

# On the NP-Hardness of Bounded Distance Decoding of Reed-Solomon Codes

Venkata Gandikota  
Purdue University

Badih Ghazi  
MIT

Elena Grigorescu  
Purdue University

**Abstract**—Guruswami and Vardy (IEEE Trans. Inf. Theory, 2005) show that given a Reed-Solomon code over a finite field  $\mathbb{F}$ , of length  $n$  and dimension  $t$ , and given a target vector  $v \in \mathbb{F}^n$ , it is NP-hard to decide if there is a codeword that disagrees with  $v$  on at most  $n - t - 1$  coordinates. Understanding the complexity of this Bounded Distance Decoding problem as the amount of error in the target decreases is an important open problem in the study of Reed-Solomon codes. In this work, we extend the result of Guruswami and Vardy by proving that it is NP-hard to decide the existence of a codeword that disagrees with  $v$  on  $n - t - 2$ , and on  $n - t - 3$  coordinates. No other NP-hardness results were known before for an amount of error  $< n - t - 1$ . The core of our proofs is showing the NP-hardness of a parameterized generalization of the Subset-Sum problem to higher degrees (called Moments Subset-Sum) that may be of independent interest.

## I. INTRODUCTION

A linear error-correcting code of length  $n$  and dimension  $t$ , over a finite field alphabet  $\mathbb{F}$ , is a vector space  $\mathcal{C} \subseteq \mathbb{F}^n$  of dimension  $t$ . The Hamming distance between  $x, y \in \mathbb{F}^n$  is  $\delta(x, y) := |\{i \in [n] : x_i \neq y_i\}|$ , and the minimum distance of  $\mathcal{C}$  is  $\lambda(\mathcal{C}) := \min_{x \neq y \in \mathcal{C}} \delta(x, y)$ . In the Bounded Distance Decoding problem with parameter  $d$ , we are given a code  $\mathcal{C} \subseteq \mathbb{F}^n$  and a target vector  $v \in \mathbb{F}^n$ , and we are asked to decide whether there exists a codeword  $c \in \mathcal{C}$  such that the Hamming distance  $\delta(v, c) \leq e(d, \lambda(\mathcal{C}))$ , where  $e(d, \cdot)$  is an error-weight function of interest. This is a fundamental question in the study of general error-correcting codes, and its complexity is not fully understood even for well-studied codes such as Reed-Solomon (RS) codes. More precisely, RS codes still exhibit a wide gap between the setting of small error-weight, where the problem is solvable in polynomial-time [Sud97], [GS99], and the setting of large error-weight, where the problem is known to be NP-hard [GV05]. In this work, we improve this gap by generalizing the decade-old NP-hardness results of [GV05] to smaller error parameters.

A Reed-Solomon code of length  $n$ , dimension  $t$ , defined over a finite field  $\mathbb{F}$ , and specified by an evaluation set  $D = \{x_1, x_2, \dots, x_n\} \subseteq \mathbb{F}$  is the

set  $RS_{D,t} = \{\langle f(x_1), \dots, f(x_n) \rangle : x_1, \dots, x_n \in D, f(x) \in \mathbb{F}[x], \deg(f(x)) \leq t - 1\}$ . Its minimum distance is  $\lambda = n - t + 1$ . We are interested in the Bounded Distance Decoding problem for RS codes with parameter  $d$  (denoted  $RS\text{-BDD}_d$ ) corresponding to the error-weight function  $e(d, \lambda) = \lambda - d - 1 = n - t - d$ .

When  $e \leq n - \sqrt{nt}$  (i.e., the ‘‘Johnson radius’’), the list-decoding algorithms of [Sud97], [GS99] can solve  $RS\text{-BDD}_{\sqrt{nt-t}}$  in polynomial time. At the other end of the spectrum, if  $e = n - t$  one can easily interpolate a degree  $t - 1$  polynomial that agrees with the target on a set of  $t$  arbitrary entries. The famous result of [GV05] shows that right below this value, i.e., when  $e = n - t - 1$ ,  $RS\text{-BDD}_1$  is NP-hard. Their proof is via a reduction from the 3D-matching problem, and it does not imply any hardness results for smaller values of the error-weight parameter. In the range where  $e < n - t - 1$ , the only known hardness result is due to [CW10] who show a reduction from the Discrete Log problem, when the evaluation set is  $D = \mathbb{F}_q^*$ , for  $e \geq \frac{2}{3}(n - t + 1)$ . Note that the hardness of Discrete Log is a stronger complexity-theoretic assumption than  $P \neq NP$ . In particular, the Discrete Log problem over finite fields is not believed to be NP-hard and [Sho97] gives a polynomial-time quantum algorithm solving it.

In this work, we prove that for every  $d \in \{2, 3\}$ , the  $RS\text{-BDD}_d$  problem is NP-hard. We note that, as is the case of [GV05], our reductions require that  $n$  be polylogarithmic in  $|\mathbb{F}|$ . Our proof starts with the observation that an extension of the Subset-Sum problem up to  $d$  moments (called Moments Subset-Sum with parameter  $d$ ) is polynomial-time reducible to  $RS\text{-BDD}_d$ . The crux of our proof is showing the NP-hardness of the Moments Subset-Sum problem, formally defined next.

**Definition 1** (Moments Subset-Sum:  $MSS_d$ ). *Given a set  $A = \{a_1, a_2, \dots, a_n\} \subseteq \mathbb{F}^n$ , integer  $k$ , and  $m_1, m_2, \dots, m_d \in \mathbb{F}$ , decide if there exists a subset  $S \subseteq A$  of size  $k$ , satisfying  $\sum_{s \in S} s^i = m_i$  for all  $i \in [d]$ .*

Note that  $MSS_1$  is the usual Subset-Sum problem, which is in fact used by [CM07] to show an alternate

proof of [GV05]. The NP-hardness of Subset-Sum can be shown by a well-known reduction from 3-SAT. Surprisingly, it turns out to be much more difficult to prove NP-hardness for  $MSS_d$  for values of  $d \geq 2$ . Intuitively, this is because the reduction from 3-SAT to Subset-Sum encodes satisfiability of a 3-SAT formula in the decimal representation of the integers of the Subset-Sum instance, and it becomes more difficult to control the decimal representations when constraints involving squares and higher powers are introduced.

In fact, the current barrier to extending our proof approach to show the NP-hardness of  $MSS_d$  for larger values of  $d$  (and thereby obtain the NP-hardness of  $RS-BDD_d$  for larger values of  $d$ ) is the following intriguing algebraic question.

**Question 1.** *Given a finite field  $\mathbb{F}$ ,  $a, b \in \mathbb{F}$ , and  $d \in \mathbb{N}$ , does there exist  $t = t(d)$  and explicit  $x_1, x_2, \dots, x_t, y_1, y_2, \dots, y_t \in \mathbb{F}$  satisfying  $x_1 + x_2 + \dots + x_t = y_1 + y_2 + \dots + y_t$  and  $a^i + \sum_{j=1}^t x_j^i = b^i + \sum_{j=1}^t y_j^i$  for every  $i \in \{2, \dots, d\}$  ?<sup>1</sup>*

For  $d \in \{2, 3\}$ , we are able to answer this question in the affirmative and to provide explicit constructions over domains that are either finite fields of large characteristic, the rational numbers or the ring of integers. These constructions form the starting point of our NP-hardness proofs.

## II. REDUCTION FROM MOMENTS SUBSET SUM

We start by reformulating the  $RS-BDD$  problem as a polynomial reconstruction question as follows.

**Definition 2** ( $RS-BDD_d$ ). *Given a set of  $n$  distinct points in  $\mathbb{F}^2$ ,  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ , and an integer  $t < n$ , decide if there exists a polynomial  $f(x) \in \mathbb{F}[x]$  of degree at most  $t - 1$  passing through at least  $t + d$  points in  $D$ .*

Here, the set of evaluation points of the code is  $D' = \{x_1, \dots, x_n\}$  and the target vector is  $y = (y_1, \dots, y_n) \in \mathbb{F}^n$ . Note that if a polynomial  $f(x)$  passes through at least  $t + d$  points in  $D$ , then the corresponding codeword  $\langle f(x_1), \dots, f(x_n) \rangle$  is at a Hamming distance of at most  $e = n - t - d$  from  $y$ . We will reduce from the following problem which can be easily seen to be equivalent to  $MSS_d$  over large prime finite fields  $\mathbb{F}$  using Newton's identities [Sta99]. We note that this connection has been previously made (e.g. [LW08]).

