NP-hardness of testing equivalence to sparse polynomials and to constant-support polynomials

Omkar Baraskar
University of Waterloo
obaraska@uwaterloo.ca

Agrim Dewan
Indian Institute of Science
agrimdewan@iisc.ac.in

Chandan Saha[*]
Indian Institute of Science
chandan@iisc.ac.in

Pulkit Sinha[†]
University of Waterloo
psinha@uwaterloo.ca

Abstract

An s-sparse polynomial has at most s monomials with nonzero coefficients. The Equivalence Testing problem for sparse polynomials (ETsparse) asks to decide if a given polynomial f is equivalent to (i.e., in the orbit of) some s-sparse polynomial. In other words, given $f \in \mathbb{F}[\mathbf{x}]$ and $s \in \mathbb{N}$, ETsparse asks to check if there exist $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$ such that $f(A\mathbf{x} + \mathbf{b})$ is s-sparse. We show that ETsparse is NP-hard over any field $\mathbb{F}$, if f is given in the sparse representation, i.e., as a list of nonzero coefficients and exponent vectors. This answers a question posed by Gupta, Saha and Thankey (SODA 2023) and also, more explicitly, by Baraskar, Dewan and Saha (STACS 2024). The result implies that the Minimum Circuit Size Problem (MCSP) is NP-hard for a *dense* subclass of depth-3 arithmetic circuits if the input is given in sparse representation. We also show that approximating the smallest s_0 such that a given s-sparse polynomial f is in the orbit of some s_0-sparse polynomial to within a factor of $s^{\frac{1}{3}-\epsilon}$ is NP-hard for any $\epsilon > 0$; observe that s-factor approximation is trivial as the input is s-sparse. Finally, we show that for any constant $\sigma \geq 6$, checking if a polynomial (given in sparse representation) is in the orbit of some support-σ polynomial is NP-hard. Support of a polynomial f is the maximum number of variables present in any monomial of f. These results are obtained via direct reductions from the 3-SAT problem.

[*] Partially supported by a MATRICS grant of the Science and Engineering Research Board, DST, India.

[†] Partially supported by a Mike and Ophelia Lazardis Fellowship, and partially supported by NSERC Canada.

A preliminary version of this paper will appear in the proceedings of the 51st International Colloquium on Automata, Languages and Programming (ICALP), 2024.

Contents

1 Introduction

The Polynomial Equivalence (PE) problem asks to decide if two polynomials, given as lists of coefficients, are equivalent. Polynomials $f, g \in \mathbb{F}[\mathbf{x}]$ are *equivalent*, denoted as $f \sim g$, if there is an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and a $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$ such that $f = g(A\mathbf{x} + \mathbf{b})$. Equivalent polynomials represent the same function up to a change of the coordinate system.[1] The PE problem is thus regarded as the algebraic analog of the graph isomorphism (GI) problem. PE is at least as hard as GI [AS05, Kay11], but we do not know if it is much harder than GI. There is, in fact, a cryptographic authentication scheme based on the presumed average-case hardness of PE [Pat96]. Is PE NP-hard? Over finite fields, PE is not NP-hard unless the polynomial hierarchy collapses [Sax06, Thi98]. In contrast, PE is not even known to be decidable over $\mathbb{Q}$. With the aim of gaining more insight into the complexity of testing polynomial equivalence, a natural variant of PE has been studied in the literature. This variant is known as *equivalence testing*.

In the following discussion, whenever we write "circuit(s)" and "formula(s)", we mean arithmetic circuit(s) and arithmetic formula(s), respectively, unless mentioned otherwise. [2]

Equivalence testing. Equivalence testing (ET) comes in two flavors – ET for polynomial families and ET for circuit classes. ET for a polynomial family $\mathscr{F}$ is defined as follows: given a *single* polynomial f, check if it is equivalent to some $g \in \mathscr{F}$. This variant of PE was introduced in [Kay12a, Kay11], wherein randomized polynomial-time ET algorithms were provided for the permanent, determinant, and elementary and power symmetric polynomial families. Subsequently, efficient ET algorithms were given for various other important polynomial families, such as the iterated matrix multiplication (IMM) family [KNST19] (see Section 1.4). These algorithms are efficient even if f is provided as a circuit or a black-box.[3] ET for a circuit class $\mathscr{C}$ (a.k.a testing equivalence to $\mathscr{C}$) is defined similarly: given a polynomial f, decide if it is equivalent to some polynomial g that is computable by a circuit in $\mathscr{C}$. Recently, efficient ET algorithms have been given for read-once formulas [GST23] and a special subclass of sparse polynomials, namely t-design polynomials for constant t [BDS24]. Sparse polynomials are depth-2 circuits.[4] It is natural to ask whether or not ET can be solved efficiently for *general* sparse polynomials. This question was posed in [GST23] and also, more explicitly, in [BDS24].

Before proceeding to discuss ET for sparse polynomials, we point out a subtle difference between ET for polynomial families and that for circuit classes. The polynomial families for which ET has been studied so far are such that if f is equivalent to some g in the family, then g is unique and it can be readily identified from f. For example, if f is equivalent to some determinant polynomial[5], then we know which one simply from the number of variables of f. Moreover, polynomials in most of these families admit well-known polynomial-size circuits. So, a circuit for g can be derived once it is identified. Thus, if f is also given as a circuit, then ET for such a family reduces to PE with the input polynomials given as circuits. Over finite fields, this version of PE is in $\mathsf{AM} \cap \mathsf{coAM}$ and hence unlikely to be NP-hard. On the other hand, in the case of ET for a circuit class, if f is equivalent to some circuit C in the class, then C need *not* be unique, and further, C may not be easily deducible from f. This leaves us with the prospect of proving that ET is hard for some natural circuit class. Do sparse polynomials form such a class?

[1] Over $\mathbb{R}$, an invertible map $\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b}$ is simply a combination of rotation, reflection, scaling, and translation.

[2] An *arithmetic circuit* is like a Boolean circuit but with AND and OR replaced by $\times$ and $+$ gates, and with edges labelled by $\mathbb{F}$-elements. It computes a polynomial over $\mathbb{F}$. A *formula* is a circuit whose underlying graph is a tree.

[3] Black-box access to f means oracle access to f, we get $f(\mathbf{a})$ from a query point $\mathbf{a}$ in one unit time. It is as if f is given as a "hidden" circuit and the only operation we are allowed to do is evaluate the circuit at chosen points.

[4] We assume that a depth-2 circuit has a $+$ gate on top and a bottom layer of $\times$ gates. If the top gate is a $\times$ gate, then ET can be solved efficiently using polynomial factorization algorithms [KT90].

[5] The n^2-variate determinant polynomial is the determinant of the matrix $(x_{i,j})_{i,j \in [n]}$ of formal variables.

ET for sparse polynomials. An n-variate, degree-d polynomial is *s-sparse* if it has at most s monomials with nonzero coefficients. An s-sparse polynomial is computable by a depth-2 circuit having top fan-in s. Sparse polynomials have been extensively studied in algebraic complexity, particularly with regard to identity testing [KS01, LV03], interpolation [BT88, GKS90, KS01, BJ14], and factorization [vzGK85, BSV20] (see the tutorial [Roc18] and the references therein for more algorithms involving sparse polynomials). ET provides yet another avenue to understand these "basic" polynomials better. ET for sparse polynomials asks to check if a given polynomial is sparse in some coordinate system. More formally, given a polynomial f as an arithmetic circuit and an $s \in \mathbb{N}$, decide if there is an s-sparse polynomial g such that $f \sim g$. This problem was studied in [GK93] over $\mathbb{Q}$, wherein an exponential in n^4 time algorithm was provided. There has not been any significant progress on this problem since that work. The lack of improvements in the complexity for over three decades makes one wonder:

Is ET for sparse polynomials NP-hard?

In this work, we answer this question in the affirmative over *any* field (see the first part of Theorem 1) even if the input f is provided as a depth-2 circuit. The result answers the question posed in [GST23, BDS24]. To our knowledge, the theorem gives the first example of a natural circuit class for which ET is provably hard.

Although ET for sparse polynomials (ETsparse) is a fairly natural problem, there is a deeper reason to study ETsparse that originates from the expressive power of affine projections of sparse polynomials and the *Minimum Circuit Size Problem* (MCSP) for depth-3 circuits. We discuss this reason below to motivate ETsparse when the input is a *homogeneous* polynomial.

1.1 ETsparse and MCSP for depth-3 circuits

First, we need a few definitions: A polynomial g is an *affine projection* of f if $g = f(A\mathbf{x} + \mathbf{b})$ for some $A \in \mathbb{F}^{|\mathbf{x}|\times|\mathbf{x}|}$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$. If $\mathbf{b} = 0$, we say g is a *linear projection* of f; additionally, if $A \in \mathrm{GL}(|\mathbf{x}|)$, we say g is in the *orbit* of f, denoted as $\mathrm{orb}(f)$. Depth-3 circuits form a highly expressive class [GKKS16, Tav15]. A depth-3 ($\Sigma\Pi\Sigma$) circuit is a circuit with a $+$ gate on top, a middle layer of $\times$ gates, and a bottom layer of $+$ gates. A depth-3 circuit with a top fan-in of s is an affine projection of an s-sparse polynomial. Thus, the problem of deciding if a given f is an affine projection of an s-sparse polynomial is closely related to MCSP for depth-3 circuits. We say "closely related to" instead of "the same as" because the size of a depth-3 circuit is determined by not only its top fan-in but also its formal degree.

MCSP. The complexity of MCSP for Boolean circuits has baffled researchers for over six decades. MCSP for a Boolean circuit class $\mathscr{C}$ ($\mathscr{C}$-MCSP) takes input the truth table of an n-variate Boolean function f and a parameter $s \in \mathbb{N}$ and asks to check if f is computable by a circuit in $\mathscr{C}$ of size at most s. There are intriguing connections between MCSP and several other areas such as cryptography [KC00, AD17], learning theory [CIKK16], average-case complexity [Hir18], and proof complexity [PS19]. Whether or not MCSP for general Boolean circuits is NP-hard is a long-standing open question. It is known that MCSP is NP-hard for DNF [Mas79, AHM$^+$06] and DNF $\circ$ XOR formulas [HOS18]. But no NP-hardness result is known (under deterministic polynomial-time reductions) for more general circuit models such as AC^0 circuits.[6] This is

[6] However, strong hardness results are known for several powerful circuit models under randomized or quasi-polynomial time or subexponential time reductions [Ila20, ILO20, Ila21, Hir22].

not too surprising as [KC00] showed that NP-hardness of $\mathscr{C}$-MCSP under *natural*[7] deterministic polynomial-time reductions implies a $2^{\Omega(n)}$ lower bound for $\mathscr{C}$, unless NP $\subseteq$ SUBEXP. Unfortunately, such strong lower bounds are not known even for depth-3 Boolean circuits. However, a $2^{\Omega(n)}$ lower bound is known for XOR $\circ$ AND $\circ$ XOR formulas [Raz87], which are depth-3 arithmetic circuits over $\mathbb{F}_2$ and are like DNF $\circ$ XOR formulas but with the top OR gate replaced by an XOR gate. In fact, a $2^{\Omega(n)}$ lower bound is known for depth-3 arithmetic circuits over any fixed finite field [GR98]. This raises hope that we will be able to prove the hardness of MCSP for depth-3 arithmetic circuits over finite fields. But how is the input given in the case of MCSP for arithmetic circuits? And what about depth-3 circuits over fields of characteristic 0?

MCSP for arithmetic circuits: Input representation and model of computation. In the Boolean setting of MCSP, one of the main reasons for assuming that the input is a truth table is that the assumption puts MCSP in NP. Analogously, in the algebraic setting, we could assume that the polynomial is given in the dense representation as a list of $\binom{n+d}{n}$ coefficients. But observe that even if the input is given as an arithmetic circuit, MCSP is in the complexity class MA over finite fields. This is because verifying if two circuits compute the same polynomial is the polynomial identity testing problem, which admits a randomized polynomial-time algorithm [DL78, Zip79, Sch80]. Furthermore, class MA equals NP, assuming a widely believed circuit lower bound [IW97]. A succinct input representation also opens up the possibility of proving NP-hardness of MCSP for models, such as depth-3 circuits over fields of characteristic 0, for which strong exponential lower bounds are unknown (the MCSP hardness to lower bound implication in [KC00] needs the input in the dense format). The current best lower bound for depth-3 circuits over fields of characteristic 0 is quasi-polynomial in n [LST21, AGK$^+$23].

It is, therefore, reasonable to assume that the input polynomial is given succinctly as a circuit which should only facilitate our efforts in proving NP-hardness of MCSP for arithmetic circuit classes. For example, there is an instance in the Boolean setting wherein succinct representation of the input helped prove NP-hardness of MCSP long before such a hardness result was shown with respect to the dense representation – it is the case of the *partial* MCSP problem [HJLT96, Hir22]. In this work, we assume that the input is given as a depth-2 circuit, i.e., as a list of nonzero coefficients, and exponent vectors in unary – this is the *sparse representation*.[8]

A few remarks are in order concerning the model of computation. Over finite fields, we assume the Turing machine model. However, over arbitrary fields of characteristic 0, it is natural to consider an arithmetic model of computation (similar to the Blum-Shub-Smale machine model [BSS89]) that allows us to store a field element in unit space and perform an arithmetic operation in unit time. Over $\mathbb{Q}$, it is not clear if MCSP for arithmetic circuits is even decidable in the Turing machine model. But, if we confine our search to size-s circuits whose field constants are $s^{O(1)}$ bit rational numbers, then we can work with the Turing machine model.

MCSP for homogeneous depth-3 circuits. The size of a $\Sigma\Pi\Sigma$ circuit is primarily determined by its formal degree and its top fan-in, whereas the size of a homogeneous depth-3 (hom-$\Sigma\Pi\Sigma$) circuit is mainly decided by its top fan-in (the formal degree of a $\Sigma\Pi\Sigma$ circuit is the maximum fan-in of the middle layer of $\times$ gates). MCSP for $\Sigma\Pi\Sigma$ circuits can be defined as follows: given f and $D, s \in \mathbb{N}$, decide if there is a $\Sigma\Pi\Sigma$ circuit with formal degree bounded by D and top fan-in bounded by s that computes f. Similarly, MCSP for hom-$\Sigma\Pi\Sigma$ circuits is defined as:

[7] i.e., the size of the output of the reduction and the output parameter s depend only on the size of the input instance. Almost all reductions that show NP-hardness of problems are natural.

[8] Sparse representations of polynomials are also used in computer algebra systems wherein the exponent vector is given in binary. As the degree is $n^{O(1)}$ in this work (except on one occasion; see the remark following Theorem 3), whether or not the exponent vector is given in unary or binary makes little difference.

given a homogeneous f and $s \in \mathbb{N}$, check if there is a hom-$\Sigma\Pi\Sigma$ circuit with top fan-in at most s that computes f. In order to prove NP-hardness of $\Sigma\Pi\Sigma$-MCSP, it is *necessary* to prove NP-hardness of hom-$\Sigma\Pi\Sigma$-MCSP. The reason is: a polynomial $f(x_1, x_2, \ldots, x_n)$ has a $\Sigma\Pi\Sigma$ circuit with formal degree bounded by D and top fan-in bounded by s if and only if the homogeneous polynomial $z^D f(x_1 z^{-1}, x_2 z^{-1}, \ldots, x_n z^{-1})$ has a hom-$\Sigma\Pi\Sigma$ circuit with top fan-in bounded by s. Also, if the reduction in a hypothetical proof of NP-hardness of hom-$\Sigma\Pi\Sigma$-MCSP has a certain simple feature, then it would imply NP-hardness of $\Sigma\Pi\Sigma$-MCSP (see the last remark following Proposition 3.4). Hence, it is natural to study the hardness of hom-$\Sigma\Pi\Sigma$-MCSP first.

NP-hardness of MCSP is known for two interesting subclasses of hom-$\Sigma\Pi\Sigma$ circuits, namely depth-3 powering circuits [Shi16] and set-multilinear $\Sigma\Pi\Sigma$ circuits [Hås90]; the top fan-in's of circuits in these two classes correspond to Waring rank and tensor rank, respectively. Perhaps an appealing evidence in favor of NP-hardness of hom-$\Sigma\Pi\Sigma$-MCSP is a proof of NP-hardness of MCSP for a “dense" subclass of hom-$\Sigma\Pi\Sigma$ circuits. Intuitively, $\mathscr{C}$ is a *dense* subclass of hom-$\Sigma\Pi\Sigma$ circuits if every hom-$\Sigma\Pi\Sigma$ circuit can be approximated “infinitesimally closely" by circuits in $\mathscr{C}$.[9] Unfortunately, depth-3 powering circuits and set-multilinear $\Sigma\Pi\Sigma$ circuits are *not* dense inside hom-$\Sigma\Pi\Sigma$ circuits.[10] On the other hand, *orbits of homogeneous sparse polynomials* form a dense subclass of hom-$\Sigma\Pi\Sigma$ circuits.[11] It is natural to ask:

Is MCSP for orbits of homogeneous sparse polynomials NP-hard?

MCSP for orbits of homogeneous sparse polynomials is exactly the ETsparse problem on inputs that are homogeneous polynomials. The second part of Theorem 1 answers the question positively over any field.

Approximating the sparse-orbit complexity. Call the smallest s_0 such that f is in the orbit of an s_0-sparse polynomial, the *sparse-orbit complexity* of f. Theorem 1 shows that sparse-orbit complexity is hard to compute in the worst case.

Is sparse-orbit complexity easy to approximate?

In Theorem 2, we show that approximating the sparse-orbit complexity of a given s-sparse polynomial (homogeneous or not) to within a $s^{1/3-\epsilon}$ factor is NP-hard for any $\epsilon \in (0, 1/3)$. As the input is s-sparse, approximating the sparse-orbit complexity to within a factor s is trivial.

1.2 ET for constant-support polynomials

ET is efficiently solvable for two special sparse polynomial families, namely the power symmetric polynomial $\mathsf{PSym} := x_1^d + \ldots + x_n^d$ [Kay11] and the sum-product polynomial $\mathsf{SP} := \sum_{i\in[s]} \prod_{j\in[d]} x_{i,j}$ [MS21, Kay11]. What makes ET easy for these sparse polynomials? Explanations were provided in [GST23, BDS24]: SP is a read-once formula; it is also a 1-design polynomial. PSym is a 1-design polynomial, but it is also a support-1 polynomial.

[9] Formally, a subclass $\mathscr{C}$ of hom-$\Sigma\Pi\Sigma$ circuits is *dense* if there are polynomial functions $p, q : \mathbb{N} \to \mathbb{N}$ such that the following holds: For $n, d, s \in \mathbb{N}$, the coefficient vector of every n-variate degree-d polynomial computable by a size-s hom-$\Sigma\Pi\Sigma$ circuit is in the *Zariski closure* of the set of coefficient vectors of $p(nds)$-variate degree-d polynomials computable by size-$q(nds)$ circuits in $\mathscr{C}$. Here, “size" means “top fan-in".

[10] Circuits of these two classes have small read-once algebraic branching programs (ROABPs), and the class ROABP is closed under Zariski closure [For16]. So, the closures of these two classes are also contained inside ROABPs. But, there are explicit $O(n)$ size hom-$\Sigma\Pi\Sigma$ circuits that require $2^{\Omega(n)}$ size ROABPs [ST21, KNS20].

[11] Every n-variate degree-d hom-$\Sigma\Pi\Sigma$ circuit of size-s is a linear projection of an s-sparse degree-d homogeneous polynomial in at most sd variables. It is well known that linear projections of f are contained in the Zariski closure of the orbit of f over fields of characteristic 0 (see [ST21] for a proof of this fact).

Is ET easy for constant-support polynomials?

In Theorem 3, we show that checking if a given f is in the orbit of a support-6 polynomial is NP-hard; this answers the question in the negative.

1.3 Our results

We now state our results formally. The ETsparse problem is defined as follows.

Problem 1.1 (ETsparse)**.** Given a polynomial $f \in \mathbb{F}[\mathbf{x}]$ in its sparse representation and an integer s, check if there exist an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and a $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$ such that $f(A\mathbf{x} + \mathbf{b})$ is s-sparse.

Our first result, Theorem 1, shows the NP-hardness of ETsparse over any field.

Theorem 1 (ETsparse is NP-hard)**.** 1. *Let $\mathbb{F}$ be any field. There is a deterministic polynomial-time many-one reduction from* 3-SAT *to* ETsparse *over* $\mathbb{F}$.

2. *Let $\mathbb{F}$ be any field. There is a deterministic polynomial-time many-one reduction from* 3-SAT *to* ETsparse *over* $\mathbb{F}$ *where the input polynomial to the* ETsparse *problem is homogeneous.*

Remarks. 1. Part 2 of the theorem subsumes part 1. We state parts 1 and 2 separately because of two reasons: One, part 1 has a simpler proof. Two, the degree parameters in the proof of part 1 have a better upper bound in comparison to that in the proof of part 2.

2. The reduction is *natural*[12] and has the feature that a satisfying assignment can be mapped to a sparsifying invertible $A \in \{-1, 0, 1\}^{|\mathbf{x}| \times |\mathbf{x}|}$ and vice versa. So, ETsparse is NP-hard even when A is restricted to having only $\{-1, 0, 1\}$ entries.

3. The authors of [CGS23] showed the undecidability over $\mathbb{Z}$ of testing if a given f is shift equivalent to some sparse polynomial (f is shift equivalent to a polynomial g, if there exists a $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$ s.t $f = g(\mathbf{x} + \mathbf{b})$). However, their result does not imply the intractability of ETsparse as testing shift equivalence to a sparse polynomial is a special case of ETsparse when A is the identity map.

4. The authors of [BDS24] gave a randomized polynomial-time ET algorithm for *random*[13] sparse polynomials, assuming black-box access to the input. Such average-case results for hard problems are not unusual in both algebraic and Boolean settings. In the algebraic setting, MCSP is NP-hard for depth-3 powering circuits [Shi16] and for set-multilinear depth-3 circuits [Hås90].[14] Yet, [KS19] gave average-case learning algorithms for both these circuit models. In the Boolean setting, [DF89] gave polynomial-time algorithms for average cases of NP-hard problems like Graph 3-colorability.

5. Depth-3 power circuits, set-multilinear depth-3 circuits, and shifted sparse polynomials are all contained inside ROABPs. So, these models admit polynomial-time (improper) learning algorithms [BBB$^+$00, KS06] and quasi-polynomial-time hitting sets [AGKS15, FS13]. Orbits of sparse polynomials require exponential size ROABPs [ST21]; we cannot expect to improperly learn them via ROABPs. Theorem 1 suggests that proper learning orbits of sparse polynomials is likely hard. Nonetheless, there is a quasi-polynomial time hitting set for orbits of sparse polynomials [MS21, ST21].

[12] unless $\mathrm{char}(\mathbb{F}) = 2$. See the remark following Observation 3.6 in Section 3.5.1.

[13] A random s-sparse degree-d polynomial in their work was defined to be a polynomial where each monomial is formed independently of the others by selecting d variables uniformly at random from the variable set; the coefficients are allowed to be arbitrary.

[14] In a depth-3 powering circuit, each term is a power of a linear form. In a set-multilinear depth-3 circuit, the variable set is partitioned into d sets such that each term is a product of d linear forms, the i^{th} linear form being a linear form in the i^{th} set.

We prove Theorem 1 in Section 3. Next, we define the gap version of ETsparse.

Problem 1.2 (α-gap-ETsparse)**.** Let $\alpha > 1$ be a parameter. Given a polynomial $f \in \mathbb{F}[\mathbf{x}]$ in its sparse representation and an integer s_0, output:

- YES, if there exist an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}$ such that $f(A\mathbf{x} + \mathbf{b})$ is s_0-sparse.
- NO, if for all $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}$, $f(A\mathbf{x} + \mathbf{b})$ has sparsity at least αs_0.

Our second result, Theorem 2, shows that α-gap-ETsparse is NP-hard for $\alpha = s^{\frac{1}{3}-\epsilon}$, where s is the sparsity of the input polynomial f and $\epsilon \in (0, \frac{1}{3})$ is an arbitrary constant. Theorem 2 is proven in Section 4. From Theorem 2, we get Corollary 1.1 which states that $s^{\frac{1}{3}-\epsilon}$ factor approximation of the sparse-orbit complexity of an s-sparse polynomial is NP-hard.

Theorem 2 ($s^{\frac{1}{3}}$-gap-ETsparse is NP-hard)**.** *Let $\epsilon \in (0, \frac{1}{3})$ be an arbitrary constant.*

1. *Let $\mathbb{F}$ be any field. There exists a deterministic polynomial-time many-one reduction from* 3-SAT *to* $s^{\frac{1}{3}-\epsilon}$-gap-ETsparse *over $\mathbb{F}$ where the input polynomial in* $s^{\frac{1}{3}-\epsilon}$-gap-ETsparse *is s-sparse.*
2. *Let $\mathbb{F}$ be any field. There exists a deterministic polynomial-time many-one reduction from* 3-SAT *to* $s^{\frac{1}{3}-\epsilon}$-gap-ETsparse *over $\mathbb{F}$ where the input polynomial in* $s^{\frac{1}{3}-\epsilon}$-gap-ETsparse *is homogeneous and s-sparse.*

Remarks. 1. Like Theorem 1, part 2 of Theorem 2 subsumes part 1. We state parts 1 and 2 separately because of two reasons. One, part 1 has a simpler proof. Two, the degree parameters in the proof of part 1 have a better bound in comparison to that in the proof of part 2.

2. It may be possible to improve the constant $\frac{1}{3}$ in $s^{\frac{1}{3}-\epsilon}$ using a more careful analysis.
3. Interestingly, the above results are obtained without invoking the celebrated PCP theorem [AS98, ALM$^+$98, Din07].

Corollary 1.1. Let $0 < \epsilon < \frac{1}{3}$ be an arbitrary constant.

1. Let $\mathbb{F}$ be any field. It is NP-hard to compute $s^{\frac{1}{3}-\epsilon}$ factor approximation of the sparse-orbit complexity when the input is an s-sparse polynomial over $\mathbb{F}$.
2. Let $\mathbb{F}$ be any field. It is NP-hard to compute $s^{\frac{1}{3}-\epsilon}$ factor approximation of the sparse-orbit complexity when the input is an s-sparse homogeneous polynomial over $\mathbb{F}$.

Remarks. Thus, approximating the sparse-orbit complexity within a certain super-constant factor is NP-hard over any field. In contrast, [SWZ17, BIJL18, Swe18] showed that approximating the tensor rank (which corresponds to the smallest top fan-in of a set-multilinear depth-3 circuit) within a $1 + \delta$ factor, where $\delta \approx 0.0005$, is NP-hard over any field. We do not know of any hardness of approximation result for the Waring rank (which corresponds to the smallest top fan-in of a depth-3 powering circuit).

Now, we formally define the support of a polynomial.

Definition 1.1 (Support of a polynomial)**.** For a monomial $\mathbf{x}^{\boldsymbol{\alpha}}$, where $\boldsymbol{\alpha}$ is the exponent vector, the support of $\mathbf{x}^{\boldsymbol{\alpha}}$, $\mathrm{Supp}(\mathbf{x}^{\boldsymbol{\alpha}})$, is the number of variables with non-zero exponent. The support of a polynomial f, $\mathrm{Supp}(f)$, is the maximum support size over all the monomials of f.

Thus, a polynomial has support σ if there exists a monomial with support σ and no other monomial has support $> \sigma$. The ET problem for constant-support polynomials and a stronger version of it are defined next (henceforth, σ is assumed to be a constant).

Problem 1.3 (ETsupport)**.** Given a polynomial $f \in \mathbb{F}[\mathbf{x}]$ in its sparse representation and an integer σ, check if there exists an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that $\mathrm{Supp}(f(A\mathbf{x})) \leq \sigma$.

Problem 1.4 ($(\sigma + 1)$-to-σ ETsupport)**.** Given a polynomial $f \in \mathbb{F}[\mathbf{x}]$ with support $\sigma + 1$ in its sparse representation, check if there exists an $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$ such that $\mathrm{Supp}(f(A\mathbf{x})) \leq \sigma$.

Remarks. 1. Unlike ETsparse, checking if f is in the *orbit* of a constant-support polynomial is the same as checking if f is equivalent to a constant-support polynomial. This follows from the observation that $\mathrm{Supp}(f(\mathbf{x})) = \mathrm{Supp}(f(\mathbf{x} + \mathbf{b}))$ for any $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$.

Our third and last result, Theorem 3, shows that ETsupport and $(\sigma + 1)$-to-σ ETsupport are NP-hard. We prove Theorem 3 in Section 5.

Theorem 3 (ETsupport is NP-hard)**.** *Let $\sigma \geq 6$ be a constant and $\mathbb{F}$ be a field with* $\mathrm{char}(\mathbb{F}) = 0$ *or* $> \sigma + 1$*. There is a deterministic polynomial-time many-one reduction from* 3-SAT *to* ETsupport *over* $\mathbb{F}$*. In particular,* 3-SAT *reduces to* $(\sigma + 1)$-to-σ ETsupport *in deterministic polynomial time.*

Remarks. 1. Over fields of finite characteristic, it is assumed that the exponent vectors corresponding to the monomials of the input polynomial are given in binary.

We prove Theorems 1, 2 and 3 by direct reductions from 3-SAT, and at the beginning of Sections 3, 4 and 5, we give proof sketches of the respective reductions.

