Algebras with Polynomial Identities and Computing the Determinant

Steve Chien†
Microsoft Research
1065 La Avenida
Mountain View, CA 94043, U.S.A.
schien@microsoft.com

Alistair Sinclair‡
Computer Science Division
University of California, Berkeley
Berkeley, CA 94720-1776, U.S.A.
sinclair@cs.berkeley.edu

Abstract

In [18], Nisan proved an exponential lower bound on the size of an algebraic branching program (ABP) that computes the determinant of a matrix in the non-commutative "free algebra" setting, in which there are no non-trivial relationships between the matrix entries. By contrast, when the matrix entries commute there are polynomial size ABPs for the determinant.

This paper extends Nisan's result to a much wider class of non-commutative algebras, including all non-trivial matrix algebras over any field of characteristic 0, group algebras of all non-abelian finite groups over algebraically closed fields of characteristic 0, the quaternion algebra and the Clifford algebras. As a result, we obtain more compelling evidence for the essential role played by commutativity in the efficient computation of the determinant.

The key to our approach is a characterization of non-commutative algebras by means of the polynomial identities that they satisfy. Extending Nisan's lower bound framework, we find that any reduction in complexity compared to the free algebra must arise from the ability of the identities to reduce the rank of certain naturally associated matrices. Using results from the theory of algebras with polynomial identities, we are able to show that none of the identities of the above classes of algebras is able to achieve such a rank reduction.

† Part of this work was done while at the Computer Science Division, UC Berkeley, supported by NSF grants CCR-9820951 and CCR-0121555.

‡ Supported in part by NSF ITR grant CCR-0121555. Part of this work was done while the author was on sabbatical leave at Microsoft Research, Redmond, and at Université Paris Sud and Ecole Polytechnique, Paris.

1. Introduction

All known polynomial time algorithms for computing the determinant of a matrix appear to rely on the fact that multiplication in the underlying field (in which the matrix entries reside) is commutative. How hard is it to compute the determinant in a non-commutative setting? This question is motivated by the broader aim of understanding the computational power of commutativity [21], as well as by recent algorithmic applications of determinants over non-commutative algebras to approximating the permanent [3, 5].

In a landmark paper [18], Nisan pioneered the study of non-commutative computation. His main result was an exponential lower bound (of the form $\Omega(2^n)$) for the size of any algebraic branching program (ABP) that computes the determinant of an $n \times n$ matrix, viewed as a formal algebraic expression whose indeterminates $\{x_{11}, x_{12}, \ldots, x_{nn}\}$ do not commute. Since the determinant can be computed by an ABP of size $O(n^3)$ in the commutative setting [9, 16, 17, 22], this provides intriguing evidence of the computational power of commutativity.† Very recently, the ABP model has been used by Raz and Shpilka [21] to give an efficient deterministic algorithm for polynomial identity testing over non-commutative formulas.

The main limitation of Nisan's result, and of the ABP model as used to date, is that it is restricted to the free algebra $\mathbb{F}\langle x_{11}, \ldots, x_{nn}\rangle$ over a field $\mathbb{F}$, in which not only does commutativity fail to hold, but there is no structure whatsoever (i.e., no non-trivial relations hold between the indeterminates). However, it remains quite conceivable that, in some specific non-commutative algebra (e.g.,

† Nisan in fact stated his result in terms of formula size, rather than ABP complexity. His exponential lower bound on ABP size translates directly to a similar bound on formula size, which he contrasted with the fact that the determinant can be computed by a formula of size $n^{O(\log n)}$ in the commutative setting.

the quaternions), the determinant can be computed efficiently. It seems important that any comparison of commutative and non-commutative computation consider all non-commutative algebras, not only the free algebra.

In this paper, we address this issue by proposing an approach that allows the algebras to have much more structure than the free algebra, without being commutative. We characterize an algebra over $\mathbb{F}$ by the *polynomial identities* that it satisfies. In this view, commutative $\mathbb{F}$-algebras (when $\mathbb{F}$ has characteristic 0) are characterized by the polynomial identity $x_1x_2 - x_2x_1 = 0$. By contrast, the free algebra $\mathbb{F}\langle x_1, x_2, \ldots\rangle$ satisfies no non-trivial identities. Adding polynomial identities (without including the commutative identity) creates a spectrum of algebras between these two extremes. The study of polynomial identities has been an active topic in algebra for the past fifty years (see the book [7] for a survey). Our aim is to use the machinery of that field to study the power of non-commutativity in a manner that is more sensitive to the structure of the algebra; in particular, lower bounds for algebras that admit a rich class of identities give more compelling evidence for the importance of commutativity than that provided by Nisan's result.

Our first step is to extend Nisan's lower bound framework for ABPs, based on the combinatorial structure of the monomials of the function being computed, as expressed in the rank of certain naturally associated matrices. This simple step leads to a useful tool for comparing the ABP complexities in the free algebra and in the algebra of interest: essentially, any reduction in complexity corresponds to an ability of the polynomial identities to reduce the rank of the associated matrices.

We then go on to apply this framework to obtain exponential lower bounds for the ABP complexity of the determinant over a range of natural non-commutative algebras. The first class we consider are *matrix algebras*, whose elements are $d \times d$ matrices over a field $\mathbb{F}$. This is probably the most natural and most widely studied family of non-commutative algebras, and has also arisen in connection with approximating the permanent [3, 5]. We note that the algebra of $d \times d$ matrices satisfies the symmetric polynomial identity $s_{2d}(x_1, \ldots, x_{2d}) = 0$, where $s_k(x_1, \ldots, x_k)$ is defined to be $\sum_\sigma \mathrm{sgn}(\sigma) \prod_{i=1}^k x_{\sigma i}$, with the sum ranging over all k-permutations σ. Thus s_2 is the commutative identity, and s_k for $k > 2$ can be viewed as higher-order generalizations of commutativity. We show an exponential ABP lower bound for the determinant over any non-trivial matrix algebra:

Theorem 1.1 *Any ABP for computing the determinant of an $n \times n$ matrix whose entries belong to the algebra of $d \times d$ matrices, $d \geq 2$, over a field $\mathbb{F}$ of characteristic 0 has size at least 2^n.*

We then go on to apply the above result, together with some additional observations, to deduce a similar lower bound for computing the determinant over several other classes of non-commutative algebras. These results are summarized in the following theorem:

Theorem 1.2 *The same lower bound as in Theorem 1.1 holds for ABPs computing the determinant over any of the following algebras:*

1. *The algebra of $d \times d$ upper triangular matrices over a field $\mathbb{F}$ of characteristic 0, for any $d \geq 2$.*
2. *The quaternion algebra, and indeed all higher Clifford algebras.*‡
3. *The group algebra of any finite, non-abelian group G over any algebraically closed field $\mathbb{F}$ of characteristic 0. (I.e., the algebra whose elements are $\mathbb{F}$-vectors indexed by group elements, with multiplication inherited from the group.)*

To conclude this introduction, we briefly mention some related work on lower bounds for the determinant and permanent in other restricted models of computation. In addition to the non-commutative case studied here and in [18], exponential lower bounds are known for other restricted models including formulas of depth 3 (over finite fields) [10, 11] and various restricted classes of multilinear formulas [19, 21]. In a recent breakthrough, Raz [20] obtained a super-polynomial bound (of the form $n^{\Omega(\log n)}$) on the size of an arbitrary multilinear formula for the permanent or the determinant (over any field). By contrast, the best known lower bound for the size of general arithmetic formulas for the determinant is $\Omega(n^3)$ [12].