**Definition 3** (Symmetric Subset-Sum ( $SSS_d$ )). *Given a set of  $n$  distinct elements of  $\mathbb{F}$ ,  $A = \{a_1, a_2, \dots, a_n\}$ ,*

<sup>1</sup>We note that the same question is also of interest when  $a, b, x_1, x_2, \dots, x_t, y_1, y_2, \dots, y_t$  are rational numbers or integers.

*integer  $k$ , and  $B_1, B_2, \dots, B_d \in \mathbb{F}$ , decide if there exists a subset  $S$  of  $A$  of size  $k$ , such that for every  $i \in [d]$  the elementary symmetric sums of the elements of  $S = \{s_1, \dots, s_k\}$  satisfy  $E_i(S) = \sum_{1 \leq j_1 < j_2 < \dots < j_i \leq k} s_{j_1} \dots s_{j_i} = B_i$ .*

We defer the proof of the reduction from  $SSS_d$  (and hence from  $MSS_d$ ) to  $RS-BDD_d$  to Lemma 6 in Appendix A.

## III. NP-HARDNESS OF $MSS_2$

**Theorem 4.**  *$MSS_2$  is NP-hard.*

*Proof:* We present the proof for the case where the domain is the ring of integers, and then observe that the proof easily extends to any sufficiently large prime field (see the end of the proof for the details).

We give a polynomial-time reduction from the 1-in-3-SAT problem in which we are given a 3-SAT formula  $\phi$  on  $n$  variables and  $m$  clauses and are asked to determine if there exists an assignment  $x \in \{0, 1\}^n$  satisfying exactly one literal in each clause. It is known that this problem is NP-hard even for  $m = O(n)$  [Sch78]. We start by recalling the reduction from 1-in-3-SAT to Subset-Sum which will be used in our reduction to  $MSS_2$ . In that reduction, each variable  $(x_i, \bar{x}_i)$  is mapped to 2 integers  $a_i$  (corresponding to  $x_i$ ) and  $b_i$  (corresponding to  $\bar{x}_i$ ). The integers  $a_i$  and  $b_i$  and the target  $B$  are defined in terms of their length- $(n + m)$   $\Sigma$ -ary representation (for some sufficiently large even constant, say  $\Sigma = 4$ ) as follows:

- The  $\Sigma$ -ary representations of  $a_i$  and  $b_i$  consist of 2 parts: a variable region consisting of the leftmost  $n$  digits and a clause region consisting of the (remaining) rightmost  $m$  digits.
- In the variable region,  $a_i$  and  $b_i$  have a 1 at the  $i$ th digit and 0's at the other digits.
- In the clause region, for every  $j \in [m]$ ,  $a_i$  (resp.  $b_i$ ) has a 1 at the  $j$ th location if  $x_i$  (resp.  $\bar{x}_i$ ) appears in clause  $j$ , and a 0 otherwise.
- The target  $B$  is set to the integer whose  $\Sigma$ -ary representation is the all 1's.

Those  $\Sigma$ -ary representations are illustrated in Figure 1.

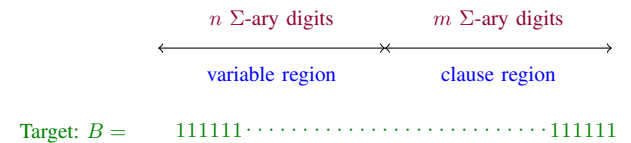


Fig. 1.  $\Sigma$ -ary representations in the original reduction from 1-in-3-SAT to Subset-Sum.

This reduction to Subset-Sum can be seen to be complete and sound. We now use it to give a reduction to  $MSS_2$ . In this reduction, each variable  $(x_i, \bar{x}_i)$  is mapped to 6 integers  $a_{i,1}, a_{i,2}, a_{i,3}$  (corresponding to  $x_i$ ) and  $b_{i,1}, b_{i,2}, b_{i,3}$  (corresponding to  $\bar{x}_i$ ). Let  $\{a_i, b_i : i \in [n]\}$  be the integers produced by the above reduction to Subset-Sum. We denote by  $a_i^v$  (resp.  $a_i^c$ ) the  $\Sigma$ -ary representation of the variable (resp. clause) region of  $a_i$ . For any  $\Sigma$ -ary representation  $x$  of a natural number, let  $(x)_{\Sigma'}$  be the natural number whose  $\Sigma'$ -representation is  $x$ . For example, if  $x$  is the  $\Sigma$ -ary number (10011), then  $(x)_{\Sigma'} = \Sigma'^4 + \Sigma' + 1$ . Denote by  $1^l$  the concatenation of  $l$  ones. Let  $\nu$  and  $h$  be natural numbers to be specified later on. The integers  $a_{i,1}, a_{i,2}, a_{i,3}, b_{i,1}, b_{i,2}, b_{i,3}$  and the targets  $B_1$  and  $B_2$  are defined as follows:

- $a_{i,1} = \Sigma^{\nu+m+h-1}(a_i^v)_{\Sigma^h} + \Sigma^{\nu}(a_i^c)_{\Sigma}$
- $b_{i,1} = \Sigma^{\nu+m+h-1}(b_i^v)_{\Sigma^h} + \Sigma^{\nu}(b_i^c)_{\Sigma}$
- $a_{i,2} = \frac{a_{i,1}^2 - b_{i,1}^2}{4\Sigma^i} - \frac{\Sigma^i}{2}$  and  $a_{i,3} = -a_{i,2}$ .
- $b_{i,2} = \frac{a_{i,1}^2 - b_{i,1}^2}{4\Sigma^i} + \frac{\Sigma^i}{2}$  and  $b_{i,3} = -b_{i,2}$ .
- $B_1 = \sum_{i=1}^n \Sigma^{\nu+m+h-1}(1^n)_{\Sigma^h} + \Sigma^{\nu}(1^m)_{\Sigma}$
- $B_2 = \sum_{i=1}^n a_{i,1}^2 + a_{i,2}^2 + a_{i,3}^2$ .

Note that  $a_{i,1}$  (resp.  $a_{i,2}$ ) defined above can be seen as obtained by inserting  $h-1$  consecutive 0's in the variable region of the  $\Sigma$ -ary representation of  $a_i$  (resp.  $b_i$ ), and then inserting  $\nu$  zeros at the right. Those  $\Sigma$ -ary representations are illustrated in Figure 2. Moreover, observe that  $a_{i,2}, a_{i,3}, b_{i,2}, b_{i,3}$  are defined above in such a way that  $a_{i,2} + a_{i,3} = b_{i,2} + b_{i,3} = 0$  and  $a_{i,1}^2 + a_{i,2}^2 + a_{i,3}^2 = b_{i,1}^2 + b_{i,2}^2 + b_{i,3}^2$ .

Note that this reduction runs in time polynomial in  $n, m, h$  and  $\nu$ . We next prove the correctness of the reduction.

