^d_{\geq 0}$ such that NICG(X, b);*
2) *If $\lambda_1, \lambda_2 : X \to \mathbb{Z}_{\geq 0}$ are non-negative integer coefficients for vectors in X such that*
$$\sum_{x \in X} \lambda_1(x)x = \sum_{x \in X} \lambda_2(x)x$$
and $\lambda_1(x) \in \{0,1\}$ for all $x \in X$, then $\lambda_2 = \lambda_1$.
3) *For $\{x_1, \ldots, x_n\} = X$ (for $x_1, \ldots, x_n$ distinct), the system of d equations expressed in vector form as*
$$\lambda(x_1)x_1 + \ldots + \lambda(x_n)x_n = \sum X \tag{5}$$
has $(\lambda(x_1), \ldots, \lambda(x_n)) = (1, \ldots, 1)$ as the unique solution in $\mathbb{Z}^n_{\geq 0}$.
4) *NICG(X).*

Proof. 1) → 2): This is Lemma 1.

2) → 3): Assume 2) and let $\lambda_1(x_i) = 1$ for $1 \leq i \leq n$. For any solution λ_2 we then have $\sum_{x \in X} \lambda_1(x)x = \sum_{x \in X} \lambda_2(x)x$, so $\lambda_2 = \lambda_1$. Therefore, λ_1 is the unique solution.

3) → 4): Assume 3). Clearly $\sum X \in \text{int\_cone}(X)$; it remains to prove that X is minimal. Let $y \in X$. For the sake of contradiction, suppose $\sum X \in \text{int\_cone}(X \setminus \{y\})$. Then there exists a solution $\lambda(x)$ for (5) with $\lambda(y) = 0 \neq 1$, a contradiction with the uniqueness of the solution.

4) → 1): Take $b = \sum X$.

Corollary 1 is used in [10] to establish the bound on the size of X with NICG(X). We obtain it directly from Lemma 1 taking $\lambda_2(x) \in \{0,1\}$.

Corollary 1. *If NICG(X) then for $Y_1, Y_2 \subseteq X$, $Y_1 \neq Y_2$ we have $\sum Y_1 \neq \sum Y_2$.*

The following lemma says that it suffices to establish bounds on the cardinality of X such that NICG(X), because they give bounds on all X.

Lemma 2. *If $b \in \text{int\_cone}(X)$, then there exists a subset $\tilde{X} \subseteq X$ such that $b \in \text{int\_cone}(\tilde{X})$ and NICG$(\tilde{X}, b)$.*

Proof. If $b \in \text{int\_cone}(X)$ then by definition $b \in \text{int\_cone}(X_0)$ for a finite $X_0 \subseteq X$. If not NICG(X_0, b), then $b \in \text{int\_cone}(X_1)$ where X_1 is a proper subset of X_0. Continuing in this fashion we obtain a sequence $X_0 \supset X_1 \supset \ldots \supset X_k$ where $k \leq |X_0|$. The last element X_k satisfies NICG(X_k, b).

Moreover, the property NICG(X) is hereditary, i.e. it applies to all subsets of a set that has it.[1]

[1] The reader familiar with matroids [32] might be interested to know that, for $d \geq 4$, the family of sets $\{X \subseteq \{0,1\}^d \mid \text{NICG}(X)\}$ is not a matroid, because it contains multiple subset-maximal elements of different cardinality.

Lemma 3. *If NICG(X) and $Y \subseteq X$, then NICG(Y).*

Proof. Suppose that NICG(X) and $Y \subseteq X$ but not NICG$(Y, \sum Y)$. Because $\sum Y \in \text{int\_cone}(X)$, there is $z \in Y$ such that $\sum Y \in \text{int\_cone}(Y \setminus \{z\})$. Then also $\sum Y \in \text{int\_cone}(X \setminus \{z\})$, contradicting Lemma 1.

The following theorem gives our bounds on $|X|$. As in [10], we only use Corollary 1 instead of the stronger Lemma 1, suggesting that the bound is not tight.

Theorem 2. *Let $X \subseteq \{0,1\}^d$ and NICG(X). Then*

$$|X| \leq (1 + \varepsilon(d))(d \log d) \tag{6}$$

where $\varepsilon(d) \leq 1$ for all $d \geq 1$, and $\lim_{d\to\infty} \varepsilon(d) = 0$.