1.4 Related work

Results on ET. As mentioned in Section 1, the study of ET was initiated in [Kay11] where efficient ET algorithms were given for the power symmetric and the elementary symmetric polynomials. Following this, efficient ET algorithms were given for several other important polynomial families and circuit classes such as the permanent [Kay12a], the determinant [Kay12a, Gro12, GGKS19], the iterated matrix multiplication (IMM) polynomial [KNST19, MNS20], the continuant polynomial [MS21], read-once formulas (ROFs) [GST23], and design polynomials [BDS24, GS19]. ET algorithms have also been used to give efficient reconstruction algorithms; for example, [KNS19] gave an efficient average-case reconstruction algorithm for low-width ABPs based on ET for the determinant.

The sum-product polynomial $\mathrm{SP} := \sum_{i \in [s]} \prod_{j \in [d]} x_{i,j}$ is a rare example for which three different ET algorithms are known. The SP polynomial can be computed by an ROF. So, the ET algorithm for ROFs [GST23, Kay11], which is based on analyzing the Hessian determinant, gives ET for SP. Also, SP is a design polynomial, so the ET algorithm of [BDS24], which uses the vector space decomposition framework of [KS19, GKS20], holds for SP. The authors of [MS21] also observed that ET for SP follows from the reconstruction algorithm in [KS19]. A third ET algorithm for SP can be designed by analyzing its Lie algebra. Observe that the orbit of SP is a dense subclass of homogeneous depth-3 circuits. However, as ET for SP is easy, it does not provide any supporting evidence for the hardness of MCSP for homogeneous depth-3 circuits.

Results on PE. Quadratic form equivalence can be solved in polynomial time over $\mathbb{R}, \mathbb{C}$, finite fields and $\mathbb{Q}$ (assuming access to integer factoring oracle) [Sax06, Wal13]. These algorithms are based on well-known classification of quadratic forms [Lam04, Ara11]. In contrast, [AS05] showed that cubic form equivalence (CFE) is at least as hard as graph isomorphism. The authors of [GQ23] showed that CFE is polynomial time equivalent to several other problems like group isomorphism for p-groups, algebra isomorphism, trilinear form equivalence, etc.

A variant of PE is the shift equivalence problem, where given two n-variate polynomials f and g, one needs to check if there exists $\mathbf{b} \in \mathbb{F}^n$ such that $f(\mathbf{x}) = g(\mathbf{x}+\mathbf{b})$. The author of [Gri97] gave a deterministic algorithm over characteristic 0 fields, a randomized algorithm over prime residue fields and a quantum algorithm over characteristic 2 fields for shift equivalence testing. All these algorithms have running time polynomial in the dense representation of the input, that is, for n-variate, degree-d polynomials given in the verbose representation as input, the running time is $\mathrm{poly}(\binom{n+d}{d})$. The authors of [DOS14] gave a randomized algorithm for shift equivalence testing assuming black-box access to n-variate polynomials f and g with degree bound d and circuit size bound s. Their algorithm runs in $\mathrm{poly}(n, d, s)$ time. Another randomized polynomial-time shift equivalence test is given in [Kay12a].

A variant of the shift equivalence problem, call it the sparse shift equivalence problem, is where a single n-variate polynomial $f(\mathbf{x})$ and a positive integer t are given as inputs, and the objective is to decide if there exists $\mathbf{b} \in \mathbb{F}^n$ such that $f(\mathbf{x}+\mathbf{b})$ is t-sparse. The authors of [LS95] studied this problem for univariate polynomials over $\mathbb{Q}$ and gave sufficient conditions for the uniqueness and rationality of a t-sparsifying shift. The authors of [GL00] extended these conditions to multivariate polynomials and gave two algorithms for computing t-sparsifying shifts for n-variate, degree-d polynomials, one where the input polynomial has finitely many t-sparsifying shifts and the other for polynomials without any finiteness restriction on the number of t-sparsifying shifts. The running time of the first algorithm is $(dt)^{O(n)}$ without randomization and $t^{O(n)}$ with randomization, while that of the second one is $(nt)^{O(n^2)}$. The authors of [CGS23] showed that the sparse shift equivalence problem is undecidable over $\mathbb{Z}$ by showing a reduction from polynomial solvability over $\mathbb{Z}$ to the sparse shift equivalence problem. They also showed the NP-hardness of a gap version of the sparse shift equivalence problem over $\mathbb{R}, \mathbb{Q}$ and finite fields.

The scaling equivalence problem is yet another variant of PE, which involves checking for given n-variate polynomials f and g whether there exists a diagonal matrix $S \in \mathrm{GL}(n, \mathbb{F})$ such that $f(\mathbf{x}) = g(S\mathbf{x})$. The authors of [BRS17] gave a randomized polynomial-time algorithm for the scaling equivalence problem over $\mathbb{R}$.

Hardness results. The author of [Kay12a] showed that the problem of checking if a polynomial is an affine projection of another polynomial is NP-hard via a reduction from Graph 3-Colorability. Computing the tensor rank (which is MCSP for depth-3 set multilinear circuits) is NP-hard [Hås90], so is computing the Waring rank for a polynomial (which is MCSP for depth-3 powering circuits) [Shi16]. In the Boolean world, [KS08] showed that there is no polynomial-time algorithm to $n^{1-\delta}$-approximate, where $\delta > 0$ is an arbitrarily small constant, a DNF with minimum number of terms for any n-variate Boolean function given as a truth table, unless NP is decidable in quasi-polynomial time. It is also known that $(1+\delta)$-approximate MCSP, where $\delta \approx 0.0005$ is a constant, is NP-hard for set-multilinear depth three circuits [SWZ17, Swe18, BIJL18]. In [KS09], it was shown that depth-3 arithmetic circuits cannot be PAC-learned in polynomial time unless the length of a shortest nonzero vector of an n-dimensional lattice can be approximated to within a factor of $\tilde{O}(n^{1.5})$ in polynomial time by a quantum algorithm. This means it is hard to PAC-learn the class of Boolean functions that match the output of depth-3 arithmetic circuits on the Boolean hypercube.

Hitting sets and lower bounds for orbits of sparse polynomials. The authors of [MS21] gave a quasi-polynomial time hitting set[15] construction for the orbits of sparse polynomials. Orbits of sparse polynomials form a subclass of homogeneous depth-3 circuits. The authors of [NW97] showed that any homogeneous depth-3 circuit computing the n-variate elementary symmetric polynomial of degree $2d$ has size $\Omega((\frac{n}{4d})^d)$. The authors of [KST16] showed the existence of an explicit polynomial family in n variables and degree d, with $d \geq n$, for which any homogeneous depth-3 circuit computing it must be of size at least $2^{\Omega(n)}$.

1.5 Roadmap of the paper

In Section 2, we state a few useful observations and claims, the proofs of which appear in Section B of the appendix. The proof of part one of Theorem 1 for fields of characteristic zero is given in Sections 3.1-3.3. Section 3.4 has the proof of part two of the same theorem for characteristic zero fields. In Section 3.5, we prove Theorem 1 for fields of finite characteristics. Similarly, the proofs of parts one and two of Theorem 2 for characteristic zero fields appear in Sections 4.1 and 4.2, respectively. Section 4.3 contains the proof of Theorem 2 over fields of finite characteristics. In Section 5, we prove Theorem 3. For simplicity, we ignore the effect of translation vectors in the above-mentioned sections. In Section A of the appendix, we show how to handle translation vectors. The missing proofs of the observations, claims, lemmas, and propositions in Sections 3, 4 and 5 appear in Sections C, D and E of the appendix, respectively.

2 Preliminaries

2.1 Definitions and notations

For $n, a, b \in \mathbb{N}$, $[n]$ denotes the set $\{1, 2 \ldots, n\}$ and $[a, b]$ denotes the integers from a to b, both inclusive. A polynomial is *homogeneous* if all its monomials have the same total degree. The set of invertible linear transforms in n variables over a field $\mathbb{F}$ is denoted by $\mathrm{GL}(n, \mathbb{F})$. For a polynomial $f \in \mathbb{F}[\mathbf{x}]$, the action of a linear transform $A \in \mathbb{F}^{|\mathbf{x}| \times |\mathbf{x}|}$ on its variables is denoted by $f(A\mathbf{x})$ as well as by $A(f)$. The *sparsity* of a polynomial f, denoted as $\mathcal{S}(f)$, is the number of monomials in f with non-zero coefficients. For a polynomial f, $\mathrm{var}(f)$ denotes the set of variables that occur in at least one monomial of f. We have used the notation $f \sim g$ earlier to denote $f = g(A\mathbf{x} + \mathbf{b})$. Henceforth, we will ignore the translation vector $\mathbf{b}$ in the main body of the discussion for simplicity but mention the necessary changes in the proofs or point to appropriate sections when translations are involved. Thus, for polynomials f and g, $f \sim g$ will mean $f(\mathbf{x}) = g(A\mathbf{x})$ where $A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})$. Similarly, the *orbit* of a polynomial f will denote the set $\{f(A\mathbf{x}), A \in \mathrm{GL}(|\mathbf{x}|, \mathbb{F})\}$. The *degree* of a monomial is its total degree, and the degree of a polynomial f is the maximum degree amongst all monomials in f. The x-degree of a monomial is the degree of the variable x in the monomial.

Definition 2.1 (Degree separated polynomials)**.** Polynomials f and g are *degree separated* if no monomial of f has the same degree as a monomial of g. Similarly, f and g are degree separated *with respect to a variable* x if no monomial of f has the same x-degree as a monomial of g.

The *set of degrees* of a polynomial is the set of distinct degrees of all the monomials in the polynomial. For example, the set of degrees of $f(x_1, x_2) = x_1^2 + x_1 x_2 + 4x_2$ is $\{2, 1\}$. A *linear form* is a homogeneous degree one polynomial. An *affine form* is a degree one polynomial.

[15] A hitting set for a circuit class $\mathscr{C}$ is a set $S \subseteq \mathbb{F}^{|\mathbf{x}|}$ such that for every non-zero polynomial $f(\mathbf{x})$ computable by a circuit $C \in \mathscr{C}$, $f(\mathbf{a}) \neq 0$ for some $\mathbf{a} \in S$.

2.2 Algebraic preliminaries

The proofs of the observations and claims stated in this section can be found in Appendix B.

Observation 2.1. Let f and g be polynomials such that $f \sim g$. Then, f and g have the same set of degrees for the monomials. Thus, if f and g are degree separated, then $f \not\sim g$.

Observation 2.2. If f and g are degree separated (or degree separated with respect to some variable), then $\mathcal{S}(f+g) = \mathcal{S}(f) + \mathcal{S}(g)$.

Observation 2.3. If f and g are degree separated, $f_1 \sim f$ and $g_1 \sim g$, then $\mathcal{S}(f_1 + g_1) = \mathcal{S}(f_1) + \mathcal{S}(g_1)$.

Observation 2.4 analyzes the sparsity of powers of linear forms. Observation 2.5 is a special case of Observation 2.4 and is stated separately because it is simpler and is invoked many times. Observation 2.6 analyzes the sparsity of powers of affine forms.

Observation 2.4. Let ℓ be a linear form in m variables and $d \in \mathbb{N}$. If $\text{char}(\mathbb{F}) = 0$, $\mathcal{S}(\ell^d) = \binom{d+m-1}{m-1}$, and if $\text{char}(\mathbb{F}) = p$, $\mathcal{S}(\ell^d) = \prod_{i=0}^{k} \binom{e_i+m-1}{m-1}$, where $d = \sum_{i=0}^{k} e_i p^i$, $e_i \in [0, p-1]$.

Observation 2.5. If $\text{char}(\mathbb{F}) = 0$ and ℓ be a linear form in exactly two variables, then $\mathcal{S}(\ell^d) = d+1$. The result holds for characteristic p fields if $p > d$ or if $d = p^k - 1$ for some $k \in \mathbb{N}$. Further, if ℓ is a linear form in more than two variables and d is as before, then $\mathcal{S}(\ell^d) \geq d+1$.

Observation 2.6. Let $h = \ell + c_0$, where ℓ is a linear form in at least one variable and $c_0 \in \mathbb{F}\backslash\{0\}$, then $\mathcal{S}(h^d) \geq \mathcal{S}(\ell^d) + 1$. More precisely, $\mathcal{S}(h^d) \geq d+1$ holds if $\text{char}(\mathbb{F}) = 0$ or if $\text{char}(\mathbb{F}) = p$ and $p > d$ or $d = p^k - 1$ for some $k \in \mathbb{N}$.

Claim 2.1 analyzes the sparsity of polynomials divisible by a power of some linear form in at least two variables and is used to prove part two of Theorems 1 and 2. Claim 2.2 analyzes the support of monomials under invertible linear transforms and is used to prove Theorem 3.

Claim 2.1. Let $\text{char}(\mathbb{F}) = 0$. If $f \in \mathbb{F}[\mathbf{x}]$ is a non-zero polynomial divisible by ℓ^d for some linear form ℓ in at least two variables, then $\mathcal{S}(f) \geq d+1$. The claim also holds for characteristic p fields, where the degree of f is less than p.

Claim 2.2. Let $\sigma, d, n \in \mathbb{N}$, $d \geq \sigma$, $f = (x_1 \cdots x_n)^d$, and $\ell_1, \ldots, \ell_n$ be linearly independent linear forms in $x_1, \ldots, x_n$. If $|\cup_{i=1}^{n} \text{var}(\ell_i)| \geq \sigma$ and $g := f(\ell_1 \cdots \ell_n)$, then $\text{Supp}(g) \geq \sigma$. The claim holds if $\text{char}(\mathbb{F}) = 0$, or $\text{char}(\mathbb{F}) = p$ with $p > d$, or $p > \sigma$ and $d = p^k - 1$ for some $k \in \mathbb{N}$.

3 NP-hardness of ETsparse

In this section, we prove Theorem 1. We first show the reduction over characteristic 0 fields in the non-homogeneous case without considering translations for ease of understanding. Section 3.4 shows the reduction over characteristic 0 fields in the homogeneous case. In Section 3.5, the reduction is shown to hold over finite characteristic fields for both the non-homogeneous and the homogeneous case. In Appendix C, we prove the lemmas and the observations of this section. Appendix A.1 shows how the reduction holds while also considering translations.[16]

[16] Note that for two homogeneous polynomials f and g, $f(\mathbf{x}) = g(A\mathbf{x} + \mathbf{b})$ implies $f(\mathbf{x}) = g(A\mathbf{x})$, where $A \in \text{GL}(|\mathbf{x}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{x}|}$. Hence, it suffices to prove part 2 of Theorem 1 without translations.

Proof sketch. The reduction maps each variable and clause of a 3-CNF[17] ψ to distinct degree separated polynomials which, summed together, give the polynomial f. As the summands are degree separated, the sparsity of f under invertible transforms can be analyzed by doing so for individual polynomials. The degrees are chosen such that f is equivalent to an s-sparse polynomial (for a suitable sparsity parameter s) if and only if $\psi \in$ 3-SAT.

3.1 Constructing f and s

Let ψ be a 3-CNF in variables $\mathbf{x} := \{x_1, x_2 \dots x_n\}$ and m clauses:

$$\psi = \wedge_{k=1}^{m} \vee_{j \in C_k} (x_j \oplus a_{k,j}),$$

where C_k denotes the set of indices of the variables in the k^{th} clause and $a_{k,j} \in \{0,1\}$. Let $\mathbf{y} := \{y_1, y_2 \dots y_n\}$, x_0 be a new variable and $\mathbf{z} := \{x_0\} \sqcup \mathbf{x} \sqcup \mathbf{y}$. For $d_1, d_2, d_3, d_4 \in \mathbb{N}$, consider the following polynomials:

- Corresponding to variable x_i, where $i \in [n]$, define $Q_i(\mathbf{z})$ as:

  $$Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z}), \text{ where}$$
  $$Q_{i,1}(\mathbf{z}) := x_0^{(3i-2)d_1} x_i^{d_2},\ Q_{i,2}(\mathbf{z}) := x_0^{(3i-1)d_1}(y_i + x_i)^{d_3} \text{ and } Q_{i,3}(\mathbf{z}) := x_0^{3id_1}(y_i - x_i)^{d_3}.$$

  Intuitively, $Q_{i,2}$ and $Q_{i,3}$ correspond to assigning 0 and 1, respectively, to x_i in ψ. $Q_{i,1}$ is used to establish a mapping between satisfying assignments and sparsifying transforms.

- For the k^{th} clause, $k \in [m]$, define $R_k(\mathbf{z}) := x_0^{(3n+k)d_1} \prod_{j \in C_k} (y_j + (-1)^{a_{k,j}} x_j)^{d_4}$.

Define $s := 1 + n(3 + d_3) + m(d_4 + 1)^2$ and the polynomial f as:

$$f(\mathbf{z}) := x_0^{d_1} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{1}$$

The following conditions are imposed on the d_i's:

$$d_1 \geq \max(s, d_2 + 1),\ d_2 \geq 2d_3,\ d_3 \geq m(d_4 + 1)^2 + 1, \text{ and } d_4 \geq m. \tag{2}$$

For characteristic 0 fields, the inequalities of (2) can be converted to equalities. Thus, we get

$$\begin{aligned} &d_4 = m,\ d_3 = m(m+1)^2 + 1 = O(m^3) \implies s = O(nm^3) \\ &d_2 = 2m(m+1)^2 + 2 = O(m^3),\ d_1 = 1 + n(4 + m(m+1)^2) + m(m+1)^2 = O(nm^3). \end{aligned} \tag{3}$$

Note, for the above choices of d_3 and d_2, $s \geq d_2 + 1$. Hence, d_1 is set to s. Under the conditions of (2) the following observations hold.

Observation 3.1. For all $i \in [n], k \in [m]$, the polynomials $x_0^{d_1}$, $Q_{i,1}(\mathbf{z})$, $Q_{i,2}(\mathbf{z})$, $Q_{i,3}(\mathbf{z})$ and $R_k(\mathbf{z})$ are degree separated from one another. Also, $Q_i(\mathbf{z})$ is degree separated from other $Q_j(\mathbf{z})$'s, for $i, j \in [n]$ and $i \neq j$. Similarly, $R_k(\mathbf{z})$ is degree separated from $R_l(\mathbf{z})$ for $k, l \in [m]$ and $k \neq l$.

Observation 3.2. The degree of f is $(3n + m)d_1 + 3d_4 = (mn)^{O(1)}$.

Observation 3.3. $\mathcal{S}(f(\mathbf{z})) = 1 + n(2d_3 + 3) + m(d_4 + 1)^3$ and $\text{Supp}(f) = 7$.

[17] We assume, without loss of generality, that each clause of a 3-CNF has 3 distinct variables. This can be achieved by introducing extra variables for clauses with < 3 variables.

3.2 The forward direction

Proposition 3.1 shows how a satisfiable ψ implies the existence of an invertible A, such that $\mathcal{S}(f(A\mathbf{z})) \leq s$ by constructing A from a satisfying assignment $\mathbf{u} \in \{0,1\}^n$ of ψ.

Proposition 3.1. Let $\mathbf{u} = (u_1, \ldots, u_n) \in \{0,1\}^n$ be such that $\psi(\mathbf{u}) = 1$. Then $\mathcal{S}(f(A\mathbf{z})) \leq s$, where A is as:

$$A : x_0 \mapsto x_0, x_i \mapsto x_i, y_i \mapsto y_i + (-1)^{u_i} x_i, \quad \forall i \in [n]. \tag{4}$$

Proof. It follows from the definition of f in (1), Observations 3.1 and 2.3 that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

Thus, it suffices to analyze the sparsity of $A(x_0^{d_1})$, $Q_i(A\mathbf{z})$'s and $R_k(A\mathbf{z})$'s. Now, $\mathcal{S}(A(x_0^{d_1})) = 1$ as $A(x_0^{d_1}) = x_0^{d_1}$. We now analyze $\mathcal{S}(Q_i(A\mathbf{z}))$ for $i \in [n]$. If $u_i = 0$, then

$$Q_{i,1}(A\mathbf{z}) = x_0^{(3i-2)d_1} x_i^{d_2},\ Q_{i,2}(A\mathbf{z}) = x_0^{(3i-1)d_1}(y_i + 2x_i)^{d_3} \text{ and } Q_{i,3}(A\mathbf{z}) = x_0^{3id_1} y_i^{d_3}.$$

If $u_i = 1$, then

$$Q_{i,1}(A\mathbf{z}) = x_0^{(3i-2)d_1} x_i^{d_2},\ Q_{i,2}(A\mathbf{z}) = x_0^{(3i-1)d_1} y_i^{d_3} \text{ and } Q_{i,3}(A\mathbf{z}) = x_0^{3id_1}(y_i - 2x_i)^{d_3}.$$

By Observation 2.5 (for linear forms in two variables over characteristic 0 fields), if $u_i = 0$ then $\mathcal{S}(Q_{i,2}(A\mathbf{z})) = d_3 + 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z})) = 1$ and, if $u_i = 1$ then $\mathcal{S}(Q_{i,2}(A\mathbf{z})) = 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z})) = d_3 + 1$. In either case, by Observations 3.1 and 2.3,

$$\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z})) = d_3 + 3.$$

For the k^{th} clause, $k \in [m]$, the action of A on the corresponding polynomial R_k is:

$$R_k(A\mathbf{z}) = x_0^{(3n+k)d_1} \prod_{j \in C_k} (y_j + ((-1)^{a_{k,j}} + (-1)^{u_j})x_j)^{d_4}.$$

As the multiplicands in $R_k(A\mathbf{z})$ do not share any variables, $\mathcal{S}(R_k(A\mathbf{z}))$ is the product of the sparsity of the multiplicands. Since $\psi(\mathbf{u}) = 1$, therefore in the k^{th} clause there exists $j \in C_k$ such that $a_{k,j} \neq u_j$. For that j, $(y_j + ((-1)^{a_{k,j}} + (-1)^{u_j})x_j)^{d_4} = y_j^{d_4}$. As at least one literal is true in every clause under $\mathbf{u}$, $\mathcal{S}(R_k(A\mathbf{z})) \leq (d_4 + 1)^2$ using Observation 2.5. Thus,

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})) \leq 1 + n(d_3 + 3) + m(d_4 + 1)^2 = s.$$

□

3.3 The reverse direction

Now, we show that $(f, s) \in$ ETsparse implies $\psi \in$ 3-SAT by showing that the permuted and scaled versions of the transform of (4) form all the viable sparsifying invertible linear transforms. This is where the constraints on the d_i's are used. So, let $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ be such that $\mathcal{S}(f(A\mathbf{z})) \leq s$. Lemma 3.1 shows that $A(x_0)$ is just a variable by leveraging $d_1 \geq s$.

Lemma 3.1. Without loss of generality, $A(x_0) = x_0$.

The proof of Lemma 3.2 uses $d_2 \geq 2d_3$ while that of Lemma 3.3 uses $d_3 \geq m(d_4+1)^2+1$.

Lemma 3.2. For any invertible A and $i \in [n]$:

$$\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z})) \geq d_3 + 3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined in Section 3.1. Equality holds if and only if under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (-1)^{u_i} X_i$$

for some scaled variables $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0,1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z})) \neq d_3+3$, then $\mathcal{S}(Q_i(A\mathbf{z})) \geq 2d_3+3$.

Lemma 3.3. Under the given A, $\mathcal{S}(Q_i(A\mathbf{z})) = d_3+3$ holds for all $i \in [n]$.

Lemmas 3.1, 3.2 and 3.3 together show that A is a permuted scaled version of the transform of (4). We can assume A to be as described in (4) without loss of generality as permutation and non-zero scaling of variables do not affect the sparsity of a polynomial. Proposition 3.2 shows how a satisfying assignment can be derived from A using $d_4 \geq m$.

Proposition 3.2. With A as described in (4), $\mathbf{u} = (u_1, \ldots, u_n)$ is a satisfying assignment for ψ.

Proof. Suppose not; then there exists $k \in [m]$ such that the k^{th} clause, $\vee_{j \in C_k}(x_j \oplus a_{k,j})$, in ψ is unsatisfied. Since this clause is unsatisfied, $u_j = a_{k,j}$ for all $j \in C_k$. Thus, $R_k(A\mathbf{z}) = x_0^{(3n+k)d_1} \prod_{j \in C_k}(y_j \pm 2x_j)^{d_4}$, where R_k is as defined in Section 3.1, and $\mathcal{S}(R_k(A\mathbf{z})) = (d_4+1)^3 \geq (m+1)(d_4+1)^2$ by Observation 2.5, the fact that $R_k(A\mathbf{z})$ is a product of linear forms not sharing variables, and the condition $d_4 \geq m$. By the definition of f and s in Section 3.1, Observations 3.1 and 2.3, it holds that

$$\begin{aligned}\mathcal{S}(f(A\mathbf{z})) &\geq \mathcal{S}(A(x_0)^{d_1}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \mathcal{S}(R_k(A\mathbf{z})) \\ &\geq 1 + n(3+d_3) + m(d_4+1)^2 + (d_4+1)^2 = s + (d_4+1)^2 > s,\end{aligned}$$

a contradiction. Thus, $\mathbf{u}$ is a satisfying assignment for ψ. □

3.4 The homogeneous case

We show a modification of the construction in Section 3.1 which, along with arguments similar to those in Sections 3.2 and 3.3, can be used to prove Theorem 1 for homogeneous polynomials over characteristic 0 fields. Because the polynomials are homogeneous, we cannot use degree separation like in the non-homogeneous case. Instead, we introduce a new variable y_0, a new degree parameter $d_5 \in \mathbb{N}$, and redefine $Q_i(\mathbf{z})$ and $R_k(\mathbf{z})$ of Section 3.1 along with modified constraints on the d_i's so that:

1. Each polynomial is homogeneous with the same degree and is divisible by $x_0^{d_1}$ and $y_0^{d_2}$.

2. Each polynomial has a distinct x_0 degree.

The divisibility condition ensures that both x_0 and y_0 map to scaled variables under a sparsifying invertible linear transform (see Lemma 3.4 and its proof). Due to the second condition, the polynomials are degree separated with respect to x_0 (see Observation 3.4). This fact is used to show that, under a sparsifying invertible linear transform, the polynomials are degree separated with respect to x_0 (see Lemma 3.5 and its proof). Formally, let x_0, $\mathbf{x}$ and $\mathbf{y}$ be as defined in Section 3.1 and y_0 be a new variable. Define $\mathbf{z} := \mathbf{x} \sqcup \mathbf{y} \sqcup \{x_0\} \sqcup \{y_0\}$. Let $d_1, d_2, d_3, d_4, d_5 \in \mathbb{N}$. Consider the following polynomials:

1. For each variable x_i, $i \in [n]$, define $Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z})$, where

$$\begin{aligned} Q_{i,1}(\mathbf{z}) &:= x_0^{d_1+(3i-2)(d_3+1)} y_0^{d_2+(3n+m-3i+3)(d_3+1)-d_3} x_i^{d_3}, \\ Q_{i,2}(\mathbf{z}) &:= x_0^{d_1+(3i-1)(d_3+1)} y_0^{d_2+(3n+m-3i+2)(d_3+1)-d_4} (y_i + x_i)^{d_4}, \\ Q_{i,3}(\mathbf{z}) &:= x_0^{d_1+3i(d_3+1)} y_0^{d_2+(3n+m-3i+1)(d_3+1)-d_4} (y_i - x_i)^{d_4}. \end{aligned}$$

2. For the k^{th} clause, $k \in [m]$, define

$$R_k(\mathbf{z}) := x_0^{d_1+(3n+k)(d_3+1)} y_0^{d_2+(m-k+1)(d_3+1)-3d_5} \prod_{j \in C_k} (y_j + (-1)^{a_{k,j}} x_j)^{d_5}.$$

Define $s := 1 + n(d_4+3) + m(d_5+1)^2$ and impose the following conditions on the d_i's:

$$\begin{aligned} &d_1 \geq d_2 + (3n+m+1)(d_3+1) + 1,\ d_2 \geq \max((3n+m+1)(d_3+1), s) + 1, \\ &d_3 \geq 2d_4,\ d_4 \geq m(d_5+1)^2 + 1,\ d_5 \geq m. \end{aligned} \tag{5}$$

For characteristic 0 fields, the inequalities of (5) can be converted to equalities to get

$$\begin{aligned} &d_5 = m,\ d_4 = m(d_5+1)^2 + 1,\ d_3 = 2d_4, \\ &d_2 = \max((3n+m+1)(d_3+1), s) + 1, \\ &d_1 = d_2 + (3n+m+1)(d_3+1) + 1. \end{aligned} \tag{6}$$

For these choices, $s = 1 + n(d_4+3) + m(d_5+1)^2 = O(nm^3)$, while $(3n+m+1)(d_3+1) = \Theta((n+m)m^3)$. Hence,

$$\begin{aligned} &d_5 = O(m),\ d_4 = O(m^3),\ d_3 = O(m^3), \\ &d_2 = O((n+m)m^3),\ d_1 = O((n+m)m^3). \end{aligned}$$

Using the conditions in (5), it is easy to verify that the individual degree of x_0 and y_0 in every polynomial defined above is at least d_1 and d_2, respectively. Define f as:

$$f(\mathbf{z}) := x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{7}$$

Clearly, f is a homogeneous polynomial of degree $d_1 + d_2 + (3n+m+1)(d_3+1)$ and is divisible by $x_0^{d_1}$ and $y_0^{d_2}$. Further, the following observations hold under the constraints of (5).

