On Finding Small 2-Generating Sets

Isabelle Fagnot, Guillaume Fertin, Stéphane Vialette

On Finding Small 2-Generating Sets

Isabelle Fagnot[1], Guillaume Fertin[2], and Stéphane Vialette[3]

[1] IGM-LabInfo, CNRS UMR 8049, Université Paris-Est,
5 Bd Descartes 77454 Marne-la-Vallée, France
and Université Paris Diderot, Paris 7, France
fagnot@univ-mlv.fr
[2] Laboratoire d'Informatique de Nantes-Atlantique (LINA), UMR CNRS 6241
Université de Nantes, 2 rue de la Houssinière, 44322 Nantes Cedex 3 - France
fertin@lina.univ-nantes.fr
[3] IGM-LabInfo, CNRS UMR 8049, Université Paris-Est,
5 Bd Descartes 77454 Marne-la-Vallée, France
vialette@univ-mlv.fr

Abstract. Given a set of positive integers S, we consider the problem of finding a minimum cardinality set of positive integers X (called a *minimum 2-generating set of* S) s.t. every element of S is an element of X or is the sum of two (non-necessarily distinct) elements of X. We give elementary properties of 2-generating sets and prove that finding a minimum cardinality 2-generating set is hard to approximate within ratio $1+\varepsilon$ for any $\varepsilon > 0$. We then prove our main result, which consists in a representation lemma for minimum cardinality 2-generating sets.

1 Introduction

In this paper, we consider the problem of 2-generating a set of positive integers S with a minimum cardinality set of integers X, where X is said to 2-*generate* S if every element of S is an element of X or is the sum of two (non-necessarily distinct) elements of X. We note that, in this context, X does not have to be a subset of S. We refer to this problem as Minimum 2-Generating Set.

Minimum 2-Generating Set is a simple restriction of Minimum Generating Set (a natural problem in number theory) [4]. The Minimum Generating Set problem is defined as follows: Given a set of positive integers S, find a minimum cardinality set of integers X such that every element of S is the sum of a subset of X. Minimum Generating Set has been shown to be **NP**-hard [4], and is related, among other things, to planning radiation therapy: elements of S represent radiation dosages required at various points, while an element of X represents a dose delivered simultaneously to multiple points. Note also that both Minimum 2-Generating Set and Minimum Generating Set can be seen as natural extensions of the Postage Stamp problem [13].

Strongly related to our work are minimum sum covers of finite Abelian groups as investigated in [9,7]. A subset X of a finite Abelian group G is said to be a *sum cover* of G if $\{x + x' : x, x' \in X\} = G$, a *strict sum cover* of G if $\{x + x' : x, x' \in X \ \wedge x \neq x'\} = G$, and a *difference cover* of G if $\{x - x' : x, x' \in$

$X\} = G$. Swanson [19] gives some constructions and computational results for maximum difference packings of cyclic groups. Haanpää, Huima, and Östergård compute maximum sum and strict sum packings of cyclic groups [10]. Fitch and Jamison [7] give minimum sum and strict sum covers of small cyclic groups, and Wiedemann [20] determines minimum difference covers for cyclic groups of order at most 133.

Another area of research related to our work is the problem of covering a set of strings S with a set X of substrings in S, where X is said to *cover* S if every string in S can be written as a concatenation of the substrings in X [12,2] (see also [14] and [3] for a more general treatment of the combinatorial rank). Covering a set of strings S with a set X of substrings in S is indeed the Minimum Generating Set problem for unary alphabet under the unary encoding scheme. To narrow the context, notice that, given a set of binary strings S, finding a minimum cardinality set X of substrings in S such that every string in S can be written as a concatenation of *at most* two substrings in X is **NP**-complete (the proof being an easy binary alphabet encoding of the result of Néraud [14]). Finally, Hajiaghayi *et al.* [11] considered the Minimum Multicolored Subgraph problem, which can be seen as a generalization of Minimum 2-Generating Set when every integer in the input set is bounded by a polynomial in the length of the input.

This paper is organized as follows: we first recall basic definitions in Section 2, and we then formally introduce the considered problem. In Section 3, we give some elementary properties of 2-generating sets. Section 4 is devoted to prove hardness of Minimum 2-Generating Set and we prove in Section 5 a representation lemma. Notice that some proofs are omitted due to space constraints.