Proof. Let $X \subseteq \{0,1\}^d$, NICG(X) and $N = |X|$. We first prove $2^N \leq (N+1)^d$. Suppose that, on the contrary, $2^N > (N+1)^d$. If $\sum Y = (x^1, \ldots, x^d)$ for $Y \subseteq X$, then $0 \leq x^j \leq N$ because $Y \subseteq \{0,1\}^d$ and $|Y| \leq N$. Therefore, there are only $(N+1)^d$ possible sums $\sum Y$. Because there are 2^N subsets $Y \subseteq X$, there exist two distinct subsets $U, V \in 2^X$ such that $\sum U = \sum V$. This contradicts Corollary 1. Therefore, $2^N \leq (N+1)^d$, so $N \leq d \log(N+1)$.

We first show that this implies $N \leq 2d \log(2d)$. We show the contrapositive. Suppose $N > 2d \log(2d)$. Then $\frac{N}{2d} > \log(2d)$ from which we have:

$$1 < \frac{2^{\frac{N}{2d}}}{2d} \tag{7}$$

Moreover, $d \geq 1$ so $\frac{N}{2d} > \log(2d) \geq \log 2 = 1$, which implies

$$\log(1 + \frac{N}{2d}) \leq \frac{N}{2d} \tag{8}$$

From (7) and (8) we have, similarly to [10],

$$\begin{aligned} d \log(N+1) &< d \log(N \tfrac{2^{\frac{N}{2d}}}{2d} + 1) = d \log(2^{\frac{N}{2d}}(\tfrac{N}{2d} + 2^{-\frac{N}{2d}})) < d \log(2^{\frac{N}{2d}}(\tfrac{N}{2d} + 1)) \\ &= d(\tfrac{N}{2d} + \log(1 + \tfrac{N}{2d})) < d(\tfrac{N}{2d} + \tfrac{N}{2d}) = N. \end{aligned}$$

By contraposition, from $N \leq d \log(N+1)$ we conclude $N \leq 2d \log(2d)$. Substituting this bound on N back into $N \leq d \log(N+1)$ we obtain

$$\begin{aligned} N &\leq d \log(N+1) \leq d \log(2d \log(2d) + 1) = d \log(2d(\log(2d) + \tfrac{1}{2d})) \\ &= d(1 + \log d + \log(\log(2d) + \tfrac{1}{2d})) = d \log d(1 + \tfrac{1 + \log(\log(2d) + \frac{1}{2d})}{\log d}) \end{aligned}$$

so we can let

$$\varepsilon(d) = \frac{1 + \log(\log d + 1 + \frac{1}{2d})}{\log d}.$$

It may be of interest for problems arising in practice that, for $d \leq 23170$ we have $\varepsilon(d) \leq \frac{5}{\log d}$ and thus $N \leq d(\log d + 5)$.

We can now define the function whose bounds we are interested in computing.

Definition 3. $N(d) = \max\{|X| \ \mid \ X \subseteq \{0,1\}^d, \mathit{NICG}(X)\}$

Theorem 2 implies $N(d) \leq (1+\varepsilon(d))(d \log d)$.

5 Notes on Lower Bounds and Set Algebra with Real Measures

While we currently do not have a tight lower bound on $N(d)$, in this section we show, in sequence, the following:

1. $d \leq N(d)$ for all d;
2. $N_R(d) = d$ if we use real variables instead of integer variables;
3. $N(d) = d$ for $d \in \{1,2,3\}$;
4. for $d + \lfloor \frac{d}{4} \rfloor \leq N(d)$ for $4 \leq d$.

We first show $d \leq N(d)$.

Lemma 4. *Let* $X = \{(x_i^1, \ldots, x_i^d) \mid 1 \leq i \leq n\}$ *and*

$$X^+ = \{(x_i^1, \ldots, x_i^d, 0) \mid 1 \leq i \leq n\} \cup \{(0, \ldots, 0, 1)\}$$

Then $\mathit{NICG}(X)$ *if and only if* $\mathit{NICG}(X^+)$.

Corollary 2. $N(d) + 1 \leq N(d+1)$ *for all* $d \geq 1$.

Proof. Let $X \subseteq \{0,1\}^d$, NICG(X), and $|X| = N(d)$. Then NICG(X^+) by Lemma 4 and $|X^+| = N(d) + 1$, which implies $N(d+1) \geq N(d) + 1$.

Note that we have $N(1) = 1$ because there is only one non-zero $\{0,1\}$ vector in one dimension. From Corollary 2 we obtain our lower bound, with standard basis as NICG.