Observation 3.4. For all $i \in [n], k \in [m]$, the polynomials $x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)}$, $Q_{i,1}(\mathbf{z})$, $Q_{i,2}(\mathbf{z})$, $Q_{i,3}(\mathbf{z})$ and $R_k(\mathbf{z})$ are degree separated with respect to x_0 from one another. Also, $Q_i(\mathbf{z})$ is degree separated with respect to x_0 from other $Q_j(\mathbf{z})$'s, for $i, j \in [n]$ and $i \neq j$. Similarly, $R_k(\mathbf{z})$ is degree separated with respect to x_0 from $R_l(\mathbf{z})$ for $k, l \in [m]$ and $k \neq l$.

Observation 3.5. $\mathcal{S}(f(\mathbf{z})) = 1 + n(2d_4+3) + m(d_5+1)^3$ and $\text{Supp}(f) = 8$.

The forward direction. Let $\mathbf{u} \in \{0,1\}^n$ be such that $\psi(\mathbf{u}) = 1$ and f, as described in (7), be the polynomial corresponding to ψ. Proposition 3.3 shows how $\mathbf{u}$ can be used to construct a sparsifying transform. The proof of Proposition 3.3 is very similar to that of Proposition 3.1.

Proposition 3.3. $\mathcal{S}(f(A\mathbf{z})) \leq s$ where $A \in \text{GL}(|\mathbf{z}|, \mathbb{F})$ is as follows:

$$A : y_0 \mapsto y_0,\ x_0 \mapsto x_0,\ x_i \mapsto x_i,\ y_i \mapsto y_i + (-1)^{u_i} x_i \ \ i \in [n]. \tag{8}$$

The reverse direction. Let $\mathcal{S}(f(A\mathbf{z})) \leq s$ for some $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. Lemma 3.4, the proof of which requires Claim 2.1, shows that $A(x_0)$ and $A(y_0)$ have only one variable each. With this established, Lemma 3.5 shows that the summands of $f(A\mathbf{z})$ must be degree separated with respect to x_0.

Lemma 3.4. Without loss of generality, $A(x_0) = x_0$ and $A(y_0) = y_0$.

Lemma 3.5. For all $i \in [n]$, $k \in [m]$, the polynomials $x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)}$, $Q_{i,1}(A\mathbf{z})$, $Q_{i,2}(A\mathbf{z})$, $Q_{i,3}(A\mathbf{z})$ and $R_k(A\mathbf{z})$ are degree separated from one another with respect to x_0. Also, $Q_i(A\mathbf{z})$ is degree separated with respect to x_0 from other $Q_j(A\mathbf{z})$'s, for $i, j \in [n]$ and $i \neq j$. Similarly, $R_k(A\mathbf{z})$ is degree separated with respect to x_0 from $R_l(A\mathbf{z})$ for $k, l \in [m]$ and $k \neq l$.

$$\therefore \;\; \mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})), \text{ by Lemma 3.5.}$$

Lemmas 3.6 and 3.7 are modified versions of Lemmas 3.2 and 3.3 respectively and have similar proofs as the original lemmas. Together, Lemmas 3.4, 3.6 and 3.7 show that A is a permuted scaled version of the transform of (8). Proposition 3.4 then shows how to obtain a satisfying assignment from A and can be proved similarly as Proposition 3.2.

Lemma 3.6. For any invertible A and $i \in [n]$:

$$\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z})) \geq d_4 + 3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined earlier. Equality holds if and only if under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (-1)^{u_i} X_i$$

for some scaled variables $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0, 1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z})) \neq d_4 + 3$, then $\mathcal{S}(Q_i(A\mathbf{z})) \geq 2d_4 + 3$.

Lemma 3.7. Under the given A, $\mathcal{S}(Q_i(A\mathbf{z})) = d_4 + 3$ for all $i \in [n]$.

Proposition 3.4. With A as described in (8), $\mathbf{u} = (u_1, \ldots, u_n)$ is a satisfying assignment for ψ.

Remarks. 1. In the definition of f in Section 3.1 and this section, an extra summand is present besides Q_i's and R_k's. We can drop the summand by suitably modifying f, the current parameters and arguments to make the reduction work. In particular, for an f divisible by a suitable power of x_0 (and y_0 for the homogeneous case), Lemmas 3.1 and 3.5 can be proved using Claim 2.1 (over characteristic 0 fields) or by an argument as in Section C.12 (over finite characteristic fields). We preserve the extra summand here for two reasons: One, it leads to a simpler argument and better bounds on the d_i's for the non-homogeneous case over finite characteristic fields. Two, it proves useful in showing the reduction when also considering translations (see Appendix A.1).

2. A simpler construction of f for the homogeneous case is possible with four degree parameters d_1, d_2, d_3 and d_4 under the constraints of (2). In this construction, the extra summand, Q_i's and R_k's are defined very similarly as in Section 3.1, with each polynomial multiplied by an appropriate power of y_0 such that f is homogeneous and is divisible by $x_0^{d_1}$ and $y_0^{d_1}$. The arguments presented in this section go through with some changes for the simpler construction. The reason we present the current construction is to have a single construction with which the reduction goes through for finite characteristic fields (as shown in Section 3.5.3) and characteristic 0 fields.

3. A feature of our reduction is that we can easily alter the output polynomial to $w^D f(\mathbf{z})$, where $w \notin \mathbf{z}$. This can be achieved by multiplying the output polynomial f of the current reduction by w^D, where D is greater than the sparsity parameter s in the reduction. If a proof of NP-hardness of hom-$\Sigma\Pi\Sigma$-MCSP has this feature, then it would imply NP-hardness of $\Sigma\Pi\Sigma$-MCSP (via a homogenization trick).

3.5 Extension to finite characteristic fields

In this section, we will show how the construction of Sections 3.1 and 3.4, with some changes for appropriate cases, can be used to show the NP-hardness of ETsparse over finite characteristic fields for the non-homogeneous case (in the following section) and the homogeneous case (in Section 3.5.3), respectively.

3.5.1 The non-homogeneous case

We first show how the construction of Section 3.1 also proves the reduction over fields $\mathbb{F}$ where the characteristic is greater than 2 and then give a modified construction to prove the reduction over characteristic 2 fields. Note that the degrees are chosen in Section 3.5.2 to satisfy (2) in any finite characteristic field.

So, let the characteristic be p, where $p > 2$. In this case, the polynomial f and the parameter s of Section 3.1 remain the same with the d_i's chosen as specified in Section 3.5.2. The overall argument in both directions of the reduction is highly similar to the characteristic 0 case, with the main differences being the choice of the d_i's and that in the proofs of the observations, lemmas and propositions in Sections 3.1, 3.2 and 3.3 wherever Observation 2.5 is used, then it is invoked for the finite characteristic case. Thus, Observations 3.1 and 3.2 hold without any change while Observation 3.3 holds by using Observation 2.5 for the finite characteristic case.

The forward direction. If $\mathbf{u} \in \{0,1\}^n$ is such that $\psi(\mathbf{u}) = 1$ and f, as described in (1), is the polynomial corresponding to ψ, then Proposition 3.1 shows that for the transform A of (4), defined using $\mathbf{u}$, $\mathcal{S}(f(A\mathbf{z})) \leq s$ holds.

The reverse direction. If $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ is such that $\mathcal{S}(f(A\mathbf{z})) \leq s$, then the analysis of Section 3.3 continues to hold in this case with little changes. Thus, Lemma 3.1 shows that $A(x_0) = x_0$ without loss of generality. Then, Lemmas 3.2 and 3.3 analyse the sparsity of $Q_i(A\mathbf{z})$, where $i \in [n]$ and Q_i is as defined in Section 3.1. Together, Lemmas 3.1, 3.2 and 3.3 show that A is a permuted scaled version of the transform of (4). Finally, Proposition 3.2 shows that a satisfying assignment for ψ can be extracted from A.

<u>**Construction for characteristic 2 fields**</u>

Over characteristic 2 fields, the polynomial $y_i + x_i$ is the same as $y_i - x_i$. Due to this, the definition of Q_i and that of R_k in Section 3.1 need to be changed. Moreover, the sparsifying transform will also be slightly different. Formally, let ψ, $\mathbf{x}, x_0, \mathbf{y}$ and $\mathbf{z}$ be as denoted in Section 3.1. Let $d_1, d_2, d_3, d_4 \in \mathbb{N}$. The construction of Section 3.1 is modified as follows:

- For all $i \in [n]$, define $Q_i(\mathbf{z})$ as:

$$Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z}), \text{ where}$$
$$Q_{i,1}(\mathbf{z}) := x_0^{(3i-2)d_1} x_i^{d_2},\ Q_{i,2}(\mathbf{z}) := x_0^{(3i-1)d_1}(y_i + x_i)^{d_3} \text{ and } Q_{i,3}(\mathbf{z}) := x_0^{3id_1} y_i^{d_3}.$$

- For the k^{th} clause, $k \in [m]$, define $R_k := x_0^{(3n+k)d_1} \prod_{j \in C_k} (y_j + a_{k,j} x_j)^{d_4}$.

Define $s := 1 + n(d_3 + 3) + m(d_4 + 1)^2$ as before. Set the d_i's as specified in Section 3.5.2 with the conditions of (2) imposed. Define f as:

$$f(\mathbf{z}) := x_0^{d_1} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{9}$$

Observations 3.1 and 3.2 hold with little change. Observation 3.6 analyses the sparsity and support of f.

Observation 3.6. $\mathcal{S}(f) \leq 1 + n(d_3+3) + m(d_4+1)^3$ and $4 \leq \text{Supp}(f) \leq 7$.

Remarks. Over characteristic 2 fields, the sparsity of the polynomial output by the reduction depends on the number of variables which are complemented within a clause. Hence, for the same number of variables n and the same number of clauses m, the output polynomial corresponding to two different ψ's may have different sparsity. Thus, the reduction is not natural over characteristic 2 fields.

The forward direction. Let $\mathbf{u} \in \{0,1\}^n$ be such that $\psi(\mathbf{u}) = 1$ and f, as described in (9), be the polynomial corresponding to ψ. Proposition 3.5 shows how $\mathbf{u}$ can be used to construct a sparsifying transform. The proof of Proposition 3.5 is very similar to that of Proposition 3.1.

Proposition 3.5. $\mathcal{S}(f(A\mathbf{z})) \leq s$ where $A \in \text{GL}(|\mathbf{z}|, \mathbb{F})$ is as follows:

$$A : x_0 \mapsto x_0,\ x_i \mapsto x_i,\ y_i \mapsto y_i + (1-u_i)x_i \ \ \forall i \in [n]. \tag{10}$$

The reverse direction. Let $A \in \text{GL}(|\mathbf{z}|, \mathbb{F})$ be such that $\mathcal{S}(f(A\mathbf{z})) \leq s$. The analysis of Section 3.3 holds with some changes. Formally, Lemma 3.1 holds without any change in its proof. Thus, $A(x_0) = x_0$ without loss of generality. Lemma 3.8 analyses $\mathcal{S}(Q_i(A\mathbf{z}))$, $i \in [n]$, and its proof is similar to that of Lemma 3.2.

Lemma 3.8. For any invertible A and $i \in [n]$:

$$\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z})) \geq d_3 + 3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined in this subsection. Equality holds if and only if under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (1-u_i)X_i$$

for some scaled $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0,1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z})) \neq d_3 + 3$, then $\mathcal{S}(Q_i(A\mathbf{z})) \geq 2d_3 + 3$.

Lemma 3.3 also holds with the same proof as before. Lemmas 3.1, 3.8 and 3.3 together show that A is a permuted scaled version of the transform described in (10). Proposition 3.6 then holds and can be proven similarly to Proposition 3.2.

Proposition 3.6. With A as described in (10), $\mathbf{u} = (u_1, \ldots, u_n)$ is a satisfying assignment for ψ.

3.5.2 Setting of parameters in the non-homogeneous case

Let the characteristic be $p > 0$. If $p > d_1$, where the value of d_1 is as set in (3) for characteristic 0 fields, then the d_i's are chosen to be the same as in (3). Otherwise, p must be $O(nm^3)$. When $p = O(nm^3)$, we choose d_1, d_2, d_3 and d_4, to be of form $p^j - 1$, $j \in \mathbb{N}$, while satisfying the inequalities of (2) along with $d_1 > d_2 > d_3 > d_4$. This is done so that Observation 2.5 can be used for characteristic p fields with $p = O(nm^3)$. It is possible to choose d_i's in this way because, for any $k \in \mathbb{N}$, there is exactly one number of form $p^j - 1$, $j \in \mathbb{N}$, in $[k, pk]$. The bounds on d_1, d_2, d_3 and d_4 are as follows:

$$\begin{aligned} &d_4 \leq pm,\ d_3 \leq pm(d_4+1)^2 + p = O(p^3m^3) \implies s = O(nm^3p^3), \\ &d_2 = pd_3 + (p-1) = O(p^4m^3),\ d_1 = \max(r, pd_2 + p - 1) \end{aligned}$$

where $r \in [s, ps]$ is of form $p^j - 1$, $j \in \mathbb{N}$. Thus, $r \leq ps = O(nm^3p^4)$, while $pd_2 + p - 1 = O(p^5m^3)$. As $p = O(nm^3)$,

$$d_1 = O(n^5m^{18}),\ d_2 = O(n^4m^{15}),\ d_3 = O(n^3m^{12}),\ d_4 = O(nm^4) \text{ and } s = O(n^4m^{12}).$$

3.5.3 The homogeneous case

Like in the non-homogeneous case, we first show how the construction of Section 3.4 can be used to prove the reduction over fields where the characteristic is greater than 2 and then give a modification of this construction to prove the reduction over characteristic 2 fields. Note that the degrees are chosen in Section 3.5.4 to satisfy (5) in any finite characteristic field.

So, let the characteristic be p, where $p > 2$. We consider the polynomial f and parameter s as defined in Section 3.4. Note that the degree of f is $d_1 + d_2 + (3n + m + 1)(d_3 + 1)$. We choose d_i's in Section 3.5.4 such that d_3, d_4 and d_5 are of form $p^k - 1$ for some $k \in \mathbb{N}$, while d_1 and d_2 are of form $p^l(p^t - 1)$, for some $t, l \in \mathbb{N}$. For this choice of the d_i's, Observations 3.4 and 3.5 continue to hold for f. While the forward direction is proved similarly to the characteristic 0 case, the reverse direction requires some change. More precisely, Lemma 3.4, which was proven earlier using Claim 2.1, requires a different proof. This is because Claim 2.1 holds for fields with characteristic 0 or p, with p being "large enough". If $p > d_1 + d_2 + (3n + m + 1)(d_3 + 1)$ for d_i's as chosen in (6), then Claim 2.1 holds and so does Lemma 3.4 along with the rest of the argument in the reverse direction of Section 3.4. Thus, we consider the case when $p \leq d_1 + d_2 + (3n + m + 1)(d_3 + 1) = O((n + m)m^3)$ and prove Lemma 3.4 by a different argument. Then, the rest of the argument in the reverse direction of Section 3.4 continues to hold in the same way as before.

The forward direction. Let $\mathbf{u} \in \{0, 1\}^n$ be such that $\psi(\mathbf{u}) = 1$ and f, as described in Section 3.4, be the polynomial corresponding to ψ. Proposition 3.3, with the same proof as before, shows how $\mathbf{u}$ can be used to construct a sparsifying transform.

The reverse direction. Let $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s$. We prove Lemma 3.4 (refer Section C.12 for its proof), which shows that $A(x_0)$ and $A(y_0)$ have only one variable each, by showing that for an appropriate choice of d_1 and d_2 (see Section 3.5.4), and the characteristic being finite, the following holds

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1}))\mathcal{S}(A(y_0^{d_2}))\mathcal{S}(g(A\mathbf{z})).$$

Here $g(\mathbf{z})$ is a polynomial of degree $(3n + m + 1)(d_3 + 1)$. Using Observation 2.4 and the choice of d_i's, $A(x_0)$ and $A(y_0)$ are shown to be single variables. With Lemma 3.4 proven, Lemma 3.5, with the same proof as before, shows that all the summands in $f(A\mathbf{z})$ are degree separated from one another with respect to x_0. From Lemmas 3.4 and 3.5, it follows that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(x_0^{d_1}y_0^{d_2+(3n+m+1)(d_3+1)}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

Then, Lemmas 3.6 and 3.7 hold just as in the characteristic 0 case because d_3, d_4 and d_5 are of form $p^k - 1$. Together, Lemmas 3.4, 3.6 and 3.7 show that A is a permuted scaled version of the transform of (8). Proposition 3.4 shows how to derive a satisfying assignment for ψ from A.

Construction for characteristic 2 fields

Similar to the non-homogeneous case, we modify the construction of Section 3.4 to make the reduction work over characteristic 2 fields. Consider the following polynomials:

- For all $i \in [n]$, define $Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z})$ as:

$$\begin{aligned} Q_{i,1}(\mathbf{z}) &:= x_0^{d_1+(3i-2)(d_3+1)} y_0^{d_2+(3n+m-3i+3)(d_3+1)-d_3} x_i^{d_3}, \\ Q_{i,2}(\mathbf{z}) &:= x_0^{d_1+(3i-1)(d_3+1)} y_0^{d_2+(3n+m-3i+2)(d_3+1)-d_4} (y_i + x_i)^{d_4}, \\ Q_{i,3}(\mathbf{z}) &:= x_0^{d_1+3i(d_3+1)} y_0^{d_2+(3n+m-3i+1)(d_3+1)-d_4} y_i^{d_4}. \end{aligned}$$

- For the k^{th} clause, $k \in [m]$, define

$$R_k(\mathbf{z}) := x_0^{d_1+(3n+k)(d_3+1)} y_0^{d_2+(m-k+1)(d_3+1)-3d_5} \prod_{j \in C_k} (y_j + a_{k,j} x_j)^{d_5}.$$

Define $s := 1 + n(d_4 + 3) + m(d_5 + 1)^2$ as before. Set the d_i's as specified in Section 3.5.4 to satisfy the conditions in (5). Define f as:

$$f(\mathbf{z}) := x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{11}$$

Observation 3.4 holds with little change. Observation 3.7 analyses $\mathcal{S}(f)$ and $\operatorname{Supp}(f)$.

Observation 3.7. $\mathcal{S}(f) \le 1 + n(d_4 + 3) + m(d_5 + 1)^3$ and $5 \le \operatorname{Supp}(f) \le 8$.

Remarks. Like the non-homogeneous case, over characteristic 2 fields, the sparsity of the homogeneous polynomial output by the reduction depends on the number of variables which are complemented within a clause. Hence, for the same number of variables n and same number of clauses m, the output polynomial corresponding to two different ψ's may have different sparsity. Hence, the reduction is not natural over characteristic 2 fields.

The forward direction. Let $\mathbf{u} \in \{0,1\}^n$ be such that $\psi(\mathbf{u}) = 1$ and f, as described in (11), be the polynomial corresponding to ψ. Proposition 3.7 shows how $\mathbf{u}$ can be used to construct a sparsifying transform. The proof of Proposition 3.7 is similar to that of Proposition 3.3.

Proposition 3.7. $\mathcal{S}(f(A\mathbf{z})) \le s$ where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ is as follows:

$$A : x_0 \mapsto x_0,\ y_0 \mapsto y_0,\ x_i \mapsto x_i,\ y_i \mapsto y_i + (1 - u_i)x_i\ \ \forall i \in [n]. \tag{12}$$

The reverse direction. Let $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ be such that $\mathcal{S}(f(A\mathbf{z})) \le s$. Lemmas 3.4 and 3.5 hold with little change in the arguments presented in the finite characteristic case. Thus, $A(x_0) = x_0$ and $A(y_0) = y_0$ without loss of generality. Lemma 3.9 analyses $\mathcal{S}(Q_i(A\mathbf{z}))$, $i \in [n]$, and its proof is similar to that of Lemma 3.6.

Lemma 3.9. For any invertible A and $i \in [n]$:

$$\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z})) \ge d_4 + 3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined in this subsection. Equality holds if and only if under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (1 - u_i)X_i$$

for some scaled $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0,1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z})) \ne d_4 + 3$, then $\mathcal{S}(Q_i(A\mathbf{z})) \ge 2d_4 + 3$.

Lemma 3.7 also holds with the same proof as before. Lemmas 3.4, 3.9 and 3.7 together show that A is a permuted scaled version of the transform described in (12). Proposition 3.8 then holds and can be proven similarly to Proposition 3.4.

Proposition 3.8. With A as described in (12), $\mathbf{u} = (u_1, \ldots, u_n)$ is a satisfying assignment for ψ.

3.5.4 Setting of parameters in the homogeneous case

Let the characteristic be $p > 0$. If $p > d_1 + d_2 + (3n+m+1)(d_3+1)$, where d_1, d_2 and d_3 are as chosen in (6) for the characteristic 0 case, then the same setting of d_i's holds. Otherwise, $p = O((n+m)m^3)$. Similar to the non-homogeneous case, we choose d_3, d_4 and d_5 so that $d_3 > d_4 > d_5$ and they are of form $p^k - 1$, $k \in \mathbb{N}$. We have the following bounds:

$$\begin{aligned} d_5 &\le pm \implies d_5 = O(pm) = O((n+m)m^4), \\ d_4 &\le pm(d_5+1)^2 + p \implies d_4 = O(p^3m^3) = O((n+m)^3m^{12}), \\ \therefore s &= O(p^3nm^3) = O((n+m)^3nm^{12}), \\ d_3 &= pd_4 + p - 1 \implies d_3 = O(p^4m^3) = O((n+m)^4m^{15}). \end{aligned} \tag{13}$$

This choice of d_3, d_4 and d_5 ensures they are $(mn)^{O(1)}$, satisfy (5) and that Observation 2.5 can be used for characteristic p fields with $p = O((n+m)m^3)$. Now, let $k_1 := \lfloor \log_p(s) \rfloor + 1, k_2 := \lfloor \log_p((3n+m+1)(d_3+1)) \rfloor + 1$, then set

$$d_2 = \sum_{i=k_2}^{k_1+k_2-1} (p-1)p^i = p^{k_1+k_2} - p^{k_2}. \tag{14}$$

For this choice, $d_2 > s$ and $d_2 > (3n+m+1)(d_3+1)$. Lastly, let $k_3 := \lfloor \log_p(d_2 + (3n+m+1)(d_3+1)) \rfloor + 1$, then set

$$d_1 = \sum_{i=k_3}^{k_3+k_1-1} (p-1)p^i = p^{k_1+k_3} - p^{k_3} > d_2 + (3n+m+1)(d_3+1). \tag{15}$$

For this choice of d_1 and d_2 it holds that,

$$d_2 = O(p^2s(3n+m+1)(d_3+1)) = O(p^9nm^6(n+m)) = O((n+m)^{10}nm^{33})$$

$$d_1 = O(p^2s(d_2 + (3n+m+1)(d_3+1))) = O(p^{14}n^2m^9(n+m)) = O((n+m)^{15}n^2m^{51}).$$

4 NP-hardness of α-gap-ETsparse

In this section, we prove Theorem 2. We first prove parts 1 and 2 of Theorem 2 over characteristic 0 fields without considering translations in Sections 4.1 and 4.2, respectively. Section 4.3 extends both parts to finite characteristic fields. Appendix D contains the proofs of the lemmas in this section. Appendix A.2 proves part 1 of Theorem 2 while considering translations.[18]

[18] For part 2 of Theorem 2, translations are not considered; see footnote 16 for an explanation.

Proof sketch. For a 3-CNF ψ, we carefully analyze the sparsity of the corresponding polynomial f as defined in Section 3, for the non-homogeneous and the homogeneous case. For unsatisfiable ψ's, we do a slightly deeper analysis on $\mathcal{S}(f(A\mathbf{z}))$, for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$, to show a lower bound. For satisfiable ψ's, $\mathcal{S}(f(A\mathbf{z}))$ has already been upper bounded for an appropriate $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. The degree parameters are also chosen differently so that the gap between the lower bound and the upper bound is significant. Comparing the sparsities for satisfiable and unsatisfiable ψ's proves Theorem 2. Note that throughout this section, we assume $\epsilon \in (0, 1/3)$ to be an arbitrary constant.

4.1 Analyzing the gap: the non-homogeneous case

For a 3-CNF ψ, consider the polynomial f as defined in (1). Choose the d_i's to satisfy the following while also being $(mn)^{O(1)}$:

$$d_4 \geq \max(4mn, (mn)^{O(1/\epsilon)}),\ d_3 = m(d_4+1)^2 + 1,\ d_2 = d_3^2 + 1,\ d_1 = d_2 + 1. \tag{16}$$

Note under these constraints, the conditions in (2) are also satisfied. Let $s := \mathcal{S}(f)$. Then, $s = 1 + n(2d_3 + 3) + m(d_4 + 1)^3$ by Observation 3.3. For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.1 shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$ for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. In the lemma, Item 1 is essentially Lemma 3.1, Items 2 and 3 are a slightly deeper analysis of that in Lemmas 3.2 and 3.3, and Item 4 is the analysis in Proposition 3.2. Thus, Lemma 4.1 encapsulates the analysis of Section 3.3. For $\psi \in$ 3-SAT, by Proposition 3.1, there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2$. Proposition 4.1 shows α-gap-ETsparse is NP-hard by comparing the sparsity for satisfiable and unsatisfiable ψ's, and uses Lemma 4.1 and the conditions in (16).

Lemma 4.1. Let $\psi \in \overline{\text{3-SAT}}$, f be as defined in (1) corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_1 + 1$.
2. If A is not as in item 1 and $A(x_j)$ is a linear form in at least 2 variables for some $j \in [n]$, then $\mathcal{S}(f(A\mathbf{z})) \geq d_2 + 1$.
3. If A is not as in items 1 and 2 and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j - x_j)$ is a linear form in at least 3 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq \frac{d_3^2 + 3d_3 + 2}{2}$.
4. If A is not of the form described in the previous three items, then $\mathcal{S}(f(A\mathbf{z})) \geq (d_4 + 1)^3$.

Proposition 4.1. Let $\mathrm{char}(\mathbb{F}) = 0$. If the input in α-gap-ETsparse is an s-sparse polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4). If $\psi \in \overline{\text{3-SAT}}$, then it follows from Lemma 4.1 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \geq \min\left(d_1 + 1, d_2 + 1, \frac{d_3^2 + 3d_3 + 2}{2}, (d_4 + 1)^3\right).$$

The constraints imposed in (16) ensure that $(d_4 + 1)^3$ is the minimum. As $d_3 = m(d_4 + 1)^2 + 1$, therefore $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2 \leq 3nd_3 = 3mn(d_4 + 1)^2 + 3n \leq 4mn(d_4 + 1)^2$. Thus, the gap in the sparsities of the YES instances and the NO instances is

$$\frac{(d_4 + 1)^3}{s_0} \geq \frac{(d_4 + 1)^3}{4mn(d_4 + 1)^2} = \frac{d_4 + 1}{4mn}.$$

Also, as $d_4 \geq 4mn$, $\mathcal{S}(f) = s \leq 2m(d_4+1)^3 \implies d_4 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{d_4+1}{4mn} \geq \frac{s^{1/3}}{2^{1/3}4m^{4/3}n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_4^{3\epsilon} \geq 2^{1/3}4m^{4/3}n \implies \frac{s^{1/3}}{2^{1/3}4m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

4.2 Analyzing the gap: the homogeneous case

Consider the polynomial f as defined in (7) for ψ. Choose the d_i's to satisfy the following constraints while also being $(mn)^{O(1)}$.

$$\begin{aligned} &d_5 \geq \max((4mn, (mn)^{O(1/\epsilon)}),\ d_4 = m(d_5+1)^2+1,\ d_3 = d_4^2+1, \\ &d_2 = \max((3n+m+1)(d_3+1), s) + 1, \\ &d_1 = d_2 + (3n+m+1)(d_3+1)+1. \end{aligned} \tag{17}$$

Under these constraints, the conditions in (5) are also satisfied. Let $s := \mathcal{S}(f)$. Then, $s = 1 + n(2d_4+3) + m(d_5+1)^3$ by Observation 3.5. For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.2, proved using Claim 2.1, shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$, for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. In the lemma, Items 1 and 2 are essentially Lemma 3.4, Items 3 and 4 are a deeper analysis of that in Lemmas 3.6 and 3.7, and Item 5 is the analysis in Proposition 3.4. For $\psi \in$ 3-SAT, by Proposition 3.3, there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_4+3) + m(d_5+1)^2$. Proposition 4.2 proves α-gap-ETsparse is NP-hard using Lemma 4.2 and the conditions in (17).

Lemma 4.2. Let $\psi \in \overline{\text{3-SAT}}$, $f(\mathbf{z})$ be the polynomial as defined in (7) corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_1 + 1$.
2. If A is not as in item 1 and $A(y_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_2 + 1$.
3. If A is not as in items 1 and 2, and $A(x_j)$ is a linear form in at least 2 variables for some $j \in [n]$, then $\mathcal{S}(f(A\mathbf{z})) \geq d_3 + 1$.
4. If A is not as in items 1, 2 and 3, and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j - x_j)$ is a linear form in at least 3 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq \frac{d_4^2+3d_4+2}{2}$.
5. If A is not of the form described in the previous four items, then $\mathcal{S}(f(A\mathbf{z})) \geq (d_5+1)^3$.