2 Preliminaries

We use $\mathbb{N}^*$ to refer to the set of all natural numbers excluding zero, *i.e.*, $\mathbb{N}^* = \{1, 2, \ldots\}$. Let $S = \{s_1, s_2, \ldots, s_n\} \subset \mathbb{N}^*$. For any $k \in \mathbb{N}^*$, we write kS for the set of all integers that can be expressed as the sum of *exactly* k *non necessarily distinct* integers of S, *i.e.*, $kS = \{s_{i_1} + s_{i_2} + \ldots + s_{i_k} : s_{i_1}, s_{i_2}, \ldots, s_{i_k} \in S\}$. According to this definition, for any set S, $S = 1S$. A set $X \subset \mathbb{N}^*$ is a *k-generating set* of S (or *k-generates* S) if $S \subseteq \bigcup_{i=1}^{k} iX$. (Notice here that we do not require the additional constraint $\bigcup_{i=1}^{k} iX \subseteq S$.) It is called a *minimum* k-generating set of S if X is a k-generating set of S of minimum cardinality. The *k-rank* of S, denoted $\mathrm{rk}_k(S)$, is the cardinality of a minimum k-generating set of S. A set $S \subset \mathbb{N}^*$ is *k-elementary* if $\mathrm{rk}_k(S) = |S|$. Let $\min(S)$ and $\max(S)$ stand for $\min\{s_i : s_i \in S\}$ and $\max\{s_i : s_i \in S\}$, respectively. The *length* of S, denoted $\mathrm{len}(S)$, is defined to be $\mathrm{len}(S) = \max(S) - \min(S)$.

We are now in position to define the Minimum k-Generating Set problem we are interested in: Given a set $S \subset \mathbb{N}^*$, find a k-generating set X of S of minimum cardinality. Actually, our main interest here is in finding minimum 2-generating sets, and hence we shall be concerned in this paper with Minimum 2-Generating Set only. Of particular importance, we assume hereafter any

reasonable (e.g. binary) encoding of any instance of MINIMUM 2-GENERATING SET so that the input is in $O(n \log m)$ space, where $n = |S|$ and $m = \max(S)$.

We assume readers have basic knowledge about graph theory [5] and we shall only recall basic notations. We write $G = (V, E)$ for a graph with *vertex set* V and *edge set* E. For a graph G and a vertex $u \in V$, we write $d_G(u)$ for the *degree* of u in G. A graph is *bipartite* if it does not contain an odd cycle.

3 Elementary properties

Generalities. To fix the context, we begin by giving easy bounds for $\mathrm{rk}_2(S)$.

Lemma 1. *For any $S \subset \mathbb{N}^*$ of cardinality n, $\left\lceil \frac{1}{2}(\sqrt{8n+9}-3) \right\rceil \leq \mathrm{rk}_2(S) \leq n$.*

Proof. The upper bound is trivial. To prove the lower bound, let $X \subset \mathbb{N}^*$ be a 2-generating set of S, and let k stand for $|X|$. For one, $|X \cup 2X| \leq \binom{k}{2} + 2k$. For another, $|X \cup 2X| \geq n$ since X 2-generates S. Combining the two inequalities yields the claimed lemma. □

Combinatorial properties of intervals [8] will prove to be a simple but powerful tool for 2-generating sets. We write $[i : i+j]$ for the set of consecutive integers (*i.e.*, interval) $\{i, i+1, \ldots, i+j\}$. For any interval system $\mathcal{I}$, the *matching number* of $\mathcal{I}$, denoted $\nu(\mathcal{I})$, is the maximum number of pairwise disjoint intervals of $\mathcal{I}$. Let $S = \{s_i : 1 \leq i \leq n\} \subset \mathbb{N}^*$. Define the 2-*generating interval system* of S, in symbols $\mathcal{I}_2(S)$, to be $\mathcal{I}_2(S) = \{[\lceil s_i/2 \rceil \ : \ s_i] : s_i \in S\}$.

Lemma 2. *Let $S \subset \mathbb{N}^*$ and $X \subset \mathbb{N}^*$ be a 2-generating set of S. Then, for every $s \in S$, $|X \cap [\lceil s/2 \rceil \ : \ s]| \neq \emptyset$.*

Proof. Suppose the lemma is false. Then some $s \in S$ is obtained by summing at most 2 integers of X, each upper-bounded by $\lceil s/2 \rceil - 1$. But $2(\lceil s/2 \rceil - 1) < 2((s/2+1)-1) = s$ which yields the desired contradiction. □

Corollary 1. *For any $S \subset \mathbb{N}^*$, $\nu(\mathcal{I}_2(S)) \leq \mathrm{rk}_2(S)$.*

It follows from Lemma 1 that if $\nu(\mathcal{I}_2(S)) = |S|$ then S is 2-elementary. The converse is false as shown by $S = \{7, 8, 9\}$. The following application of Corollary 1 will prove useful in the sequel.