Lemma 5. $d \leq N(d)$. *Specifically,* $\mathit{NICG}(\{e_1, \ldots, e_d\})$.

Note that for $X = \{e_1, \ldots, e_d\}$ we have int_cone$(X) = \mathbb{Z}^d_{\geq 0}$, which implies that X is a *maximal* NICG, in the sense that no proper superset $W \supset X$ for $W \subseteq \{0,1\}^d$ has the property NICG(W).

Real-valued relaxation of QFBAPA. It is interesting to observe that, for a variation of the QFBAPA problem over *real numbers*, which we call QFBALA (Quantifier-Free Boolean Algebra with Linear Arithmetic), we have $N'(d) = d$ as a lower *and upper* bound for every d.

We define QFBALA similarly as QFBAPA, but we use real (or rational) linear arithmetic instead of integer linear arithmetic and we interpret $|A|$ is some real-valued measure of the set A. A possible application of QFBALA are generalizations of probability consistency problems such as [5, Page 385, Example 8.3]. Set algebra operations then correspond to the σ-algebra of events, and the measure of the set is the probability of the event. Another model of QFBALA is to interpret sets as finite disjoint unions of intervals contained in $[0, 1]$, and let $|A|$ be the sum of the lengths of the disjoint intervals making up A.

The conditions we are using on the models are 1) for two disjoint sets A, B, we have $|A \cup B| = |A| + |B|$, and 2) if $|C| = p$ and $0 \leq q \leq p$, then there exists $B \subseteq C$ such that $|B| = q$. (In addition, if the model allows $|A| = 0$ for $A \neq \emptyset$, then we introduce an additional propositional variable for each Venn region variable to track its emptiness.)

We can reduce the satisfiability of QFBALA to the satisfiability of a quantifier-free linear arithmetic formula over reals and a formula of the form (2) but with l_β non-negative real values instead of integer values. We then reduce formula (2) to a formula of the form (2). The question is then, what can we use as the bound $N'(d)$ for QFBALA problems? This question reduces to following. Define convex cone generated by a set of vectors by

$$\text{cone}(X) = \{\lambda_1 x_1 + \ldots + \lambda_t x_t \mid t \geq 0 \wedge x_1, \ldots, x_t \in X \wedge \lambda_1, \ldots, \lambda_n \geq 0\}$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are non-negative real coefficients. If $b \in \text{cone}(X)$, what bound can we put on the cardinality of a subset $\tilde{X} \subseteq X$ such that $X \in \text{cone}(\tilde{X})$? Note that d is a lower bound, using the same example of unit vectors as X. In the case of real numbers, Carathéodory's theorem [8] states that d is an upper bound as well: $b \in \text{cone}(\tilde{X})$ for some $\tilde{X}$ of cardinality at most d. We can also explain that $N'(d) = d$ using the terminology of linear programming [30]. The equations (2) along with $l_\beta \geq 0$ for $\beta \in \{0, 1\}^e$ determine a polytope in $\mathbb{R}^{2^e}$, so if they have a solution, they have a solution that is a vertex of the polytope. The vertex in $\mathbb{R}^{2^e}$ is the intersection of 2^e hyperplanes, of which at most d are given by (2), so the remaining ones must be hyperplanes of the form $l_\beta = 0$. This implies that at least $2^e - d$ coordinates of the vertex are zero and at most d of them can be non-zero.

Note that QFBALA is a relaxation of QFBAPA, and can be used as a sound (but incomplete) method for proving the absence of solutions of a QFBAPA formula.

$N(d) = d$ for $d \in \{1, 2, 3\}$. We next show that for $d \in \{1, 2, 3\}$ not only $d \leq N(d)$ but also $N(d) \leq d$.

Lemma 6. *$N(d) = d$ for $d \in \{1, 2, 3\}$.*

Proof. By Corollary 2, if $N(d + 1) = d + 1$, then $N(d) + 1 \leq d + 1$ so $N(d) \leq n$. Therefore, $N(d) = 3$ implies $N(2) = 2$ as well, so we can take $d = 3$.

If $N(d) > d$, then there exists a set X with $\text{NICG}(X)$ and $|X| > d$. From Lemma 3, a subset $X_0 \subseteq X$ with $|X| = d+1$ also satisfies $\text{NICG}(X_0)$. Therefore, $N(3) = 3$ is equivalent to showing that there is no set $X \subseteq \{0, 1\}^3$ with $\text{NICG}(X)$ and $|X| = 4$.