The remainder of the paper is organized as follows. In Section 2 we provide necessary background on ABPs and polynomial identities. In Section 3 we review Nisan's framework for lower bounds based on rank, and extend it to general algebras with polynomial identities. We apply this framework to obtain an exponential lower bound for matrix algebras (Theorem 1.1) in Section 4, and for several other algebras (as enumerated in Theorem 1.2) in Section 5. We conclude with a discussion of some limitations of our results, and some suggestions for future work, in Section 6.

2. Background

2.1. Algebras with polynomial identities

Let $\mathbb{F}$ be a field and $\mathcal{A}$ an associative algebra over $\mathbb{F}$, or an $\mathbb{F}$-*algebra* (i.e., $\mathcal{A}$ is a vector space over $\mathbb{F}$ together with a distributive multiplication operation). Note that we will always assume that multiplication in $\mathcal{A}$ is associative, but it

‡ See, e.g., [14] for a definition. These algebras were used in [5] in connection with approximating the permanent.

need *not* be commutative. We also assume the existence of a multiplicative unity. Familiar examples of $\mathbb{F}$-algebras are the following:

1. $\mathbb{F}\langle X\rangle$, the *free algebra* over $\mathbb{F}$ generated by a countable set of indeterminates $X = \{x_1, x_2, \ldots\}$, corresponding to all polynomials with coefficients in $\mathbb{F}$ in which no non-trivial relationships exist between the indeterminates.
2. $\mathbb{F}[X]$, the standard *polynomial ring* over $\mathbb{F}$, corresponding to polynomials over $\mathbb{F}$ in which the indeterminates commute.
3. The *matrix algebra* $\mathrm{Mat}_d(\mathbb{F})$, consisting of all $d \times d$ matrices with entries in $\mathbb{F}$.
4. For any group G, the *group algebra* $\mathbb{F}G$, whose elements are vectors of the form $\sum_{g\in G} c_g g$ with $c_g \in \mathbb{F}$, with multiplication defined by $(\sum_{g\in G} a_g g)(\sum_{h\in G} b_h h) = \sum_{g,h\in G} a_g b_h gh$.

We now introduce the central concept of *polynomial identities*, which we shall use to characterize different $\mathbb{F}$-algebras. We shall limit our treatment to the essentials; for more background on this topic, see the monograph [7].

Definition 2.1 *Let $\mathcal{A}$ be any algebra over $\mathbb{F}$. A polynomial $z(x_1, \ldots, x_m) \in \mathbb{F}\langle X\rangle$ is a* polynomial identity *of $\mathcal{A}$ if and only if $z(a_1, \ldots, a_m) = 0$ for all $a_1, \ldots, a_m \in \mathcal{A}$.*

It is well known (see, e.g., [7]) that the set of all polynomial identities of a given algebra $\mathcal{A}$ forms a *T-ideal* of $\mathbb{F}\langle X\rangle$, i.e., a two-sided ideal that is closed under all endomorphisms of $\mathcal{A}$.[§] Thus if $z(x_1, \ldots, x_m)$ is an identity of $\mathcal{A}$, then substitution for each of the x_i by an arbitrary element of $\mathbb{F}\langle X\rangle$ also yields an identity. We denote this T-ideal by $T(\mathcal{A})$. We say that a set of identities $B \subseteq \mathbb{F}\langle X\rangle$ is a *basis* or *generating set* of $T(\mathcal{A})$ if every element of $T(\mathcal{A})$ can be expressed as a linear combination of the form

$$\sum_{\ell} \alpha_\ell z_\ell(g_{1\ell}, \ldots, g_{m\ell})\beta_\ell,$$

with $z_\ell \in B$ and $\alpha_\ell, \beta_\ell, g_{i\ell} \in \mathbb{F}\langle X\rangle$.

For example, the free algebra $\mathbb{F}\langle X\rangle$ has no non-trivial identities, while any commutative algebra over a field $\mathbb{F}$ of characteristic 0 has a generating set consisting of the single identity $x_1x_2 - x_2x_1$. The study of polynomial identities has been an active topic in algebra for the past fifty years (see [7] for a survey), but it remains an important open problem to find (minimal) generating sets for most widely studied algebras.[¶] A celebrated theorem of Amitsur and Levitzki [1] says that the matrix algebra $\mathrm{Mat}_d(\mathbb{F})$ satisfies the identity s_{2d}, where s_k is defined as

$$s_k(x_1, \ldots, x_k) = \sum_{\sigma\in\mathcal{S}_k} \mathrm{sgn}(\sigma) \prod_{i=1}^{k} x_{\sigma i}$$

and is known as the *standard identity* of degree k. (Note that the commutative identity $x_1x_2 - x_2x_1$ is the standard identity s_2; the standard identities can be viewed as natural, progressively weaker generalizations of commutativity.) Moreover, $\mathrm{Mat}_d(\mathbb{F})$ satisfies no identities of lower degree. For $d = 2$ (the 2×2 matrix algebra), it has been shown relatively recently by Drensky [6] that, if $\mathbb{F}$ has characteristic 0, then s_4 together with the *Hall identity*

$$h(x_1, x_2) = [[x_1, x_2]^2, x_1],$$

(where $[a, b]$ denotes the "commutator" $ab - ba$) form a minimal generating set. This fact will be crucial to the present paper. Generating sets (let alone minimal ones) for $\mathrm{Mat}_d(\mathbb{F})$ are not known for any $d > 2$; indeed, this remains one of the central open questions in the area of polynomial identities. (See [4] for recent computer-assisted efforts in this direction.) Note, however, that any identity satisfied by $\mathrm{Mat}_d(\mathbb{F})$ for $d > 2$ is also satisfied by $\mathrm{Mat}_2(\mathbb{F})$ (but not conversely).

Another fact that will be useful later is that, if $\mathbb{F}$ has characteristic 0, then for any $\mathbb{F}$-algebra $\mathcal{A}$ the T-ideal $T(\mathcal{A})$ of polynomial identities of $\mathcal{A}$ has a generating set consisting entirely of *multilinear* identities. (See [7, Thm 4.3.2] for a proof; a polynomial $f(x_1, \ldots, x_n)$ is considered multilinear if x_i has degree 1 in each monomial of f for all $1 \le i \le n$.) From this it is straightforward to deduce that, if z is a polynomial identity of $\mathcal{A}$, then all the homogeneous components of z are themselves identities.