Lemma 3. *Let $A = \{a_i : 1 \leq i \leq n\} \subset \mathbb{N}^*$ be such that (i) $a_1 \geq 4$ and (ii) $a_{i+1} > 4a_i - 3$, $1 \leq i \leq n-1$. Then, the set $S = \{2a_i - 1 : 1 \leq i \leq n\} \cup \{4a_i - 3 : 1 \leq i \leq n\} \subset \mathbb{N}^*$ is 2-elementary.*

Integer arithmetic sequences. An *integer arithmetic sequence* is a sequence of integers such that the difference of any two successive members of the sequence is a constant.

Lemma 4. *Let $S \subset \mathbb{N}^*$ be an integer arithmetic sequence of length n. Then $\mathrm{rk}_2(S) = \Theta(\sqrt{n})$.*

Proof. Write $S = \{s_0 + ic : 0 \leq i \leq n-1\}$ for some $s_0 \in \mathbb{N}^*$ and $c \in \mathbb{N}^*$. Define $X = X_1 \cup X_2$, where $X_1 = \{s_0 + ic\lceil\sqrt{n}\rceil : 0 \leq i \leq \lceil\sqrt{n}\rceil - 1\}$, and $X_2 = \{ic : 1 \leq i \leq \lceil\sqrt{n}\rceil - 1\}$. An easy check shows that $S \subseteq X \cup 2X$, and hence X is a 2-generating set of S. Clearly, $|X_1| = \lceil\sqrt{n}\rceil$ and $|X_2| = \lceil\sqrt{n}\rceil - 1$. Therefore, $|X| \leq 2\lceil\sqrt{n}\rceil - 1 \leq 2(\sqrt{n}+1) - 1 = 2\sqrt{n} + 1$. Combining this with Lemma 1 yields the claimed result. □

In case S is an arithmetic sequence of length $n = k^2$, the above lemma reduces to $\mathrm{rk}_2(S) \leq 2\sqrt{n} - 1$. We also observe that Lemma 4 could be an issue for dealing with dense sets. Define a set $S \subset \mathbb{N}^*$ to be ε-dense if $|S| = \varepsilon\,\mathrm{len}(S)$ for some $\varepsilon > 0$. The following result is an immediate consequence of Lemma 4. We also note that the (easy) proof can be turned into an approximation algorithm with performance ratio $O(\sqrt{\varepsilon})$ for ε-dense sets.

Corollary 2. *Let $S \subset \mathbb{N}^*$ be an ε-dense set of cardinality n. Then $\mathrm{rk}_2(S) = O(\sqrt{n/\varepsilon})$.*

Integer geometric sequences. An *integer geometric sequence* is a sequence of numbers where each term, except the first one, is found by multiplying the previous one by a fixed integer $r \geq 2$ called the *common ratio*. Results turn out to be more precise compared to arithmetic sequences.

Lemma 5. *Let $S \subset \mathbb{N}^*$ be an integer geometric sequence of length n with common ratio $r \geq 2$. Then, (i) $\mathrm{rk}_2(S) = \lceil n/2 \rceil$ if $r = 2$ and (ii) $\mathrm{rk}_2(S) = n$ if $r > 2$.*

Proof. A straightforward application of Corollary 1 proves (ii). To prove (i), write $S = \{s_i : 1 \leq i \leq n\}$ and $S_{\mathrm{odd}} = \{s_i : s_i \in S \ \wedge \ i \equiv 1 \pmod 2\}$. For one, $X = S_{\mathrm{odd}}$ is a 2-generating set of S, and hence $\mathrm{rk}_2(S) \leq |S_{\mathrm{odd}}| = \lceil n/2 \rceil$. For another, $\nu(\mathcal{I}_2(S)) \geq |S_{\mathrm{odd}}|$ since $S_{\mathrm{odd}} \subseteq S$ and $\nu(\mathcal{I}_2(S_{\mathrm{odd}})) = |S_{\mathrm{odd}}|$. (The latter point follows from the fact that $s_{i+2}/2 = 2s_i > s_i$ for $1 \leq i \leq n-2$.) Combining this with Corollary 1 yields $\mathrm{rk}_2(S) \geq |S_{\mathrm{odd}}| = \lceil n/2 \rceil$. □

Expansion and contraction. Let $S \subset \mathbb{N}^*$. For any $c \in \mathbb{N}^*$, we write $S \times c$ for the set $\{s_i\, c : s_i \in S\}$ and we refer to $S \times c$ as the *c-expansion* of S. Similarly, for any $c \in \mathbb{N}^*$ common divisor of S, we write S/c for the set $\{s_i/c : s_i \in S\}$ and we refer to S/c as the *c-contraction* of S.

Lemma 6 (c-expansion). *Let $S \subset \mathbb{N}^*$ and $c \in \mathbb{N}^*$. Then $\mathrm{rk}_2(S \times c) \leq \mathrm{rk}_2(S)$.*

Replacing S by S/c in Lemma 6 yields a formulation well-suited for contraction considerations.