Consider a possible counterexample $X = \{x_1, x_2, x_3, x_4\} \subseteq \{0, 1\}^3$ with $b \in X$. By previous argument on real-value relaxation, $N'(3) = 3$, so b is in convex cone of some three vectors from X, say $b \in \text{cone}(\{x_1, x_3, x_3\})$. On the other hand, $b \notin \text{int\_cone}(\{x_1, x_3, x_3\})$. If we consider a system $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = b$ this implies that such system has solution over non-negative reals, but not over non-negative integers. This can only happen if in the process of Gaussian elimination we obtain coefficients whose absolute value is more than 1. The only set of three vectors for which this can occur is $X_1 = \{(0,1,1),(1,0,1),(1,1,0)\}$ We then consider all possibilities for the fourth vector in X, which, modulo symmetry of coordinates are $(0,0,0)$, $(1,1,1)$, $(1,1,0)$, and $(1,0,0)$. However, adding any of these vectors violates the uniqueness of the solution to $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4 = \sum X$, so NICG$(X)$ does not hold by Theorem 1, condition 3).

$N = \frac{5}{4}d - \frac{3}{4}$ lower bound. I next show that there exists an example $X_5 \subseteq \{0,1\}^4$ with NICG(X_5) and $|X_5| = 5$. From this it follows that $N(d) > d$ for all $d \geq 4$.

Consider the following system of 4 equations with 5 variables, where all variable coefficients are in $\{0,1\}$. (I found this example by narrowing down the search using the observations on minimal counterexamples in the proof of Lemma 6.)

$$\begin{aligned}
\lambda_1 + \lambda_2 + \lambda_3 \qquad\qquad &= 3 \\
\lambda_2 + \lambda_3 + \lambda_4 \qquad &= 3 \\
\lambda_1 \qquad + \lambda_3 + \lambda_4 + \lambda_5 &= 4 \\
\lambda_1 + \lambda_2 \qquad + \lambda_4 + \lambda_5 &= 4
\end{aligned} \tag{9}$$

Performing Gaussian elimination yields an equivalent upper-triangular system

$$\begin{aligned}
\lambda_1 + \lambda_2 + \lambda_3 \qquad\qquad &= 3 \\
\lambda_2 + \lambda_3 \ + \lambda_4 \qquad &= 3 \\
\lambda_3 + 2\lambda_4 \ + \lambda_5 &= 4 \\
3\lambda_4 + 2\lambda_5 &= 5
\end{aligned}$$

From this form it easy to see that the system has $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = (1,1,1,1,1)$ as *the only solution* in the space of non-negative integers. Note that all variables are non-zero in this solution. (In contrast, as discussed above, because the system is satisfiable, it must have a solution in non-negative reals where at most 4 coordinates are non-zero; an example of such solution is $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = (0, 1.5, 1.5, 0, 2.5)$.) The five columns of the system (9) correspond to the set of vectors $X_5 = \{(1,0,1,1),(1,1,0,1),(1,1,1,0),(0,1,1,1),(0,0,1,1)\}$ such that NICG(X_5). The set X_5 is also a maximal NICG, because adding any of the remaining 9 non-zero vectors in $\{0,1\}^4 \setminus X_5$ results in a set that is not NICG.

Using k identical copies of X_5 (with 4 equations in a group mentioning a disjoint set of 5 variables) we obtain systems of $4k$ equations with $5k$ variables such that the only solution is a vector $(1,\ldots,1)$ of all ones. By adding p unit vector columns for $1 \leq p \leq 3$, we also obtain systems of $4k+p$ equations with $5k+p$ variables, with

$$N = \frac{5}{4}d - \frac{p}{4} = d + \left\lfloor \frac{d}{4} \right\rfloor \geq \frac{5}{4}d - \frac{3}{4}$$

which, in particular, shows that $N = d$ upper bound is invalid for all $d \geq 4$.

This argument shows that there exist maximal NICG of size larger than d for $d \geq 4$. As we have remarked before, the set of d unit vectors is a maximal NICG for every d, which means that, unlike linearly independent sets of vectors over a field or other independent sets in a matroid [32], there are maximal NICG sets of different cardinality.

Note also that X_5 is not a Hilbert basis [31]. Namely, we have that $(1, 1, 1, 1) \in \mathrm{cone}(X_5) \setminus \mathrm{int\_cone}(X_5)$ because

$$(1, 1, 1, 1) = 1/3((1, 0, 1, 1) + (1, 1, 0, 1) + (1, 1, 1, 0) + (0, 1, 1, 1)).$$

This illustrates why previous results on Hilbert bases do not directly apply to the notion of NICG.