*a) Completeness:* Assume that the starting 3-SAT formula has a satisfying 1-in-3-SAT assignment  $x$ . Consider the subset  $S = \cup_{i \in [n]: x_i=1} \{a_{i,1}, a_{i,2}, a_{i,3}\} \cup \cup_{i \in [n]: x_i=0} \{b_{i,1}, b_{i,2}, b_{i,3}\}$  of size  $k = 3n$ . We now check that the first 2 moments of  $S$  equal  $B_1$  and  $B_2$ .

$$\begin{aligned}
\sum_{y \in S} y &= \sum_{\substack{i \in [n]: \\ x_i=1}} a_{i,1} + a_{i,2} + a_{i,3} + \sum_{\substack{i \in [n]: \\ x_i=0}} b_{i,1} + b_{i,2} + b_{i,3} \\
&= \sum_{i \in [n]: x_i=1} a_{i,1} + \sum_{i \in [n]: x_i=0} b_{i,1} \\
&= \sum_{i \in [n]: x_i=1} \Sigma^{\nu+m+h-1}(a_i^v)_{\Sigma^h} + \Sigma^{\nu}(a_i^c)_{\Sigma} \\
&+ \sum_{i \in [n]: x_i=0} \Sigma^{\nu+m+h-1}(b_i^v)_{\Sigma^h} + \Sigma^{\nu}(b_i^c)_{\Sigma} \\
&= \Sigma^{\nu+m+h-1} \left( \sum_{i \in [n]: x_i=1} a_i^v + \sum_{i \in [n]: x_i=0} b_i^v \right)_{\Sigma^h}
\end{aligned}$$

$$\begin{aligned}
&+ \Sigma^{\nu} \left( \sum_{i \in [n]: x_i=1} a_i^c + \sum_{i \in [n]: x_i=0} b_i^c \right)_{\Sigma} \\
&= \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^{\nu} (1^m)_{\Sigma} \\
&= B_1
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\sum_{y \in S} y^2 &= \sum_{i \in [n]: x_i=1} a_{i,1}^2 + a_{i,2}^2 + a_{i,3}^2 \\
&+ \sum_{i \in [n]: x_i=0} b_{i,1}^2 + b_{i,2}^2 + b_{i,3}^2 \\
&= \sum_{i \in [n]} a_{i,1}^2 + a_{i,2}^2 + a_{i,3}^2 \\
&= B_2
\end{aligned}$$

where we used the fact that  $a_{i,1}^2 + a_{i,2}^2 + a_{i,3}^2 = b_{i,1}^2 + b_{i,2}^2 + b_{i,3}^2$  for every  $i \in [n]$ .

*b) Soundness:* Assume that  $S$  is a satisfying subset of the  $MSS_2$  instance. We will use  $S$  in order to construct a 1-in-3-SAT satisfying assignment to the starting 3-SAT instance  $\phi$ . Let's first sketch the high-level intuition of the proof. First, we argue that  $S$  should contain the same number of positive and negative integers from the set  $\{a_{1,2}, b_{1,2}, a_{1,3}, b_{1,3}\}$ . This is because these four integers have roughly the same absolute value, which turns out to be much larger than  $B_1$  and than any other possible integer in  $S$ . Thus, unless  $S$  contains the same number of positive and negative elements from  $\{a_{1,2}, b_{1,2}, a_{1,3}, b_{1,3}\}$ , the integers in  $S$  cannot add up to  $B_1$ . Then, we show using an inductive argument that for each  $i \in [n]$ ,  $S$  contains the same number of positive and negative elements from  $\{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}$  (i.e., Lemma 1). We next prove that, in fact, the elements of  $S$  from the set  $\cup_{i=1}^n \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}$  exactly cancel each other out, i.e., they add up to 0. Hence, we obtain that the integers of  $S$  from  $\cup_{i=1}^n \{a_{i,1}, b_{i,1}\}$  add up to  $B_1$ , at which point we can run the soundness analysis of the original reduction from 1-in-3-SAT to Subset Sum in order to construct an assignment satisfying exactly one literal in each clause. We now give the formal proof, which starts with the following lemma whose proof appears in Appendix B. We assume henceforth that  $\nu \geq 10(m + nh)$  and  $h \geq 10m$ .

**Lemma 1.** *For every  $i \in [n]$ ,  $|S \cap \{a_{i,2}, b_{i,2}\}| = |S \cap \{a_{i,3}, b_{i,3}\}|$ .*

Lemma 1 implies that for every  $i \in [n]$ ,  $\sum_{y \in S \cap \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} y = r_i \times \Sigma^i$  for some  $r_i \in \{-1, 0, 1\}$ . The next lemma, whose proof appears in Appendix B, argues that in fact each  $r_i$  equals 0.

**Lemma 2.** *For every  $i \in [n]$ ,  $r_i = 0$ .*

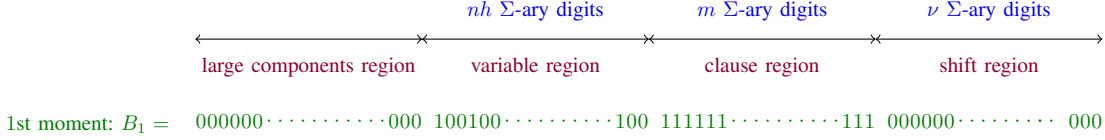


Fig. 2.  $\Sigma$ -ary representations in the reduction from 1-in-3-SAT to  $MSS_2$ . In this figure,  $h = 3$ . The “large components region” only contains zeros in  $\{a_{i,1}, b_{i,1} : i \in [n]\}$  but contains non-zeros in  $\{|a_{i,2}|, |b_{i,2}| : i \in [n]\}$ .

Lemma 2 implies that  $\sum_{y \in S \cap \cup_{i=1}^n \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} y = 0$  and hence that  $\sum_{y \in S \cap \cup_{i=1}^n \{a_{i,1}, b_{i,1}\}} y = B_1$ . We can

now carry out the soundness analysis of the original reduction from 1-in-3-SAT to Subset Sum in order to construct an assignment that satisfies exactly one literal in each clause of the given 3-SAT formula. This concludes the soundness analysis of our reduction. To sum up, setting  $h = 10m$ ,  $\nu = 10(m + nh)$  and  $\Sigma$  to be a sufficiently large even constant (say 4) yields a polynomial-time reduction from 1-in-3-SAT to  $MSS_2$ . Hence,  $MSS_2$  is NP-hard.

We now extend the proof to any sufficiently large prime field. Let  $p \geq 2^{cn^2}$  be a prime number with  $c$  some absolute constant. This ensures that the magnitudes of all integers appearing in the above proof are  $< p/2$ . Performing the above computations modulo  $p$ , we can carry out all the arguments as in the case of the integers, and hence the proof holds over  $\mathbb{F}_p$ . ■

#### IV. NP-HARDNESS OF $MSS_3$

**Theorem 5.**  $MSS_3$  is NP-hard.

*Proof:* We present the proof over the field of rational numbers, and then observe that the proof easily extends to any sufficiently large prime field (see the end of the proof for the details).

We give a polynomial-time reduction from the 1-in-3-SAT problem in which we are given a 3-SAT formula  $\phi$  on  $n$  variables and  $m$  clauses and are asked to determine whether there exists an assignment  $x \in \{0, 1\}^n$  that satisfies exactly one literal in each clause. It is known that this problem is NP-hard even for  $m = O(n)$  [Sch78]. Recall the known reduction from 1-in-3-SAT to Subset-Sum that is described at the beginning of the proof of Theorem 4. Let  $a_i$  and  $b_i$  be the positive integers corresponding to variable  $(x_i, \bar{x}_i)$  and let  $B$  be the target positive integer in that reduction. As before the  $\Sigma$ -ary representations of the  $a_i$ 's, the  $b_i$ 's and  $B$  are illustrated in Figure 1. We now use this reduction to reduce 1-in-3-SAT to  $MSS_3$ . In the new reduction, each variable  $(x_i, \bar{x}_i)$  is mapped to 6 rational numbers  $a'_{i,1}, a'_{i,2}, a'_{i,3}$  (corresponding to  $x_i$ ) and  $b'_{i,1}, b'_{i,2}, b'_{i,3}$  (corresponding to  $\bar{x}_i$ ). The symbols  $a_i^c$ ,

$a_i^c$  and  $(x)_{\Sigma^c}$  are defined as in the proof of Theorem 4. Let  $\nu$  and  $h$  be natural numbers to be specified later on. The rational numbers  $a'_{i,1}, a'_{i,2}, a'_{i,3}, b'_{i,1}, b'_{i,2}, b'_{i,3}$  and the targets  $B_1, B_2$  and  $B_3$  are defined as follows:

- $a'_{i,1} = \Sigma^{\nu+m+h-1} (a_i^c)_{\Sigma^h} + \Sigma^{\nu} (a_i^c)_{\Sigma}$
- $b'_{i,1} = \Sigma^{\nu+m+h-1} (b_i^c)_{\Sigma^h} + \Sigma^{\nu} (b_i^c)_{\Sigma}$
- Define  $\beta_i := \Sigma^i$ ,  $\gamma_i := \frac{a'^2_{i,1} - b'^2_{i,1}}{2\beta_i}$  and  $\alpha_i := \frac{a'^3_{i,1} - b'^3_{i,1}}{6\beta_i\gamma_i}$ .
- $a'_{i,2} = \alpha_i + \frac{\gamma_i}{2} - \frac{\beta_i}{2}$ .
- $a'_{i,3} = \alpha_i - \frac{\gamma_i}{2} + \frac{\beta_i}{2}$ .
- $b'_{i,2} = \alpha_i + \frac{\gamma_i}{2} + \frac{\beta_i}{2}$ .
- $b'_{i,3} = \alpha_i - \frac{\gamma_i}{2} - \frac{\beta_i}{2}$ .
- $B_1 = \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^{\nu} (1^m)_{\Sigma} + \sum_{i=1}^n a'_{i,2} + a'_{i,3}$ .
- $B_2 = \sum_{i=1}^n a'^2_{i,1} + a'^2_{i,2} + a'^2_{i,3}$ .
- $B_3 = \sum_{i=1}^n a'^3_{i,1} + a'^3_{i,2} + a'^3_{i,3}$ .

Note that  $a'_{i,1}$  and  $b'_{i,1}$  are positive integers defined as in the proof of Theorem 4 and, in particular, their  $\Sigma$ -ary representations are identical to those illustrated in Figure 2 for the reduction from 1-in-3-SAT to  $MSS_2$ .

Note that this reduction runs in time polynomial in  $n, m, h$  and  $\nu$ . We next prove the correctness of the reduction.

*c) Completeness:* The completeness analysis uses the next proposition whose proof appears in Appendix C.

**Proposition 1.** For every  $i \in [n]$ , the quantities  $a'_{i,2}, a'_{i,3}, b'_{i,2}, b'_{i,3}$  defined above satisfy

$$\begin{aligned} a'_{i,2} + a'_{i,3} &= b'_{i,2} + b'_{i,3} \\ a'^2_{i,1} + a'^2_{i,2} + a'^2_{i,3} &= b'^2_{i,1} + b'^2_{i,2} + b'^2_{i,3} \\ a'^3_{i,1} + a'^3_{i,2} + a'^3_{i,3} &= b'^3_{i,1} + b'^3_{i,2} + b'^3_{i,3} \end{aligned}$$

Assume that the starting 3-SAT formula has a satisfying 1-in-3-SAT assignment  $x$ . Consider the subset  $S = \cup_{i \in [n]: x_i=1} \{a'_{i,1}, a'_{i,2}, a'_{i,3}\} \cup \cup_{i \in [n]: x_i=0} \{b'_{i,1}, b'_{i,2}, b'_{i,3}\}$  of size  $k = 3n$ . In Appendix C, we use Proposition 1 to check that the first 3 moments of  $S$  equal  $B_1, B_2$  and  $B_3$  respectively.

d) *Soundness*: Assume that  $S$  is a satisfying subset of the  $MSS_3$  instance. We will use  $S$  in order to construct a 1-in-3-SAT satisfying assignment to the starting 3-SAT formula  $\phi$ . Let's first sketch the high-level intuition of the proof. First, we prove using an inductive argument (i.e., Lemma 3) that, for each  $i \in [n]$ ,  $S$  should contain the same number of positive and negative rational numbers from the set  $\{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}$ . This is because these four numbers have roughly the same absolute value, which turns out to be much larger than  $B_1$  and than any other possible number in  $S$ . Thus, unless  $S$  contains the same number of positive and negative elements from  $\{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}$ , the elements of  $S$  cannot add up to  $B_1$ . Next, we show using another inductive argument (i.e., Lemma 4) that, for every  $i \in [n]$ ,  $S$  should contain exactly two elements from the set  $\{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}$ . This is because otherwise, the sum of squares of the elements in  $S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}$  would be very far from  $a'^2_{i,2} + a'^2_{i,3}$ , which would imply that the sum of squares of the elements in  $S$  is very far from  $B_2$ . We then show (in Lemma 5) that, in fact, the elements of  $S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}$  exactly add up to  $2\alpha_i$ ; namely, the  $\beta_i$  and  $\gamma_i$  terms cancel each other out. Hence, we obtain that the elements of  $S$  from  $\cup_{i=1}^n \{a'_{i,1}, b'_{i,1}\}$  add up to  $B_1 - 2 \sum_{i \in [n]} \alpha_i = \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^\nu (1^m)_\Sigma$ , at

which point we can run the soundness analysis of the original reduction from 1-in-3-SAT to Subset Sum in order to construct an assignment satisfying exactly one literal in each clause. We now give the formal proof which uses the following lemmas whose proofs appear in Appendix C. We set  $\nu = 10(m + nh)$  and  $h = 10m$ .

**Lemma 3.**  $\forall i \in [n]$ ,  $|S \cap \{a'_{i,2}, b'_{i,2}\}| = |S \cap \{a'_{i,3}, b'_{i,3}\}|$ .

**Lemma 4.**  $\forall i \in [n]$ ,  $|S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}| = 2$ .

**Lemma 5.** 
$$\sum_{y \in S \cap \cup_{i=1}^n \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y = 2 \sum_{i=1}^n \alpha_i$$

Lemmas 3, 4 and 5 imply that  $\sum_{y \in S \cap \cup_{i=1}^n \{a'_{i,1}, b'_{i,1}\}} y = \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^\nu (1^m)_\Sigma$ . We now carry out the soundness analysis of the original reduction from 1-in-3-SAT to Subset Sum to construct an assignment that satisfies exactly one literal in each clause of the given 3-SAT formula. This concludes our soundness analysis.

Thus, setting  $\Sigma$  to be a sufficiently large even constant (say 4) yields a polynomial-time reduction from 1-in-3-SAT to  $MSS_3$ . Hence,  $MSS_3$  is NP-hard.

The proof also holds over the ring of integers by first multiplying the numbers  $\{a'_{i,j}, b'_{i,j}\}$ ,  $B_1$ ,  $B_2$ ,  $B_3$  by the least common multiple of the denominators of

$\{\alpha_i, \gamma_i/2, \alpha_i/2\}$ . This preserves the moment constraints since they are homogeneous equations. Let  $c = O(1)$  and  $p > 2^{cn^3}$  be a prime number such that all integers appearing in this proof are  $< p/2$ . The proof extends to  $\mathbb{F}_p$  by performing all computations modulo  $p$ . ■

## V. CONCLUSION

In this work, we proved that  $RS\text{-}BDD_2$  and  $RS\text{-}BDD_3$  are NP-hard. The main open problem raised by our work is to understand the limits of our approach. In particular, can one design explicit constructions satisfying the properties in Question 1, and use them to prove the NP-hardness of  $RS\text{-}BDD_d$  for larger values of  $d$ ? Finally, we remark that a positive answer to Question 1 could yield to a better understanding of the structure of BCH codes over large fields, and to their local testability properties (e.g., [GK12].)

## ACKNOWLEDGMENT

The authors would like to thank Amir Abboud, Swastik Kopparty, Madhu Sudan and Andrew Sutherland for very helpful discussions and pointers.