Corollary 3. *Let $S \subset \mathbb{N}^*$ and $c \in \mathbb{N}^*$ be a common divisor of S. Then $\mathrm{rk}_2(S) \leq \mathrm{rk}_2(S/c)$.*

Lemma 7 (c-contraction). *Let $S \subset \mathbb{N}^*$ and $c \in \mathbb{N}^*$ be a common divisor of S. Then, $\mathrm{rk}_2(S/c) = \mathrm{rk}_2(S)$ if c is odd and $\mathrm{rk}_2(S) \leq \mathrm{rk}_2(S/c) \leq 2\,\mathrm{rk}_2(S)$ if c is even.*

To complement Lemma 7, we observe that we may have $\mathrm{rk}_2(S/2c) < 2\,\mathrm{rk}_2(S)$ for even c as shown in the following example.

Example 1. For any $c \in \mathbb{N}^*$, let $S = \{14c, 16c, 18c\}$. Clearly, $X = \{7c, 9c\}$ is a 2-generating set of S, and hence $\mathrm{rk}_2(S) = 2$. But $S/2c = \{7, 8, 9\}$ has no smaller 2-generating set than itself, and hence $\mathrm{rk}_2(S/2c) = \mathrm{card}(S/2c) = 3$.

The upper-bound $\mathrm{rk}_2(S/c) \leq 2\,\mathrm{rk}_2(S)$ in Lemma 7 is, however, not overestimated, as shown by the following lemma.

Lemma 8. *For any $n \in \mathbb{N}^*$, there exists a set $S \subseteq \mathbb{N}^*$ of cardinality n such that*

$$\frac{\mathrm{rk}_2(S/2)}{\mathrm{rk}_2(S)} = 2 - \frac{1}{n+1}.$$

Proof. Let $b > 8$ be some fixed even integer. For any $n \in \mathbb{N}^*$, let $S = \{2\} \cup \{b^i + 2 : 1 \leq i \leq n\} \cup \{2b^i + 2 : 1 \leq i \leq n\}$. We can show, using Lemma 3, that $\mathrm{rk}_2(S) = n + 1$ and $\mathrm{rk}_2(S/2) = 2n + 1$ (proof omitted due to space constraints.) □

4 Hardness

Minimum Generating Set (*i.e.*, given a set a positive integers S, find a minimum cardinality set of integers X such that every element of S is the sum of a subset of X) was proved to be **NP**-complete in [4]. We complement this result by showing that Minimum 2-Generating Set is **APX**-hard, *i.e.*, hard to approximate within ratio $1 + \varepsilon$ for any $\varepsilon > 0$.

Proposition 1. Minimum 2-Generating Set *is* **APX**-*hard.*

Proof. We propose an L-reduction [16] from Vertex Cover for cubic graphs: Given a cubic graph $G = (V, E)$, find a minimum cardinality vertex cover of G, *i.e.*, a subset $V' \subseteq V$ such that, for each edge $\{u, v\} \in E$, at least one of u and v belongs to V'. Minimum Vertex Cover for cubic graphs is **APX**-complete [1,17]. Assume, wlog, that $V = \{1, 2, \ldots, n\}$. Define the corresponding instance of Minimum 2-Generating Set by defining $S \subset \mathbb{N}^*$ to be $S = \{b^0\} \cup \{b^i : 1 \leq i \leq n\} \cup \{2b^i : 1 \leq i \leq n\} \cup \{b^0 + b^i : 1 \leq i \leq n\} \cup \{b^0 + b^i + b^j : \{i, j\} \in E\}$ for some even constant $b > 4$. We claim that there exists a vertex cover of G of cardinality at most k if and only if there exists a 2-generating set for S of cardinality at most $n + k + 1$.

($\Rightarrow$) Suppose that there exists a vertex cover $V' \subseteq V$ of cardinality k of G. Define $X \subset \mathbb{N}^*$ (actually $X \subset S$) to be $X = \{b^0\} \cup \{b^i : 1 \leq i \leq n\} \cup \{b^0 + b^i : i \in V'\}$. We claim that X is a 2-generating set for S. Since $X \subset S$ and $b^0 \in X$, it is enough to prove that, for each $\{i, j\} \in E$, $b^0 + b^i + b^j$ is 2-generated by X. Indeed, since V' is a vertex cover of G, we have $i \in V'$ or $j \in V'$ (possibly both), and if we let $\ell = i$ if $i \in V'$ and $\ell = j$ if $i \notin V'$, we have $(b^0 + b^\ell) \in X$. Therefore

$b^0 + b^i + b^j$ is 2-generated by X as $(b^0 + b^\ell) + b^{\ell'}$, where $\ell' = j$ if $\ell = i$ and $\ell' = i$ otherwise.