6 A decision procedure for QFBAPA

Using Theorem 2 we obtain a non-deterministic polynomial-time algorithm for checking QFBAPA satisfiability. For formulas generated from verification, it is likely that a QFBAPA decision procedure implementation can effectively use bounds smaller than $(1+\varepsilon(d))d \log d$ to find counterexamples and to prove their absence, as follows.

1. Attempt to find counterexamples for small N. If a counterexample for any N is found, it is a valid counterexample. One could expect that such counterexamples would often be found by "small scope hypothesis" [13] for typical formulas arising in software verification.
2. If no counterexample is found for small N, then the decision procedure can use the bound $N = d$ with real linear arithmetic and try to prove the absence of solutions. No solutions found means that the original QFBAPA problem has no solutions either. The examples from [18] and the experience from [9, Section 8] suggest that this approach would often succeed in proving the absence of solutions for unsatisfiable QFBAPA formulas.
3. Finally, if a solution is found in real numbers but not for small N in integers, then the system can use the bound $N = (1 + \varepsilon(d))d \log d$, which gives a definite answer thanks to Theorem 2.

The first two steps can be viewed as heuristics for finding the answer faster in common cases; their usefulness remains to be experimentally evaluated.

7 Related Work

To our knowledge, our result is the only decision procedure for a logic with sets and cardinality constraints that does not explicitly construct all set partitions. Using a new form of small model property, the "small number of non-zero variables property", we obtained a non-deterministic polynomial-time algorithm that can be solved by producing polynomially large quantifier-free Presburger arithmetic formulas. A polynomial bound sufficient for our result can be derived from [10]. In addition to slight improvements in the bounds, we introduced the notion of nonredundant integer cone generators and proved additional results that may help us understand their properties and eventually establish tight bounds on their size. We note that previous results such as [31] consider matroids and Hilbert bases. In contrast, nonredundant integer cone generators are the natural notion for our problem. As we remark in Section 5, the sets of vectors X with $\mathrm{NICG}(X)$ do not form a matroid, and maximal $\mathrm{NICG}(X)$ need not be a Hilbert basis. Note also that the equations generated from QFBAPA problems are more difficult than set packing and set partitioning problems [2] because integer variables are not restricted to be $\{0, 1\}$.

Presburger arithmetic. The original result on decidability of PA is [28]. The space bound for PA was shown in [12]. The matching lower and upper bounds for PA were shown in [4], see also [14, Lecture 24].

Reasoning about Sets. The first results on decidability of BA of sets are from [22], [1, Chapter 4] and use quantifier elimination, from which one can derive small model property. [14] gives the complexity of the satisfiability problem for arbitrary BA. [25] study unification in Boolean rings. The quantifier-free fragment of BA is shown NP-complete in [24]; see [20] for a generalization of this result using the parameterized complexity of the Bernays-Schönfinkel-Ramsey class of first-order logic [6, Page 258]. [7] gives an overview of several fragments of set theory including theories with quantifiers but no cardinality constraints and theories with cardinality constraints but no quantification over sets. The decision procedure for quantifier-free fragment with cardinalities in [7, Chapter 11] introduces exponentially many integer variables to reduce the problem to PA.

References

1. W. Ackermann. *Solvable Cases of the Decision Problem*. North Holland, 1954.
2. Egon Balas and Manfred W. Padberg. Set partitioning: A survey. *SIAM Review*, 18(4):710–760, 1976. http://links.jstor.org/sici?sici=0036-1445%28197610%2918%3A4%3C710%3ASPAS%3E2.0.CO%3B2-4.
3. Clark Barrett and Sergey Berezin. CVC Lite: A new implementation of the cooperating validity checker. In *Proc. 16^{th} Int. Conf. on Computer Aided Verification (CAV '04)*, volume 3114 of *Lecture Notes in Computer Science*, pages 515–518, 2004.
4. Leonard Berman. The complexity of logical theories. *Theoretical Computer Science*, 11(1):71–77, 1980.
5. Dimitris Bertsimas and John N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, Belmont, Massachusetts, 1997.
6. Egon Börger, Erich Grädel, and Yuri Gurevich. *The Classical Decision Problem*. Springer-Verlag, 1997.
7. Domenico Cantone, Eugenio Omodeo, and Alberto Policriti. *Set Theory for Computing*. Springer, 2001.
8. W. J. Cook, J. Fonlupt, and A. Schrijver. An integer analogue of Carathéodory's theorem. *Journal of Combinatorial Theory, Series B*, 40(63–70), 1986.