## REFERENCES

- [CM07] Qi Cheng and Elizabeth Murray. On deciding deep holes of Reed-Solomon codes. In *TAMC 2007, Shanghai, China*, pages 296–305, 2007.
- [CW10] Qi Cheng and Daqing Wan. Complexity of decoding positive-rate primitive Reed-Solomon codes. *IEEE Transactions on Information Theory*, 56(10):5217–5222, 2010.
- [GK12] Elena Grigorescu and Tali Kaufman. Explicit low-weight bases for BCH codes. *IEEE Transactions on Information Theory*, 58(1):78–81, 2012.
- [GS99] Venkatesan Guruswami and Madhu Sudan. Improved decoding of reed-solomon and algebraic-geometry codes. *IEEE Transactions on Information Theory*, 45(6):1757–1767, 1999.
- [GV05] Venkatesan Guruswami and Alexander Vardy. Maximum-likelihood decoding of Reed-Solomon codes is NP-hard. *IEEE Transactions on Information Theory*, 51(7):2249–2256, 2005.
- [LW08] Jiyou Li and Daqing Wan. On the subset sum problem over finite fields. *Finite Fields and Their Applications*, 14(4):911–929, 2008.
- [Sch78] Thomas J Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, pages 216–226. ACM, 1978.
- [Sho97] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.
- [Sta99] R P Stanley. *Enumerative Combinatorics, vol.2*. Cambridge University Press, 1999.
- [Sud97] Madhu Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. *J. Complexity*, 13(1):180–193, 1997.

APPENDIX

A. Proofs from Section II

**Lemma 6.** *SSS<sub>d</sub> is polynomial-time reducible to RS-BDD<sub>d</sub>.*

*Proof:* Henceforth, let  $q := |\mathbb{F}|$ . Given an instance  $\langle A, k, B_1, B_2, \dots, B_d \rangle$  of SSS<sub>d</sub>, we construct an instance  $\langle D, t \rangle$  of RS-BDD<sub>d</sub> such that there exists a polynomial  $f(x) \in \mathbb{F}_q[X]$  of degree at most  $t - 1$  which agrees with at least  $t + d$  points of  $D$  if and only if there is a solution to the given instance of SSS<sub>d</sub>. Here,  $A = \{a_1, a_2, \dots, a_n\}$  is a set of distinct, non-zero elements of  $\mathbb{F}_q$ ,  $B_1, B_2, \dots, B_d \in \mathbb{F}_q$  and  $k \in \mathbb{Z}$ . Let  $t = k - d + 1$ . Define the degree  $d$  polynomial  $p(x) := x^d - B_1x^{d-1} + \dots + (-1)^{d-1}B_{d-1}x$ . For each  $a_i$  of  $A$ , define an element of  $\mathbb{F}_q$  as  $y_i = -p(a_i)$ . The set  $D$  is then given by  $D = \{(a_i^{-1}, y_i) \text{ for all } a_i \in A\} \cup \{(0, (-1)^d B_d)\}$ . Note that  $\langle D, t \rangle$  is an instance of RS-BDD<sub>d</sub> which can be constructed in polynomial time given the instance  $\langle A, k, B_1, \dots, B_d \rangle$  of SSS<sub>d</sub>.

Let  $S$  be the solution to SSS<sub>d</sub>. We now show that there exists a polynomial of degree at most  $k - d$  which agrees with  $D$  in at least  $k + 1$  points. Define the following degree  $k$  polynomial,

$$g(x) := \prod_{a_i \in S} (x - a_i) = c_0 + c_1x + \dots + c_{k-1}x^{k-1} + x^k$$

The coefficients of this polynomial are the symmetric sums of the roots of  $g(x)$ . Therefore,  $c_{k-d} = (-1)^d B_d, \dots, c_{k-2} = B_2$ , and  $c_{k-1} = -B_1$ . Now define,

$$\begin{aligned} f(x) &= (x^k g(1/x) - x^d p(1/x))/x^d \\ &= c_0 x^{k-d} + c_1 x^{k-d-1} + \dots + c_{k-d} \end{aligned}$$

and note that  $f(x)$  has degree  $k - d = t - 1$ . Also, the constant term of this polynomial is  $c_{k-d} = (-1)^d B_d$ . Hence,  $f(0) = (-1)^d B_d$  and since  $g(a_i) = 0$ , for all  $a_i \in S$ , it follows that  $f(a_i^{-1}) = -p(a_i) = y_i$  for all  $a_i \in S$ . Therefore,  $f(x)$  agrees with  $k + 1 = t + d$  points in  $D$ .

Conversely, we now show that if there is a polynomial  $f(x)$  of degree at most  $t - 1$  which agrees with  $t + d$  points in  $D$ , then there is a solution to SSS<sub>d</sub>. We first observe that if a degree  $t - 1$  polynomial passes through  $t + d$  points of  $D$ , then it has to pass through  $(0, (-1)^d B_d)$ . To show this, assume  $f(x)$  agrees with  $t + d$  points of the form  $(a_i^{-1}, y_i) \in D$ . Let  $g(x)$  be a  $t + d - 1$  degree polynomial defined as,

$$g(x) = x^{t-1}(f(1/x) + p(x))$$

Therefore, if  $f(x) = c_0 + c_1x + \dots + c_{t-1}x^{t-1}$ ,  $g(x)$  can be written as

$$\begin{aligned} g(x) &= x^{t+d-1} + B_1x^{t+d-2} + \dots + (-1)^{d-1}B_{d-1}x^t \\ &\quad + c_0x^{t-1} + c_1x^{t-2} + \dots + c_{t-1} \end{aligned}$$

Since, we know that  $f(a_i^{-1}) = y_i = -p(a_i)$  for  $t + d$  points, we have by definition that  $g(a_i) = 0$  for these  $t + d$   $a_i$ 's. This is a contradiction since  $g(x)$  has degree at most  $t + d - 1$  and it cannot have  $t + d$  roots. Therefore,  $f(0) = c_0 = (-1)^d B_d$ . Also,  $g(x)$  has  $t + d - 1 = k$  roots which have their first  $d$  symmetric sums equal to  $B_1, B_2, \dots, B_d$  respectively. Hence, there exists a solution to the given instance of SSS<sub>d</sub>. ■

B. Proofs from Section III

We now prove Lemma 1. To do so, we need the following lemma whose proof is deferred to the end of the section.

**Lemma 7.** *Let  $j \geq 1$  be an integer and assume that  $|S \cap \{a_{j,2}, b_{j,2}\}| \neq |S \cap \{a_{j,3}, b_{j,3}\}|$ . Furthermore, assume that if  $j \geq 2$ , then  $|S \cap \{a_{i,2}, b_{i,2}\}| = |S \cap \{a_{i,3}, b_{i,3}\}|$  for all  $i < j$ . Then, we have the following two inequalities:*

$$\begin{aligned} \text{(i)} \quad & \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| \geq \Omega(\Sigma^{2\nu-n}). \\ \text{(ii)} \quad & \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| \geq \Sigma^{\Omega(\min(\nu, h))} \times \\ & \left| \sum_{y \in S \setminus \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right|. \end{aligned}$$

We are now ready to prove Lemma 1.

*Proof of Lemma 1:* We proceed by induction on  $i \in [n]$ . At the  $j$ th step of the induction (with  $j \geq 2$ ), we assume that  $|S \cap \{a_{i,2}, b_{i,2}\}| = |S \cap \{a_{i,3}, b_{i,3}\}|$  for all  $i < j$  and we need to show that  $|S \cap \{a_{j,2}, b_{j,2}\}| = |S \cap \{a_{j,3}, b_{j,3}\}|$ . Assume for the sake of contradiction that  $|S \cap \{a_{j,2}, b_{j,2}\}| \neq |S \cap \{a_{j,3}, b_{j,3}\}|$ . Then, Lemma 7 and the fact that for any  $a, b \in \mathbb{R}$ ,  $|a - b| \geq \max(|a|, |b|) - \min(|a|, |b|)$  imply that

$$\begin{aligned} & \left| \sum_{y \in S} y \right| \\ &= \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y + \sum_{y \in S \setminus \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| \\ &\geq \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| - \left| \sum_{y \in S \setminus \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| \end{aligned}$$

$$\geq \left(1 - \frac{1}{\Sigma^{\Omega(\min(\nu, h))}}\right) \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right|$$

$$= \Omega(\Sigma^{2\nu-n})$$

On the other hand, we have that  $|B_1| = B_1 \leq \Sigma^{nh+m+\nu}$ . Hence, we get a contradiction since  $\nu \geq 10(m+nh)$ .