2.2. Algebraic branching programs

Algebraic branching programs were introduced by Nisan [18] as an algebraic analog of standard (arithmetic) branching programs.

Definition 2.2 *An* algebraic branching program (ABP) *is a directed acyclic graph with one source and one sink. The vertices of the graph are partitioned into levels numbered from 0 to d (the* degree *of the ABP), and edges may only go from level i to level $i+1$. The source is the only vertex at level 0 and the sink is the only vertex at level d. Each edge is labeled with a homogeneous linear polynomial in indeterminates x_i (i.e., a function of the form $\sum_i c_i x_i$, with coefficients $c_i \in \mathbb{F}$). The* size *of an ABP is the number of vertices.*

An ABP computes the degree-d homogeneous polynomial in $\mathbb{F}\langle X\rangle$ that is the sum, over all paths from the source to the sink, of the product of the linear functions associated

§ Indeed, there is a 1-1 correspondence between the T-ideals of $\mathbb{F}\langle X\rangle$ and the varieties of algebras satisfying a given set of polynomial identities.

¶ It is known that any algebra over a field of characteristic 0 has a finite generating set for its polynomial identities [13].

with the edges along that path. Figure 1(a) shows a toy example of a branching program that computes $f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2$ in $\mathbb{R}\langle X\rangle$. Note that the order of multiplication is important and follows the order of the paths.

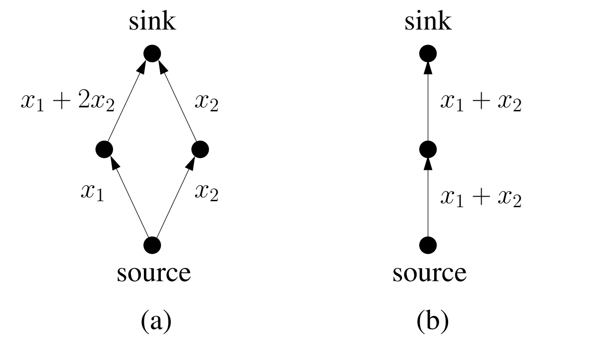

Figure 1. **(a)** An ABP that computes $x_1^2+2x_1x_2+x_2^2$ over $\mathbb{R}\langle X\rangle$; and **(b)** an ABP that computes the same function over $\mathbb{R}$.

For a homogeneous polynomial $f \in \mathbb{F}\langle X\rangle$ over n variables $x_1, \ldots, x_n$, the *branching program complexity* $B(f)$ is defined as the size of a smallest ABP that computes f. (We shall see in the next section that $B(f) = 4$ for the above polynomial f, so the ABP in Figure 1(a) is optimal.)

We stress that the original concepts of ABPs and branching program complexity refer to computation in the *free algebra* $\mathbb{F}\langle X\rangle$, where there are no non-trivial relations among the indeterminates x_i. However, we can generalize them rather naturally to describe computation over any $\mathbb{F}$-algebra $\mathcal{A}$:

Definition 2.3 *An algebraic branching program computes a function f over $\mathcal{A}$ if and only if, for all substitutions of the indeterminates x_i by values $a_i \in \mathcal{A}$, the output of the ABP is $f(a_i)$.*

The branching program complexity of f over an algebra $\mathcal{A}$ will be denoted by $B_{\mathcal{A}}(f)$, and is defined as the size of the smallest ABP that computes f over $\mathcal{A}$. (Note that the unadorned notation $B(f)$ is reserved for the free algebra complexity of f.)

These definitions are intended to capture the idea that, in a specific algebra $\mathcal{A}$, branching programs may be able to take advantage of polynomial identities in $\mathcal{A}$ to reduce the complexity of computing some functions. As an example, suppose we are working in $\mathbb{R}$ (viewed trivially as an $\mathbb{R}$-algebra). Then the branching program shown in Figure 1(b) computes the same toy polynomial f as above by making use of the identity $x_1x_2 - x_2x_1 = 0$ to instead compute the equivalent polynomial $x_1^2 + x_1x_2 + x_2x_1 + x_2^2$. Clearly since f has degree 2 this ABP must be optimal, and so $B_{\mathbb{R}}(f) = 3$.

Although it is not a major concern of this paper, we note that the measure $B(f)$ can be related to other measures such as formula size $F(f)$. For example, for any homogeneous polynomial f of degree d, we have (see [18])

$$B(f) \le d(F(f)+1); \qquad F(f) \le B(f)^{O(\log d)}.$$

2.3. ABPs for the determinant

We conclude this section by discussing ABPs in the context of the determinant function, which will be our main application. The *determinant* of an $n \times n$ matrix with entries $x_{11}, \ldots, x_{nn}$ is defined by

$$\det_n(x_{11}, \ldots, x_{nn}) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} x_{i,\sigma i}.$$

As explained in the introduction, we are interested in the branching program complexity of $\det_n$ (as a function of n) over various $\mathbb{F}$-algebras $\mathcal{A}$. Nisan showed in [18] that $B(\det_n) = 2^n$ over the free (real) algebra $\mathbb{R}\langle X\rangle$. Recall that this algebra is non-commutative, and indeed has no interesting structure. At the other extreme, we may consider the situation in which we are working over a commutative algebra, such as $\mathbb{R}$ itself. In this setting there exist *polynomial* size ABPs for $\det_n$:

Theorem 2.4 *Let $\mathcal{A}$ be any commutative algebra over a field $\mathbb{F}$. Then $B_{\mathcal{A}}(\det_n) \le O(n^3)$.*

Note that some well-known determinant algorithms, such as Gaussian elimination, cannot be formulated as ABPs. However, there are a number of polynomial-time combinatorial algorithms (e.g., [9, 16, 17, 22]) that can be used to prove Theorem 2.4. For completeness, and because it was not originally phrased as an ABP, we provide in Appendix A a sketch of one of these, due to Mahajan and Vinay [16].

In the remainder of the paper, our goal will be to understand the complexity of computing the determinant in algebras between these two extremes.

3. A framework for lower bounds

In [18] Nisan introduced a characterization of the ABP complexity $B(f)$ (over the free algebra $\mathbb{F}\langle X\rangle$) in terms of the ranks of certain matrices describing the combinatorial structure of the monomials of f. In this section we first briefly describe Nisan's framework, and then extend it to computation over general $\mathbb{F}$-algebras.

Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial of degree d. For each $0 \le k \le d$, $M_k(f)$ denotes an $n^k \times n^{d-k}$ matrix with entries in $\mathbb{F}$ as follows. Each row of $M_k(f)$ corresponds to an (ordered) monomial of degree k, and each column corresponds to a monomial of degree $d-k$; the matrix entry at position $(x_{i_1}\cdots x_{i_k}, x_{j_1}\cdots x_{j_{d-k}})$ is the coefficient of the combined monomial $x_{i_1}\cdots x_{i_k}x_{j_1}\cdots x_{j_{d-k}}$ in f.