($\Leftarrow$) Conversely, let X be a 2-generating set of S. We first note that, by integrality, $b^0 \in X$. Consider any integer $1 \leq i \leq n$, and let I_i be the interval $[b^i/2 : 2b^i]$. According to Lemma 2, $|X \cap [b^i/2 : b^i]| \geq 1$ and $|X \cap [b^i : 2b^i]| \geq 1$ since $b^i \in S$ and $2b^i \in S$. Then it follows that $|X \cap I_i| \geq 1$, and $b^i \in X$ if the inequality holds as equality. As $b > 4$, we have $2b^i < b^{i+1}/2$, $1 \leq i < n$. Then it follows that the intervals I_i, $1 \leq i \leq n$, are pairwise disjoint, and hence $|X| \geq n + 1$. Now, let $k \in \mathbb{N}^*$ be such that $|X| = n + k + 1$, and let $V' \subseteq V$ be such that $|X \cap I_i| > 1$ for every $i \in V'$. According to the above, we have $|V'| \leq k$. We now claim that V' is a vertex cover of G. Indeed, assume, aiming at a contradiction, that there exists $\{i, j\} \in E$ such that $|X \cap I_i| = 1$ and $|X \cap I_j| = 1$, and, to shorten notation, set $s = b^0 + b^i + b^j$. Then it follows that $X \cap I_i = \{b^i\}$ and $X \cap I_j = \{b^j\}$. But $s \in S$, and hence $|X \cap [s/2 : s]| \geq 1$ (Lemma 2). Furthermore, if we assume $i > j$, we have $b^i/2 < s/2$ and $s < 2b^i$, and hence $[s/2 : s] \subset I_i$, *i.e.*, $[s/2 : s]$ is a subinterval of I_i. But $X \cap I_i = \{b^i\}$, and hence we must have $(b^0 + b^j) \in X$. This is the desired contradiction since $(b^0 + b^j) \in I_j$ and $X \cap I_j = \{b^j\}$.

We omit the easy proof that the described reduction is indeed an L-reduction. (Crucial is the fact that $|V| \leq 2\,|V'|$ for any vertex cover V' since G is a cubic graph.) □

It remains open whether Minimum 2-Generating Set is **NP**-complete if every integer in S is bounded by a polynomial in the length of the input. Indeed, neither Proposition 1 nor the **NP**-hardness result of [4] rule out the existence of a pseudo-polynomial algorithm for Minimum 2-Generating Set. Observe that this question reduces to 2-covering a set of strings S for an unary alphabet with a set X of substrings in S, where X is said to 2-*cover* S if every string in S can be written as a concatenation of at most two substrings in X [12]. Approximation issues of Minimum 2-Generating Set are completely unexplored yet. Notice, however, that, as long as every integer in S is not bounded by a polynomial in the length of the input, none of the approximation results of [11,12] applies.

5 Put the blame on $\mathrm{rk}_2(S)$ only

Let S be any instance of Minimum 2-Generating Set. Write $n = |S|$, $m = \max(S)$ and $k = \mathrm{rk}_2(S)$. This section is devoted to finding a minimum cardinality 2-generating set of S (from an effective computational point of view [6,15]).

As a first attempt, let us consider the brute-force approach: generate all k-subsets X of $\{1, 2, \ldots, m\}$ and check for each of them whether it 2-generates S, *i.e.*, $S \subseteq X \cup 2X$. Correctness of this algorithm is of course immediate. There are $\binom{m}{k}$ such subsets and each subset X can be identified as a 2-generating set of S in $O(k^2 \log k)$ time (assuming a unit-cost RAM model with $\log m$ word size). Therefore, the brute-force algorithm is, as a whole, a $O(m^k k^2 \log k)$ time procedure. But m (and even $\log m$) can be arbitrarily large compared to $n = O(k^2)$ and this

naturally leads us to the problem of trying to confine the seemingly inevitable combinatorial explosion of computational difficulty to a function of k only [6,15]. We prove here that such an algorithm does exist for finding a minimum cardinality 2-generating set of S. Surprisingly enough, the time complexity of the proposed algorithm turns out to be even independent of $m = \max(S)$ (again assuming a unit-cost RAM model with $\log m$ word size). The main result of this paper can be stated as follows.