The base case of the induction corresponds to setting  $j = 1$  and it follows along the same lines as above except that we invoke Lemma 7 with  $j = 1$ . This concludes the proof of Lemma 1. ■

We now prove Lemma 2.

*Proof of Lemma 2:* To see this, assume for the sake of contradiction that there exists  $i \in [n]$  s.t.  $r_i \in \{-1, 1\}$  and let  $i^* \in [n]$  be the smallest such  $i$ . Since the  $\Sigma$ -ary representations of each of  $B_1$  and  $\{a_{i,1}, b_{i,1} : i \in [n]\}$  have zeros in the rightmost  $\nu$  digits, the first moment constraint  $\sum_{y \in S} y = B_1$  is equivalent

$$\text{to } \sum_{i=i^*+1}^{\ell} c_i \Sigma^i + r_{i^*} \Sigma^{i^*} = 0 \text{ for some } \ell \in \mathbb{N} \text{ and}$$

$$c_i \in \{-1, 0, 1, 2\} \text{ for all } i \in \{i^*+1, \dots, \ell\}. \text{ Dividing}$$

by  $\Sigma^{i^*}$ , we get that  $\sum_{i=i^*+1}^{\ell} c_i \Sigma^{i-i^*} = -r_{i^*}$ . Since  $\Sigma$  is assumed to be an even integer and since  $r_{i^*} \in \{-1, 1\}$ , the left-hand side is an even integer whereas the right-hand side is an odd integer; a contradiction. ■

The only remaining part is to prove Lemma 7.

*Proof of Lemma 7:* We first derive in the next proposition some inequalities that are satisfied by the integers  $\{a_{i,1}, a_{i,2}, a_{i,3}, b_{i,1}, b_{i,2}, b_{i,3} : i \in [n]\}$  and that will be used to prove Lemma 7.

**Proposition 2.** *Let  $j \in [n]$ .*

- 1)  $|a_{j,1} - b_{j,1}| \geq \Sigma^\nu$ .
- 2) *If  $\nu \geq 10(m+nh)$ , then*

$$\forall i \in [n] : \frac{|a_{j,2}|}{a_{i,1}} \geq \Sigma^{\Omega(\nu)} \text{ and } \frac{|a_{j,2}|}{b_{i,1}} \geq \Sigma^{\Omega(\nu)}$$

- 3) *If  $h \geq 10m$  and  $j < n$ , then*

$$\forall i \in \{j+1, \dots, n\} : \frac{|a_{j,2}|}{|a_{i,2}|} \geq \Sigma^{\Omega(h)} \text{ and}$$

$$\frac{|a_{j,2}|}{|b_{i,2}|} \geq \Sigma^{\Omega(h)}$$

*The same inequalities also hold if we replace  $a_{j,2}$  by  $b_{j,2}$ .*

*Proof of Proposition 2:* The first part of the proposition follows from the assumption (which can be

made without loss of generality) that for each  $j \in [n]$ , there exists a SAT clause that contains exactly one of  $x_j$  and  $\bar{x}_j$ .

We next prove the second part of the proposition. Fix  $j, i \in [n]$ . Without loss of generality, assume that  $a_{j,1} > b_{j,1}$ . Then, the first part of the proposition yields that  $a_{j,1} - b_{j,1} \geq \Sigma^\nu$ . Then, we have that

$$a_{j,2} = \frac{a_{j,1}^2 - b_{j,1}^2}{4\Sigma^j} - \frac{\Sigma^j}{2}$$

$$= \frac{(a_{j,1} - b_{j,1})(a_{j,1} + b_{j,1})}{4\Sigma^j} - \frac{\Sigma^j}{2}$$

$$\geq \frac{\Sigma^\nu \Sigma^\nu}{4\Sigma^j} - \frac{\Sigma^j}{2}$$

$$\geq \Sigma^{2\nu-3n}$$

On the other hand, we have that  $a_{i,1} \leq \Sigma^{\nu+m+nh}$  and  $b_{i,1} \leq \Sigma^{\nu+m+nh}$ . Using the assumption that  $\nu \geq 10(m+nh)$ , we conclude that  $\frac{|a_{j,2}|}{a_{i,1}} \geq \Sigma^{\Omega(\nu)}$  and  $\frac{|a_{j,2}|}{b_{i,1}} \geq \Sigma^{\Omega(\nu)}$ .

We now prove the third part of the proposition. Fix  $j < n$  and  $i \in \{j+1, \dots, n\}$ . Without loss of generality, assume that  $a_{j,1} > b_{j,1}$  and  $a_{i,1} > b_{i,1}$ . Then, we have that

$$a_{j,2} = \frac{(a_{j,1} - b_{j,1})(a_{j,1} + b_{j,1})}{4\Sigma^j} - \frac{\Sigma^j}{2}$$

$$\geq \frac{\Sigma^\nu \times 2 \times \Sigma^{(n-j)h+m+\nu}}{4\Sigma^j} - \frac{\Sigma^j}{2}$$

$$\geq \Sigma^{2\nu+(n-j)h+m-j}$$

On the other hand, we have that

$$a_{i,2} = \frac{(a_{i,1} - b_{i,1})(a_{i,1} + b_{i,1})}{4\Sigma^i} - \frac{\Sigma^i}{2}$$

$$\leq \frac{\Sigma^{\nu+m} \times 2 \times \Sigma^{(n-i)h+m+\nu+1}}{4\Sigma^i}$$

$$= 2 \times \Sigma^{2\nu+m+(n-i)h-i+1}$$

Therefore,

$$\frac{|a_{j,2}|}{|a_{i,2}|} \geq \frac{1}{2} \Sigma^{(i-j)h-m+(i-j)-1} \geq \frac{1}{2} \Sigma^{h-m} = \Sigma^{\Omega(h)}$$

where the last equality above follows from the assumption that  $h \geq 10m$ . The proof that  $\frac{|a_{j,2}|}{|b_{i,2}|} \geq \Sigma^{\Omega(h)}$  follows along the same lines. ■

Observe that for any  $j \geq 1$ , if  $|S \cap \{a_{j,2}, b_{j,2}\}| \neq |S \cap \{a_{j,3}, b_{j,3}\}|$ , then  $S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}$  is one of the following possible sets:

$$\{a_{j,2}\}, \{b_{j,2}\}, \{a_{j,3}\}, \{b_{j,3}\}, \{a_{j,2}, b_{j,2}, a_{j,3}\},$$

$$\{a_{j,2}, b_{j,2}, b_{j,3}\}, \{a_{j,2}, a_{j,3}, b_{j,3}\}, \{b_{j,2}, a_{j,3}, b_{j,3}\}$$

In each of these cases, the following inequality is satisfied

$$\left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| \geq \min(|a_{j,2}, b_{j,2}|) - \Sigma^j \quad (1)$$

We now prove the first part of Lemma 7. Since  $|S \cap \{a_{j,2}, b_{j,2}\}| \neq |S \cap \{a_{j,3}, b_{j,3}\}|$  and assuming that  $a_{j,1} > b_{j,1}$  without loss of generality, Eq. (1) gives that

$$\begin{aligned} \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| &\geq \frac{(a_{j,1} - b_{j,1})(a_{j,1} + b_{j,1})}{4\Sigma^j} \\ &\quad - \frac{3\Sigma^j}{2} \\ &\geq \frac{\Sigma^\nu \times 2 \times \Sigma^{\nu+m+(n-j)h}}{4\Sigma^j} \\ &\quad - \frac{3\Sigma^j}{2} \\ &= \Omega(\Sigma^{2\nu-n}) \end{aligned}$$

We now prove the second part of Lemma 7, which we first show for the case where  $j \geq 2$  and  $|S \cap \{a_{i,2}, b_{i,2}\}| = |S \cap \{a_{i,3}, b_{i,3}\}|$  for all  $i < j$ . These

equalities imply that  $\left| \sum_{y \in S \cup_{i=1}^{j-1} \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} y \right| \leq \Sigma^{O(n)}$  and hence

$$\begin{aligned} \min(|a_{j,2}, b_{j,2}|) &\geq \Sigma^{2\nu-3n} \\ &\geq \Sigma^{\Omega(\nu)} \left| \sum_{y \in S \cup_{i=1}^{j-1} \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} y \right| \quad (2) \end{aligned}$$