9. David Detlefs, Greg Nelson, and James B. Saxe. Simplify: A theorem prover for program checking. Technical Report HPL-2003-148, HP Laboratories Palo Alto, 2003.
10. Friedrich Eisenbrand and Gennady Shmonina. Carathéodory bounds for integer cones. *Operations Research Letters*, 34(5):564–568, September 2006. http://dx.doi.org/10.1016/j.orl.2005.09.008.
11. S. Feferman and R. L. Vaught. The first order properties of products of algebraic systems. *Fundamenta Mathematicae*, 47:57–103, 1959.
12. Jeanne Ferrante and Charles W. Rackoff. *The Computational Complexity of Logical Theories*, volume 718 of *Lecture Notes in Mathematics*. Springer-Verlag, 1979.
13. Daniel Jackson. *Software Abstractions: Logic, Language, & Analysis*. MIT Press, 2006.
14. Dexter Kozen. Complexity of boolean algebras. *Theoretical Computer Science*, 10:221–247, 1980.
15. Dexter Kozen. *Theory of Computation*. Springer, 2006.
16. Viktor Kuncak. *Modular Data Structure Verification*. PhD thesis, EECS Department, Massachusetts Institute of Technology, February 2007.
17. Viktor Kuncak, Hai Huu Nguyen, and Martin Rinard. An algorithm for deciding BAPA: Boolean Algebra with Presburger Arithmetic. In *20th International Conference on Automated Deduction, CADE-20*, Tallinn, Estonia, July 2005.
18. Viktor Kuncak, Hai Huu Nguyen, and Martin Rinard. Deciding Boolean Algebra with Presburger Arithmetic. *J. of Automated Reasoning*, 2006. http://dx.doi.org/10.1007/s10817-006-9042-1.
19. Viktor Kuncak and Martin Rinard. The first-order theory of sets with cardinality constraints is decidable. Technical Report 958, MIT CSAIL, July 2004.
20. Viktor Kuncak and Martin Rinard. Decision procedures for set-valued fields. In *1st International Workshop on Abstract Interpretation of Object-Oriented Languages (AIOOL 2005)*, 2005.
21. Shuvendu K. Lahiri and Sanjit A. Seshia. The UCLID decision procedure. In *CAV'04*, 2004.
22. L. Loewenheim. Über Mögligkeiten im Relativkalkül. *Math. Annalen*, 76:228–251, 1915.
23. Bruno Marnette, Viktor Kuncak, and Martin Rinard. On algorithms and complexity for sets with cardinality constraints. Technical report, MIT CSAIL, August 2005.
24. Kim Marriott and Martin Odersky. Negative boolean constraints. Technical Report 94/203, Monash University, August 1994.
25. Ursula Martin and Tobias Nipkow. Boolean unification: The story so far. *Journal of Symbolic Computation*, 7(3):275–293, 1989.
26. Hans Jürgen Ohlbach and Jana Koehler. How to extend a formal system with a boolean algebra component. In W. Bibel P.H. Schmidt, editor, *Automated Deduction. A Basis for Applications*, volume III, pages 57–75. Kluwer Academic Publishers, 1998.
27. Christos H. Papadimitriou. On the complexity of integer programming. *J. ACM*, 28(4):765–768, 1981.
28. M. Presburger. Über die vollständigkeit eines gewissen systems der aritmethik ganzer zahlen, in welchem die addition als einzige operation hervortritt. In *Comptes Rendus du premier Congrès des Mathématiciens des Pays slaves, Warsawa*, pages 92–101, 1929.

29. Peter Revesz. Quantifier-elimination for the first-order theory of boolean algebras with linear cardinality constraints. In *Proc. Advances in Databases and Information Systems (ADBIS'04)*, 2004.
30. Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.
31. András Sebö. Hilbert bases, Caratheodory's theorem and combinatorial optimization. In R. Kannan and W. Pulleyblank, editors, *Integer Programming and Combinatorial Optimization I*. University of Waterloo Press, 1990.
32. H. Whitney. On the abstract properties of linear independence. *American Journal of Mathematics*, 57:509–533, 1935.
33. Calogero G. Zarba. A quantifier elimination algorithm for a fragment of set theory involving the cardinality operator. In *18th International Workshop on Unification*, 2004.
34. Calogero G. Zarba. Combining sets with cardinals. *J. of Automated Reasoning*, 34(1), 2005.