Proposition 4.2. Let $\mathrm{char}(\mathbb{F}) = 0$. If the input in α-gap-ETsparse is an s-sparse homogeneous polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (8). If $\psi \in \overline{\text{3-SAT}}$, then it follows from Lemma 4.2 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \geq \min\left(d_1+1, d_2+1, d_3+1, \frac{d_4^2+3d_4+2}{2}, (d_5+1)^3\right).$$

The constraints imposed in (17) ensure that $(d_5+1)^3$ is the minimum. As $d_4 = m(d_5+1)^2+1$, therefore $s_0 = 1+n(d_4+3)+m(d_5+1)^2 \leq 3nd_5 = 3mn(d_5+1)^2+3n \leq 4mn(d_5+1)^2$. Thus, the gap in the sparsities of the YES instances and the NO instances is

$$\frac{(d_5+1)^3}{s_0} \geq \frac{(d_5+1)^3}{4mn(d_5+1)^2} = \frac{d_5+1}{4mn}.$$

Also, as $d_5 \geq 4mn$, $\mathcal{S}(f) = s \leq 2m(d_5+1)^3 \implies d_5+1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_5+1)^3}{s_0} \geq \frac{d_5+1}{4mn} \geq \frac{s^{1/3}}{2^{1/3}4m^{4/3}n}.$$

Finally, note that $s \geq d_5^3$. Thus, for $d_5^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_5^{3\epsilon} \geq 2^{1/3}4m^{4/3}n \implies \frac{s^{1/3}}{2^{1/3}4m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

Remarks. As mentioned in the second remark near the end of Section 3.4, a simpler construction of a homogeneous f with four degree parameters is possible. This simpler construction can be used to show the NP-hardness of α-gap-ETsparse. The reason we use the construction currently defined in Section 3.4 is that it allows us to prove part 2 of Theorem 2 for homogeneous polynomials over finite characteristic fields and characteristic 0 fields without using separate constructions.

4.3 Extension to finite characteristic fields

In this section, we will show how the construction of Sections 3.1 and 3.4, with some changes for appropriate cases, can be used to show the NP-hardness of α-gap-ETsparse over finite characteristic fields for the non-homogeneous case (in the following section) and homogeneous case (in Section 4.3.2), respectively.

4.3.1 The non-homogeneous case

Let the characteristic be p, where $p > 2$. If $p > d_1$, where d_1 is as chosen in Section 4.1, then the argument of that section holds. Hence, it is assumed that $p \leq d_1 = (mn)^{O(1)}$. Consider the polynomial f defined in (1). Let $\epsilon \in (0, 1/3)$ be an arbitrary constant. Choose the d_i's to be of form $p^k - 1$ for some $k \in \mathbb{N}$ to satisfy the following inequalities, along with those of (2):

$$d_4 \geq \max(3pmn, (mn)^{O(1/\epsilon)}),\ d_3 > m(d_4+1)^2,\ d_2 > (d_3+1)^2,\ d_1 > d_2. \tag{18}$$

Note that we can get $d_3 = O(pm(d_4+1)^2)$, $d_2 = O(p(d_3+1)^2)$ and $d_1 = O(pd_2)$. Let $s := \mathcal{S}(f)$. Then, $s = 1+n(2d_3+3)+m(d_4+1)^3$ by Observation 3.3. For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.3 shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$ for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. Like Lemma 4.1, Lemma 4.3 is a slightly deeper analysis of that in the reverse direction of Section 3.5.1. For $\psi \in$ 3-SAT, by Proposition 3.1 there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1+n(d_3+3)+m(d_4+1)^2$. Proposition 4.3 shows the NP-hardness of α-gap-ETsparse using Lemma 4.3 and the inequalities in (18).

Lemma 4.3. Let $\psi \in \overline{\text{3-SAT}}$, f, as defined in (1), be the polynomial corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_1 + 1$.
2. If A is not as in item 1 and for some $j \in [n]$, $A(x_j)$ is a linear form in at least 2 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq d_2 + 1$.
3. If A is not as in item 1 and 2 and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j - x_j)$ is a linear form in at least 3 variables, $\mathcal{S}(f(A\mathbf{z})) \geq (d_3 + 1)^{1.63}$.
4. If A is not of the form described in the previous three cases, then $\mathcal{S}(f(A\mathbf{z})) \geq (d_4 + 1)^3$.

Proposition 4.3. Let $\mathrm{char}(\mathbb{F}) = p > 2$. If the input in α-gap-ETsparse is an s-sparse polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. The proof is similar to that of Proposition 4.1. If ψ is satisfiable, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4) and $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2$. For unsatisfiable ψ, it follows from Lemma 4.3 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \geq \min(d_1 + 1, d_2 + 1, (d_3 + 1)^{1.63}, (d_4 + 1)^3).$$

As $d_3 > m(d_4 + 1)^2$, therefore $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2 \leq 3nd_3 \leq 3pmn(d_4 + 1)^2$. The conditions imposed in (18) ensure that $d_1 + 1 > d_2 + 1 > (d_3 + 1)^{1+\log_p((p+1)/2)} > (d_4 + 1)^3 > s_0$. Thus, the gap in the sparsities of the YES instances and NO instances is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{(d_4+1)^3}{3pmn(d_4+1)^2} = \frac{d_4+1}{3pmn}.$$

Also, note as $d_4 \geq 3pmn$, therefore $s \leq 2m(d_4 + 1)^3 \implies d_4 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{d_4+1}{3pmn} \geq \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_4^{3\epsilon} \geq p2^{1/3}3m^{4/3}n \implies \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

Analysis over characteristic 2 fields

In this case, we consider the construction of Section 3.5.1. For a 3-CNF ψ, let f be the corresponding polynomial as defined in (9). Let $s := \mathcal{S}(f)$. Over characteristic 2 fields, the value of s depends on the number of variables complemented in a clause. To show the hardness of α-gap-ETsparse, we require $s \geq d_4^3$ (see the proof of Proposition 4.4). This can be achieved if there exists a clause with all variables complemented. Hence, we assume, without loss of generality, that there is such a clause in ψ.[19] The d_i's are chosen in the same way as in Section 4.3.1 to satisfy (18) with p set to 2. In particular, they satisfy the following inequalities.

$$d_4 \geq \max(6mn, (mn)^{O(1/\epsilon)}),\ d_3 > m(d_4+1)^2,\ d_2 > (d_3+1)^2,\ d_1 > d_2. \tag{19}$$

[19] To have some clause, say the first one, contain only complemented variables, every uncomplemented variable x in the clause can be replaced by $\neg x$ followed by complementing each occurrence of x in the remaining clauses.

By Observation 3.6 and the assumption on ψ, it holds that

$$1 + n(d_3 + 3) + (d_4 + 1)^3 \leq s \leq 1 + n(d_3 + 3) + m(d_4 + 1)^3.$$

For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.4 shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$ where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. For $\psi \in$ 3-SAT, by Proposition 3.5 there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2$. Proposition 4.4 shows the NP-hardness of α-gap-ETsparse using Lemma 4.4 and the inequalities in (19).

Lemma 4.4. Let $\psi \in \overline{\text{3-SAT}}$, f, as defined in (9), be the polynomial corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_1 + 1$.
2. If A is not as in item 1 and for some $j \in [n]$, $A(x_j)$ is a linear form in at least 2 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq d_2 + 1$.
3. If A is not as in item 1 and 2 and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j)$ is a linear form in at least 3 variables, $\mathcal{S}(f(A\mathbf{z})) \geq (d_3 + 1)^{1.58}$.
4. If A is not of the form described in the previous three cases, then $\mathcal{S}(f(A\mathbf{z})) \geq (d_4 + 1)^3$.

Proposition 4.4. Let $\mathrm{char}(\mathbb{F}) = 2$. If the input in α-gap-ETsparse is an s-sparse polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. The proof is similar to that of Proposition 4.3. If ψ is satisfiable, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4) and $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2$. For unsatisfiable ψ, it follows from Lemma 4.3 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \geq \min(d_1 + 1, d_2 + 1, (d_3 + 1)^{1.58}, (d_4 + 1)^3).$$

As $d_3 > m(d_4 + 1)^2$, therefore $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2 \leq 3nd_3 \leq 6mn(d_4 + 1)^2$. The conditions imposed in (19) ensure that $d_1 + 1 > d_2 + 1 > (d_3 + 1)^{1.58} > (d_4 + 1)^3 > s_0$. Consequently, the gap in the sparsities of the YES instances and NO instances is

$$\frac{(d_4 + 1)^3}{s_0} \geq \frac{(d_4 + 1)^3}{6mn(d_4 + 1)^2} = \frac{d_4 + 1}{6mn}.$$

Also, note as $d_4 \geq 6mn$, therefore $s \leq 2m(d_4 + 1)^3 \implies d_4 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4 + 1)^3}{s_0} \geq \frac{d_4 + 1}{6mn} \geq \frac{s^{1/3}}{2^{4/3} 3 m^{4/3} n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_4^{3\epsilon} \geq 2^{4/3} 3 m^{4/3} n \implies \frac{s^{1/3}}{2^{4/3} 3 m^{4/3} n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

4.3.2 The homogeneous case

Let the characteristic be p, where $p > 2$. If $p > d_1$, where d_1 is as chosen in Section 4.2, then the argument of that section holds. Hence, assume $p \le d_1 = O((mn)^{O(1)})$. Consider the polynomial f defined in (7). Choose d_3, d_4 and d_5 to be of form $p^k - 1$ for some $k \in \mathbb{N}$ while also satisfying:

$$d_5 \ge \max(3pmn, (mn)^{O(1/\epsilon)}),\ d_4 > m(d_5+1)^2,\ d_3 > (d_4+1)^2. \tag{20}$$

Thus, $d_4 = O(pm(d_5+1)^2)$ and $d_3 = O(p(d_4+1)^2)$. Now, let $k_1 := \lfloor \log_p(d_3+2) \rfloor + 1, k_2 := \lfloor \log_p((3n+m+1)(d_3+1)) \rfloor + 1$, then set

$$d_2 := \sum_{i=k_2}^{k_1+k_2-1} (p-1)p^i = p^{k_1+k_2} - p^{k_2}. \tag{21}$$

Lastly, let $k_3 := \lfloor \log_p(d_2 + (3n+m+1)(d_3+1)) \rfloor + 1$, then set

$$d_1 := \sum_{i=k_3}^{k_3+k_1-1} (p-1)p^i = p^{k_1+k_3} - p^{k_3}. \tag{22}$$

For this choice of the d_i's, the conditions in (5) are satisfied. Let $s := \mathcal{S}(f)$. By Observation 3.5, $s = 1 + n(2d_4+3) + m(d_5+1)^3$. For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.5 shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$ for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. Like Lemma 4.2, Lemma 4.5 is a slightly deeper analysis of that in the reverse direction of Section 3.5.3 For $\psi \in$ 3-SAT, by Proposition 3.3 there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \le s_0$, where $s_0 = 1 + n(d_4+3) + m(d_5+1)^2$. Proposition 4.5 shows the NP-hardness of α-gap-ETsparse using Lemma 4.5 and the setting of the d_i's in this section.

Lemma 4.5. Let $\psi \in \overline{\text{3-SAT}}$, f, as defined in (7), be the polynomial corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ or $A(y_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \ge d_3 + 2$.
2. If A is not as in item 1 and $A(x_j)$ is a linear form in at least 2 variables for some $j \in [n]$, then $\mathcal{S}(f(A\mathbf{z})) \ge d_3 + 1$.
3. If A is not as in items 1 and 2 and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j - x_j)$ is a linear form in at least 3 variables, $\mathcal{S}(f(A\mathbf{z})) \ge (d_4+1)^{1.63}$.
4. If A is not of the form described in the previous three cases, then $\mathcal{S}(f(A\mathbf{z})) \ge (d_5+1)^3$.

Proposition 4.5. Let $\mathrm{char}(\mathbb{F}) = p > 2$. If the input in α-gap-ETsparse is an s-sparse homogeneous polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. The proof is similar to that of Proposition 4.2. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \le s_0$ where A is as described in (8). If $\psi \in \overline{\text{3-SAT}}$, then it follows from Lemma 4.3 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \ge \min\left(d_3+2, d_3+1, (d_4+1)^{1.63}, (d_5+1)^3\right).$$

The constraints imposed in (16) ensure that $(d_5+1)^3$ is the minimum. As $d_4 > m(d_5+1)^2$, therefore $s_0 = 1 + n(d_4+3) + m(d_5+1)^2 < 3nd_4 \le 3pmn(d_5+1)^2$. Thus, the gap in the sparsities of the YES instances and the NO instances is

$$\frac{(d_5+1)^3}{s_0} \ge \frac{(d_5+1)^3}{3pmn(d_5+1)^2} = \frac{d_5+1}{3pmn}.$$

Also, as $d_5 \geq 3pmn$, $\mathcal{S}(f) = s \leq 2m(d_5+1)^3 \implies d_5 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_5+1)^3}{s_0} \geq \frac{d_5+1}{3pmn} \geq \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n}.$$

Finally, note that $s \geq d_5^3$. Thus, for $d_5^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_5^{3\epsilon} \geq p2^{1/3}3m^{4/3}n \implies \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

Analysis over characteristic 2 fields

For characteristic 2 fields, consider the polynomial f as defined in (11). Let $s := \mathcal{S}(f)$. Like in the characteristic 2 construction for the non-homogeneous case, the value of s depends on the number of variables complemented within a clause. To show the NP-hardness of α-gap-ETsparse, $s \geq d_5^3$ is required (see the proof of Proposition 4.6), which can be achieved if there is at least one clause where all variables are complemented. Therefore, we assume that such a clause exists (see footnote 19). Set the d_i's as specified in the beginning of Section 4.3.2 with $p = 2$. For this setting of the d_i's, the constraints in (5) are satisfied. By Observation 3.7 and the assumption on ψ, it holds that

$$1 + n(d_4+3) + (d_5+1)^3 \leq s \leq 1 + n(d_4+3) + m(d_5+1)^3.$$

For $\psi \in \overline{\text{3-SAT}}$, Lemma 4.6 shows lower bounds on $\mathcal{S}(f(A\mathbf{z}))$ for all $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$. For $\psi \in$ 3-SAT, by Proposition 3.7 there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_4+3) + m(d_5+1)^2$. Proposition 4.6 shows the NP-hardness of α-gap-ETsparse using Lemma 4.6 and the setting of the d_i's in this section.

Lemma 4.6. Let $\psi \in \overline{\text{3-SAT}}$, f, as defined in (11), be the polynomial corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0)$ or $A(y_0)$ is a linear form in at least 2 variables, $\mathcal{S}(f(A\mathbf{z})) \geq d_3 + 2$.
2. If A is not as in item 1 and $A(x_j)$ is a linear form in at least 2 variables for some $j \in [n]$, then $\mathcal{S}(f(A\mathbf{z})) \geq d_3 + 1$.
3. If A is not as in items 1 and 2 and for some $j \in [n]$, $A(y_j + x_j)$ or $A(y_j - x_j)$ is a linear form in at least 3 variables, $\mathcal{S}(f(A\mathbf{z})) \geq (d_4+1)^{1.58}$.
4. If A is not of the form described in the previous three cases, then $\mathcal{S}(f(A\mathbf{z})) \geq (d_5+1)^3$.

Proposition 4.6. Let $\mathrm{char}(\mathbb{F}) = 2$. If the input in α-gap-ETsparse is an s-sparse homogeneous polynomial, then α-gap-ETsparse is NP-hard for $\alpha = s^{1/3-\epsilon}$.

Proof. The proof is similar to that of Proposition 4.5. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (12). If $\psi \in \overline{\text{3-SAT}}$, then it follows from Lemma 4.3 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$:

$$\mathcal{S}(f(A\mathbf{z})) \geq \min\left(d_3+2, d_3+1, (d_4+1)^{1.58}, (d_5+1)^3\right).$$

The constraints imposed in (16) ensure that $(d_5+1)^3$ is the minimum. As $d_4 > m(d_5+1)^2$, therefore $s_0 = 1 + n(d_4+3) + m(d_5+1)^2 < 3nd_4 \leq 6mn(d_5+1)^2$. Thus, the gap in the sparsities of the YES instances and the NO instances is

$$\frac{(d_5+1)^3}{s_0} \geq \frac{(d_5+1)^3}{6mn(d_5+1)^2} = \frac{d_5+1}{6mn}.$$

Also, as $d_5 \geq 6mn$, $s \leq 2m(d_5+1)^3 \implies d_5 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_5+1)^3}{s_0} \geq \frac{d_5+1}{6mn} \geq \frac{s^{1/3}}{2^{4/3}3m^{4/3}n}.$$

Finally, note that $s \geq d_5^3$. Thus, for $d_5^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_5^{3\epsilon} \geq 2^{4/3}3m^{4/3}n \implies \frac{s^{1/3}}{2^{4/3}3m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

5 NP-hardness of ETsupport

In this section, we prove Theorem 3. All lemmas and observations are proved in Appendix E.

Proof sketch. We map ψ, a 3-CNF, to a polynomial f, which is the sum of degree separated polynomials with at least one polynomial of support $\sigma+1$ and the rest of support σ (where σ is a constant). As the summands are degree separated, $\mathrm{Supp}(f) = \sigma + 1$ and for any invertible linear transform A, $\mathrm{Supp}(A(f))$ is equal to the maximum support size among the transformed summands. Claim 2.2 is used to show that $\psi \in$ 3-SAT iff there exists an invertible linear transform A, such that $\mathrm{Supp}(A(f)) \leq \sigma$. Thus, the reduction also holds for $(\sigma+1)$-to-σ ETsupport. For characteristic p fields, we assume $p > \sigma + 1$ so that Claim 2.2 holds.

5.1 Construction of f

Let $\sigma \geq 6$ be an even integer constant and ψ be as denoted in Section 3.1. For odd σ, we describe the changes in the construction/argument at the appropriate points. As σ is a constant, let $n \geq \sigma + 4$. The proofs of Lemmas 5.1 and 5.2 use $n \geq \sigma + 4$. To ensure $\mathrm{Supp}(f) = \sigma + 1$, we assume that all the variables in the first clause are complemented (see footnote 19). Let $\mathbf{x} := \{x_1, \ldots, x_n\}$, $\mathbf{y} := \{y_1, \ldots, y_n\}$ and $\mathbf{z} := \{z_1, \ldots, z_{\sigma-5}\}$ and $\mathbf{w} := \mathbf{x} \sqcup \mathbf{y} \sqcup \mathbf{z}$. By $(w_1 \cdots w_l)^{\star}$, where $w_i \in \mathbf{w}$ and $l, \star \in \mathbb{N}$, we denote a power of $w_1 \cdots w_l$. Consider the polynomials:

- First, introduce $\binom{n+\sigma-5}{\sigma}$ many monomials defined by the set

$$P := \{(w_1 \cdots w_\sigma)^{\star} \mid w_1, \ldots, w_\sigma \in \mathbf{z} \sqcup \mathbf{x} \text{ and are pairwise distinct}\}.$$

- Then, introduce $\binom{n}{\frac{\sigma}{2}}$ many monomials defined by the set

$$Q := \{((x_{i_1}y_{i_1}) \cdots (x_{i_{\frac{\sigma}{2}}}y_{i_{\frac{\sigma}{2}}}))^{\star} \mid i_1, \ldots, i_{\frac{\sigma}{2}} \in [n] \text{ and are pairwise distinct}\}.$$

  Note: For odd σ, the monomials are of form $((x_{i_1}y_{i_1}) \cdots (x_{i_{\frac{\sigma-1}{2}}}y_{i_{\frac{\sigma-1}{2}}})x_{i_{\frac{\sigma+1}{2}}})^{\star}$. Thus $|Q| = \binom{n}{\frac{\sigma+1}{2}}\frac{\sigma+1}{2}$.

- Let $R := \{R_k(\mathbf{w}) \mid k \in [m]\}$, where $R_k(\mathbf{w})$ is defined corresponding to the k^{th} clause as:

$$R_k(\mathbf{w}) := (\prod_{j\in C_k}(y_j - a_{k,j}x_j))^2(z_1 \cdots z_{\sigma-5})^{\star}.$$

Define $f(\mathbf{w}) := \sum_{g\in P} g(\mathbf{w}) + \sum_{h\in Q} h(\mathbf{w}) + \sum_{k=1}^{m} R_k(\mathbf{w})$. The powers, denoted by $\star$, must be such that the following hold:

1. All the polynomials in $P \sqcup Q \sqcup R$ are degree separated, with the powers being at least $\sigma + 1$. The assumption $\sigma \geq 6$ ensures that each R_k has a monomial in $\mathbf{z}$, due to which the R_k's are degree separated.
2. Over characteristic p fields, where $p > 0$, all the powers are less than p or are of the form $p^k - 1$ for some $k \in \mathbb{N}$.

In Section E.4, we choose the powers to satisfy the above conditions over any field. Based on these conditions, Observation 5.1 holds.

Observation 5.1. $\mathcal{S}(f(\mathbf{w})) = O(n^{\sigma} + m)$ and $\text{Supp}(f(\mathbf{w})) = \sigma + 1$.

5.2 The forward direction

Proposition 5.1 shows how a satisfying assignment for ψ implies the existence of an invertible A, such that $\text{Supp}(f(A\mathbf{w})) = \sigma$ by constructing A from the satisfying assignment.

Proposition 5.1. Let $\psi \in$ 3-SAT with $(u_1, \ldots, u_n) \in \{0,1\}^n$ a satisfying assignment. Then, $\text{Supp}(f(A\mathbf{w})) = \sigma$, where the transform A is defined as

$$A : z_j \mapsto z_j,\ x_i \mapsto x_i,\ y_i \mapsto y_i + (1 - u_i)x_i \quad i \in [n], j \in [\sigma - 5]. \tag{23}$$

Proof. As Condition 1 is satisfied, it suffices to analyse the action of A on individual polynomials. Clearly, for $g(\mathbf{w}) \in P$, $g(A\mathbf{w}) = g(\mathbf{w})$. Let $h(\mathbf{w}) \in Q$. Then $h(\mathbf{w})$ is of form $((x_{t_1}y_{t_1}) \cdots (x_{t_{\frac{\sigma}{2}}}y_{t_{\frac{\sigma}{2}}}))^{\star}$, $t_j \in [n]$, and A acts on h as:

$$A : ((x_{t_1}y_{t_1}) \cdots (x_{t_{\frac{\sigma}{2}}}y_{t_{\frac{\sigma}{2}}}))^{\star} \mapsto ((x_{t_1})(y_{t_1} + (1-u_{t_1})x_{t_1}) \cdots (x_{t_{\frac{\sigma}{2}}})(y_{t_{\frac{\sigma}{2}}} + (1-u_{t_{\frac{\sigma}{2}}})x_{t_{\frac{\sigma}{2}}}))^{\star}.$$

Note $|\cup_{t=1}^{\sigma} \text{var}(\ell_t)| = \sigma$, where $\ell_t = A(w)$ and $w \in \text{var}(h(\mathbf{w}))$. Thus, $\text{Supp}(h(A\mathbf{w})) \leq \sigma$. When σ is odd, a similar argument holds for the modified construction of Q. For $k \in [m]$, A acts on $R_k(\mathbf{w})$ as:

$$A : (\prod_{j\in C_k}(y_j - a_{k,j}x_j))^2 \cdot (z_1 \cdots z_{\sigma-5})^{\star} \mapsto (\prod_{j\in C_k}(y_j + (1 - a_{k,j} - u_j)x_j)^2 \cdot (z_1 \cdots z_{\sigma-5})^{\star}.$$

If $a_{k,j} \neq u_j$, then $a_{k,j} = 1 - u_j$. Since ψ is satisfiable, therefore for all $k \in [m]$, $a_{k,j} \neq u_j$ for some $j \in C_k$. Hence, $\text{Supp}(R_k(A\mathbf{w})) \leq (\sigma - 5) + 5 = \sigma$ for all $k \in [m]$. Thus, $\text{Supp}(f(A\mathbf{w})) = \sigma$. □

5.3 The reverse direction

Now, we show that if $\text{Supp}(f(A\mathbf{w})) \leq \sigma$ for $A \in \text{GL}(|\mathbf{w}|, \mathbb{F})$, then a satisfying assignment can be recovered for ψ. Lemmas 5.1 and 5.2, proved using Claim 2.2, together show that A is as:

$$A : z_j \mapsto z_j,\ x_i \mapsto x_i,\ y_i \mapsto y_i + c_i x_i \quad c_i \in \mathbb{F},\ j \in [\sigma - 5], i \in [n]$$

without loss of generality.[20] Proposition 5.2 derives a satisfying assignment for ψ from A.

[20] as permutation and non-zero scaling of variables do not affect the support.

Lemma 5.1. If $\mathrm{Supp}(f(A\mathbf{w})) \leq \sigma$, then $\forall w \in \mathbf{z} \sqcup \mathbf{x}$, $A(w) = W$, for scaled variable $W \in \mathbf{w}$.

Lemma 5.2. If $\mathrm{Supp}(f(A\mathbf{w})) \leq \sigma$, then $A(x_i) = X_i$ and $A(y_i) = Y_i + c_i X_i$, for scaled variables $Y_i, X_i \in \mathbf{w}$ and $c_i \in \mathbb{F}$.

Proposition 5.2. A satisfying assignment $\mathbf{u}$ for ψ can be extracted from A.

Proof. The action of A on R_k, where $k \in [m]$, is:

$$(\prod_{j \in C_k} (y_j - a_{k,j} x_j))^2 \cdot (z_1 \cdots z_{\sigma-5})^\star \mapsto (\prod_{j \in C_k} (y_j + (c_j - a_{k,j}) x_j))^2 \cdot (z_1 \cdots z_{\sigma-5})^\star.$$

Thus, $\mathrm{Supp}(R_k(A\mathbf{w})) \leq \sigma$ iff for some $j \in C_k$, $c_j = a_{k,j}$. By assumption $\mathrm{Supp}(R_k(A\mathbf{w})) \leq \sigma$ for all $k \in [m]$. Hence, for each $R_k(\mathbf{w})$, there exists $j \in C_k$ such that $c_j \in \{0,1\}$. Construct $\mathbf{u} \in \{0,1\}^n$ by setting $u_j := 1 - c_j$, for appropriate $j \in C_k$ and the remaining u_i's to arbitrary values in $\{0,1\}$. From the definition of $\mathbf{u}$, it follows that for the k^{th} clause, there exists $j \in C_k$ such that $u_j \neq a_{k,j}$. As k is arbitrary, all clauses are satisfied. □

6 Conclusion

In this work, we show that ET for sparse polynomials is NP-hard. Particularly, we show the NP-hardness of MCSP for orbits of homogeneous sparse polynomials (a dense subclass of hom-ΣΠΣ circuits) over characteristic 0 fields. We also define a gap version of ET for sparse polynomials and show it is NP-hard, which implies the NP-hardness of $s^{\frac{1}{3}-\epsilon}$-factor approximation of the sparse-orbit complexity of s-sparse polynomials. Lastly, we also show that ET for constant-support polynomials is NP-hard. In all three cases, we reduce 3-SAT to the respective problems. We end by listing some problems whose solutions we do not know:

1. **Hardness of** ETsparse **for constant degree polynomials:** In the reduction of Theorem 1, can the degree of the output polynomial be made constant? Currently, the degree is polynomial in the number of clauses and variables.

2. **Improving the gap in Theorem 2:** Can α-gap-ETsparse be shown NP-hard for $\alpha = s^{1-\epsilon}$, where s is the sparsity of the input polynomial and $\epsilon > 0$ is an arbitrary constant?

3. **Hardness of** ETsupport **for** $\sigma = 2$: Is checking if a given polynomial is in the orbit of a support-2 polynomial NP-hard? Theorem 3 shows that ETsupport for $\sigma \geq 6$ is NP-hard.

4. **Hardness of MCSP for** hom-ΣΠΣ **circuits:** Is MCSP for hom-ΣΠΣ circuits NP-hard?

Acknowledgments

We thank the anonymous reviewers for their detailed and constructive feedback, which has helped us improve the presentation of this work. In particular, we thank one of the reviewers for pointing out some inaccuracies in the original proofs of Lemmas 5.1 and 5.2; simpler proofs for both lemmas came up in the process of fixing these inaccuracies.