The following theorem gives a precise relationship between $B(f)$ and the matrices $M_k(f)$:

Theorem 3.1 (Nisan [18]) *For any homogeneous polynomial $f \in \mathbb{F}\langle X\rangle$ of degree d,*

$$B(f) = \sum_{k=0}^{d} \operatorname{rank}(M_k(f)).$$

By applying this theorem to the determinant function, Nisan shows that $B(\det_n) = 2^n$. (Indeed, this result applies to any polynomial that is *weakly equivalent* to the determinant, where two polynomials f and g are weakly equivalent if for each monomial in f there exists a monomial in g consisting of the same variables (possibly in a different multiplicative order and with a different non-zero coefficient), and vice versa. The permanent is an example of such a function.)

We now extend Nisan's results to handle not just $\mathbb{F}\langle X\rangle$, but all $\mathbb{F}$-algebras for fields $\mathbb{F}$ of characteristic 0. In particular, we characterize $B_{\mathcal{A}}(f)$ in terms of the polynomial identities of $\mathcal{A}$ as follows:

Theorem 3.2 *Let $\mathbb{F}$ be a field of characteristic 0, $f \in \mathbb{F}\langle X\rangle$ be a homogeneous polynomial of degree d, and $\mathcal{A}$ be any $\mathbb{F}$-algebra. Then if $f \notin T(\mathcal{A})$ (i.e., if f is not identically zero over $\mathcal{A}$), the ABP complexity of f over $\mathcal{A}$ is given by*

$$B_{\mathcal{A}}(f) = \inf_{z(\cdot)\equiv 0} B(f+z) = \inf_{z(\cdot)\equiv 0} \sum_{k=0}^{d} \operatorname{rank}(M_k(f+z)),$$

where the infimum is over the zero polynomial and all degree-d homogeneous polynomial identities z of $\mathcal{A}$.

Proof: Let P be a branching program that computes $f(x_1,\ldots,x_n)$ over $\mathcal{A}$. By definition P computes a homogeneous formal polynomial $g \in \mathbb{F}\langle X\rangle$; we need to show that g is of the form $f+z$ for some homogeneous degree-d polynomial identity z of $\mathcal{A}$.

This is almost immediate. Since P computes f over $\mathcal{A}$, we have by definition that $f(a_1,\ldots,a_n) = g(a_1,\ldots,a_n)$ for all instantiations $a_i \in \mathcal{A}$. Hence if we define the function $z(x_1,\ldots,x_n) = g(x_1,\ldots,x_n) - f(x_1,\ldots,x_n)$, we have that $z(a_1,\ldots,a_n) = 0$ for all $a_i \in \mathcal{A}$, and thus z is either the zero polynomial or a polynomial identity. But $z = g - f$ with f, g homogeneous; so, recalling from Section 2.1 that each of the homogeneous components of z is also a polynomial identity, and using the assumption that f is not itself an identity, we may conclude that z is in fact homogeneous of degree d. Since P computes $g = f+z$, we are done. The second equality in the theorem is a direct application of Nisan's result (Theorem 3.1). $\square$

We can see a very simple application of this general framework by referring to the examples in Figure 1, where we are computing $f(x_1,x_2) = x_1^2 + 2x_1x_2 + x_2^2$. For the case of the free algebra $\mathbb{R}\langle X\rangle$, we apply Nisan's theorem and find that

$$M_0(f) = \begin{array}{cc} & \begin{array}{ccc} x_1^2 & x_1x_2 & x_2^2 \end{array} \\ 1 & \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \end{array}, \qquad M_1(f) = \begin{array}{cc} & \begin{array}{cc} x_1 & x_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \end{array}$$

and $M_2(f) = M_0(f)^{\mathrm{T}}$, which implies $B(f) = \operatorname{rank}(M_0(f)) + \operatorname{rank}(M_1(f)) + \operatorname{rank}(M_2(f)) = 1+2+1 = 4$. Hence the ABP shown in Figure 1(a) is minimal.

When working over $\mathbb{R}$ rather than $\mathbb{R}\langle X\rangle$, we can add the identity $z(x_1,x_2) = x_2x_1 - x_1x_2$ to obtain $g(x_1,x_2) = f(x_1,x_2) + z(x_1,x_2) = x_1^2 + x_1x_2 + x_2x_1 + x_2^2$. We then observe that

$$M_0(g) = \begin{array}{cc} & \begin{array}{cccc} x_1^2 & x_1x_2 & x_2x_1 & x_2^2 \end{array} \\ 1 & \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \end{array}, \quad M_1(g) = \begin{array}{cc} & \begin{array}{cc} x_1 & x_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{array}$$

and $M_2(g) = M_0(g)^{\mathrm{T}}$, so the ABP complexity of f over $\mathbb{R}$ is $B_{\mathbb{R}}(f) \le 1+1+1 = 3$. Since this is clearly minimal (f has degree 2), equality must hold and we confirm that the ABP in Figure 1(b) is optimal.

In principle, then, given any function f, we can compare its branching program complexity over any given algebra $\mathcal{A}$ to its free-algebra complexity by determining all of the polynomial identities of $\mathcal{A}$ and checking if any of them are able to reduce the rank of the matrices $M_k(f)$.

We conclude this section by noting that as a consequence of the ABP model, the only identities we need to consider in Theorem 3.2 are homogeneous (of the same degree as f). This might be seen as a weakness of our framework, as one may conceivably be able to exploit non-homogeneous identities to reduce the complexity of f. (The ABP model is fully general for computation in the free algebra, for if f is homogeneous then there is no advantage in allowing the ABP to be non-homogeneous.) However, even allowing non-homogeneous ABPs, we can obtain the bound

$$\hat{B}_{\mathcal{A}}(f)^2 \ge C \inf_{z(\cdot)\equiv 0} B(f+z)$$

for a universal constant C, where $\hat{B}$ is the size of the smallest *general* ABP that computes f over $\mathcal{A}$, and the infimum is over the zero polynomial and all homogeneous identities z of degree d. (For more precise definitions and a proof, see Appendix B.) Thus at the cost of a square we can assume all identities are homogeneous. Since our lower bounds will be exponential, this square will affect only the constant in the exponent.