Using Equation (1) and Proposition 2, we then have that

$$\begin{aligned} \left| \sum_{y \in S \cap \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right| &\geq \min(|a_{j,2}, b_{j,2}|) - \Sigma^j \\ &\geq \frac{1}{\Theta(n)} \sum_{j=1}^{\Theta(n)} \min(|a_{j,2}, b_{j,2}|) - \Sigma^j \\ &\geq \frac{1}{\Theta(n)} \left( \sum_{y \in S \cup_{i=1}^n \{a_{i,1}, b_{i,1}\}} \Sigma^{\Omega(\nu)} y \right. \\ &\quad \left. + \sum_{y \in S \cup_{i=j+1}^n \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} \Sigma^{\Omega(h)} |y| \right. \\ &\quad \left. + \Sigma^{\Omega(\nu)} \left| \sum_{y \in S \cup_{i=1}^{j-1} \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}} y \right| \right) - \Sigma^j \end{aligned}$$

$$\geq \Sigma^{\Omega(\min(\nu, h))} \times \left| \sum_{y \in S \setminus \{a_{j,2}, b_{j,2}, a_{j,3}, b_{j,3}\}} y \right|$$

where the last inequality above follows from the triangle inequality.

The proof for the case where  $j = 1$  follows along the same lines as above except that the quantity

$\sum_{y \in S \cup_{i=1}^{j-1} \{a_{i,2}, b_{i,2}, a_{i,3}, b_{i,3}\}}$   $y$  is replaced by 0 and Equation (2) is not needed. ■

### C. Proofs from Section IV

*Proof of Proposition 1:* For the first identity, we have that  $a'_{i,2} + a'_{i,3} = 2\alpha_i = b'_{i,2} + b'_{i,3}$ . For the second identity, we have that

$$\begin{aligned} &a'_{i,1}{}^2 + a'_{i,2}{}^2 + a'_{i,3}{}^2 - (b'_{i,1}{}^2 + b'_{i,2}{}^2 + b'_{i,3}{}^2) \\ &= a'_{i,1}{}^2 - b'_{i,1}{}^2 + (a'_{i,2} - b'_{i,2})(a'_{i,2} + b'_{i,2}) \\ &\quad + (a'_{i,3} - b'_{i,3})(a'_{i,3} + b'_{i,3}) \\ &= a'_{i,1}{}^2 - b'_{i,1}{}^2 - \beta_i(2\alpha_i + \gamma_i) + \beta_i(2\alpha_i - \gamma_i) \\ &= a'_{i,1}{}^2 - b'_{i,1}{}^2 - 2\beta_i\gamma_i \\ &= 0 \end{aligned}$$

For the third identity, we have that

$$\begin{aligned} &a'_{i,1}{}^3 + a'_{i,2}{}^3 + a'_{i,3}{}^3 - (b'_{i,1}{}^3 + b'_{i,2}{}^3 + b'_{i,3}{}^3) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 + (a'_{i,2} - b'_{i,2})(a'_{i,2} + a'_{i,2}b'_{i,2} + b'_{i,2}{}^2) \\ &\quad + (a'_{i,3} - b'_{i,3})(a'_{i,3} + a'_{i,3}b'_{i,3} + b'_{i,3}{}^2) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 - \beta_i(a'_{i,2}{}^2 + a'_{i,2}b'_{i,2} + b'_{i,2}{}^2) \\ &\quad + \beta_i(a'_{i,3}{}^2 + a'_{i,3}b'_{i,3} + b'_{i,3}{}^2) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 - \beta_i \left( (a'_{i,2} - a'_{i,3})(a'_{i,2} + a'_{i,3}) \right. \\ &\quad \left. + a'_{i,2}b'_{i,2} - a'_{i,3}b'_{i,3} + (b'_{i,2} - b'_{i,3})(b'_{i,2} + b'_{i,3}) \right) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 - \beta_i \left( (\gamma_i - \beta_i)2\alpha_i + a'_{i,2}b'_{i,2} \right. \\ &\quad \left. - a'_{i,3}b'_{i,3} + (\gamma_i + \beta_i)2\alpha_i \right) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 - 4\alpha_i\beta_i\gamma_i - \beta_i(a'_{i,2}b'_{i,2} - a'_{i,3}b'_{i,3}) \\ &= a'_{i,1}{}^3 - b'_{i,1}{}^3 - 6\alpha_i\beta_i\gamma_i \\ &= 0 \end{aligned}$$

where the penultimate equality uses the fact that  $a'_{i,2}b'_{i,2} - a'_{i,3}b'_{i,3} = 2\alpha_i\gamma_i$ . ■

We now use Proposition 1 in order to check that the first 3 moments of  $S$  are equal to  $B_1$ ,  $B_2$  and



$B_3$  respectively, and thereby finish the completeness analysis of Theorem 5.

$$\begin{aligned}
\sum_{y \in S} y &= \sum_{i \in [n]: x_i=1} a'_{i,1} + a'_{i,2} + a'_{i,3} \\
&+ \sum_{i \in [n]: x_i=0} b'_{i,1} + b'_{i,2} + b'_{i,3} \\
&= \sum_{i \in [n]: x_i=1} a'_{i,1} + \sum_{i \in [n]: x_i=0} b'_{i,1} \\
&+ \sum_{i \in [n]: x_i=1} a'_{i,2} + a'_{i,3} + \sum_{i \in [n]: x_i=0} b'_{i,2} + b'_{i,3} \\
&= \sum_{i \in [n]: x_i=1} a'_{i,1} + \sum_{i \in [n]: x_i=0} b'_{i,1} \\
&+ \sum_{i \in [n]} a'_{i,2} + a'_{i,3} \\
&= \sum_{i \in [n]: x_i=1} \Sigma^{\nu+m+h-1}(a_i^v)_{\Sigma^h} + \Sigma^{\nu}(a_i^c)_{\Sigma} \\
&+ \sum_{i \in [n]: x_i=0} \Sigma^{\nu+m+h-1}(b_i^v)_{\Sigma^h} + \Sigma^{\nu}(b_i^c)_{\Sigma} \\
&+ \sum_{i \in [n]} a'_{i,2} + a'_{i,3} \\
&= \Sigma^{\nu+m+h-1} \left( \sum_{i \in [n]: x_i=1} a_i^v + \sum_{i \in [n]: x_i=0} b_i^v \right)_{\Sigma^h} \\
&+ \Sigma^{\nu} \left( \sum_{i \in [n]: x_i=1} a_i^c + \sum_{i \in [n]: x_i=0} b_i^c \right)_{\Sigma} \\
&+ \sum_{i \in [n]} a'_{i,2} + a'_{i,3} \\
&= \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^{\nu} (1^m)_{\Sigma} \\
&+ \sum_{i \in [n]} a'_{i,2} + a'_{i,3} \\
&= B_1
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\sum_{y \in S} y^2 &= \sum_{i \in [n]: x_i=1} a'^2_{i,1} + a'^2_{i,2} + a'^2_{i,3} \\
&+ \sum_{i \in [n]: x_i=0} b'^2_{i,1} + b'^2_{i,2} + b'^2_{i,3} \\
&= \sum_{i \in [n]} a'^2_{i,1} + a'^2_{i,2} + a'^2_{i,3} \\
&= B_2
\end{aligned}$$

Finally,

$$\begin{aligned}
\sum_{y \in S} y^3 &= \sum_{i \in [n]: x_i=1} a'^3_{i,1} + a'^3_{i,2} + a'^3_{i,3} \\
&+ \sum_{i \in [n]: x_i=0} b'^3_{i,1} + b'^3_{i,2} + b'^3_{i,3}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i \in [n]} a'^3_{i,1} + a'^3_{i,2} + a'^3_{i,3} \\
&= B_3
\end{aligned}$$

We now turn to the proofs of Lemmas 3, 4 and 5. We will use the following proposition whose proof is similar to that of Proposition 2 (because of the fact that  $|a'_{i,\ell}| = \Theta(|a_{i,\ell}|)$  and  $|b'_{i,\ell}| = \Theta(|b_{i,\ell}|)$  for every  $i \in [n]$  and  $\ell \in [3]$ ).