References

[AD17] Eric Allender and Bireswar Das. Zero knowledge and circuit minimization. *Inf. Comput.*, 256:2–8, 2017. Conference version appeared in the proceedings of MFCS 2014. 2

[AGK+23] Prashanth Amireddy, Ankit Garg, Neeraj Kayal, Chandan Saha, and Bhargav Thankey. Low-depth arithmetic circuit lower bounds: Bypassing set-multilinearization. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming, ICALP 2023, July 10-14, 2023, Paderborn, Germany*, volume 261 of *LIPIcs*, pages 12:1–12:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. 3

[AGKS15] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-Sets for ROABP and Sum of Set-Multilinear Circuits. *SIAM J. Comput.*, 44(3):669–697, 2015. 5

[AHM+06] Eric Allender, Lisa Hellerstein, Paul McCabe, Toniann Pitassi, and Michael E. Saks. Minimizing DNF Formulas and $\mathrm{AC}^0{}_\mathrm{d}$ Circuits Given a Truth Table. In *21st Annual IEEE Conference on Computational Complexity (CCC 2006), 16-20 July 2006, Prague, Czech Republic*, pages 237–251. IEEE Computer Society, 2006. 2

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof Verification and the Hardness of Approximation Problems. *J. ACM*, 45(3):501–555, 1998. Conference version appeared in the proceedings of FOCS 1992. 6

[Ara11] Manuel Araújo. Classification of quadratic forms. https://www.math.tecnico.ulisboa.pt/~ggranja/manuel.pdf, 2011. 8

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Conference version appeared in the proceedings of FOCS 1992. 6

[AS05] Manindra Agrawal and Nitin Saxena. Automorphisms of Finite Rings and Applications to Complexity of Problems. In *23rd Annual Symposium on Theoretical Aspects of Computer Science, STACS 2005*, pages 1–17, 2005. 1, 8

[BBB+00] Amos Beimel, Francesco Bergadano, Nader H. Bshouty, Eyal Kushilevitz, and Stefano Varricchio. Learning functions represented as multiplicity automata. *J. ACM*, 47(3):506–530, 2000. Conference version appeared in the proceedings of FOCS 1996. 5

[BDS24] Omkar Baraskar, Agrim Dewan, and Chandan Saha. Testing Equivalence to Design Polynomials. In Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, and Daniel Lokshtanov, editors, *41st International Symposium on Theoretical Aspects of Computer Science, STACS 2024, March 12-14, 2024, Clermont-Ferrand, France*, volume 289 of *LIPIcs*, pages 9:1–9:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024. 1, 2, 4, 5, 7

[BIJL18] Markus Bläser, Christian Ikenmeyer, Gorav Jindal, and Vladimir Lysikov. Generalized matrix completion and algebraic natural proofs. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1193–1206, 2018. 6, 8

[BJ14] Markus Bläser and Gorav Jindal. A new deterministic algorithm for sparse multivariate polynomial interpolation. In Katsusuke Nabeshima, Kosaku Nagasaka, Franz Winkler, and Ágnes Szántó, editors, *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, pages 51–58. ACM, 2014. 2

[BRS17] Markus Bläser, B. V. Raghavendra Rao, and Jayalal Sarma. Testing Polynomial Equivalence by Scaling Matrices. In *Proceedings of 21st International Symposium on Fundamentals of Computation Theory (FCT), France*, volume 10472, pages 111–122, 2017. 8

[BSS89] Lenore Blum, Mike Shub, and Steve Smale. On a Theory of Computation and Complexity over the Real Numbers: NP-completeness, Recursive Functions and Universal Machines. *Bulletin of the American Mathematical Society*, 21(1):1–46, 1989. 3

[BSV20] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Deterministic Factorization of Sparse Polynomials with Bounded Individual Degree. *J. ACM*, 67(2):8:1–8:28, 2020. Conference version appeared in the proceedings of FOCS 2018. 2

[BT88] Michael Ben-Or and Prasoon Tiwari. A Deterministic Algorithm for Sparse Multivariate Polynominal Interpolation (Extended Abstract). In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 301–309. ACM, 1988. 2

[CGS23] Suryajith Chillara, Coral Grichener, and Amir Shpilka. On hardness of testing equivalence to sparse polynomials under shifts. In Petra Berenbrink, Patricia Bouyer, Anuj Dawar, and Mamadou Moustapha Kanté, editors, *40th International Symposium on Theoretical Aspects of Computer Science, STACS 2023, March 7-9, 2023, Hamburg, Germany*, volume 254 of *LIPIcs*, pages 22:1–22:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. 5, 8

[CIKK16] Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. Learning Algorithms from Natural Proofs. In Ran Raz, editor, *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, volume 50 of *LIPIcs*, pages 10:1–10:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. 2

[DF89] M.E Dyer and A.M Frieze. The solution of some random np-hard problems in polynomial expected time. *Journal of Algorithms*, 10(4):451–489, 1989. 5

[Din07] Irit Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007. Conference version appeared in the proceedings of STOC 2006. 6

[DL78] Richard A. DeMillo and Richard J. Lipton. A Probabilistic Remark on Algebraic Program Testing. *Inf. Process. Lett.*, 7(4):193–195, 1978. 3

[DOS14] Zeev Dvir, Rafael Oliveira, and Amir Shpilka. Testing equivalence of polynomials under shifts. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, volume 8572 of *Lecture Notes in Computer Science*, pages 417–428. Springer, 2014. 8

[Fis94] Ismor Fischer. Sums of Like Powers of Multivariate Linear Forms. *Mathematics Magazine*, 67(1):59–61, 1994.

[For16] Michael A. Forbes. Some concrete questions on the border complexity of polynomials, 2016. `https://www.youtube.com/watch?v=1HMogQIHT6Q`. 4

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-Time Identity Testing of Non-commutative and Read-Once Oblivious Algebraic Branching Programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252. IEEE Computer Society, 2013. 5

[GGKS19] Ankit Garg, Nikhil Gupta, Neeraj Kayal, and Chandan Saha. Determinant equivalence test over finite fields and over Q. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Greece*, volume 132 of *LIPIcs*, pages 62:1–62:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 7

[GK93] Dima Grigoriev and Marek Karpinski. A Zero-Test and an Interpolation Algorithm for the Shifted Sparse Polynominals. In Gérard D. Cohen, Teo Mora, and Oscar Moreno, editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 10th International Symposium, AAECC-10, San Juan de Puerto Rico, Puerto Rico, May 10-14, 1993, Proceedings*, volume 673 of *Lecture Notes in Computer Science*, pages 162–169. Springer, 1993. 2

[GKKS16] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic Circuits: A Chasm at Depth 3. *SIAM J. Comput.*, 45(3):1064–1079, 2016. Conference version appeared in the proceedings of FOCS 2013. 2

[GKS90] Dima Grigoriev, Marek Karpinski, and Michael F. Singer. Fast Parallel Algorithms for Sparse Multivariate Polynomial Interpolation over Finite Fields. *SIAM J. Comput.*, 19(6):1059–1063, 1990. 2

[GKS20] Ankit Garg, Neeraj Kayal, and Chandan Saha. Learning sums of powers of low-degree polynomials in the non-degenerate case. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 889–899. IEEE, 2020. 7

[GL00] Dima Grigoriev and Yagati N. Lakshman. Algorithms for computing sparse shifts for multivariate polynomials. *Appl. Algebra Eng. Commun. Comput.*, 11(1):43–67, 2000. conference version appeared in the proceedings of ISSAC 1995. 8

[GQ23] Joshua A. Grochow and Youming Qiao. On the complexity of isomorphism problems for tensors, groups, and polynomials I: tensor isomorphism-completeness. *SIAM J. Comput.*, 52(2):568–617, 2023. Conference version appeared in the proceedings of ITCS 2021. 8

[GR98] Dima Grigoriev and Alexander A. Razborov. Exponential Complexity Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions Over Finite Fields. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 269–278. IEEE Computer Society, 1998. 3

[Gri97] Dima Grigoriev. Testing shift-equivalence of polynomials by deterministic, probabilistic and quantum machines. *Theor. Comput. Sci.*, 180(1-2):217–228, 1997. 8

[Gro12] Joshua A. Grochow. *Symmetry and equivalence relations in classical and geometric complexity theory*. PhD thesis, University of Chicago, Chicago, IL, 2012. 7

[GS19] Nikhil Gupta and Chandan Saha. On the symmetries of and equivalence test for design polynomials. In *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 53:1–53:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 7

[GST23] Nikhil Gupta, Chandan Saha, and Bhargav Thankey. Equivalence Test for Read-Once Arithmetic Formulas. In Nikhil Bansal and Viswanath Nagarajan, editors, *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 4205–4272. SIAM, 2023. 1, 2, 4, 7

[Hås90] Johan Håstad. Tensor Rank is NP-Complete. *J. Algorithms*, 11(4):644–654, 1990. Conference version appeared in the proceedings of ICALP 1989. 4, 5, 8

[HG18] Qiao-Long Huang and Xiao-Shan Gao. Deterministic interpolation of sparse black-box multivariate polynomials using kronecker type substitutions, 2018.

[Hir18] Shuichi Hirahara. Non-Black-Box Worst-Case to Average-Case Reductions within NP. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 247–258. IEEE Computer Society, 2018. 2

[Hir22] Shuichi Hirahara. NP-Hardness of Learning Programs and Partial MCSP. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 968–979. IEEE, 2022. 2, 3

[HJLT96] Thomas R. Hancock, Tao Jiang, Ming Li, and John Tromp. Lower Bounds on Learning Decision Lists and Trees. *Inf. Comput.*, 126(2):114–122, 1996. 3

[HOS18] Shuichi Hirahara, Igor C. Oliveira, and Rahul Santhanam. NP-hardness of Minimum Circuit Size Problem for OR-AND-MOD Circuits. In Rocco A. Servedio, editor, *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, volume 102 of *LIPIcs*, pages 5:1–5:31. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 2

[Ila20] Rahul Ilango. Constant Depth Formula and Partial Function Versions of MCSP are Hard. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 424–433. IEEE, 2020. 2

[Ila21] Rahul Ilango. The Minimum Formula Size Problem is (ETH) Hard. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 427–432. IEEE, 2021. 2

[ILO20] Rahul Ilango, Bruno Loff, and Igor C. Oliveira. NP-Hardness of Circuit Minimization for Multi-Output Functions. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 169 of *LIPIcs*, pages 22:1–22:36. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 2

[IW97] Russell Impagliazzo and Avi Wigderson. $P = BPP$ if E Requires Exponential Circuits: Derandomizing the XOR Lemma. In Frank Thomson Leighton and Peter W. Shor, editors, *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 220–229. ACM, 1997. 3

[Kay11] Neeraj Kayal. Efficient algorithms for some special cases of the polynomial equivalence problem. In Dana Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 1409–1421. SIAM, 2011. 1, 4, 7

[Kay12a] Neeraj Kayal. Affine projections of polynomials: extended abstract. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 643–662, 2012. 1, 7, 8

[Kay12b] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.

[KC00] Valentine Kabanets and Jin-yi Cai. Circuit minimization problem. In F. Frances Yao and Eugene M. Luks, editors, *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 73–79. ACM, 2000. 2, 3

[KNS19] Neeraj Kayal, Vineet Nair, and Chandan Saha. Average-case linear matrix factorization and reconstruction of low width algebraic branching programs. *Comput. Complex.*, 28(4):749–828, 2019. 7

[KNS20] Neeraj Kayal, Vineet Nair, and Chandan Saha. Separation Between Read-once Oblivious Algebraic Branching Programs (ROABPs) and Multilinear Depth-three Circuits. *ACM Trans. Comput. Theory*, 12(1):2:1–2:27, 2020. Conference version appeared in the proceedings of STACS 2016. 4

[KNST19] Neeraj Kayal, Vineet Nair, Chandan Saha, and Sébastien Tavenas. Reconstruction of Full Rank Algebraic Branching Programs. *ACM Trans. Comput. Theory*, 11(1):2:1–2:56, 2019. Conference version appeared in the proceedings of CCC 2017. 1, 7

[KS01] Adam R. Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 216–223, 2001. 2

[KS06] Adam R. Klivans and Amir Shpilka. Learning Restricted Models of Arithmetic Circuits. *Theory of Computing*, 2(10):185–206, 2006. Conference version appeared in the proceedings of COLT 2003. 5

[KS08] Subhash Khot and Rishi Saket. Hardness of Minimizing and Learning DNF Expressions. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 231–240. IEEE Computer Society, 2008. 8

[KS09] Adam R. Klivans and Alexander A. Sherstov. Cryptographic hardness for learning intersections of halfspaces. *J. Comput. Syst. Sci.*, 75(1):2–12, 2009. 8

[KS19] Neeraj Kayal and Chandan Saha. Reconstruction of non-degenerate homogeneous depth three circuits. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 413–424. ACM, 2019. 5, 7

[KST16] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An almost cubic lower bound for depth three arithmetic circuits. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, volume 55 of *LIPIcs*, pages 33:1–33:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. 9

[KT90] Erich Kaltofen and Barry M. Trager. Computing with Polynomials Given By Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators. *J. Symb. Comput.*, 9(3):301–320, 1990. Conference version appeared in the proceedings of FOCS 1988. 1

[Lam04] T. Y. Lam. *Introduction To Quadratic Forms Over Fields*. American Mathematical Society, 2004. 8

[LS95] Yagati N. Lakshman and B. David Saunders. Sparse polynomial interpolation in nonstandard bases. *SIAM J. Comput.*, 24(2):387–397, 1995. 8

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 804–814. IEEE, 2021. 3

[Luc78] Edouard Lucas. Théorie des fonctions numériques simplement périodiques. *American Journal of Mathematics*, 1(2):184–196, 1878. 49

[LV03] Richard J. Lipton and Nisheeth K. Vishnoi. Deterministic identity testing for multivariate polynomials. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA*, pages 756–760. ACM/SIAM, 2003. 2

[Mas79] W. J. Masek. Some NP-complete set covering problems. *Unpublished Manuscript*, 1979. 2

[MNS20] Janaky Murthy, Vineet Nair, and Chandan Saha. Randomized Polynomial-Time Equivalence Between Determinant and Trace-IMM Equivalence Tests. In *45th International Symposium on Mathematical Foundations of Computer Science, MFCS 2020, August 24-28, 2020, Prague, Czech Republic*, volume 170 of *LIPIcs*, pages 72:1–72:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 7

[MS21] Dori Medini and Amir Shpilka. Hitting sets and reconstruction for dense orbits in $\mathrm{VP_e}$ and $\Sigma\Pi\Sigma$ circuits. In *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 19:1–19:27. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 4, 5, 7, 9

[NW97] Noam Nisan and Avi Wigderson. Lower Bounds on Arithmetic Circuits Via Partial Derivatives. *Computational Complexity*, 6(3):217–234, 1997. Conference version appeared in the proceedings of FOCS 1995. 9

[Pat96] Jacques Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In *Advances in Cryptology*

- *EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, pages 33–48, 1996. 1

[PS19] Ján Pich and Rahul Santhanam. Why are Proof Complexity Lower Bounds Hard? In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 1305–1324. IEEE Computer Society, 2019. 2

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Math. Notes*, 41:333–338, 1987. 3

[Roc18] Daniel S. Roche. What Can (and Can't) we Do with Sparse Polynomials? In Manuel Kauers, Alexey Ovchinnikov, and Éric Schost, editors, *Proceedings of the 2018 ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2018, New York, NY, USA, July 16-19, 2018*, pages 25–30. ACM, 2018. 2

[Sax06] Nitin Saxena. *Morphisms of rings and applications to complexity*. PhD thesis, Indian Institute of Technology, Kanpur, 2006. 1, 8

[Sch80] Jacob T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, 1980. 3

[Shi16] Yaroslav Shitov. How hard is the tensor rank?, 2016. 4, 5, 8

[ST21] Chandan Saha and Bhargav Thankey. Hitting Sets for Orbits of Circuit Classes and Polynomial Families. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 50:1–50:26. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 4, 5

[Swe18] Joseph Swernofsky. Tensor rank is hard to approximate. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2018, August 20-22, 2018 - Princeton, NJ, USA*, volume 116 of *LIPIcs*, pages 26:1–26:9. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 6, 8

[SWZ17] Zhao Song, David P. Woodruff, and Peilin Zhong. Relative Error Tensor Low Rank Approximation. *CoRR*, abs/1704.08246, 2017. 6, 8

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. Conference version appeared in the proceedings of MFCS 2013. 2

[Thi98] Thomas Thierauf. The Isomorphism Problem for Read-Once Branching Programs and Arithmetic Circuits. *Chicago J. Theor. Comput. Sci.*, 1998, 1998. 1

[vzGK85] Joachim von zur Gathen and Erich L. Kaltofen. Factoring Sparse Multivariate Polynomials. *J. Comput. Syst. Sci.*, 31(2):265–287, 1985. 2

[Wal13] Lars Ambrosius Wallenborn. Computing the hilbert symbol, quadratic form equivalence and integer factoring. Diploma thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn, 2013. 8

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation, EUROSAM '79, An International Symposium on Symbolic and Algebraic Computation, Marseille, France, June 1979, Proceedings*, pages 216–226, 1979. 3

A Handling translations

A.1 Extending Theorem 1 for translations

In this section, we modify the construction of Section 3.1 to show NP-hardness of ETsparse when translations are also involved. The idea is to choose the degree parameters for the polynomial f such that if $f(A\mathbf{z}+\mathbf{b})$ is s-sparse, where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$, then $\mathbf{b}$ must be $\mathbf{0}$ (the all 0s vector in $\mathbb{F}^{|\mathbf{z}|}$) and A must be as described in (4). We first show the reduction over any field of characteristic not equal to 2 and give a separate construction for characteristic 2 fields.

Formally, let ψ, $\mathbf{x}, x_0, \mathbf{y}$ and $\mathbf{z}$ be as denoted in Section 3.1. For $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$, $\mathbf{b}_{|w}$ denotes the component of $\mathbf{b}$ corresponding to the variable $w \in \mathbf{z}$. Let d_1, d_2, d_3, $d_4 \in \mathbb{N}$. Consider the following polynomials:

- Corresponding to x_i, where $i \in [n]$, define $Q_i(\mathbf{z})$ as:

$$Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z}),\ Q_{i,1}(\mathbf{z}) := x_0^{(3i-2)(d_2+1)} x_i^{d_2},$$
$$Q_{i,2}(\mathbf{z}) := x_0^{(3i-1)(d_2+1)} (y_i + x_i)^{d_3} \text{ and } Q_{i,3}(\mathbf{z}) := x_0^{3i(d_2+1)} (y_i - x_i)^{d_3}.$$

- For the k^{th} clause, $k \in [m]$, define $R_k(\mathbf{z}) := x_0^{k(3d_4+1)} \prod_{j \in C_k} (y_j + (-1)^{a_{k,j}} x_j)^{d_4}$.

Define $s := 1 + n(d_3+3) + m(d_4+1)^2$. Impose the following conditions on the d_i's:

$$d_1 \geq 6n(d_2+1) + 2d_3 + 2,\ d_2 \geq 2d_3,\ d_3 \geq m(d_4+1)^2 + 1,\ d_4 \geq m. \tag{24}$$

Finally, define $f(\mathbf{z})$ as:

$$f(\mathbf{z}) := x_0^{d_1} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{25}$$

The d_i's are chosen in Section A.1.1 such that they are $(mn)^{O(1)}$ and also satisfy the inequalities of (24) over any field. Observations A.1, A.2 and A.3 hold under the conditions of (24).

Observation A.1. For all $i \in [n], k \in [m]$, the polynomials $x_0^{d_1}$, $Q_{i,1}(\mathbf{z})$, $Q_{i,2}(\mathbf{z})$, $Q_{i,3}(\mathbf{z})$ and $R_k(\mathbf{z})$ are degree separated from one another. Also, $Q_i(\mathbf{z})$ is degree separated from other $Q_j(\mathbf{z})$'s, for $i, j \in [n]$ and $i \neq j$. Similarly, $R_k(\mathbf{z})$ is degree separated from $R_l(\mathbf{z})$ for $k, l \in [m]$ and $k \neq l$.

Proof. Let $i \in [n]$. Note that $Q_{i,1}(\mathbf{z})$ has degree $(3i-2)(d_2+1) + d_2$, $Q_{i,2}(\mathbf{z})$ has degree $(3i-1)(d_2+1) + d_3$ and $Q_{i,3}(\mathbf{z})$ has degree $3i(d_2+1) + d_3$. Clearly,

$$3i(d_2+1) + d_3 > (3i-1)(d_2+1) + d_3 > (3i-2)(d_2+1) + d_2.$$

Thus, $Q_{i,1}(\mathbf{z})$, $Q_{i,2}(\mathbf{z})$ and $Q_{i,3}(\mathbf{z})$ are degree separated and Q_i is a sum of three degree separated polynomials and has degree $3i(d_2+1) + d_3$.

Now, let $i \in [n]$ and $k \in [m]$. The lowest degree of any monomial of Q_i is $(3i-2)(d_2+1)+d_2 > 2d_2$, while the degree of any monomial of R_k is $k(3d_4+1)+3d_4 \leq m(3d_4+1)+3d_4$. As $d_2 \geq 2d_3$ and $d_3 > m(d_4+1)^2$ from (24), therefore

$$2d_2 \geq 4d_3 > 4m(d_4+1)^2 > m(3d_4+1)+3d_4$$

Thus Q_i is degree separated from R_k.

Lastly, let $i, j \in [n]$ where $i < j$ without loss of generality. The highest degree of any monomial of Q_i is $3i(d_2+1)+d_3$ while the lowest degree of any monomial of Q_j is $(3j-2)(d_2+1)+d_2$. Note that

$$(3j-2)(d_2+1)+d_2 \geq (3i+1)(d_2+1)+d_2 > 3i(d_2+1)+d_3$$

because $j \geq i+1$ and $d_2 \geq 2d_3$ from (24). Therefore Q_i and Q_j are degree separated. That $R_k(\mathbf{z})$ is degree separated from $R_l(\mathbf{z})$ for $k, l \in [m]$ and $k \neq l$ can be observed from the fact that the degree of $R_k(\mathbf{z})$ is $k(3d_4+1)+3d_4$. Clearly, $x_0^{d_1}$ is degree separated from the rest of the polynomials because $d_1 \geq 6n(d_2+1)+2d_3+2$ while the highest degree polynomial among Q_i's and R_k's is Q_n of degree $3n(d_2+1)+d_3$. □

Observation A.2. The degree of f is d_1 with $\frac{d_1}{2} > s$.

Proof. By Observations A.1 and 2.2 the degree of f is the maximum degree among $x_0^{d_1}$, Q_i's and R_k's, where $i \in [n]$ and $k \in [m]$. As observed in the proof of Observation A.3, the degree of Q_i is $3i(d_2+1)+d_3$, that of R_k is $k(3d_4+1)+3d_4$ and $3i(d_2+1)+d_3 > k(3d_4+1)+3d_4$. Further, $d_1 \geq 6nd_2+6n+2d_3+2$. Hence the degree of f is d_1. Finally, note that under the conditions of (24), $\frac{d_1}{2} > 3nd_2 > 3nd_3 > s$. □

Observation A.3. $\mathcal{S}(f(\mathbf{z})) = 1 + n(2d_3+3) + m(d_4+1)^3$ and $\mathrm{Supp}(f) = 7$.

The proof of Observation A.3 is similar to that of Observation 3.3 and uses Observations A.1, 2.2 and 2.5.

The forward direction. If $\mathbf{u} \in \{0,1\}^n$ is such that $\psi(\mathbf{u}) = 1$ and f, as described in (25), is the polynomial corresponding to ψ, then Proposition 3.1, with some changes to its statement and proof, shows that for $\mathbf{b} = \mathbf{0}$ and the transform A as described in (4), $\mathcal{S}(f(A\mathbf{z})) \leq s$ holds.

The reverse direction. We leverage the constraints in (24) to show that if $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \leq s$ for some $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$, then A is as described in (4) and $\mathbf{b} = \mathbf{0}$. Lemma A.1 shows that $A(x_0) = x_0$, without loss of generality, and $\mathbf{b}_{|x_0} = 0$. Then, Lemma A.2 shows that the summands of $f(A\mathbf{z}+\mathbf{b})$ are degree separated with respect to x_0.

Lemma A.1. Without loss of generality, $A(x_0) + \mathbf{b}_{|x_0} = x_0$.

Proof. Let $A(x_0) + \mathbf{b}_{|x_0} = \ell_0 + b_0$, where $\ell_0 = A(x_0)$. If $b_0 = 0$ and ℓ_0 is a linear form in at least two variables, then by the choice of d_1 and Observation 2.5, $\mathcal{S}((\ell_0+b_0)^{d_1}) \geq d_1 + 1 > s$. If $b_0 \neq 0$, then by the binomial theorem

$$(\ell_0+b_0)^{d_1} = \sum_{i=0}^{d_1} \binom{d_1}{i} b_0^i \ell_0^{d_1-i}.$$

The summands in the above expansion are degree separated and $\binom{d_1}{i} \neq 0$ for all $i \in [0, d_1]$ because of the choice of d_1 (and Lucas's Theorem if the characteristic is finite). Thus, $(\ell_0+b_0)^{d_1}$

contains at least one monomial of degree i, for all $i \in [0, d_1]$. Since $\frac{d_1}{2} > 3n(d_2+1)+d_3 > s$ therefore at least s monomials in $(\ell_0+b_0)^{d_1}$ are degree separated from $Q_i(A\mathbf{z}+\mathbf{b})$'s and $R_k(A\mathbf{z}+\mathbf{b})$'s. Hence $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) > s$, a contradiction. Thus, $\mathbf{b}_{|x_0} = b_0$ must be 0 and $A(x_0) = \ell_0$ must have only one variable, which we can assume to be x_0 without loss of generality. □

Lemma A.2. *For all $i \in [n]$, $k \in [m]$, $x_0^{d_1}$, $Q_{i,1}(A\mathbf{z}+\mathbf{b})$, $Q_{i,2}(A\mathbf{z}+\mathbf{b})$, $Q_{i,3}(A\mathbf{z}+\mathbf{b})$ and $R_j(A\mathbf{z}+\mathbf{b})$ are degree separated from one another with respect to x_0. Also, $Q_i(A\mathbf{z}+\mathbf{b})$ is degree separated with respect to x_0 from other $Q_j(A\mathbf{z}+\mathbf{b})$'s, for $i,j \in [n]$ and $i \neq j$. Similarly, $R_k(A\mathbf{z}+\mathbf{b})$ is degree separated with respect to x_0 from $R_l(A\mathbf{z}+\mathbf{b})$ for $k,l \in [m]$ and $k \neq l$.*

Proof. Let $i \in [n]$. For $Q_{i,1}(A\mathbf{z}+\mathbf{b})$, $Q_{i,2}(A\mathbf{z}+\mathbf{b})$ and $Q_{i,3}(A\mathbf{z}+\mathbf{b})$, the respective range of x_0-degree of a monomial of respective polynomials is $[(3i-2)(d_2+1), (3i-2)(d_2+1)+d_2]$, $[(3i-1)(d_2+1), (3i-1)(d_2+1)+d_3]$ and $[3i(d_2+1), 3i(d_2+1)+d_3]$. As $d_2 > d_3$, it can be observed that these ranges are disjoint, implying $Q_{i,1}(A\mathbf{z}+\mathbf{b}), Q_{i,2}(A\mathbf{z}+\mathbf{b})$ and $Q_{i,3}(A\mathbf{z}+\mathbf{b})$ are degree separated from one another with respect to x_0.

Let $i,j \in [n]$ and $i < j$ without loss of generality. For $Q_i(A\mathbf{z}+\mathbf{b})$ and $Q_j(A\mathbf{z}+\mathbf{b})$, the respective range of x_0-degree of a monomial of respective polynomials is $[(3i-2)(d_2+1), 3i(d_2+1)+d_3]$ and $[(3j-2)(d_2+1), 3j(d_2+1)+d_3]$. As $d_2 > d_3$ and $j \geq i+1$, therefore $(3j-2)(d_2+1) > 3i(d_2+1)+d_3$. Hence, $Q_i(A\mathbf{z}+\mathbf{b})$ is degree separated from $Q_j(A\mathbf{z}+\mathbf{b})$ with respect to x_0.

Now, let $k,l \in [m]$ and $k < l$ without loss of generality. For $R_k(A\mathbf{z}+\mathbf{b})$ and $R_l(A\mathbf{z}+\mathbf{b})$, the respective range of x_0-degree of a monomial of respective polynomials is $[k(3d_4+1), k(3d_4+1)+3d_4]$ and $[l(3d_4+1), l(3d_4+1)+3d_4]$. As $l \geq k+1$, therefore $l(3d_4+1) > k(3d_4+1)+3d_4$. Hence, $Q_i(A\mathbf{z}+\mathbf{b})$ is degree separated from $Q_j(A\mathbf{z}+\mathbf{b})$ with respect to x_0.

Lastly, let $i \in [n]$ and $k \in [m]$. The highest x_0-degree of a monomial in $R_k(A\mathbf{z}+\mathbf{b})$ is $k(3d_4+1)+3d_4 \leq m(3d_4+1)+3d_4$, while the lowest x_0-degree of any monomial in $Q_i(A\mathbf{z}+\mathbf{b})$ is $(3i-2)(d_2+1) > d_2$. Now,

$$d_2 \geq 2d_3 > 2m(d_4+1)^2 > m(3d_4+1)+3d_4$$

where the inequalities follow from the conditions in (24). Therefore $Q_i(A\mathbf{z}+\mathbf{b})$ is degree separated from $R_k(A\mathbf{z}+\mathbf{b})$'s with respect to x_0. Clearly, $x_0^{d_1}$ is degree separated with respect to x_0 from $Q_i(A\mathbf{z}+\mathbf{b})$ and $R_k(A\mathbf{z}+\mathbf{b})$ because $d_1 > 3n(d_2+1)+d_3$, the highest x_0-degree of any monomial among the $Q_i(A\mathbf{z}+\mathbf{b})'s$ and $R_k(A\mathbf{z}+\mathbf{b})$'s. □

$$\therefore \mathcal{S}(f(A\mathbf{z}+\mathbf{b})) = \mathcal{S}(x_0^{d_1}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z}+\mathbf{b})) \text{ by Lemma A.2.}$$

Lemma A.3 analyses the sparsity of $Q_i(A\mathbf{z}+\mathbf{b})$. The proof of Lemma A.3 is similar to that of Lemma 3.2 and uses Observations 2.5 and 2.6.

Lemma A.3. *For any invertible $A, \mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$ and $i \in [n]$:*

$$\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) = \mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b})) \geq d_3+3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined in this section. Equality holds if and only if $\mathbf{b}_{|x_i} = 0, \mathbf{b}_{|y_i} = 0$ and under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (-1)^{u_i} X_i$$

for some scaled $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0,1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \neq d_3+3$, then $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \geq 2d_3+3$.