4. Computing the determinant over matrix algebras

As mentioned earlier, Nisan [18] showed that the ABP complexity of the determinant over the free algebra [‖] $\mathbb{F}\langle X\rangle$ satisfies

$$B(\det_n) = 2^n. \tag{1}$$

In this section we prove a similar lower bound for a much wider class of non-commutative algebras, namely, for all *matrix algebras* over any field of characteristic 0. This is Theorem 1.1 from the introduction, which we restate here:

Theorem 4.1 *For any $d \geq 2$, the ABP complexity of computing the determinant over the $d \times d$ matrix algebra over any field $\mathbb{F}$ of characteristic 0 is given by*

$$B_{\mathrm{Mat}_d(\mathbb{F})}(\det_n) = 2^n.$$

Proof: We observe that Nisan proves (1) via Theorem 3.1, by showing that each of the matrices $M_k(\det_n)$ has rank exactly $\binom{n}{k}$. (This actually follows rather easily as the matrices have a very special form.) Following our generalized framework of Theorem 3.2, our task is to show that, for any homogeneous, degree-n polynomial identity z of the matrix algebra $\mathrm{Mat}_d(\mathbb{F})$, the rank of $M_k(\det_n + z)$ remains at least $\binom{n}{k}$. In other words, we have to show that no identity of the matrix algebras can reduce the rank of $M_k(\det_n)$. Note that since any polynomial identity of $\mathrm{Mat}_d(\mathbb{F})$ for $d > 2$ is also an identity for $\mathrm{Mat}_2(\mathbb{F})$, it is sufficient to show this for the identities of $\mathrm{Mat}_2(\mathbb{F})$.

In fact, for each value of k, we will examine only a submatrix of $M_k(\det_n + z)$ and show that this submatrix already contains $\binom{n}{k}$ linearly independent rows. We define our submatrix as follows: for each subset $S = \{a_1, \ldots, a_k\}$ of size k, $1 \leq a_1 < \ldots < a_k \leq n$, we keep those rows that correspond to monomials of the form $\prod_{j=1}^{k} x_{\sigma j, a_{\sigma j}}$ for all permutations $\sigma \in \mathcal{S}_k$; for each such subset, we assume that its rows are contiguous in the submatrix. Similarly, for each subset $S' = \{b_{k+1}, \ldots, b_n\}$ of size $n-k$, $1 \leq b_{k+1} < \ldots < b_n \leq n$, we keep those columns that correspond to monomials of the form $\prod_{j=1}^{n-k} x_{k+\sigma j, b_{k+\sigma j}}$ for all permutations $\sigma \in \mathcal{S}_{n-k}$; these columns are also assumed to be contiguous. Hence each pair of subsets (S, S') defines a "block" of size $k! \times (n-k)!$ in the submatrix.

We will denote this submatrix $\widetilde{M_k}(\det_n + z)$, and also denote by $\widetilde{M_k}(\det_n)$ and $\widetilde{M_k}(z)$ the submatrices created by restricting $M_k(\det_n)$ and $M_k(z)$ to the same sets of rows and columns. Note that $\widetilde{M_k}(\det_n + z) = \widetilde{M_k}(\det_n) + \widetilde{M_k}(z)$.

‖ In fact, Theorem 3.2 shows that Nisan's lower bound also holds for any algebra (over a field of characteristic 0) that does not satisfy any polynomial identities.

We now analyze the structure of $\widetilde{M_k}(\det_n)$ and $\widetilde{M_k}(z)$ in each of the (S, S') blocks. Note that S and S' correspond naturally to two sets of variables, $\Gamma(S) = \{x_{1a_1}, \ldots, x_{ka_k}\}$ and $\Gamma(S') = \{x_{k+1,b_{k+1}}, \ldots, x_{nb_n}\}$, and that a non-zero entry in an (S, S') block corresponds to a monomial whose variables are exactly those in $\Gamma(S) \cup \Gamma(S')$. The analysis of $\widetilde{M_k}(\det_n)$ is straightforward: if S and S' are disjoint (equivalently, if their union is $[n]$), then the (S, S') block in $\widetilde{M_k}(\det_n)$ will contain a single non-zero entry (either 1 or -1); otherwise, it is entirely zero.

The key to the rest of the proof is the following claim about the structure of $\widetilde{M_k}(z)$:

Claim 4.2 *For any $z \in T(\mathrm{Mat}_2(\mathbb{F}))$, the sum of the entries of $\widetilde{M_k}(z)$ in any (S, S') block is zero.*

Before proving the claim, we use it to complete the proof of the theorem. We can think of $\widetilde{M_k}(\det_n + z)$ as a large matrix divided into a $\binom{n}{k} \times \binom{n}{k}$ grid of (S, S') blocks. In this grid, there is a diagonal of (S, S') blocks whose entries sum to ± 1, but all other (S, S') blocks have an entry sum of 0. From this it is easy to see that $\mathrm{rank}(\widetilde{M_k}(\det_n + z)) \geq \binom{n}{k}$, as required.

To conclude, we supply the proof of the above claim.

Proof of Claim 4.2: Recall from Section 2.1 that any $z \in T(\mathrm{Mat}_2(\mathbb{F}))$ must be a sum of identities generated by the standard identity $s_4(x_1, x_2, x_3, x_4) = \sum_{\sigma \in \mathcal{S}_4} \mathrm{sgn}(\sigma) \prod_{i=1}^{4} x_{\sigma i}$ and the Hall identity in two variables $h(x_1, x_2) = [[x_1, x_2]^2, x_1]$, as these form a basis for $T(\mathrm{Mat}_2(\mathbb{F}))$. In fact, we will prove a stronger version of the claim by allowing z to take on a more general form. Notice that both s_4 and h can be generated by the polynomial $t(x_1, x_2, x_3, x_4) = [x_1, x_2][x_3, x_4]$; specifically, $s_4(x_1, x_2, x_3, x_4)$ is the sum of six terms of the form $t(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4})$ with appropriate signs, while $h(x_1, x_2) = t(x_1, x_2, x_1, x_2)x_1 - x_1 t(x_1, x_2, x_1, x_2)$. Hence any identity z in $T(\mathrm{Mat}_2(\mathbb{F}))$ can be written in the form

$$z = \sum_{\ell} \alpha_\ell t(g_{1\ell}, g_{2\ell}, g_{3\ell}, g_{4\ell}) \beta_\ell$$

where $\alpha_\ell, \beta_\ell, g_{i\ell} \in \mathbb{F}\langle X\rangle$. Note that in order for z to be homogeneous of degree n, each of the α_ℓ, β_ℓ and $g_{i\ell}$ must also be homogeneous. Furthermore, since t is multilinear, we can assume without loss of generality that α_ℓ, β_ℓ, and the $g_{i\ell}$ are all monomials.