**Proposition 3.** *Let  $j \in [n]$ .*

- 1)  $|a'_{j,1} - b'_{j,1}| \geq \Sigma^{\nu}$ .
- 2) *If  $\nu \geq 10(m + nh)$ , then*  
 $\forall i \in [n] : \frac{|a'_{j,2}|}{a'_{i,1}} \geq \Sigma^{\Omega(\nu)}$  *and*  $\frac{|a'_{j,2}|}{b'_{i,1}} \geq \Sigma^{\Omega(\nu)}$
- 3) *If  $h \geq 10m$  and  $j < n$ , then*  
 $\forall i \in \{j+1, \dots, n\} : \frac{|a'_{j,2}|}{|a'_{i,2}|} \geq \Sigma^{\Omega(h)}$
- 4) *If  $\nu \geq 10(m + nh)$ ,  $j \geq 2$ ,  $i < j$  and  $|S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}| = 2$ , then  $a'^2_{j,2}/\Sigma^{\Omega(\nu)}$  is at least*

$$\left| a'^2_{i,2} + a'^2_{i,3} - \sum_{y \in S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right|$$

*The same inequalities also hold if we replace  $a'_{j,2}$  by one of  $\{b'_{j,2}, a'_{j,3}, b'_{j,3}\}$  or if we replace  $a'_{i,2}$  by one of  $\{b'_{i,2}, a'_{i,3}, b'_{i,3}\}$*

*Proof of Lemma 3:* Follows along the same lines as the proof of Lemma 1 but using Proposition 3 instead of Proposition 2.  $\blacksquare$

The proof of Lemma 4 uses the following lemma.

**Lemma 8.** *Let  $j \geq 1$  be an integer. Furthermore, assume that if  $j \geq 2$ , then  $|S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}| = 2$  for all  $i < j$ . Then,  $\min_{y \in \{a'^2_{j,2}, b'^2_{j,2}, a'^2_{j,3}, b'^2_{j,3}\}} y^2$  is lower*

*bounded by:*

$$\begin{aligned}
\text{(i)} \quad & \Sigma^{\Omega(\min(\nu, h))} \times \left( \sum_{i=1}^n a'^2_{i,1} + \sum_{i=j+1}^n (a'^2_{i,2} + a'^2_{i,3}) + \sum_{i=1}^{j-1} \left| a'^2_{i,2} + a'^2_{i,3} - \sum_{y \in S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right| \right) \\
\text{(ii)} \quad & \Sigma^{\Omega(\min(\nu, h))} \times \left( \sum_{y \in S \cup_{i=j+1}^n \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 + \sum_{y \in S \cup_{i=1}^n \{a'_{i,1}, b'_{i,1}\}} y^2 \right)
\end{aligned}$$

*Proof of Lemma 8:* Similar to that of Lemma 7 but uses Proposition 3 instead of Proposition 2. ■

*Proof of Lemma 4:* We proceed by induction on  $i \in [n]$ . At the  $j$ th step of the induction (with  $j \geq 2$ ), we assume that  $|S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}| = 2$  for all  $i < j$  and we need to show that  $|S \cap \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}| = 2$ . Assume for the sake of contradiction that  $|S \cap \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}| \neq 2$ . We have that

$$\left| B_2 - \sum_{y \in S \cup_{i=1}^j \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right| = |q_1 - q_2|$$

$$\text{where } q_1 := (a'_{j,2}{}^2 + a'_{j,3}{}^2) - \sum_{y \in S \cap \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2$$

and

$$\begin{aligned} q_2 := & \sum_{i=1}^{j-1} \left( \sum_{y \in S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 - a'_{i,2}{}^2 - a'_{i,3}{}^2 \right) \\ & - \sum_{i=1}^n a'_{i,1}{}^2 - \sum_{i=j+1}^n (a'_{i,2}{}^2 + a'_{i,3}{}^2) \end{aligned}$$

Then, Lemma 8 and the fact that for any  $a, b \in \mathbb{R}$ ,  $|a - b| \geq \max(|a|, |b|) - \min(|a|, |b|)$  imply that

$$\begin{aligned} & \left| B_2 - \sum_{y \in S \cup_{i=1}^j \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right| \geq |q_1| - |q_2| \\ & \geq \left| (a'_{j,2}{}^2 + a'_{j,3}{}^2) - \sum_{y \in S \cap \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2 \right| \\ & - \left( \sum_{i=1}^{j-1} \left| a'_{i,2}{}^2 + a'_{i,3}{}^2 - \sum_{y \in S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right| \right) \\ & + \sum_{i=1}^n a'_{i,1}{}^2 + \sum_{i=j+1}^n (a'_{i,2}{}^2 + a'_{i,3}{}^2) \\ & \geq \Omega \left( \min_{y \in \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2 \right) \\ & - \frac{1}{\Sigma \Omega(\min(\nu, h))} \min_{y \in \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2 \\ & = \Omega \left( \min_{y \in \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2 \right) \end{aligned}$$

On the other hand, the fact that  $\sum_{y \in S} y^2 = B_2$  and

Lemma 8 imply that

$$\begin{aligned} & \left| B_2 - \sum_{y \in S \cup_{i=1}^j \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \right| \\ & = \sum_{y \in S \cup_{i=j+1}^n \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y^2 \\ & + \sum_{y \in S \cup_{i=1}^n \{a'_{i,1}, b'_{i,1}\}} y^2 \\ & \leq \frac{1}{\Sigma \Omega(\min(\nu, h))} \min_{y \in \{a'_{j,2}, b'_{j,2}, a'_{j,3}, b'_{j,3}\}} y^2 \end{aligned}$$

which yields a contradiction. ■

*Proof of Lemma 5:* Lemmas 3 and 4 imply that for every  $i \in [n]$ ,  $\sum_{y \in S \cap \{a'_{i,2}, b'_{i,2}, a'_{i,3}, b'_{i,3}\}} y = 2\alpha_i + r_i \times \Sigma^i$  for some  $r_i \in \{-1, 0, 1\}$ . The next lemma argues that, in fact, all  $r_i$ 's are equal to 0.

**Lemma 9.** For every  $i \in [n]$ ,  $r_i = 0$ .

*Proof of Lemma 9:* To see this, assume for the sake of contradiction that there exists  $i \in [n]$  s.t.  $r_i \in \{-1, 1\}$  and let  $i^* \in [n]$  be the smallest such  $i$ . Note that the first moment constraint  $\sum_{y \in S} y =$

$$B_1 \text{ is equivalent to } \sum_{y \in S \cup_{i=1}^n \{a'_{i,1}, b'_{i,1}\}} y + \sum_{i=1}^n r_i \times \Sigma^i$$

$= \Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^{\nu} (1^m)_{\Sigma}$ . Since the  $\Sigma$ -ary representations of each of the positive integers  $\Sigma^{\nu+m+h-1} (1^n)_{\Sigma^h} + \Sigma^{\nu} (1^m)_{\Sigma}$  and  $\{a'_{i,1}, b'_{i,1} : i \in [n]\}$  have zeros in the rightmost  $\nu$  digits, this constraint

is also equivalent to  $\sum_{i=i^*+1}^{\ell} c_i \Sigma^i + r_{i^*} \Sigma^{i^*} = 0$  for some  $\ell \in \mathbb{N}$  and  $c_i \in \{-1, 0, 1, 2\}$  for all  $i \in \{i^* + 1, \dots, \ell\}$ .

Dividing by  $\Sigma^{i^*}$ , we get that  $\sum_{i=i^*+1}^{\ell} c_i \Sigma^{i-i^*} = -r_{i^*}$ .

Since  $\Sigma$  is assumed to be an even integer and since  $r_{i^*} \in \{-1, 1\}$ , the left-hand side is an even integer whereas the right-hand side is an odd integer; a contradiction. ■

This completes the proof of Lemma 5. ■