Proof. From Lemma A.2 and Observation 2.2 it follows that $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) = \mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b}))$. The if direction of the lemma statement is easy to verify. For the only if direction consider the following cases of A and $\mathbf{b}$:

1. $\mathcal{S}(A(x_i)+\mathbf{b}_{|x_i}) \geq 2$: If $\mathbf{b}_{|x_i} = 0$ then $\mathcal{S}(A(x_i)) \geq 2$ and by Observation 2.5, $\mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) \geq d_2+1$. If $\mathbf{b}_{|x_i} \neq 0$ then by the choice of d_2 and Observation 2.6, $\mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) \geq d_2+1$. Also, $\mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) \geq 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b})) \geq 1$. Hence, $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \geq d_2+3 \geq 2d_3+3$ as $d_2 \geq 2d_3$. For the remaining cases, we consider $\mathcal{S}(A(x_i)+\mathbf{b}_{|x_i}) = 1$, meaning $A(x_i) = X_i$ for some scaled variable $X_i \in \mathbf{z}$ and $\mathbf{b}_{|x_i} = 0$.

2. $\mathcal{S}(A(x_i)+\mathbf{b}_{|x_i}) = 1$, $\mathcal{S}(A(y_i+x_i)+\mathbf{b}_{|y_i}+\mathbf{b}_{|x_i}) \geq 2$ and $\mathcal{S}(A(y_i-x_i)+\mathbf{b}_{|y_i}-\mathbf{b}_{|x_i}) \geq 2$: Like in the previous case, it can be shown using Observation 2.5 (if $\mathbf{b}_{|y_i} = 0$) or Observation 2.6 (if $\mathbf{b}_{|y_i} \neq 0$) that $\mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) \geq d_3+1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b})) \geq d_3+1$ implying $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \geq 2d_3+3$.

3. $\mathcal{S}(A(x_i)+\mathbf{b}_{|x_i}) = 1$ with $\mathcal{S}(A(y_i+x_i)+\mathbf{b}_{|y_i}+\mathbf{b}_{|x_i}) = 1$ or $\mathcal{S}(A(y_i-x_i)+\mathbf{b}_{|y_i}-\mathbf{b}_{|x_i}) = 1$: Because A is invertible, $\mathcal{S}(A(y_i-x_i)) \geq 1$ and $\mathcal{S}(A(y_i+x_i)) \geq 1$. Further, since $\mathbf{b}_{|x_i} = 0$, therefore $\mathbf{b}_{|y_i}$ must be 0. This observation and the invertibility of A imply that exactly one of $\mathcal{S}(A(y_i+x_i)) = 1$ or $\mathcal{S}(A(y_i-x_i)) = 1$ holds. Without loss of generality, let $\mathcal{S}(A(y_i+x_i)) = 1$, which implies $A(y_i) = Y_i - X_i$ for some scaled variable $Y_i \in \mathbf{z}$. Then $A(y_i-x_i) = Y_i - 2X_i$. Hence, $\mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) = 1$, $\mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) = 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b})) = d_3+1$ (by Observation 2.5) implying $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) = d_3+3$.

The first two cases show that if A and $\mathbf{b}$ are not as per the lemma statement, then $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \geq 2d_3+3$; otherwise, $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) = d_3+3$. □

Lemma 3.3 then holds as before. Lemmas A.1, A.3 and 3.3 together show that A is a permuted scaled version of the transform of (4) and that $\mathbf{b} = \mathbf{0}$. Then, Proposition 3.2 shows how a satisfying assignment for ψ can be recovered from A.

Construction for characteristic 2 fields

Since over characteristic 2 fields $y_i + x_i$ and $y_i - x_i$ are the same polynomials, we need to modify the previous construction. Moreover, the sparsifying transform will also be slightly different. Formally, let ψ, $\mathbf{x}, x_0, \mathbf{y}$ and $\mathbf{z}$ be as denoted in Section 3.1. Let d_1, d_2, d_3, $d_4 \in \mathbb{N}$. Consider the following polynomials:

- Corresponding to x_i, where $i \in [n]$, define $Q_i(\mathbf{z})$ as:

$$Q_i(\mathbf{z}) := Q_{i,1}(\mathbf{z}) + Q_{i,2}(\mathbf{z}) + Q_{i,3}(\mathbf{z}),\ Q_{i,1}(\mathbf{z}) := x_0^{(3i-2)(d_2+1)} x_i^{d_2},$$
$$Q_{i,2}(\mathbf{z}) := x_0^{(3i-1)(d_2+1)} (y_i + x_i)^{d_3} \text{ and } Q_{i,3}(\mathbf{z}) := x_0^{3i(d_2+1)} (y_i)^{d_3}.$$

- For the k^{th} clause, $k \in [m]$, define $R_k(\mathbf{z}) := x_0^{k(3d_4+1)} \prod_{j \in C_k} (y_j + a_{k,j} x_j)^{d_4}$.

Define $s := 1 + n(d_3+3) + m(d_4+1)^2$. Choose d_i's as specified in Section A.1.1 so that they are $(mn)^{O(1)}$ and satisfy the conditions of (24). Finally, define $f(\mathbf{z})$ as:

$$f(\mathbf{z}) := x_0^{d_1} + \sum_{i=1}^{n} Q_i(\mathbf{z}) + \sum_{k=1}^{m} R_k(\mathbf{z}). \tag{26}$$

Observations A.1, A.2 hold with little change. Observation A.4 analyses the sparsity and support of f and has a proof similar to that of Observations A.3 and Observation 3.6.

Observation A.4. $\mathcal{S}(f(\mathbf{z})) \leq 1 + n(d_3+3) + m(d_4+1)^3$ and $4 \leq \mathrm{Supp}(f) \leq 7$.

Remarks. Like in Section 3.5.3, the sparsity of the polynomial output by the reduction over characteristic 2 fields depends on the number of variables which are complemented within a clause. Hence, for the same number of variables n and the same number of clauses m, the output polynomial corresponding to two different ψ's may have different sparsity. Thus, the reduction is not natural over characteristic 2 fields.

The forward direction. Let $\mathbf{u} \in \{0,1\}^n$ be such that $\psi(\mathbf{u}) = 1$ and f, as described in (26), be the polynomial corresponding to ψ. Proposition 3.5 shows how $\mathbf{u}$ can be used to construct a sparsifying transform A with $\mathbf{b} = \mathbf{0}$.

The reverse direction. Let $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$ be such that $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \leq s$. The analysis of the reverse direction in the previous section holds with some changes. Formally, Lemma A.1 holds without any change in its proof while Lemma A.2 holds with some change in its statement and proof. Thus, $A(x_0) = x_0$ without loss of generality. Lemma A.4 analyses $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b}))$, where $i \in [n]$, and its proof is similar to that of Lemma A.3.

Lemma A.4. For any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F}^{|\mathbf{z}|})$, $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$ and $i \in [n]$:

$$\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) = \mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,2}(A\mathbf{z}+\mathbf{b})) + \mathcal{S}(Q_{i,3}(A\mathbf{z}+\mathbf{b})) \geq d_3 + 3,$$

where Q_i, $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$ are as defined in this section. Equality holds if and only if $\mathbf{b}_{|x_i} = 0, \mathbf{b}_{|y_i} = 0$ and under A

$$x_i \mapsto X_i \text{ and } y_i \mapsto Y_i + (1-u_i)X_i$$

for some scaled $X_i, Y_i \in \mathbf{z}$ and $u_i \in \{0,1\}$. Further, if $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \neq d_3 + 3$, then $\mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) \geq 2d_3 + 3$.

Lemma 3.3 holds with the same proof as before. Lemmas A.1, A.4 and 3.3 together show that A is a permuted scaled version of the transform of (10) and that $\mathbf{b} = \mathbf{0}$. Then, Proposition 3.6 shows how a satisfying assignment for ψ can be recovered from A.

A.1.1 Setting of parameters

For characteristic 0 fields. In this case, the inequalities in (24) can be converted to equalities. Thus

$$d_4 = m,\ d_3 = m(m+1)^2 + 1 = O(m^3) \implies s = O(nm^3),$$
$$d_2 = 2m(m+1)^2 + 2 = O(m^3),\ d_1 = 6nd_2 + 6n + 2d_3 + 2 = O(nm^3).$$

For finite characteristic fields. Let the characteristic be $p > 0$. If $p > d_1$, where the value of d_1 is as in the characteristic 0 fields case, then the d_i's are set as per the characteristic 0 fields case. Otherwise $p = O(nm^3)$. In such a case, we choose the d_i's to be of form $p^k - 1$, for some $k \in \mathbb{N}$, so that the conditions in (24) are satisfied and Observation 2.5 can be used over characteristic p fields. The lemmas and the observations in the previous section hold for this choice of the d_i's. Now, the following bounds hold on the d_i's:

$$d_4 \leq pm,\ d_3 \leq pm(d_4+1)^2 + p) = O(p^3m^3) \implies s = O(nm^3p^3),$$
$$d_2 = pd_3 + (p-1) = O(p^4m^3),\ d_1 \leq p(6nd_2 + 6n + 2d_3 + 2) = O(np^5m^3).$$

As $p = O(nm^3)$, therefore

$$d_1 = O(n^6m^{18}),\ d_2 = O(n^4m^{15}),\ d_3 = O(n^3m^{12}),\ d_4 = O(nm^4) \text{ and } s = O(n^4m^{12}).$$

A.2 Extending Theorem 2 for translations

In this section, we prove part 1 of Theorem 2 over all fields while considering translations. The proof involves a careful analysis of the sparsity of f as defined corresponding to a 3-CNF ψ. We split the proof into two cases: over characteristic 0 fields and finite characteristic fields. The reason for the split is that in the analysis, we consider the sparsity of powers of affine forms $h = \ell + c$, where $\mathcal{S}(h) \geq 3$. The sparsity of such affine forms depends on the underlying field, as shown in the following analysis using Observation 2.4 and the binomial theorem.

A.2.1 For characteristic 0 fields

Consider the polynomial f as defined in (25) for a 3-CNF ψ and choose the d_i's to be $(mn)^{O(1)}$ and also satisfy the following conditions:

$$d_4 \geq \max(4mn, (mn)^{O(1/\epsilon)}),\ d_3 = m(d_4+1)^2+1,\ d_2 = d_3^2+1,\ d_1 = 6nd_2+6n+2d_3+2. \quad (27)$$

Note that the constraints in (24) are also satisfied under (27). Let $s := \mathcal{S}(f)$. By Observation A.3, $s = 1 + n(2d_3+3) + m(d_4+1)^3$. From Section A.1, it follows that for satisfiable ψ's, there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$, $\mathbf{b} = \mathbf{0}$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_3+3) + m(d_4+1)^2$. For unsatisfiable ψ's, Lemma A.5 gives lower bounds on $\mathcal{S}(f(A\mathbf{z}+\mathbf{b}))$, where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F}^{)}$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$, and encapsulates the argument of the reverse direction of the reduction in Section A.1 with a slightly deeper analysis. Comparing the sparsities for satisfiable and unsatisfiable ψ's proves part 1 of Theorem 2 for translations. Proposition A.1 shows α-gap-ETsparse is NP-hard using Lemma A.5 and the conditions in (27).

Lemma A.5. Let $\psi \in \overline{\text{3-SAT}}$, f, as defined in (25), be the polynomial corresponding to ψ, $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$.

1. If $A(x_0) + \mathbf{b}_{|x_0}$ is a non-trivial affine form in then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \frac{d_1}{2}$.[21]
2. If A and $\mathbf{b}$ are not as in item 1 and for some $j \in [n]$, $A(x_j) + \mathbf{b}_{|x_j}$ is a non-trivial affine form, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq d_2 + 1$.
3. If A and $\mathbf{b}$ are not as in items 1 and 2 and for some $j \in [n]$, $\mathcal{S}(A(y_j + x_j) + \mathbf{b}_{|y_j}) \geq 3$ or $\mathcal{S}(A(y_j - x_j) + \mathbf{b}_{|y_j}) \geq 3$, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \frac{(d_3+1)(d_3+2)}{2}$.
4. If A and $\mathbf{b}$ are not of the form described in the previous three cases, then $\mathbf{b} = \mathbf{0}$ and $\mathcal{S}(f(A\mathbf{z})) \geq (d_4+1)^3$.

Proof.

1. The proof of this case follows from the argument in the proof of Lemma A.1. Henceforth, we assume $A(x_0) = x_0$ and $\mathbf{b}_{|x_0} = 0$. Then, by Lemma A.2, it follows that

$$\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) = \mathcal{S}(x_0^{d_1}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z}+\mathbf{b})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z}+\mathbf{b})).$$

2. By Lemma A.2 and Observation 2.6 (applied to $Q_{i,1}(A\mathbf{z}+\mathbf{b})$) it follows that

$$S(f(A\mathbf{z}+\mathbf{b})) \geq \mathcal{S}(Q_{i,1}(A\mathbf{z}+\mathbf{b})) \geq d_2 + 1.$$

[21] an affine form $\ell + c$, where ℓ is a linear form, is non-trivial if ℓ is a linear form in at least two variables or $c \in \mathbb{F}\backslash\{0\}$.

3. In this case, for $i \in [0, n]$, $A(x_i)$ is some scaled variable in $\mathbf{z}$ and $\mathbf{b}_{|x_i} = 0$. If $\mathbf{b}_{|y_j} = 0$, then this case is the same as the third case of Lemma 4.1. Otherwise, let $A(y_j + x_j) + \mathbf{b}_{|y_j} = \ell_j + b_j$, where ℓ_j is a linear form in at least two variables. Then, using the binomial theorem and Observations 2.2 and 2.4, it follows that

$$\mathcal{S}((\ell_j + b_j)^{d_3}) = \mathcal{S}\Big(\sum_{i=0}^{d_3} \binom{d_3}{i} b_j^{d_3 - i} \ell_j^i\Big) = \sum_{i=0}^{d_3} \mathcal{S}\Big(\binom{d_3}{i} b_j^{d_3 - i} \ell_j^i\Big) \geq \sum_{i=0}^{d_3} (i+1) = \frac{(d_3+1)(d_3+2)}{2}.$$

Thus,

$$S(f(A\mathbf{z} + \mathbf{b})) \geq \mathcal{S}(Q_{i,2}(A\mathbf{z} + \mathbf{b})) \geq \frac{(d_3+1)(d_3+2)}{2}.$$

Similarly, if $\mathcal{S}(A(y_j - x_j) + \mathbf{b}_{|y_j}) \geq 3$, then $\mathcal{S}(f(A\mathbf{z} + \mathbf{b})) \geq \frac{(d_3+1)(d_3+2)}{2}$.

4. In this case, for $i \in [0, n]$, $A(x_i) = X_i$ and $\mathbf{b}_{|x_i} = 0$. For $i \in [n]$, $\mathcal{S}(A(y_i + x_i) + \mathbf{b}_{|y_i}) \leq 2$ and $\mathcal{S}(A(y_i - x_i) + \mathbf{b}_{|y_i}) \leq 2$. As A is invertible, $\mathcal{S}(A(y_i + x_i)) \geq 1$ and $\mathcal{S}(A(y_i - x_i)) \geq 1$. If $\mathcal{S}(A(y_i + x_i)) = 2$ or $\mathcal{S}(A(y_i - x_i)) = 2$, then $\mathbf{b}_{|y_i} = 0$. By the invertibility of A, if $\mathcal{S}(A(y_i + x_i)) = 1$, then $\mathcal{S}(A(y_i - x_i)) \geq 2$ and vice versa. This observation also implies $\mathbf{b}_{|y_i} = 0$. Thus, $\mathbf{b}_{|y_i} = 0$ for all $i \in [n]$ implying $\mathbf{b} = \mathbf{0}$. This case can then be proved in the same way as the last case of Lemma 4.1.

□

Proposition A.1. Let the input to α-gap-ETsparse be an s-sparse polynomial. Then, for $\alpha = s^{1/3-\epsilon}$, where $\epsilon \in (0, 1/3)$ is an arbitrary constant, α-gap-ETsparse is NP-hard.

Proof. The proof is similar to that of Proposition 4.1. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4) and $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2$. If $\psi \in \overline{\text{3-SAT}}$, it follows from Lemma A.5 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$:

$$\mathcal{S}(f(A\mathbf{z} + \mathbf{b})) \geq \min\Big(\frac{d_1}{2}, d_2 + 1, \frac{d_3^2 + 3d_3 + 2}{2}, (d_4 + 1)^3\Big).$$

The conditions imposed in (16) ensure that $(d_4 + 1)^3 > s_0$ and $(d_4 + 1)^3$ is the minimum. As $d_3 = m(d_4 + 1)^2 + 1$, therefore $s_0 = 1 + n(d_3 + 3) + m(d_4 + 1)^2 \leq 3nd_3 \leq 4mn(d_4 + 1)^2$. Consequently, the gap in the sparsities of the YES instances and NO instances is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{(d_4+1)^3}{4mn(d_4+1)^2} = \frac{d_4+1}{4mn}.$$

Also, note that as $d_4 \geq 4mn$, therefore $s \leq 2m(d_4 + 1)^3 \implies d_4 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{d_4+1}{4mn} \geq \frac{s^{1/3}}{2^{1/3} 4 m^{4/3} n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_4^{3\epsilon} \geq 2^{1/3} 4 m^{4/3} n \implies \frac{s^{1/3}}{2^{1/3} 4 m^{4/3} n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$.

□

A.2.2 For finite characteristic fields

Let the characteristic be p, where $p > 2$. If $p > d_1$, where d_1 is as chosen in Section A.2.1, then the argument of that section holds. Hence, it is assumed that $p \leq d_1 = (mn)^{O(1)}$. We again consider the polynomial f as defined in (25) and impose the following constraints on the d_i's.

$$d_4 \geq \max(3pmn, (mn)^{O(1/\epsilon)}),\; d_3 > m(d_4+1)^2,\; d_2 > (d_3+1)^2,\; d_1 > 6nd_2+6n+2d_3+2. \quad (28)$$

Note, we can get $d_3 = O(pm(d_4+1)^2 + p)$, $d_2 = O(p(d_3+1)^2)$ and $d_1 = O(p(6nd_2+6n+2d_3+2))$. The conditions of (24) are also satisfied under (28). From Section A.1, it follows that for satisfiable ψ's, there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$, $\mathbf{b} = \mathbf{0}$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_3+3) + m(d_4+1)^2$. For unsatisfiable ψ's, Lemma A.6 gives lower bounds on $\mathcal{S}(f(A\mathbf{z}+\mathbf{b}))$, where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F}^{)}$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$. Proposition A.2 shows α-gap-ETsparse is NP-hard using Lemma A.6 and the constraints in (28).

Lemma A.6. Let $\psi \in \overline{\text{3-SAT}}$, $f(\mathbf{z})$ be as defined in (25) corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$.

1. If $A(x_0) + \mathbf{b}_{|x_0}$ is a non-trivial affine form in then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \frac{d_1}{2}$.

2. If A and $\mathbf{b}$ are not as in item 1 and for some $j \in [n]$, $A(x_j) + \mathbf{b}_{|x_j}$ is a non-trivial affine form, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq d_2 + 1$.

3. If A and $\mathbf{b}$ are not as in items 1 and 2 and for some $j \in [n]$, $\mathcal{S}(A(y_j + x_j) + \mathbf{b}_{|y_j}) \geq 3$ or $\mathcal{S}(A(y_j - x_j) + \mathbf{b}_{|y_j}) \geq 3$, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq (d_3+1)^{1.63}$.

4. If A and $\mathbf{b}$ are not of the form described in the previous three cases, then $\mathbf{b} = \mathbf{0}$ and $\mathcal{S}(f(A\mathbf{z})) \geq (d_4+1)^3$.

Proof. The first, second and fourth cases can be proved similarly to those of Lemma A.5. Hence, we consider the third case. Then, for $i \in [0, n]$, $A(x_i)$ is some scaled variable in $\mathbf{z}$ and $\mathbf{b}_{|x_i} = 0$. If $\mathbf{b}_{|y_j} = 0$, then this case is the same as the third case of Lemma 4.3. If $\mathbf{b}_{|y_j} \neq 0$, then, without loss of generality, $A(y_j + x_j) + \mathbf{b}_{|y_j} = \ell_j + b_j$, with ℓ_j a linear form in at least two variables. Using the binomial theorem, the fact that $d_3 = p^k - 1 = \sum_{i=0}^{k-1}(p-1)p^i$, Lucas's theorem and Observations 2.2 and 2.4 gives

$$\mathcal{S}((\ell_j + b_j)^{d_3}) = \mathcal{S}\Big(\sum_{i=0}^{d_3}\binom{d_3}{i} b_j^{d_3-i}\ell_j^i\Big) = \sum_{i=0}^{d_3}\mathcal{S}\Big(\binom{d_3}{i} b_j^{d_3-i}\ell_j^i\Big)$$

$$\geq \sum_{i=0}^{d_3}\prod_{l=0}^{k-1}\binom{e_{i,l}+2-1}{2-1} = \sum_{i=0}^{d_3}\prod_{l=0}^{k-1}(e_{i,l}+1)$$

where $i = \sum_{l=0}^{k-1} e_{i,l}p^l$ with $e_{i,l} \in [0, p-1]$. Note that for $i = rp$, where $r \in [0, p^{k-1}-1]$, the value of $e_{t,l}$, where $l \geq 1$ and $t \in [i, i+p-1]$, is the same for all t while $e_{t,0} = t - i$. Therefore, for such i's, the following holds

$$\sum_{t=i}^{i+p-1}\prod_{l=0}^{k-1}(e_{t,l}+1) = \Big(\prod_{l=1}^{k-1}(e_{rp,l}+1)\Big)\cdot\sum_{t=rp}^{rp+p-1}(t-rp+1) = \frac{p(p+1)}{2}\prod_{l=1}^{k-1}(e_{rp,l}+1).$$

Using the above observation and the fact that $d_3 = p^k - 1$,

$$\sum_{i=0}^{d_3}\prod_{l=0}^{k-1}(e_{i,l}+1) = \sum_{r=0}^{p^{k-1}-1}\sum_{t=0}^{p-1}\prod_{l=0}^{k-1}(e_{rp+t,l}+1) = \frac{p(p+1)}{2}\sum_{r=0}^{p^{k-1}-1}\prod_{l=1}^{k-1}(e_{rp,l}+1).$$

By repeating the same argument, we get

$$\sum_{i=0}^{d_3}\prod_{l=0}^{k-1}(e_{i,l}+1) = \left(\frac{p(p+1)}{2}\right)^k = (d_3+1)^{1+\log_p((p+1)/2)}.$$

Now, $\log_p((p+1)/2)$ is an increasing function for $p \geq 3$. Thus, $\log_p((p+1)/2) \geq \log_3((3+1)/2) \geq 0.63$. We can then conclude that

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,2}(A\mathbf{z}+\mathbf{b})) \geq (d_3+1)^{1.63}.$$

Similarly, if $\mathcal{S}(A(y_j - x_j) + \mathbf{b}_{|y_j}) \geq 3$, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \mathcal{S}(Q_{j,3}(A\mathbf{z}+\mathbf{b})) \geq (d_3+1)^{1.63}$. □

Proposition A.2. Let the input to α-gap-ETsparse be an s-sparse polynomial. Then, for $\alpha = s^{1/3-\epsilon}$, where $\epsilon \in (0, 1/3)$ is an arbitrary constant, α-gap-ETsparse is NP-hard.

Proof. The proof is similar to that of Proposition A.1. For the polynomial f defined in (25), $s := \mathcal{S}(f) = 1 + n(2d_3+3) + m(d_4+1)^3$. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4) and $s_0 = 1 + n(d_3+3) + m(d_4+1)^2$. If $\psi \in \overline{\text{3-SAT}}$, it follows from Lemma A.5 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$:

$$\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \min\left(\frac{d_1}{2}, d_2+1, (d_3+1)^{1.63}, (d_4+1)^3\right).$$

The conditions imposed in (28) ensure that $(d_4+1)^3 > s_0$ and $(d_4+1)^3$ is the minimum. As $d_3 = m(d_4+1)^2 + 1$, therefore $s_0 = 1 + n(d_3+3) + m(d_4+1)^2 \leq 3nd_3 \leq 3pmn(d_4+1)^2$. Consequently, the gap in the sparsities of the YES instances and NO instances is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{(d_4+1)^3}{3pmn(d_4+1)^2} = \frac{d_4+1}{3pmn}.$$

Also, note that as $d_4 \geq 3pmn$, therefore $s \leq 2m(d_4+1)^3 \implies d_4+1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{d_4+1}{3pmn} \geq \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^\epsilon \geq d_4^{3\epsilon} \geq p2^{1/3}3m^{4/3}n \implies \frac{s^{1/3}}{p2^{1/3}3m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

For characteristic 2 fields

Let the characteristic be 2. Consider the polynomial as defined in (26). Let $s := \mathcal{S}(f)$. Now, s depends on the number of variables complemented in a clause. To prove the hardness of α-gap-ETsparse, $s \geq d_4^3$ is required (see the proof of Proposition A.3), and this can be achieved if there is at least one clause where all the variables are complemented. Thus, assume, without loss of generality, that such a clause exists (see footnote 19). Choose the d_i's to satisfy (28) with p set to 2. By Observation A.4 and the assumption on ψ, it holds that

$$1 + n(d_3+3) + (d_4+1)^3 \leq s \leq 1 + n(d_3+3) + m(d_4+1)^3.$$

For $\psi \in \overline{\text{3-SAT}}$, Lemma A.7, which can be proved in the same way as Lemma A.6, shows lower bounds on $\mathcal{S}(f(A\mathbf{z}+\mathbf{b}))$ where $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$. For $\psi \in$ 3-SAT, by Proposition 3.5 there exists $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ such that $\mathcal{S}(f(A\mathbf{z})) \leq s_0$, where $s_0 = 1 + n(d_3+3) + m(d_4+1)^2$. Proposition A.3 shows the NP-hardness of α-gap-ETsparse using Lemma A.7 and the inequalities in (28).

Lemma A.7. Let $\psi \in \overline{\text{3-SAT}}$, $f(\mathbf{z})$ be as defined in (26) corresponding to ψ and $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$.

1. If $A(x_0) + \mathbf{b}_{|x_0}$ is a non-trivial affine form in then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \frac{d_1}{2}$.
2. If A and $\mathbf{b}$ are not as in item 1 and for some $j \in [n]$, $A(x_j) + \mathbf{b}_{|x_j}$ is a non-trivial affine form, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq d_2 + 1$.
3. If A and $\mathbf{b}$ are not as in items 1 and 2 and for some $j \in [n]$, $\mathcal{S}(A(y_j + x_j) + \mathbf{b}_{|y_j}) \geq 3$ or $\mathcal{S}(A(y_j) + \mathbf{b}_{|y_j}) \geq 3$, then $\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq (d_3+1)^{1.58}$.
4. If A and $\mathbf{b}$ are not of the form described in the previous three cases, then $\mathbf{b} = \mathbf{0}$ and $\mathcal{S}(f(A\mathbf{z})) \geq (d_4+1)^3$.

Proposition A.3. Let the input to α-gap-ETsparse be an s-sparse polynomial. Then, for $\alpha = s^{1/3-\epsilon}$, where $\epsilon \in (0, 1/3)$ is an arbitrary constant, α-gap-ETsparse is NP-hard.

Proof. The proof is similar to that of Proposition A.2. If $\psi \in$ 3-SAT, then $\mathcal{S}(f(A\mathbf{z})) \leq s_0$ where A is as described in (4) and $s_0 = 1 + n(d_3+3) + m(d_4+1)^2$. If $\psi \in \overline{\text{3-SAT}}$, it follows from Lemma A.7 that for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^{|\mathbf{z}|}$:

$$\mathcal{S}(f(A\mathbf{z}+\mathbf{b})) \geq \min\left(\frac{d_1}{2}, d_2+1, (d_3+1)^{1.58}, (d_4+1)^3\right).$$

The conditions imposed in (28) ensure that $(d_4+1)^3 > s_0$ and $(d_4+1)^3$ is the minimum. As $d_3 = m(d_4+1)^2 + 1$, therefore $s_0 = 1 + n(d_3+3) + m(d_4+1)^2 \leq 3nd_3 \leq 6mn(d_4+1)^2$. Consequently, the gap in the sparsities of the YES instances and NO instances is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{(d_4+1)^3}{6mn(d_4+1)^2} = \frac{d_4+1}{6mn}.$$

Also, note that as $d_4 \geq 6mn$, therefore $s \leq 2m(d_4+1)^3 \implies d_4 + 1 \geq (\frac{s}{2m})^{1/3}$. Then, the gap is

$$\frac{(d_4+1)^3}{s_0} \geq \frac{d_4+1}{6mn} \geq \frac{s^{1/3}}{2^{4/3}3m^{4/3}n}.$$

Finally, note that $s \geq d_4^3$. Thus, for $d_4^{3\epsilon} \geq (mn)^{O(1)}$ large enough,

$$s^{\epsilon} \geq d_4^{3\epsilon} \geq 2^{4/3}3m^{4/3}n \implies \frac{s^{1/3}}{2^{4/3}3m^{4/3}n} \geq s^{1/3-\epsilon}.$$

Hence, the gap is at least $s^{1/3-\epsilon}$. Therefore, 3-SAT reduces to α-gap-ETsparse for $\alpha = s^{1/3-\epsilon}$. □

B Missing proofs from Section 2

B.1 Proof of Observation 2.1

Under any invertible linear transform applied to the variables of f, every monomial of f maps to a linear combination of monomials of the same degree. Thus, no new degree can be added to the set of degrees of f under any invertible linear transform. As $f \sim g$, the set of degrees of f is contained in the set of degrees of g, and vice versa, implying the two sets are the same.