We now show that for each ℓ, the sum of the entries of $M_k(\alpha_\ell t(g_{i\ell})\beta_\ell)$ is zero over any (S, S') block. We do this by showing that any non-zero entry in an (S, S') block is canonically canceled by another entry in the same block. Suppose $\alpha_\ell t(g_{i\ell})\beta_\ell$ is non-zero somewhere in an (S, S') block. Then the set of variables used in α_ℓ, β_ℓ and the $g_{i\ell}$ is exactly $\Gamma(S) \cup \Gamma(S')$. Furthermore, there exists some ordering of $g_{1\ell}$ and $g_{2\ell}$ and of $g_{3\ell}$ and $g_{4\ell}$ (without loss of generality, say $g_{1\ell}, g_{2\ell}$ and $g_{3\ell}, g_{4\ell}$) such that the first k variables of the resulting monomial $\alpha_\ell g_{1\ell} g_{2\ell} g_{3\ell} g_{4\ell} \beta_\ell$ are exactly those in $\Gamma(S)$ and the last $n-k$ variables are exactly those in $\Gamma(S')$. This implies that (1) the variables used in α_ℓ, $g_{1\ell}$ and $g_{2\ell}$ are contained in $\Gamma(S)$; or (2) the variables used in $g_{3\ell}$,$g_{4\ell}$, and β_ℓ are contained in $\Gamma(S')$. ((1) holds if α_ℓ, $g_{1\ell}$ and $g_{2\ell}$ contain at most k variables in total, and (2) holds if $g_{3\ell}$, $g_{4\ell}$ and β_ℓ contain at most $n-k$ variables in total; note that both may hold.) If the former is true, then the reordering $g_{2\ell} g_{1\ell} g_{3\ell} g_{4\ell}$ produces the same non-zero entry with opposite sign elsewhere. Otherwise, the reordering $g_{1\ell} g_{2\ell} g_{4\ell} g_{3\ell}$ has the same effect.

This finishes the proof of the claim and theorem. □

Remark: As with Nisan's original result for $\mathbb{F}\langle X \rangle$, our theorem also holds for all polynomials that are weakly equivalent to $\det_n$ (as defined just after Theorem 3.1), using a similar proof. □

5. Other non-commutative algebras

In this section we combine our result for matrix algebras in the previous section with some additional observations to obtain similar lower bounds for several other important classes of non-commutative algebras, as claimed in Theorem 1.2 of the introduction.

5.1. Upper triangular matrices

Let $\mathrm{UMat}_d(\mathbb{F})$ denote the set of $d \times d$ upper triangular matrices over a field $\mathbb{F}$. Clearly $\mathrm{UMat}_d(\mathbb{F})$ is a subalgebra of $\mathrm{Mat}_d(\mathbb{F})$. It turns out (see [7, Thm 5.2.1(i)]) that the single identity $t(x_1, x_2, x_3, x_4) = [x_1, x_2][x_3, x_4]$ used in the proof of Theorem 4.1 is actually a generating set for $\mathrm{UMat}_2(\mathbb{F})$. Therefore, that proof actually establishes that the ABP complexity of the determinant is exponentially large even over $\mathrm{UMat}_d(\mathbb{F})$, a stronger result.

Theorem 5.1 *For any $d \geq 2$, the ABP complexity of computing the determinant over the algebra of $d \times d$ upper triangular matrices over any field $\mathbb{F}$ of characteristic 0 is given by*

$$B_{\mathrm{UMat}_d(\mathbb{F})}(\det_n) = 2^n.$$

5.2. Group algebras

Let G be a finite group. The *group algebra* of G over a field $\mathbb{F}$, denoted $\mathbb{F}G$, consists of elements of the form $\sum_{g \in G} c_g g$ for $c_g \in \mathbb{F}$, with addition defined in the natural way. Multiplication is defined according to the group operation, namely: $(\sum_{g \in G} a_g g)(\sum_{h \in G} b_h h) = \sum_{g,h \in G} a_g b_h gh$.

We will combine our result for matrix algebras in the previous section with some elementary representation theory to prove an exponential lower bound on the ABP complexity of computing the determinant over any non-commutative group algebra, whenever $\mathbb{F}$ is algebraically closed and of characteristic 0:

Theorem 5.2 *Let G be a finite non-abelian group, and $\mathbb{F}$ an algebraically closed field of characteristic 0. The complexity of computing the determinant over the group algebra $\mathbb{F}G$ is given by*

$$B_{\mathbb{F}G}(\det_n) = 2^n.$$

Proof: We show that if G is non-abelian, then any polynomial identity satisfied by $\mathbb{F}G$ is also satisfied by $\mathrm{Mat}_d(\mathbb{F})$ for some $d \geq 2$. A direct application of Theorem 4.1 then yields our result.

A classical fact from group representation theory tells us that, since G is finite and non-abelian, it must have an irreducible representation of degree at least two; i.e., there exists a homomorphism $\rho : G \to GL_d(\mathbb{F})$ for some $d \geq 2$ such that the image of G, $\{\rho(g) : g \in G\}$, has no nontrivial invariant subspaces in $\mathbb{F}^d$.

We can now apply the following result of Burnside (see, e.g., [15]):

Lemma 5.3 (Burnside) *Let H be a group of invertible $d \times d$ matrices over an algebraically closed field $\mathbb{F}$. Then H has no nontrivial invariant subspaces in $\mathbb{F}^d$ if and only if H contains d^2 linearly independent matrices, i.e., if and only if the $\mathbb{F}$-span of H in $\mathrm{Mat}_d(\mathbb{F})$ is $\mathrm{Mat}_d(\mathbb{F})$ itself.*

From this we see that the induced homomorphism on algebras $\hat{\rho} : \mathbb{F}G \to \mathrm{Mat}_d(\mathbb{F})$ is surjective, and therefore any polynomial identity satisfied by $\mathbb{F}G$ must also be satisfied by $\mathrm{Mat}_d(\mathbb{F})$. □

5.3. Quaternions and Clifford algebras

One of the most familiar examples of a non-commutative algebra is Hamilton's quaternions. This is a real algebra of dimension four, with basis elements $\{1, i, j, k\}$ and defining relations $i^2 = j^2 = k^2 = -1$ and $ij = k = -ji$. We can again apply our result on $\mathrm{Mat}_2(\mathbb{F})$ to deduce an exponential lower bound for computing the determinant of a quaternion matrix:

Theorem 5.4 *Let $\mathbb{H}$ denote Hamilton's quaternions. Then*

$$B_{\mathbb{H}}(\det_n) = 2^n.$$

Proof: We invoke the useful fact that, if $\mathbb{F}$ has characteristic 0, the polynomial identities of an $\mathbb{F}$-algebra do not change when the base field $\mathbb{F}$ is extended. This fact is folklore, but according to Drensky [8] can be traced back to [23, Lemma 2.3]. Since both $\mathbb{H}$ and $\mathrm{Mat}_2(\mathbb{R})$ become $\mathrm{Mat}_2(\mathbb{C})$ when the base field is extended from $\mathbb{R}$ to $\mathbb{C}$, we deduce that $T(\mathbb{H}) = T(\mathrm{Mat}_2(\mathbb{R}))$, i.e., the quaternions satisfy exactly the same identities as the 2×2 matrices. The theorem now follows directly from Theorem 4.1. $\square$

This theorem can be immediately extended to all higher *Clifford algebras*, as defined (e.g.) in [14]. The first three Clifford algebras are $\mathit{CL}_1 = \mathbb{R}$, $\mathit{CL}_2 = \mathbb{C}$ and $\mathit{CL}_3 = \mathbb{H}$; the mth Clifford algebra CL_m has dimension 2^{m-1} over the reals, and is isomorphic to either $\mathrm{Mat}_d(\mathbb{F})$ or $\mathrm{Mat}_d(\mathbb{F}) \oplus \mathrm{Mat}_d(\mathbb{F})$, for some d and $\mathbb{F} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. (A more operational definition is given in [5]. Note that we are abusing notation here because $\mathbb{H}$ is not a field.) We need only note that CL_m is contained in CL_{m+1} for all $m \geq 1$, and therefore $T(\mathit{CL}_{m+1}) \subseteq T(\mathit{CL}_m)$. Thus each CL_m for $m \geq 3$ inherits the lower bound for $B_{\mathbb{H}}(\det_n)$ given in Theorem 5.4.