Lemma 9 (representation). *Let $S = \{s_i : 1 \leq i \leq n\} \subset \mathbb{N}^*$ and write k for $\mathrm{rk}_2(S)$. Then, there exist rationals $\alpha_{i,j} \in \{-1, -2^{-1}, 0, 2^{-1}, 1\}$, $1 \leq i \leq k$ and $1 \leq j \leq n$, such that $X = \{\sum_{j=1}^{n} \alpha_{i,j}\, s_j : 1 \leq i \leq k\}$ is a minimum cardinality 2-generating set of S.*

Before proving Lemma 9, we need a new definition that translates the problem to elementary graph theory terms. Let $S = \{s_1, s_2, \ldots, s_n\}$ be a set of positive integers and $X = \{x_1, x_2, \ldots, x_k\}$ be a 2-generating set for S. Define an *X-realization* of S to be a bipartite graph $B = (S, X, E)$ such that $d_B(s) \in \{1, 2\}$ for all $s \in S$, and

- if $d_B(s) = 1$, say $\{s, x_i\} \in E$, then $s = x_i$ or $s = 2x_i$, and
- if $d_B(s) = 2$, say $\{s, x_i\} \in E$ and $\{s, x_j\} \in E$, $x_i \neq x_j$, then $s = x_i + x_j$.

Note that, in the above definition of an X-realization, X (resp. S) is considered as a set of integers, *and* as a set of vertices in a graph. We chose not to correct this ambiguity in the rest of the paper, in order to avoid heavy notations. Besides, the context will always be clear about the fact that we are concerned with integers or vertices.

Coming back to X-realizations, it is clear that every simple cycle of B has length at least 6. (A simple cycle of length 4, say (x_1, s_1, x_2, s_2), would result in the contradiction $s_1 = x_1 + x_2 = s_2$.) An X-realization of S is said to be *minimum* if X is a minimum cardinality 2-generating set of S. Of course, an X-realization of a set S may not be unique.

Lemma 10. *Let $S \in \mathbb{N}^*$, $B = (S, X, E)$ be a minimum X-realization of S, and let B' be any connected component of B. If $d_{B'}(s) = 2$ for every vertex $s \in S$, then there exists a simple cycle of B' of length $4\ell + 2$ for some $\ell \geq 1$.*

We are now in position to prove Lemma 9.

Proof (of Lemma 9). Write $k = \mathrm{rk}_2(S)$. Let $X = \{x_i : 1 \leq i \leq k\}$ be a minimum cardinality 2-generating set of S and $B = (S, X, E)$ be any X-realization of S. Let $B_1, B_2, \ldots, B_t$ be the connected components of B. We consider each connected component of B separately. Consider any connected component $B_i = (S_i, X_i, E_i)$ of B with $S_i \subseteq S$ and $X_i \subseteq X$. Wlog, write $S_i = \{s_1, s_2, \ldots, s_{n_i}\}$. It is enough to show that for any $x \in X_i$, there exist rationals $\alpha_j \in \{-1, -2^{-1}, 0, 2^{-1}, 1\}$, $1 \leq j \leq n_i$, such that $x = \sum_{1 \leq j \leq n_i} \alpha_j\, s_j$, *i.e.*, x is a linear combination with coefficients taken from $\{-1, -2^{-1}, 0, 2^{-1}, 1\}$ of the vertices in S_i. We need to consider two cases: *(1)* $d_{B_i}(s) = 1$ for some $s \in S_i$; or *(2)* $d_{B_i}(s) = 2$ for every vertex $s \in S_i$.

(1) $d_{B_i}(s) = 1$ ***for some*** $s \in S_i$**.** For convenience, write $s_1 = s$. Let P be a simple path from vertex s_1 to vertex x. (Such a path exists since B_i is connected.) Wlog, write $P = (s_1, x_1, s_2, x_2, \ldots, x_{p-1}, s_p, x)$. Then it follows $s_1 = \delta x_1,\ s_2 = x_1 + x_2,\ \ldots,\ s_p = x_{p-1} + x$ for some $\delta \in \{1, 2\}$, and hence $x = \frac{(-1)^{p-1}}{\delta} s_1 + \sum_{i=2}^{p} (-1)^{p-i} s_i$. Therefore there exist rationals $\alpha_i \in \{-1, -2^{-1}, 2^{-1}, 1\}$, $1 \leq i \leq p$, such that $x = \sum_{i=1}^{p} \alpha_i s_i$, *i.e*, x is a linear combination with coefficients taken from $\{-1, -2^{-1}, 2^{-1}, 1\}$ of those vertices s_i that lie on the path from s_1 to x.

(2) $d_{B_i}(s) = 2$ ***for every vertex*** $s \in S_i$**.** According to Lemma 10, there exists a simple cycle C of length $4\ell + 2$ for some $\ell \geq 1$ in B_i. Write $C = (x_p, s_{p+1}, x_{p+1}, \ldots, x_{p+q-1}, s_{p+q})$, for some $q = 2\ell+1$. Since graph B_i is bipartite, any cycle that starts at a vertex in S_i must alternate between vertices in S_i and X_i, and hence must be of even length (on return to the start vertex again).