B.2 Proof of Observation 2.2

As f and g are degree separated (or degree separated with respect to some variable), each monomial of $f + g$ is a monomial of f or g, but not both.

B.3 Proof of Observation 2.3

As $f_1 \sim f$ and $g_1 \sim g$, therefore f_1 and g_1 are degree separated by Observation 2.1. By Observation 2.2, the statement holds.

B.4 Proof of Observation 2.4

Without loss of generality, let $\ell = \sum_{i=1}^{m} c_i x_i$ where $c_i \in \mathbb{F}\backslash\{0\}$. If $\mathrm{char}(\mathbb{F}) = 0$, then $\mathcal{S}(\ell^d) = \binom{d+m-1}{m-1}$ follows from the multinomial theorem and the fact that the number of monomials of degree d in m variables is $\binom{d+m-1}{m-1}$. Suppose $\mathrm{char}(\mathbb{F}) = p$. Then, d is expressible as in the observation statement. It will be shown by induction on k that,

$$\mathcal{S}(\ell^d) = \prod_{i=0}^{k} \binom{e_i + m - 1}{m - 1}.$$

In the base case $k = 0$, $d < p$ and, like the $\mathrm{char}(\mathbb{F}) = 0$ case, it easily follows that $\mathcal{S}(\ell^d) = \binom{d+m-1}{m-1}$. Assume the statement for all $j < k$. Suppose $d = e_k p^k + \sum_{i=0}^{k-1} e_i p^i$, where $0 < e_k < p$. Then, using the fact that $(\sum_{j=1}^{m} c_j x_j)^p = \sum_{j=1}^{m} c_j^p x_j^p$ over $\mathbb{F}$,

$$\ell^d = \Big(\sum_{j=1}^{m} c_j x_j\Big)^{\sum_{i=0}^{k} e_i p^i} = \Big(\sum_{j=1}^{m} c_j^{p^k} x_j^{p^k}\Big)^{e_k} \prod\nolimits_{i=0}^{k-1} \Big(\sum_{j=1}^{m} c_j^{p^i} x_j^{p^i}\Big)^{e_i}.$$

Let $h = \prod_{i=0}^{k-1} \big(\sum_{j=1}^{m} c_j^{p^i} x_j^{p^i}\big)^{e_i}$. Note,

$$\Big(\sum_{j=1}^{m} c_j^{p^k} x_j^{p^k}\Big)^{e_k} = \sum_{\alpha_1+\cdots+\alpha_m = e_k} \binom{e_k}{\alpha_1 \ldots \alpha_m} \Big(\prod_{i=1}^{m} (c_i^{p^k} x_i^{p^k})^{\alpha_i}\Big).$$

By the inductive hypothesis, $\mathcal{S}(h) = \prod_{i=0}^{k-1} \binom{e_i+m-1}{m-1}$, while $\mathcal{S}((\sum_{j=1}^{m} c_j^{p^k} x_j^{p^k})^{e_k}) = \binom{e_k+m-1}{m-1}$, as $e_k < p$. Now,

$$\ell^d = \sum_{\alpha_1+\cdots+\alpha_m = e_k} \binom{e_k}{\alpha_1 \ldots \alpha_m} \Big(\prod_{i=1}^{m} (c_i^{p^k} x_i^{p^k})^{\alpha_i}\Big) \cdot h.$$

The degree of $h < p^k$, while any two monomials in the above expansion are degree separated by at least p^k in at least one variable. Consequently, by Observation 2.2, $\mathcal{S}(\ell^d) = \prod_{i=0}^{k} \binom{e_i+m-1}{m-1}$. The above inductive argument is similar to the multinomial version of Lucas's theorem.

B.5 Proof of Observation 2.5

Let ℓ be a linear form in exactly 2 variables. When $\mathrm{char}(\mathbb{F}) = 0$, then by Observation 2.4, $\mathcal{S}(\ell^d) = \binom{d+2-1}{2-1} = d + 1$. When $\mathrm{char}(\mathbb{F}) = p$, then $\mathcal{S}(\ell^d) = \prod_{i=0}^{k-1} \binom{e_i+2-1}{2-1} = \prod_{i=0}^{k-1} (e_i + 1)$, where $d = \sum_{i=0}^{k-1} e_i p^i$. It is easy to see the observation holds when $d < p$. When $d = p^k - 1 = \sum_{i=0}^{k-1} (p-1) p^i$, then $\prod_{i=0}^{k-1} (e_i + 1) = p^k = d + 1$. Finally, when ℓ is a linear form in $m \geq 2$ variables, then the observation follows from the fact that $\binom{c+m-1}{m-1} \geq c + 1$ for any $c \in \mathbb{N}$.

B.6 Proof of Observation 2.6

Using the binomial theorem,

$$h^d = \ell^d + c_0^d + \sum_{i=1}^{d-1} \binom{d}{i} c_0^i \ell^{d-i}.$$

As the degree of every monomial in ℓ^{d-i} is $d-i$, all the summands in the above expansion are degree separated. From Observations 2.2 and 2.5, it holds that $\mathcal{S}(h^d) \geq \mathcal{S}(\ell^d) + 1$. More precisely, $\mathcal{S}(h^d) \geq d+1$, as $\mathcal{S}(\ell^{d-i}) \geq 1$ and $\binom{d}{i} \neq 0$ for d as in the observation statement (by Lucas's theorem).

B.7 Proof of Claim 2.1

We prove this by induction on d. For the base case $d = 0$, it is easy to see that the sparsity of any non-zero polynomial is at least 1. Suppose now the result holds for all $k < d$. Let $\ell = \sum_{i=1}^{n} c_i x_i$ and $f = \ell^d h$. Without loss of generality, assume f is not divisible by any variable, for if it were divisible by some variable x_i, then x_i must not divide ℓ as ℓ contains at least two distinct variables and hence x_i divides h, in which case we can replace f and h by $\frac{f}{x_i}$ and $\frac{h}{x_i}$ respectively. Let x_j be a variable in ℓ with a non-zero coefficient and consider $\frac{\partial f}{\partial x_j}$. Now,

$$\mathcal{S}(f) \geq 1 + \mathcal{S}\left(\frac{\partial f}{\partial x_j}\right)$$

as the derivative map either sends monomials to distinct monomials or eliminates them, and by assumption some monomial in f is not divisible by x_j and will be eliminated. As $f = \ell^d h$,

$$\frac{\partial f}{\partial x_j} = c_j d \ell^{d-1} h + \ell^d \frac{\partial h}{\partial x_j}.$$

Clearly, ℓ^{d-1} divides $\frac{\partial f}{\partial x_j}$. By induction, $\mathcal{S}(\frac{\partial f}{\partial x_j}) \geq d$. Hence, $\mathcal{S}(f) \geq 1 + \mathcal{S}(\frac{\partial f}{\partial x_j}) \geq d+1$.

B.8 Proof of Claim 2.2

The claim is first proven for $n = 1$. Thus, $g = \ell^d$ where $\ell = \sum_{i=1}^{|\text{var}(\ell)|} c_i x_i$, $c_i \neq 0$, and $|\text{var}(\ell)| \geq \sigma$, without loss of generality. Note that

$$\frac{\partial^\sigma g}{\partial x_1 \cdots \partial x_\sigma} = \sigma! \binom{d}{\sigma} c_1 c_2 \cdots c_\sigma \ell^{d-\sigma}.$$

Clearly, $\sigma!\binom{d}{\sigma} \neq 0$ when $\text{char}(\mathbb{F}) = 0$. When $\text{char}(\mathbb{F}) = p$ with $p > d$, or $p > \sigma$ and $d = p^k - 1$ for some $k \in \mathbb{N}$, this follows by Lucas's Theorem [Luc78]. So the derivative is non-zero, as $d \geq \sigma$, implying there exists a monomial of support at least σ in g.

Now, for arbitrary n, $g = (\ell_1 \cdots \ell_n)^d$, where $|\cup_{i=1}^n \text{var}(\ell_i)| \geq \sigma$. Observe that

$$\frac{\partial^\sigma g}{\partial x_{i_1} \cdots \partial x_{i_\sigma}} = \sum_{\substack{j_1+\cdots+j_n=\sigma \\ j_i \geq 0}} c_{j_1,\cdots,j_n} \cdot \frac{\partial^\sigma (\ell_1^{j_1} \cdots \ell_n^{j_n})}{\partial x_{i_1} \cdots \partial x_{i_\sigma}} \cdot (\ell_1^{d-j_1} \cdots \ell_n^{d-j_n})$$

where $c_{j_1 \cdots j_n} = \binom{d}{j_1} \cdots \binom{d}{j_n}$. Clearly, when $\text{char}(\mathbb{F}) = 0$ or $> d$, $c_{j_1,\cdots,j_n} \neq 0$. When $\text{char}(\mathbb{F}) = p$ and $d = p^k - 1$ for some $k \in \mathbb{N}$, then by Lucas's Theorem, all the binomial coefficients are non-zero. Hence $c_{j_1,\cdots,j_n} \neq 0$. Observe that the elements of the set $\mathscr{M} := \{\ell_1^{d-j_1} \cdots \ell_n^{d-j_n} \mid j_1 + \cdots + j_n = \sigma, \;\; j_i \geq 0\}$ are linearly independent as the ℓ_i's are linearly independent and $d \geq \sigma$. Also, $\frac{\partial^\sigma(\ell_1^{j_1} \cdots \ell_n^{j_n})}{\partial x_{i_1} \cdots \partial x_{i_\sigma}} \in \mathbb{F}$ as $j_1 + \cdots + j_n = \sigma$. It suffices to show that for some choice of $j_1, \ldots, j_n$ and $x_{i_1}, \ldots, x_{i_\sigma}$,where $i_1, \ldots, i_\sigma$ are pairwise distinct, $\frac{\partial^\sigma(\ell_1^{j_1} \cdots \ell_n^{j_n})}{\partial x_{i_1} \cdots \partial x_{i_\sigma}} \neq 0$. As elements of $\mathscr{M}$ are linearly independent and $c_{j_1 \cdots j_n} \neq 0$, this would imply $\frac{\partial^\sigma g}{\partial x_{i_1} \cdots \partial x_{i_\sigma}} \neq 0$, indicating that $\mathrm{Supp}(g) \geq \sigma$.

For every $i \geq 1$, define $S_i := \mathrm{var}(\ell_i) \setminus \cup_{j=1}^{i-1} S_j$ if $| \cup_{j=1}^{i} S_j| < \sigma$ else choose $S_i \subseteq \mathrm{var}(\ell_i) \setminus \cup_{j=1}^{i-1} S_j$ such that $| \cup_{j=1}^{i} S_j| = \sigma$; here, $\cup_{j=1}^{i-1} S_j = \emptyset$ for $i = 1$. Note that such a collection of sets always exists as $| \cup_{i=1}^{n} \mathrm{var}(\ell_i)| \geq \sigma$. Say we choose $m \leq n$ such non-empty sets. Let $j_i := |S_i|$ and $S_i := \{x_{i1}, \ldots, x_{ij_i}\}$. Hence,

$$\frac{\partial^\sigma(\ell_1^{j_1} \cdots \ell_m^{j_m})}{(\partial x_{11} \cdots \partial x_{1j_1}) \cdots (\partial x_{m1} \cdots \partial x_{mj_m})} = \prod_{i=1}^{m} \frac{\partial^{j_i} \ell_i^{j_i}}{\partial x_{i1} \cdots \partial x_{ij_i}}.$$

As $\{x_{i1}, \ldots, x_{ij_i}\} \subseteq \mathrm{var}(\ell_i)$ is not empty and $j_i \leq \sigma < p$ (in case of finite characteristic fields), by the analysis of the $n = 1$ case, $\frac{\partial^{j_i} \ell_i^{j_i}}{\partial x_{i1} \cdots \partial x_{ij_i}} \neq 0$. Hence,

$$\frac{\partial^\sigma(\ell_1^{j_1} \cdots \ell_m^{j_m})}{(\partial x_{11} \cdots \partial x_{1j_1}) \cdots (\partial x_{m1} \cdots \partial x_{mj_m})} \neq 0$$

and $\mathrm{Supp}(g) \geq \sigma$.

C Missing proofs from Section 3

C.1 Proof of Observation 3.1

Let $i \in [n]$. Note that $Q_{i,1}(\mathbf{z})$ has degree $(3i-2)d_1 + d_2$, $Q_{i,2}(\mathbf{z})$ has degree $(3i-1)d_1 + d_3$ and $Q_{i,3}(\mathbf{z})$ has degree $3id_1 + d_3$. Clearly $3id_1 + d_3 > (3i-1)d_1 + d_3$. Also, $(3i-1)d_1 + d_3 > (3i-2)d_1 + d_2$ because $d_1 > d_2 > d_2 - d_3$ by the conditions in (2). Thus, $Q_{i,1}(\mathbf{z})$, $Q_{i,2}(\mathbf{z})$ and $Q_{i,3}(\mathbf{z})$ are degree separated and Q_i is a sum of 3 degree separated polynomials and has degree $3id_1 + d_3$.

Now, let $i \in [n]$ and $k \in [m]$. $R_k(\mathbf{z})$ is a polynomial of degree $(3n+k)d_1 + 3d_4$ while the degree of $Q_i(\mathbf{z})$ is $3id_1 + d_3$. Note that

$$(3n+k)d_1 + 3d_4 \geq (3n+1)d_1 + 3d_4 > 3nd_1 + d_3 \geq 3id_1 + d_3,$$

where the second inequality holds because $d_1 > d_3$ by the constraints in (2). Therefore, $R_k(\mathbf{z})$ and $Q_i(\mathbf{z})$ (hence also $Q_{i,1}$, $Q_{i,2}$ and $Q_{i,3}$) are degree separated from one another. Further, the degree of $R_k(\mathbf{z})$ and that of $Q_i(\mathbf{z})$ are greater than d_1 implying $x_0^{d_1}$ is degree separated from $R_k(\mathbf{z})$ and $Q_i(\mathbf{z})$.

Lastly, let $i, j \in [n]$ where $i < j$ without loss of generality. The highest degree of a monomial in $Q_i(\mathbf{z})$ is $3id_1 + d_3$, while the lowest degree of a monomial in $Q_j(\mathbf{z})$ is $(3j-2)d_1 + d_2$. Now,

$$(3j-2)d_1 + d_2 \geq (3i+1)d_1 + d_2 > 3id_1 + d_3$$

as $j \geq i+1$ and $d_1 > d_2 > d_3$ by the conditions in (2). Thus, Q_i and Q_j are degree separated. That $R_k(\mathbf{z})$ is degree separated from $R_l(\mathbf{z})$ for $k, l \in [n]$ and $k \neq l$ can be observed from the fact that the degree of $R_k(\mathbf{z})$ is $(3n+k)d_1 + 3d_4$.

C.2 Proof of Observation 3.2

From the definition of f in (1), it follows that the degree of f is the maximum of that of $x_0^{d_1}$, Q_i and R_k , where $i \in [n]$ and $k \in [m]$. As observed in the proof of Observation 3.1, degree of Q_i is $3id_1 + d_3$, degree of R_k is $(3n + k)d_1 + 3d_4$ and $(3n + k)d_1 + 3d_4 > 3id_1 + d_3 > d_1$. Since $k \leq m$, therefore the highest degree is $(3n + m)d_1 + 3d_4$. So, the degree of f is $(3n + m)d_1 + 3d_4$.

C.3 Proof of Observation 3.3

By Observation 3.1, f is a sum of the $n + m + 1$ degree separated polynomials $x_0^{d_1}$, Q_i and R_k, where $i \in [n]$ and $k \in [m]$. Applying Observation 2.5 (for linear forms in two variables over $\text{char}(\mathbb{F}) = 0$ fields) to $Q_{i,2}$ and $Q_{i,3}$ and Observation 2.2 to Q_i, we get

$$\mathcal{S}(Q_i(\mathbf{z})) = \mathcal{S}(Q_{i,1}(\mathbf{z})) + \mathcal{S}(Q_{i,2}(\mathbf{z})) + \mathcal{S}(Q_{i,3}(\mathbf{z})) = 2d_3 + 3, \ \forall i \in [n].$$

By Observation 2.5 (for linear forms in two variables over $\text{char}(\mathbb{F}) = 0$ fields) and the assumption that each clause in ψ has 3 distinct variables, we get that:

$$\mathcal{S}(R_k(\mathbf{z})) = \mathcal{S}(x_0^{(3n+k)d_1}) \prod_{j \in C_k} \mathcal{S}((y_j + (-1)^{a_{k,j}} x_j)^{d_4}) = (d_4 + 1)^3 \ \ \forall k \in [m].$$

Finally, applying Observations 3.1 and 2.2 to f gives

$$\mathcal{S}(f(\mathbf{z})) = \mathcal{S}(x_0^{d_1}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(\mathbf{z})) = 1 + n(2d_3 + 3) + m(d_4 + 1)^3.$$

Thus, $\mathcal{S}(f(\mathbf{z})) > s$ but also $(mn)^{O(1)}$.

For the support of f, note that $\text{Supp}(R_k) = 7$ for all $k \in [m]$ while $\text{Supp}(Q_i) = 3$ for all $i \in [n]$ and $\text{Supp}(x_0^{d_1}) = 1$. By Observation 3.1, $\text{Supp}(f) = \text{Supp}(R_k) = 7$.

C.4 Proof of Lemma 3.1

If $A(x_0)$ is a linear form in at least two variables, then it follows from the definition of f in (1), Observation 3.1, Observation 2.5 applied on $A(x_0^{d_1})$, Observation 2.3, and the constraint $d_1 \geq s$ in (2) that $\mathcal{S}(f(A\mathbf{z}) > \mathcal{S}(A(x_0^{d_1})) \geq d_1 + 1 > s$, a contradiction. Hence, $A(x_0)$ has only one variable. By multiplying A with a permutation and a scaling matrix, we can assume without loss of generality that $A(x_0) = x_0$. This can be assumed because permutation and non-zero scaling of variables do not affect the sparsity of a polynomial.

C.5 Proof of Lemma 3.2

It follows from Observations 3.1 and 2.3 that for any $A \in \text{GL}(|\mathbf{z}|, \mathbb{F})$, $\mathcal{S}(Q_i(A\mathbf{z})) = \mathcal{S}(Q_{i,1}(A\mathbf{z})) + \mathcal{S}(Q_{i,2}(A\mathbf{z})) + \mathcal{S}(Q_{i,3}(A\mathbf{z}))$, where Q_i is as described in Section 3.1. Now, the if direction in the lemma statement is easy to verify. For the only if direction, consider the following cases of A:

1. $\mathcal{S}(A(x_i)) \geq 2$: It follows from Observation 2.5 and $d_2 \geq 2d_3$ that $\mathcal{S}(Q_{i,1}(A\mathbf{z})) \geq d_2 + 1 \geq 2d_3 + 1$. Also, $\mathcal{S}(Q_{i,2}(A\mathbf{z})) \geq 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z})) \geq 1$. Thus, $\mathcal{S}(Q_i(A\mathbf{z})) \geq d_2 + 3 \geq 2d_3 + 3$.

2. $\mathcal{S}(A(x_i)) = 1$, $\mathcal{S}(A(y_i + x_i)) \geq 2$ and $\mathcal{S}(A(y_i - x_i)) \geq 2$: It follows from Observation 2.5 that $\mathcal{S}(Q_{i,2}(A\mathbf{z})) \geq d_3 + 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z})) \geq d_3 + 1$ implying $\mathcal{S}(Q_i(A\mathbf{z})) \geq 2d_3 + 3$.

3. $\mathcal{S}(A(x_i)) = 1$ with $\mathcal{S}(A(y_i + x_i)) = 1$ or $\mathcal{S}(A(y_i - x_i)) = 1$: Let $A(x_i) = X_i$ for some scaled variable $X_i \in \mathbf{z}$. Because A is invertible exactly one of $\mathcal{S}(A(y_i + x_i)) = 1$ or $\mathcal{S}(A(y_i - x_i)) = 1$ holds true. Let $\mathcal{S}(A(y_i + x_i)) = 1$, without loss of generality. Then, it must be that $A(y_i) = Y_i - X_i$ for some scaled variable $Y_i \in \mathbf{z}$. Thus, $A(y_i - x_i) = Y_i - 2X_i$. Hence, $\mathcal{S}(Q_{i,1}(A\mathbf{z})) = 1$, $\mathcal{S}(Q_{i,2}(A\mathbf{z})) = 1$ and $\mathcal{S}(Q_{i,3}(A\mathbf{z})) = d_3 + 1$ (by Observation 2.5) implying $\mathcal{S}(Q_i(A\mathbf{z})) = d_3 + 3$.

The first two cases show that if A is not as per the lemma statement, then $\mathcal{S}(Q_i(A\mathbf{z})) \geq 2d_3 + 3$; otherwise, $\mathcal{S}(Q_i(A\mathbf{z})) = d_3 + 3$.

C.6 Proof of Lemma 3.3

Suppose $\mathcal{S}(Q_j(A\mathbf{z})) \neq d_3 + 3$ for some $j \in [n]$. Then, $\mathcal{S}(Q_j(A\mathbf{z})) \geq 2d_3 + 3$ by Lemma 3.2. By the definition of f in (1), Observations 3.1 and 2.3 and the condition $d_3 \geq m(d_4 + 1)^2 + 1$, we get the following contradiction:

$$\begin{aligned}\mathcal{S}(f(A\mathbf{z})) > \mathcal{S}(A(x_0^{d_1})) + \sum_{i=1, i\neq j}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \mathcal{S}(Q_j(A\mathbf{z})) &\geq 1 + (n-1)(3 + d_3) + (3 + 2d_3)\\ = 1 + n(3 + d_3) + d_3 = s - m(d_4 + 1)^2 + d_3 > s.&\end{aligned}$$

C.7 Proof of Observation 3.4

The observation follows from the fact that the x_0-degree of the summands in f, as defined in (7), form an arithmetic progression with common difference $d_3 + 1$ and hence every polynomial in the observation statement has distinct x_0-degree.

C.8 Proof of Observation 3.5

By Observation 3.4, f is a sum of the $n + m + 1$ polynomials $x_0^{d_1} y_0^{d_2 + (3n+m+1)(d_3+1)}$, Q_i and R_k, where $i \in [n]$ and $k \in [m]$, which are degree separated with respect to x_0. Using arguments similar to the proof of Observation 3.3, it holds that

$$\mathcal{S}(Q_i(\mathbf{z})) = \mathcal{S}(Q_{i,1}(\mathbf{z})) + \mathcal{S}(Q_{i,2}(\mathbf{z})) + \mathcal{S}(Q_{i,3}(\mathbf{z})) = 2d_4 + 3 \;\; \forall i \in [n],$$

$$\mathcal{S}(R_k(\mathbf{z})) = \mathcal{S}(x_0^{d_1 + (3n+k)(d_3+1)} y_0^{d_2 + (m-k+1)(d_3+1) - 3d_5}) \prod_{j \in C_k} \mathcal{S}((y_j + (-1)^{a_{k,j}} x_j)^{d_5}) = (d_5 + 1)^3 \;\; \forall k \in [m],$$

and

$$\mathcal{S}(f(\mathbf{z})) = \mathcal{S}(x_0^{d_1} y_0^{d_2 + (3n+m+1)(d_3+1)}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(\mathbf{z})) = 1 + n(2d_4 + 3) + m(d_5 + 1)^3.$$

Thus, $\mathcal{S}(f(\mathbf{z})) > s$ but also $(mn)^{O(1)}$. For the support of f, note that $\mathrm{Supp}(R_k) = 8$ for all $k \in [m]$ while $\mathrm{Supp}(Q_i) = 4$ for all $i \in [n]$ and $\mathrm{Supp}(x_0^{d_1} y_0^{d_2 + (3n+m+1)(d_3+1)}) = 2$. Hence, by Observation 3.4, $\mathrm{Supp}(f) = \mathrm{Supp}(R_k) = 8$.

C.9 Proof of Lemma 3.4

Suppose one of $A(x_0)$ or $A(y_0)$ is a linear form in at least two variables. As $A(x_0)^{d_1}$ and $A(y_0)^{d_2}$ divide $f(A\mathbf{z})$, Claim 2.1 and the conditions of (5) imply $\mathcal{S}(f(A\mathbf{z})) > s$, a contradiction. So, $A(x_0)$ and $A(y_0)$ must have only one variable each. Hence, without loss of generality (i.e., after applying scaling and permutation to A), $A(x_0) = x_0$ and $A(y_0) = y_0$.

C.10 Proof of Lemma 3.5

Note that each of the $3n+m+1$ summand polynomials, as mentioned in the lemma statement, is of form $x_0^{d_1+t(d_3+1)} \cdot y_0^{d_2+v} \cdot h(\mathbf{z})$, where $t \in [0, 3n+m]$ and $v \in \mathbb{N}$. By the construction of f, each t corresponds to a unique summand polynomial. For $x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)}$ the degree of h is 0. Let $i \in [n]$. For $Q_{i,1}(A\mathbf{z})$, the degree of h is d_3 while for $Q_{i,2}(A\mathbf{z})$ and $Q_{i,3}(A\mathbf{z})$ the degree of h is d_4. Lastly, for $R_k(A\mathbf{z})$, where $k \in [m]$, the degree of h is $3d_5$. Thus, going over all the summand polynomials, the degree of h is at most the maximum of d_3, d_4 and $3d_5$. Since $d_3 > d_4$ and $d_3 > 3d_5$ by the conditions in (5), therefore the degree of h is at most d_3. As shown in Lemma 3.4, $A(x_0) = x_0$ and $A(y_0) = y_0$ while there may be monomials in $h(\mathbf{z})$ which have non-zero x_0-degree. Thus, the possible range of x_0-degree of any monomial of a summand polynomial lies in the range $[d_1 + t(d_3+1), d_1 + t(d_3+1) + d_3]$, which is clearly disjoint for distinct t. Therefore, the summand polynomials are degree separated with respect to x_0.

C.11 Proof of Observation 3.6

By Observation 3.1, f is a sum of the $n+m+1$ degree separated polynomials $x_0^{d_1}$, Q_i and R_k, where $i \in [n]$ and $k \in [m]$. Applying Observation 2.5 (for linear forms in two variables over finite characteristic fields) to $Q_{i,2}$ and Observation 2.2 to Q_i, we get

$$\mathcal{S}(Q_i(\mathbf{z})) = \mathcal{S}(Q_{i,1}(\mathbf{z})) + \mathcal{S}(Q_{i,2}(\mathbf{z})) + \mathcal{S}(Q_{i,3}(\mathbf{z})) = d_3 + 3, \ \ \forall i \in [n].$$

By Observation 2.5 (for linear forms in two variables over finite characteristic fields) and the assumption that each clause in ψ has 3 distinct variables, we get that:

$$\mathcal{S}(R_k(\mathbf{z})) = \mathcal{S}(x_0^{(3n+k)d_1}) \prod_{j \in C_k} \mathcal{S}((y_j + a_{k,j}x_j)^{d_4}) \leq (d_4+1)^3 \ \ \forall k \in [m].$$

depending on $a_{k,j} = 0$ or 1. Finally, applying Observations 3.1 and 2.2 to f gives

$$\mathcal{S}(f(\mathbf{z})) = \mathcal{S}(x_0^{d_1}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(\mathbf{z})) \leq 1 + n(d_3+3) + m(d_4+1)^3.$$

For the support of f, note that $4 \leq \mathrm{Supp}(R_k) \leq 7$ for $k \in [m]$, $\mathrm{Supp}(Q_i) = 3$ for all $i \in [n]$ and $\mathrm{Supp}(x_0^{d_1}) = 1$. By Observation 3.1, $4 \leq \mathrm{Supp}(f) \leq 7$.