Our results on Clifford algebras are relevant to recent proposals that reduce approximating the permanent of an $n \times n$ $0,1$-matrix A to computing the determinant of a related matrix $\widehat{A}$, obtained by replacing the 1-entries of A by suitable random matrices. For example, Barvinok [3] has shown that if $\det_n(\widehat{A})$ can be efficiently computed when each non-zero entry of $\widehat{A}$ is chosen independently from a standard Gaussian distribution over a high-dimensional real matrix algebra, then we obtain a polynomial time approximation algorithm for the permanent of A within ratio $(1+\epsilon)^n$ for arbitrarily small $\epsilon > 0$. More recently, Chien, Rasmussen and Sinclair [5] showed that if we can compute $\det_n(\widehat{A})$ efficiently when each non-zero entry of $\widehat{A}$ is an independent, uniformly random basis element of the Clifford algebra CL_m with $m = O(\log n)$, then we have a *fully polynomial randomized approximation scheme* for the permanent.

Our results indicate that any attempt to implement these approaches will require a determinant algorithm that cannot be cast as an algebraic branching program, or will need to use special statistical properties of the random matrix $\widehat{A}$.

6. Discussion

As stated in the introduction, our main goal in this work is to better understand the nature of non-commutative computation, and the role played by commutativity in the design of efficient algebraic algorithms. We have used the determinant as a vehicle for these investigations. In this final section, we briefly discuss some limitations of our results and some ideas for further work.

6.1. Limitations

Our current results do not fully describe the computational power of commutativity in computing determinants. While there remain some non-commutative $\mathbb{F}$-algebras $\mathcal{A}$ over which we have yet to determine $B_{\mathcal{A}}(\det_n)$, a more important limitation derives from the ABP model. As currently defined, an ABP sees an algebra only in terms of its polynomial identities, and cannot exploit any additional structure it may have. We now show two examples of this: Barvinok's symmetrized determinant and the Dieudonné determinant.

For an $n \times n$ matrix A over an $\mathbb{F}$-algebra $\mathcal{A}$, the *symmetrized determinant* [3] is defined as $\mathrm{sdet}_n(A) = \frac{1}{n!}\sum_{\sigma \in \mathcal{S}_n}\sum_{\tau \in \mathcal{S}_n} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)\prod_{i=1}^{n} a_{\sigma i, \tau i}$. Thus sdet_n can be viewed as the average of the standard (Cayley) determinant $\det_n$ over all $n!$ possible row orderings, and is in fact weakly equivalent to $\det_n$. Therefore, for (say) $\mathcal{A} = \mathrm{Mat}_2(\mathbb{R})$ (so that the entries of A are 2×2 real matrices), Theorem 4.1 implies $B_{\mathcal{A}}(\mathrm{sdet}_n) \geq 2^n$; however, Barvinok has shown that sdet_n can in fact be computed in polynomial time.** The key idea used is to operate on the four entries of each 2×2 matrix separately, something an ABP cannot do.

If A is a matrix over a division algebra $\mathcal{A}$ (such as the quaternions), we can define the *Dieudonné determinant* Ddet_n of A (see, e.g., [2]). This can be computed in polynomial time using Gaussian elimination, which is made possible by the presence of inverses in $\mathcal{A}$. While $\mathrm{Ddet}_n(A)$ and $\det_n(A)$ are not equal in general, they are similar enough for $\mathrm{Ddet}_n(A)$ to be useful in the permanent estimators discussed in Section 5.3 (see [5]). Here, it is the ability to do division that takes us outside the ABP model.

Note that neither of these examples shows that the standard determinant $\det_n$ can be efficiently computed over a specific non-commutative algebra; however, they do demonstrate the importance of considering more general models of computation.

6.2. Some open questions

Within the ABP model discussed in this paper, a number of open questions suggest themselves. Firstly, it would be interesting to complete our picture of which non-commutative algebras, if any, allow for polynomial-sized ABPs for the determinant. While our approach thus far appears ad hoc in that we have focused on specific examples of algebras, there may not in fact be too many classes of polynomial identities that we still need to examine. Note that our analysis of the degree-four polynomial identity $t(x_1, x_2, x_3, x_4) = [x_1, x_2][x_3, x_4]$ already covers a large spectrum of identities. We feel that an important gap here in both our understanding and our proof technique is the effect of polynomial identities of degree 3, such as the standard identity s_3, on the rank of the matrices M_k. Secondly, it would be interesting to extend

** More generally, if $\mathcal{A}$ is of finite dimension r, then sdet_n can be computed in time $O(n^{r+3})$ [3].

our investigation of the role of commutativity to functions other than the determinant.

Another interesting set of issues concerns the ABP model itself. To what extent can one strengthen the model while retaining similar lower bounds for determinant computation over (say) matrix algebras? One natural extension would allow an ABP computing over a matrix algebra to access the individual components of its input matrices; as we saw above, this is helpful in computing the symmetrized determinant sdet_n (though we do not know what happens for $\det_n$ itself). Similarly, in algebras with involution (such as the Clifford algebras), we might allow an ABP to use the involutions of its input variables. More ambitiously, we might consider more general models of computation, such as those capable of implementing Gaussian elimination.

Acknowledgements

We thank Vesselin Drensky for valuable help with the literature on polynomial identities, particularly for the proof of Theorem 5.4, and as always T.Y. Lam for patiently answering all our algebraic questions. We also thank Salil Vadhan and Avi Wigderson for helpful initial discussions.

References

[1] S.A. AMITSUR and J. LEVITZKI, "Minimal identities for algebras," *Proceedings of the American Mathematical Society* **2** (1950), pp. 449–463.

[2] E. ARTIN, *Geometric Algebra*, Wiley Interscience, New York, 1988.

[3] A. BARVINOK, "New permanent estimators via non-commutative determinants," Preprint, July 2000, available from www.math.lsa.umich.edu/∼barvinok/papers.html.