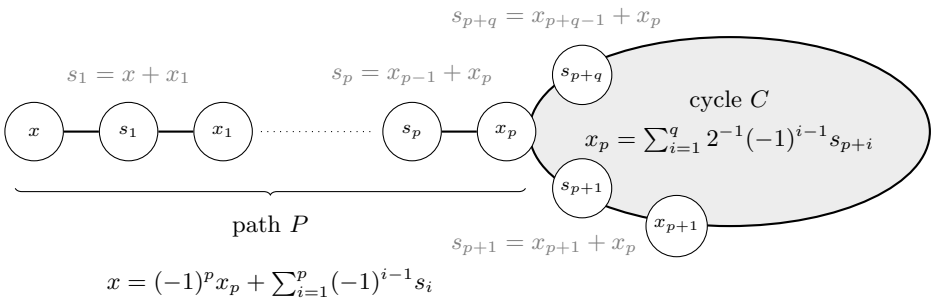

Fig. 1. For every vertex $s \in S_i$, we have $d_{B_i}(s) = 2$.

For the sake of presentation, suppose first that x does not lie on cycle C (see Figure 1 for an illustration.). Observe now that, since every vertex of S_i has degree 2 in B_i, every path leading from vertex x to cycle C intersects C at a vertex of X_i. Consider a shortest path leading from vertex x to cycle C, say $P = (x_0 = x, s_1, x_1, \ldots, x_{p-1}, s_p, x_p)$. Note that such a path exists since B_i is connected. Clearly, since vertex x does not lie on cycle C and P is a shortest path, all vertices of P but vertex x_p do not lie on cycle C. For one, we have $s_1 = x + x_1,\ s_2 = x_1 + x_2,\ \ldots,\ s_p = x_{p-1} + x_p$, and hence $x = (-1)^p x_p + \sum_{i=1}^{p} (-1)^{i-1} s_i$ **(1)**. For another, $s_{p+1} = x_p + x_{p+1},\ s_{p+2} = x_{p+1} + x_{p+2},\ \ldots,\ s_{p+q} = x_{p+q-1} + x_p$, and hence $x_p = \sum_{i=1}^{q} 2^{-1}(-1)^{i-1} s_{p+i}$ **(2)** since p is odd. Combining **(1)** and **(2)** yields $x = \sum_{i=1}^{q} 2^{-1}(-1)^{p+i-1} s_{p+i} + \sum_{i=1}^{p} (-1)^{i-1} s_i$. Therefore, there exist rationals $\alpha_i \in \{-1, -2^{-1}, 0, 2^{-1}, 1\}$, $1 \leq i \leq n$, such that $x = \sum_{i=1}^{n} \alpha_i s_i$. More precisely, x is a linear combination with coefficients taken from $\{-1, -2^{-1}, 2^{-1}, 1\}$ of those vertices s_i that lie on a shortest path leading from vertex x to a cycle C or lie on cycle C.

If x lies on cycle C, **(1)** vanishes (zero-length path), and substituting x_p by x in **(2)** yields $x = \sum_{i=1}^{q} 2^{-1}(-1)^{i-1} s_{p+i}$. Therefore, x is a linear combination

with coefficients taken from $\{-2^{-1}, 2^{-1}\}$ of those vertices s_i that lie on cycle C. □

Thanks to Lemma 9, we can prove that there exists an algorithm for Minimum 2-Generating Set that confines the combinatorial explosion of computational difficulty to a function of $k = \mathrm{rk}_2(S)$ only.

Proposition 2. *Assuming a unit-cost RAM model with* $\log m$ *word size (*$m = \max(S)$*), there exists a* $O(5^{\frac{k^2(k+3)}{2}}\ k^2 \log k)$ *time algorithm for finding a minimum cardinality* 2*-generating set of* S*, where* $k = \mathrm{rk}_2(S)$*.*

Proof. We propose a brute-force algorithm for finding a (representation of a) minimum cardinality 2-generating set of S. The basic idea is to consider the set $C(S)$ of all linear combinations $\alpha_1 s_1 + \alpha_2 s_2 + \ldots + \alpha_n s_n$ with coefficients taken from $\{-1, -2^{-1}, 0, 2^{-1}, 1\}$. Clearly, there exist 5^n such combinations. The algorithm simply tries each k-subset X of $C(S)$ and checks whether $S \subseteq X \cup 2X$. Correctness of the algorithm follows from Lemma 9. We now turn to proving its time complexity. Let N be the number of k-subsets of $C(S)$. Clearly, $N = \binom{5^n}{k} = O(5^{nk})$. But $n \leq \frac{k(k+3)}{2}$, and hence $N = O(5^{\frac{k^2(k+3)}{2}})$. Since a k-subset X of $C(S)$ can be identified as a 2-generating set of S in $O(k^2 \log k)$ time (assuming a unit-cost RAM model with $\log m$ word size), the total running time is $O(Nk^2 \log k) = O(5^{\frac{k^2(k+3)}{2}}\ k^2 \log k)$. □