C.12 Proof of Lemma 3.4 over finite characteristic fields

Note that as f is divisible by $x_0^{d_1}$ and $y_0^{d_2}$, therefore for any $A \in \mathrm{GL}(|\mathbf{z}|, \mathbb{F})$ we can write:

$$f(A\mathbf{z}) = A(x_0^{d_1}) A(y_0^{d_2}) g(A\mathbf{z}).$$

where the degree of $g(A\mathbf{z})$ is $(3n+m+1)(d_3+1)$. Let $A(x_0) = \sum_{l=1}^{|\mathbf{z}|} c_l z_l$, where $z_l \in \mathbf{z}$ and $c_l \in \mathbb{F}$. The characteristic being finite and the choice of d_1 as in (15) implies:

$$A(x_0^{d_1}) = \Big(\sum_{l=1}^{|\mathbf{z}|} c_l z_l\Big)^{\sum_{t=k_3}^{k_1+k_3-1}(p-1)p^t} = \prod_{t=k_3}^{k_1+k_3-1} \Big(\sum_{l=1}^{|\mathbf{z}|} c_l^{p^t} z_l^{p^t}\Big)^{(p-1)}.$$

Thus, the monomials of $A(x_0^{d_1})$ are of form $\prod_{l=1}^{|\mathbf{z}|} z_l^{e_l}$, where $e_l = \sum_{t=k_3}^{k_1+k_3-1} c_{l,t} p^t$ with $c_{l,t} \in [0, p-1]$, and $\sum_{l=1}^{|\mathbf{z}|} e_l = d_1$. Now, for any two monomials of $A(x_0^{d_1})$, there exists a variable

$z_l \in \mathbf{z}$ such that the difference between the z_l-degree of these two monomials is at least $p^{k_3} > d_2 + (3n+m+1)(d_3+1)$, while the degree of $A(y_0^{d_2})g(A\mathbf{z})$ is $d_2 + (3n+m+1)(d_3+1)$. This is because of the way d_1 is set in (15). Thus, with this observation and Observation 2.2, it holds that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1}))\mathcal{S}(A(y_0^{d_2})g(A\mathbf{z})).$$

Similarly, because the characteristic is finite and d_2 is as chosen in (14), for any two monomials of $A(y_0^{d_2})$ there exists a variable $z_l \in \mathbf{z}$ such that the difference between the z_l-degree of these two monomials is at least $p^{k_2} > (3n+m+1)(d_3+1)$, while the degree of $g(A\mathbf{z})$ is $(3n+m+1)(d_3+1)$. Thus, with this observation and Observation 2.2, it holds that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1}))\mathcal{S}(A(y_0^{d_2})g(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1}))\mathcal{S}(A(y_0^{d_2}))\mathcal{S}(g(A\mathbf{z})).$$

Now, suppose $A(x_0)$ is a linear form in at least 2 variables. By Observation 2.4 and the definition of d_1 in (15), it follows that

$$\mathcal{S}(A(x_0^{d_1})) \geq \prod_{k_3}^{k_1+k_3-1} \binom{p-1+2-1}{2-1} = p^{k_1} > s.$$

This implies $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(x_0^{d_1})) > s$. Thus, $A(x_0)$ must be some scaled variable in $\mathbf{z}$. Therefore,

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(y_0^{d_2})g(A\mathbf{z})).$$

Similarly, if $\mathcal{S}(A(y_0))$ is a linear form in at least 2 variables, then by Observation 2.4 and the definition of d_2 in (14)

$$\mathcal{S}(A(y_0^{d_2})) \geq \prod_{k_2}^{k_1+k_2-1} \binom{p-1+2-1}{2-1} = p^{k_1} > s.$$

This implies $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(y_0^{d_2})) > s$. Thus, $A(y_0)$ must also be some scaled variable in $\mathbf{z}$. Therefore, $A(x_0) = x_0$ and $A(y_0) = y_0$ without loss of generality by applying an appropriate permutation and scaling transform.

C.13 Proof of Observation 3.7

By Observation 3.4, f is a sum of the $n+m+1$ polynomials $x_0^{d_1}y_0^{d_2+(3n+m+1)(d_3+1)}$, Q_i and R_k, where $i \in [n]$ and $k \in [m]$, which are degree separated with respect to x_0. Applying Observation 2.5 (for linear forms in two variables over finite characteristic fields) to $Q_{i,2}$ and Observation 2.2 to Q_i, we get

$$\mathcal{S}(Q_i(\mathbf{z})) = \mathcal{S}(Q_{i,1}(\mathbf{z})) + \mathcal{S}(Q_{i,2}(\mathbf{z})) + \mathcal{S}(Q_{i,3}(\mathbf{z})) = d_4 + 3, \;\; \forall i \in [n].$$

By Observation 2.5 (for linear forms in two variables over finite characteristic fields) and the assumption that each clause in ψ has 3 distinct variables, we get that:

$$\mathcal{S}(R_k(\mathbf{z})) = \mathcal{S}(x_0^{d_1+(3n+k)(d_3+1)})\mathcal{S}(y_0^{d_2+(m-k+1)(d_3+1)-3d_5}) \prod_{j \in C_k} \mathcal{S}((y_j + a_{k,j}x_j)^{d_5}) \leq (d_5+1)^3 \;\; \forall k \in [m].$$

depending on whether $a_{k,j}$ is 0 or 1. Finally, applying Observations 3.4 and 2.2 to f gives

$$\mathcal{S}(f(\mathbf{z})) = \mathcal{S}(x_0^{d_1}y_0^{d_2+(3n+m+1)(d_3+1)}) + \sum_{i=1}^{n} \mathcal{S}(Q_i(\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(\mathbf{z})) \leq 1 + n(d_4+3) + m(d_5+1)^3.$$

For the support of f, note that $5 \leq \mathrm{Supp}(R_k) \leq 8$ for $k \in [m]$, $\mathrm{Supp}(Q_i) = 4$ for all $i \in [n]$ and $\mathrm{Supp}(x_0^{d_1}y_0^{d_2+(3n+m+1)(d_3+1)}) = 2$. By Observation 3.4, $5 \leq \mathrm{Supp}(f) \leq 8$.

D Missing proofs from Section 4

D.1 Proof of Lemma 4.1

By Observations 3.1, 2.3 and the definition of f as in (1), it follows that:

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

We now analyse $\mathcal{S}(f(A\mathbf{z}))$ under the transforms listed in the lemma statement. The list covers all possible types of transforms.

1. By Observation 2.5, $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(x_0^{d_1})) \geq d_1 + 1$ follows.

2. In this case, $A(x_0) = x_0$ without loss of generality, as permutation and non-zero scaling of variables do not influence the sparsity of the polynomial. By Observation 2.5, it follows that $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,1}(A\mathbf{z})) \geq d_2 + 1$.

3. In this case, $A(x_i)$, where $i \in [0, n]$ is some scaled variable in $\mathbf{z}$. Without loss of generality, let $A(y_j + x_j)$ be a linear form in at least 3 variables. By Observation 2.4, $\mathcal{S}(Q_{j,2}(A\mathbf{z})) \geq \binom{d_3+2}{2}$ holds. Therefore,

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_j(A\mathbf{z})) \geq \mathcal{S}(Q_{j,2}(A\mathbf{z})) \geq \binom{d_3+2}{2} = \frac{d_3^2 + 3d_3 + 2}{2}.$$

4. In this case, $A(x_i) = X_i$, where $i \in [0, n]$ and $X_i \in \mathbf{z}$ is some scaled variable. Also, $A(y_i + x_i)$ and $A(y_i - x_i)$ are linear forms in at most two variables for all $i \in [n]$. Thus, $A(y_i) = Y_i + c_i X_i$, where $c_i \in \mathbb{F}$ and $Y_i \in \mathbf{z}$ is some scaled variable. As A is invertible, the Y_i's and X_i's are distinct variables. Hence,

$$\mathcal{S}(R_k(A\mathbf{z})) = \mathcal{S}(X_0^{(3n+k)d_1}) \prod_{j \in C_k} \mathcal{S}((Y_j + (c_j + (-1)^{a_{k,j}})X_j)^{d_4}).$$

   Since ψ is unsatisfiable, for any such A, there exists $k \in [m]$ such that $\mathcal{S}(R_k(A\mathbf{z})) \geq (d_4 + 1)^3$ (by Observation 2.5). Therefore, $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(R_k(A\mathbf{z})) \geq (d_4 + 1)^3$.

D.2 Proof of Lemma 4.2

Like in the proof of Lemma 4.1, we analyse $\mathcal{S}(f(A\mathbf{z}))$ with A as listed in the lemma statement. Suppose $A(x_0)$ is a linear form in at least 2 variables. Since $x_0^{d_1}$ divides f, it follows from Claim 2.1 that $\mathcal{S}(f(A\mathbf{z})) \geq d_1 + 1$. Similarly, if $A(x_0)$ is a variable while $A(y_0)$ is a linear form in at least 2 variables, then since $y_0^{d_2}$ divides f, $\mathcal{S}(f(A\mathbf{z})) \geq d_2 + 1$ holds by Claim 2.1. For the remaining cases, $A(x_0) = x_0$ and $A(y_0) = y_0$ without loss of generality. It then follows from Lemma 3.5 that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

The last three cases then can be proved the same way as the last three cases of Lemma 4.1 in the non-homogeneous case.

D.3 Proof of Lemma 4.3

By Observations 3.1, 2.3 and the definition of f as in (1), it follows that:

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

We now analyse $\mathcal{S}(f(A\mathbf{z}))$ with A as listed in the lemma statement. The analysis for the first, second and fourth cases is similar to that of the respective cases in the proof of Lemma 4.1. In the third case, without loss of generality, let $A(y_j + x_j)$, for some $j \in [n]$, be a linear form in at least 3 variables. By Observation 2.4 and the fact that $d_3 = p^k - 1 = \sum_{i=0}^{k-1}(p-1)p^i$ for some $k \in \mathbb{N}$,

$$\mathcal{S}(A((y_j+x_j)^{d_3})) \geq \prod_{i=0}^{k-1} \binom{p-1+3-1}{3-1} = \left(\frac{p(p+1)}{2}\right)^k = (d_3+1)\left(\frac{p+1}{2}\right)^k = (d_3+1)^{1+\log_p(\frac{p+1}{2})}$$

Note, $\log_p(\frac{p+1}{2})$ is an increasing function for $p \geq 3$. Hence, $\log_p(\frac{p+1}{2}) \geq \log_3(4/2) \geq 0.63$. Therefore,

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,2}(A\mathbf{z})) \geq (d_3+1)^{1.63}.$$

Similarly, if $A(y_j - x_j)$ is a linear form in at least 3 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,3}(A\mathbf{z})) \geq (d_3+1)^{1.63}$.

D.4 Proof of Lemma 4.4

The proof of this lemma is very similar to that of Lemma 4.3. In particular, the analysis for the first, second and fourth cases is similar to that of the respective cases in the proof of Lemma 4.3. So, we consider the third case. Without loss of generality, let $A(y_j + x_j)$ be a linear form in at least 3 variables for some $j \in [n]$. Then, by Observation 2.4 and the fact that $d_3 = 2^k - 1 = \sum_{i=0}^{k-1} 2^i$ for some $k \in \mathbb{N}$,

$$\mathcal{S}(A((y_j+x_j)^{d_3})) \geq \prod_{i=0}^{k-1} \binom{1+3-1}{3-1} = 3^k = (d_3+1)^{\log_2 3} \geq (d_3+1)^{1.58}.$$

Therefore

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,2}(A\mathbf{z})) \geq (d_3+1)^{1.58}.$$

Similarly, if $A(y_j)$ is a linear form in at least 3 variables, then $\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(Q_{j,3}(A\mathbf{z})) \geq (d_3+1)^{1.58}$.

D.5 Proof of Lemma 4.5

Suppose $A(x_0)$ is a linear form in at least two variables. Applying the argument in the proof of Lemma 3.5 for the finite characteristic case (refer Section C.12) shows that

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(x_0)^{d_1}) \geq p^{k_1+1} \geq d_3+2.$$

Similarly, if $A(x_0)$ is a variable and $A(y_0)$ is a linear form in at least two variables, the same argument shows that

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(y_0)^{d_2}) \geq p^{k_1+1} \geq d_3+2.$$

For the remaining cases, $A(x_0) = x_0$ and $A(y_0) = y_0$ without loss of generality. It then follows from Lemma 3.5 that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

The last three cases can then be proved in the same way as the last three cases of Lemma 4.3.

D.6 Proof of Lemma 4.6

Suppose $A(x_0)$ is a linear form in at least two variables. Applying the argument in the proof of Lemma 3.5 for the finite characteristic case (refer Section C.12) shows that

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(x_0)^{d_1}) \geq 2^{k_1+1} \geq d_3 + 2.$$

Similarly, if $A(x_0)$ is a variable and $A(y_0)$ is a linear form in at least two variables, the same argument shows that

$$\mathcal{S}(f(A\mathbf{z})) \geq \mathcal{S}(A(y_0)^{d_2}) \geq 2^{k_1+1} \geq d_3 + 2.$$

For the remaining cases, $A(x_0) = x_0$ and $A(y_0) = y_0$ without loss of generality. It then follows from Lemma 3.5 that

$$\mathcal{S}(f(A\mathbf{z})) = \mathcal{S}(A(x_0^{d_1} y_0^{d_2+(3n+m+1)(d_3+1)})) + \sum_{i=1}^{n} \mathcal{S}(Q_i(A\mathbf{z})) + \sum_{k=1}^{m} \mathcal{S}(R_k(A\mathbf{z})).$$

The last three cases can then be proved in the same way as the last three cases of Lemma 4.4.

E Missing proofs from Section 5

E.1 Proof of Observation 5.1

As the polynomials are degree separated, by Observation 2.2,

$$\mathcal{S}(f(\mathbf{w})) = \sum_{g(\mathbf{w}) \in P} \mathcal{S}(g(\mathbf{w})) + \sum_{h(\mathbf{w}) \in Q} \mathcal{S}(h(\mathbf{w})) + \sum_{k=1}^{m} \mathcal{S}(R_k(\mathbf{w})).$$

As $g(\mathbf{w}) \in P$ and $h(\mathbf{w}) \in Q$ are monomials, thus $\mathcal{S}(g(\mathbf{w})) = 1$ and $\mathcal{S}(h(\mathbf{w})) = 1$. For $k \in [m]$,

$$\mathcal{S}(R_k(\mathbf{w})) = (\prod_{j \in C_k} \mathcal{S}(y_j - a_{k,j}x_j)^2)\mathcal{S}((z_1 \cdots z_{\sigma-5})^\star) \leq 27$$

because all the polynomials in the product are variable disjoint and $a_{k,j}$ can be non-zero for all $j \in C_k$. Lastly, as $|P| = \binom{n+\sigma-5}{\sigma}$ and $|Q| = \binom{n}{\sigma/2}$ (for odd σ, $|Q| = \binom{n}{\frac{\sigma+1}{2}}\frac{\sigma+1}{2}$), hence

$$\mathcal{S}(f) \leq \binom{n+\sigma-5}{\sigma} + \binom{n}{\sigma/2} + 27m.$$

Thus, $\mathcal{S}(f) = O(n^\sigma + m)$ as σ is a constant. Note, $\text{Supp}(R_1) = \sigma + 1$ because the characteristic is not 2 and $a_{1,j} = 1$, for all $j \in C_1$ (due to the first clause containing only complemented variables). Also, $\text{Supp}(g(\mathbf{w})) = \sigma$, $\text{Supp}(h(\mathbf{w})) = \sigma$ and $\text{Supp}(R_k(\mathbf{w})) \leq \sigma + 1$ where $g(\mathbf{w}) \in P$, $h(\mathbf{w}) \in Q$ and $k \in [2, m]$. Hence, $\text{Supp}(f(\mathbf{w})) = \text{Supp}(R_1(\mathbf{w})) = \sigma + 1$ as all the polynomials are degree separated.

E.2 Proof of Lemma 5.1

Let $g(\mathbf{w}) = (w_1 \cdots w_\sigma)^\star \in P$, where $w_j \in \mathbf{z} \sqcup \mathbf{x}$ and $\star$ represents an integer power which is at least $\sigma + 1$. Now, $g(A\mathbf{w}) = (\ell_1 \cdots \ell_\sigma)^\star$, where the ℓ_j's, with $j \in [\sigma]$, are linearly independent linear forms. If $| \cup_{j=1}^{\sigma} \mathrm{var}(\ell_j)| \geq \sigma + 1$, then by Claim 2.2, $\mathrm{Supp}(g(A\mathbf{w})) \geq \sigma + 1$ a contradiction. Thus, $| \cup_{j=1}^{\sigma} \mathrm{var}(\ell_j)| \leq \sigma$. The linear independence of these σ many linear forms implies $| \cup_{j=1}^{\sigma} \mathrm{var}(\ell_j)| \geq \sigma$. Combining these inequalities gives $| \cup_{j=1}^{\sigma} \mathrm{var}(\ell_j)| = \sigma$. Thus, there exist variables $W_1, \ldots, W_\sigma \in \mathbf{w}$ such that $\langle \{\ell_1, \ldots, \ell_\sigma\} \rangle = \langle \{W_1, \ldots, W_\sigma\} \rangle$, where $\langle S \rangle$ denotes the vector space spanned by the elements of a set S of polynomials. In particular, $\langle \{\ell_1, \ldots, \ell_\sigma\} \rangle = \langle \mathrm{var}(g(A\mathbf{w})) \rangle$.

Consider the variable $x_1 \in \mathbf{x} \sqcup \mathbf{z}$ and let $A(x_1) = \ell_1$. Now, P can be seen as a collection of σ sized subsets of $\mathbf{z} \sqcup \mathbf{x}$. Since $n + \sigma - 6 \geq 2(\sigma - 1)$ (as implied by $n \geq \sigma + 4$), there exist, without loss of generality, $g_1(\mathbf{w})$ and $g_2(\mathbf{w}) \in P$, such that $\mathrm{var}(g_1(\mathbf{w})) \cap \mathrm{var}(g_2(\mathbf{w})) = \{x_1\}$. Note that $\langle \mathrm{var}(g_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(g_2(\mathbf{w})) \rangle = \langle \mathrm{var}(g_1(\mathbf{w})) \cap \mathrm{var}(g_2(\mathbf{w})) \rangle$. Thus,

$$\dim \langle \mathrm{var}(g_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(g_2(\mathbf{w})) \rangle = \dim \langle \mathrm{var}(g_1(\mathbf{w})) \cap \mathrm{var}(g_2(\mathbf{w})) \rangle = 1.$$

The invertibility of A and the argument of the first paragraph imply there exist σ-size sets $B_1, B_2 \subseteq \mathbf{w}$, with $B_1 = \mathrm{var}(g_1(A\mathbf{w}))$ and $B_2 = \mathrm{var}(g_2(A\mathbf{w}))$, such that $\langle \mathrm{var}(g_1(\mathbf{w})) \rangle$ and $\langle B_1 \rangle$ are isomorphic and so are $\langle \mathrm{var}(g_2(\mathbf{w})) \rangle$ and $\langle B_2 \rangle$. Similarly, under A, $\langle \mathrm{var}(g_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(g_2(\mathbf{w})) \rangle$ is isomorphic to $\langle B_1 \rangle \cap \langle B_2 \rangle$. Thus, $\dim \langle B_1 \rangle \cap \langle B_2 \rangle = 1$. From the first paragraph, it is also evident that $\ell_1 \in \langle B_1 \rangle \cap \langle B_2 \rangle = \langle B_1 \cap B_2 \rangle$, where the equality follows as B_1 and B_2 are sets of variables. This implies $A(x_1) = \ell_1 = W_1$, where W_1 is some scaled variable in $\mathbf{w}$.

E.3 Proof of Lemma 5.2

The proof is similar to that of Lemma 5.1. As $\mathrm{Supp}(f(A\mathbf{w})) \leq \sigma$, Lemma 5.1 holds. Thus, for all $w \in \mathbf{x} \sqcup \mathbf{z}$, $A(w) = W$ for some scaled variable $W \in \mathbf{w}$. In particular, $A(x_i) = X_i$ where $i \in [n]$ and $X_i \in \mathbf{w}$ is some scaled variable. Let $A(y_i) = \ell_i$, where ℓ_i contains at least one variable other than those in $\mathrm{var}(A(w))$, for any $w \in \mathbf{x} \sqcup \mathbf{z}$. Without loss of generality, consider $h(\mathbf{w}) = ((x_1 y_1) \cdots (x_{\frac{\sigma}{2}} y_{\frac{\sigma}{2}}))^\star \in Q$, where $\star$ represents an integer power which is at least $\sigma + 1$. Now, $h(A\mathbf{w}) = ((X_1 \ell_1) \cdots (X_{\frac{\sigma}{2}} \ell_{\frac{\sigma}{2}}))^\star$, where $X_1 \ldots X_{\frac{\sigma}{2}}, \ell_1, \ldots \ell_{\frac{\sigma}{2}}$ are linearly independent linear forms. If $|\mathrm{var}(\ell_1) \cup \cdots \cup \mathrm{var}(\ell_{\frac{\sigma}{2}}) \cup \{X_1, \ldots, X_{\frac{\sigma}{2}}\}| \geq \sigma + 1$, then by Claim 2.2, $\mathrm{Supp}(h(A\mathbf{w})) \geq \sigma + 1$, a contradiction. Thus, $|\mathrm{var}(\ell_1) \cup \cdots \cup \mathrm{var}(\ell_{\frac{\sigma}{2}}) \cup \{X_1, \ldots, X_{\frac{\sigma}{2}}\}| \leq \sigma$. The linear independence of these σ many linear forms implies $|\mathrm{var}(\ell_1) \cup \cdots \cup \mathrm{var}(\ell_{\frac{\sigma}{2}}) \cup \{X_1, \ldots, X_{\frac{\sigma}{2}}\}| \geq \sigma$. These inequalities imply $|\mathrm{var}(\ell_1) \cup \cdots \cup \mathrm{var}(\ell_{\frac{\sigma}{2}}) \cup \{X_1, \ldots, X_{\frac{\sigma}{2}}\}| = \sigma$. Hence there exists variable set $B = \{X_1, \ldots, X_{\frac{\sigma}{2}}\} \sqcup \{Y_1, \ldots, Y_{\frac{\sigma}{2}}\}$ such that $\langle B \rangle = \langle \{\ell_1, \ldots, \ell_{\frac{\sigma}{2}}\} \sqcup \{X_1, \ldots, X_{\frac{\sigma}{2}}\} \rangle$. In particular, $\langle \{\ell_1, \ldots, \ell_{\frac{\sigma}{2}}\} \sqcup \{X_1, \ldots, X_{\frac{\sigma}{2}}\} \rangle = \langle \mathrm{var}(h(A\mathbf{w})) \rangle$.

Consider the variable $x_1 \in \mathbf{x}$. As $n - 1 \geq \sigma - 2$ (as implied by $n \geq \sigma + 4$), then for $x_1 \in \mathbf{x}$ and $y_1 \in \mathbf{y}$, there exist, $h_1(\mathbf{w}), h_2(\mathbf{w}) \in Q$ such that $\mathrm{var}(h_1(\mathbf{w})) \cap \mathrm{var}(h_2(\mathbf{w})) = \{x_1, y_1\}$. Note that $\langle \mathrm{var}(h_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(h_2(\mathbf{w})) \rangle = \langle \mathrm{var}(h_1(\mathbf{w})) \cap \mathrm{var}(h_2(\mathbf{w})) \rangle$. Thus,

$$\dim \langle \mathrm{var}(h_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(h_2(\mathbf{w})) \rangle = \dim \langle \mathrm{var}(h_1(\mathbf{w})) \cap \mathrm{var}(h_2(\mathbf{w})) \rangle = 2.$$

The invertibility of A and the argument of the first paragraph imply that there exist σ-size sets $B_1, B_2 \subseteq \mathbf{w}$, with $B_1 = \mathrm{var}(h_1(A\mathbf{w}))$ and $B_2 = \mathrm{var}(h_2(A\mathbf{w}))$, such that $\langle \mathrm{var}(h_1(\mathbf{w})) \rangle$ and $\langle B_1 \rangle$ are isomorphic, and so are $\langle \mathrm{var}(h_2(\mathbf{w})) \rangle$ and $\langle B_2 \rangle$. Similarly under A, $\langle \mathrm{var}(h_1(\mathbf{w})) \rangle \cap \langle \mathrm{var}(h_2(\mathbf{w})) \rangle$ is isomorphic to $\langle B_1 \rangle \cap \langle B_2 \rangle$. Therefore, $\dim \langle B_1 \rangle \cap \langle B_2 \rangle = 2$. As $A(x_1) = X_1$, and $A(y_1) \in \langle B_1 \rangle \cap \langle B_2 \rangle$ (by the argument in the first paragraph), therefore $A(y_1) = Y_1 + c_1 X_1$, where $Y_1 \in \mathbf{w}$ is some scaled variable and $c_1 \in \mathbb{F}$.

Note: For odd σ, the proof holds with some modification. First, $h(\mathbf{w}) \in Q$ is now of form $((x_1y_1)\cdots(x_{\frac{\sigma-1}{2}}y_{\frac{\sigma-1}{2}})x_{\frac{\sigma+1}{2}})^\star$. The argument in the first paragraph then shows that for $h(A\mathbf{w}) = ((X_1\ell_1)\cdots(X_{\frac{\sigma-1}{2}}\ell_{\frac{\sigma-1}{2}})X_{\frac{\sigma+1}{2}})^\star$, there exists a σ-size set $B \subseteq \mathbf{w}$, such that $B = \{X_1,\ldots,X_{\frac{\sigma+1}{2}}\} \sqcup \{Y_1,\ldots,Y_{\frac{\sigma-1}{2}}\}$ and $\langle B\rangle = \langle\{\ell_1,\ldots,\ell_{\frac{\sigma-1}{2}}\} \sqcup \{X_1,\ldots,X_{\frac{\sigma+1}{2}}\}\rangle$. Secondly, for a variable $x_1 \in \mathbf{x}$, the existence of $h_1(\mathbf{w}), h_2(\mathbf{w}) \in Q$ such that $\mathrm{var}(h_1(\mathbf{w})) \cap \mathrm{var}(h_2(\mathbf{w})) = \{x_1, y_1\}$ is implied by $n-1 \geq \sigma - 1$ (which itself is implied by $n \geq \sigma + 4$). With these changes, the remainder of the argument continues to hold.

E.4 Choosing the degrees

In this section, we set the powers $\star$, as denoted in Section 5.1, for all polynomials in P, Q and R such that Conditions 1 and 2, specified in that section, are satisfied over any field. Let $N := \binom{n+\sigma-5}{\sigma} + \binom{n}{\sigma/2} + m$, $i \in [m+1, N]$ and $k \in [m]$. Note for odd σ, $N := \binom{n+\sigma-5}{\sigma} + \binom{n}{\frac{\sigma+1}{2}}\frac{\sigma+1}{2} + m$.

Over characteristic 0 fields. For R_k, choose the powers as $\sigma + k$. For the polynomials in $P \sqcup Q$, arbitrarily order them and choose the powers to be of form $\sigma + i$. For this choice of the powers, every polynomial in $P \sqcup Q$ has corresponding degree $\sigma(\sigma+i)$ and is clearly degree separated from the other polynomials in $P \sqcup Q$. The degree of R_k is $6 + (\sigma+k)(\sigma-5)$. As $i \geq m+1$, $k \leq m$ and $\sigma > 1$ it can be easily observed that

$$\sigma(\sigma+i) \geq \sigma(\sigma+m+1) > 6 + (\sigma+m)(\sigma-5) \geq 6 + (\sigma+k)(\sigma-5).$$

Thus, for this choice of the powers, Condition 1 is satisfied over characteristic 0 fields. The degree of f then is $\sigma(\sigma+N) = O(n^\sigma)$.

Over finite characteristic fields. Let the characteristic be $p > 0$. We assume $p > \sigma + 1$ to ensure Claim 2.2 holds. If $p > \sigma + N$, the powers can be chosen just like in the characteristic 0 case. Otherwise if $\sigma + 1 < p \leq \sigma + N$, then we choose the powers to be of form $p^t - 1$, where $t \in \mathbb{N}$, from the following N disjoint intervals:

$$[\sigma+1, p(\sigma+1)],\ [p(\sigma+1)+1, p^2(\sigma+1)+p],\ [p^2(\sigma+1)+p+1, p^3(\sigma+1)+p^2+p], \cdots$$

For R_k, the power is chosen from the k^{th} interval. Thus, the degree of R_k is in the range

$$[6 + (p^{k-1}(\sigma+1) + \sum_{l=0}^{k-2} p^l)(\sigma-5), 6 + (p^k(\sigma+1) + \sum_{l=1}^{k-1} p^l)(\sigma-5)]$$

which is disjoint for distinct k. Order the $N - m$ polynomials in $P \sqcup Q$ arbitrarily and assign powers in each polynomial from the remaining $N - m$ intervals. Then, the degree of a polynomial in $P \sqcup Q$ lies in the range

$$[\sigma(p^{i-1}(\sigma+1) + \sum_{l=0}^{i-2} p^l), \sigma(p^i(\sigma+1) + \sum_{l=1}^{i-1} p^l)].$$

For distinct i this range is disjoint implying the polynomials in $P \sqcup Q$ are degree separated. Now, we show that R_k's are degree separated from the polynomials of $P \sqcup Q$ by showing that the lower bound on the degree of any polynomial in $P \sqcup Q$ is greater than the upper bound on the degree of any R_k. As $i \geq m+1, k \leq m$, $\sigma > 1$ and $p > 1$, it follows that

$$\sigma(p^{i-1}(\sigma+1) + \sum_{l=0}^{i-2} p^l) \geq \sigma(p^m(\sigma+1) + \sum_{l=0}^{m-1} p^l)$$

$$> 6 + (p^m(\sigma+1) + \sum_{l=1}^{m-1} p^l)(\sigma - 5) \geq 6 + (p^k(\sigma+1) + \sum_{l=1}^{k-1} p^l)(\sigma - 5).$$

Thus, for this choice of powers, Conditions 1 and 2 are satisfied. For $p > \sigma + N$, the degree of f is $\sigma(\sigma + N) = O(n^\sigma)$ and for $\sigma + 1 < p \leq \sigma + N$, the degree of f is

$$O(\sigma(p^N(\sigma+1) + \sum_{l=1}^{N-1} p^l)) = O(p^N) = O((\sigma + N)^N) = O((\sigma + n^\sigma)^{n^\sigma}).$$

Note that the degree of f can be represented in $n^{O(1)}$ many bits. Hence, as remarked just after Theorem 3, we assume that over finite characteristic fields, the exponent vectors corresponding to the monomials of the input polynomials are given in binary.