6 Conclusion

Minimum 2-Generating Set is a natural restriction of Minimum Generating Set with prospective applications (see [4]). Our representation (Lemma 9) provides a first positive algorithmic result for computing minimum 2-generating sets. We mention here some directions of interest for future works: (1) Is Minimum 2-Generating Set pseudo-polynomial time solvable ? Notice that this question is related to 2-covering a set of strings S for a unary alphabet with a set X of substrings in S, where X is said to 2-*cover* S if every string in S can be written as a concatenation of at most two substrings in X [12]. (2) For any $k > 1$, a set of integers S is said to be k-*simplifiable* if $\mathrm{rk}_k(S) < |S|$ [14]. Is there a polynomial-time algorithm for deciding whether S is 2-simplifiable ? (3) Considering the general Minimum k-Generating Set problem, is there an analog of Lemma 9 for every fixed $k \geq 2$?

Acknowledgments

The authors are thankful to Olivier Serre for helpful discussions. They are also indebted to the reviewers for a careful and thoughtful reading of the original version of this paper.

References

1. P. Alimonti and V. Kann, *Some APX-completeness results for cubic graphs*, Theoretical Computer Science **237** (2000), no. 1-2, 123–134.
2. H.L. Bodlaender, R.G. Downey, M.R. Fellows, M.T. Hallett, and H.T. Wareham, *Parameterized complexity analysis in computational biology*, Computer Applications in the Biosciences **11** (1995), 49–57.
3. C. Choffrut and J. Karhumäki, *Combinatorics of words*, Handbook of formal languages, Vol. 1: Word, language, grammar (G. Rozenberg and A. Salomaa, eds.), Springer-Verlag, 1997, pp. 329–438.
4. M.J. Collins, D. Kempe, J. Saia, and M. Young, *Nonnegative integral subset representations of integer sets*, Information Processing Letters **101** (2007), no. 3, 129–133.
5. R. Diestel, *Graph theory*, second ed., Graduate texts in Mathematics, no. 173, Springer-Verlag, 2000.
6. R. Downey and M. Fellows, *Parameterized complexity*, Springer-Verlag, 1999.
7. M.A. Fitch and R.E. Jamison, *Minimum sum covers of small cyclic groups*, Congressus Numerantium **147** (2000), 65–81.
8. A. Gyárfás, *Combinatorics of intervals, preliminary version*, Institute for Mathematics and its Applications (IMA) Summer Workshop on Combinatorics and Its Applications, 2003, available online at `http://www.math.gatech.edu/news/events/ima/newag.pdf`.
9. H. Haanpää, *Minimum sum and difference covers of abelian groups*, Journal of Integer Sequences **7** (2004), no. 2, Article 04.2.6.
10. H. Haanpää, A. Huima, and P.R.J. Östergård, *Sets in $\mathbb{Z}_n$ with distinct sums of pairs*, Discrete Applied Mathematics **138** (2004), no. 1-2, 99–106.
11. M. Hajiaghayi, K. Jain, L. Lau, A. Russell I. Mandoiu, and V. Vazirani, *Minimum multicolored subgraph problem in multiplex PCR primer set selection and population haplotyping*, Proc. 6th International Conference on Computational Science (ICCS), Part II, Atlanta, USA (V.N. Alexandrov, G. Dick van Albada, P.M.A. Sloot, and J. Dongarra, eds.), LNCS, vol. 3994, Springer, 2006, pp. 758–766.
12. D. Hermelin, D. Rawitz, R. Rizzi, and S. Vialette, *The minimum substring cover problem*, Proc. 5th International Workshop on Approximation and Online Algorithms (WAOA), Eilat, Israel (C. Kaklamanis and M. Skutella, eds.), Lecture Notes in Computer Science, no. 4927, 2007, pp. 170–183.
13. L. Moser, *On the representation of $1, 2, \ldots, n$ by sums*, Acta Arithmetica **6** (1960), 11–13.
14. J. Néraud, *Elementariness of a finite set of words is coNP-complete*, Theoretical Informatics and Applications **24** (1990), no. 5, 459–470.
15. R. Niedermeier, *Invitation to fixed parameter algorithms*, Lecture Series in Mathematics and Its Applications, Oxford University, Press, 2006.
16. C.H. Papadimitriou, *Computational complexity*, Addison-Wesley, 1994.
17. C.H. Papadimitriou and M. Yannakakis, *Optimization, approximation and complexity classes*, Journal of Computer and System Sciences **43** (1991), 425–440.
18. C.N. Swanson, *Planar cyclic difference packings*, Journal of Combinatorial Designs **8** (2000), 426–434.
19. D. Wiedemann, *Cyclic difference covers through* 133, Congressus Numerantium **90** (1992), 181–185.