[4] F. BENANTI, J. DEMMEL, V. DRENSKY and P. KOEV, "Computational approach to polynomial identities of matrices — a survey," in *Ring Theory: Polynomial Identities and Combinatorial Methods*, A. Giambruno, A. Regev and M. Zaicev, eds., Lecture Notes in Pure and Applied Mathematics Vol. 235, Dekker, 2003, pp. 141–178.

[5] S. CHIEN, L. RASMUSSEN and A. SINCLAIR, "Clifford algebras and approximating the permanent," *Journal of Computer and System Sciences* **67** (2003), pp. 263–290. (Special issue on *34th ACM STOC*, 2002, J. Reif, ed.)

[6] V. DRENSKY, "A minimal basis for the identities for a second-order matrix algebra over a field of characteristic 0," *Algebra and Logic* **20** (1981), pp. 188–194.

[7] V. DRENSKY, *Free Algebras and PI-Algebras*, Springer-Verlag, Singapore, 1999.

[8] V. DRENSKY, personal communication.

[9] D. FADEEV and V. FADEEVA, *Computational Methods in Linear Algebra*, Freeman, San Francisco, 1963.

[10] D. GRIGORIEV and M. KARPINSKI, "An exponential lower bound for depth 3 arithmetic circuits," *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, 1998, pp. 577–582.

[11] D. GRIGORIEV and A. RAZBOROV, "Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields," *Applicable Algebra in Engineering, Communication and Computing* **10** (2000), pp. 465–487.

[12] K. KALORKOTI, "The formula size of the determinant," *SIAM Journal on Computing* **14** (1995), pp. 678–687.

[13] A. KEMER, *Ideals of Identities of Associative Algebras*, Translations of Mathematical Monographs **87**, AMS, Providence RI, 1991.

[14] T.Y. LAM, *The Algebraic Theory of Quadratic Forms*, Benjamin/Addison-Wesley, Reading MA, 1973. (Reprinted with revisions, 1980.)

[15] T.Y. LAM, "A theorem of Burnside on matrix rings," *American Mathematical Monthly* **105** (1998), pp. 651–653.

[16] M. MAHAJAN and V. VINAY, "A combinatorial algorithm for the determinant," *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms*, 1997, pp. 730–738.

[17] M. MAHAJAN and V. VINAY, "Determinant: Old algorithms, new insights," *SIAM Journal on Discrete Mathematics* **12** (1999), pp. 474–490.

[18] N. NISAN, "Lower bounds for non-commutative computation," *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, 1991, pp. 410–418.

[19] N. NISAN and A. WIGDERSON "Lower bounds on arithmetic circuits via partial derivatives," *Computational Complexity* **6** (1996), pp. 217–234.

[20] R. RAZ, "Multi-linear formulas for permanent and determinant are of super-polynomial size," *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, 2004, pp. 633–641.

[21] R. RAZ and A. SHPILKA, "Deterministic polynomial identity testing in non-commutative models," *Proceedings of the 13th Conference on Computational Complexity*, 2004, pp. 215–222.

[22] L. VALIANT, "Why is boolean complexity theory difficult?" in *Boolean Function Complexity*, M. Paterson, ed., London Mathematical Society Lecture Notes Series 169, Cambridge University Press, 1992.

[23] M.R. VAUGHAN-LEE, "On varieties of Lie algebras," *Quart. J. Math. Oxford Ser. (2)* **21** (1970), pp. 297–308.

Appendix

A. An $O(n^3)$-size ABP for the determinant over commutative algebras

As claimed in Theorem 2.4, we sketch without proof an algebraic branching program of size $O(n^3)$ for computing the determinant of an $n \times n$ matrix $A = (a_{ij})$ over any commutative algebra $\mathcal{A}$, based on an algorithm of Mahajan and Vinay [16].

We first think of A as a weighted directed graph in which the vertices are labeled $1, \ldots, n$ and each edge (i, j) has weight a_{ij}. A *closed walk* on this graph is a path $(v_1, \ldots, v_k)$ (starting from v_1 and moving to v_k before returning to v_1 along edge (v_k, v_1)) in which v_1 appears only once and is the vertex with smallest label on the path (also called the *head*). A *closed walk sequence* is an ordered sequence of closed walks whose total length is n and whose heads are in strictly increasing order. The *weight* of a closed walk sequence is the product of the weights of the edges it contains.

Let $\mathcal{C}_n$ denote the set of all closed walk sequences, and for each $C \in \mathcal{C}_n$, let $w(C)$ denote the weight of C and $\mathrm{sgn}(C) = (-1)^{n+k}$, where k is the number of closed walks in C. Note that the sum of $\mathrm{sgn}(C)w(C)$ over only those closed walk sequences that are cycle covers of the graph is exactly the determinant of A. Moreover, Mahajan and Vinay show that in fact $\det_n(A) = \sum_{C \in \mathcal{C}_n} \mathrm{sgn}(C)w(C)$, where the sum is over *all* closed walk sequences. The proof of this fact relies on commutativity of the matrix entries a_{ij} to ensure that the contributions of closed walk sequences that do not correspond to cycle covers cancel. (Cancellations occur between closed walk sequences in which the same edges appear but in different orders.) Mahajan and Vinay then give a simple dynamic programming algorithm for computing this sum, which can be interpreted as an algebraic branching program as follows.

The ABP has depth n. A vertex at level i, $1 \leq i \leq n-1$, is labeled with a triple (p, h, v) with $p \in \{0, 1\}$, $h \in \{1, \ldots, n\}$, and $v \in \{1, \ldots, n\}$. The function computed at such a vertex (i.e., the function computed by the ABP having this vertex as its sink) is the sum of the weights of all valid length-i prefixes of closed walk sequences in which the parity of the number of closed walks completed so far is p, the head of the walk currently being constructed is h, and the current end vertex of this walk is v. There are edges going from level i to level $i + 1$ as follows. Vertex (p, h, v) at level $i < n - 1$ is connected to all valid (p, h, u) at level $i + 1$ by an edge with label x_{vu} (corresponding to extending the current walk to vertex u), and to all valid $(1 - p, h', h')$ by an edge with label x_{vh} (corresponding to completing the current closed walk and starting a new one with head h'). Finally, vertex (p, h, v) at level $n - 1$ is connected to the sink (the only vertex at level n) by an edge with label $(-1)^{n+p+1}x_{vh}$ (corresponding to completing the last closed walk and incorporating the sign of the closed walk sequence). This algebraic branching program has size $O(n^3)$ ($O(n^2)$ vertices at each of $O(n)$ levels), and it is straightforward to check that it computes the above sum over all closed walk sequences.

B. General algebraic branching programs

We formalize and prove the claim from Section 3 that allowing general (non-homogeneous) polynomial identities and ABPs costs us only a square in our lower bounds.

First consider a generalized setting for Theorem 3.2, in which the ABP model is extended in a natural way to allow computation of non-homogeneous polynomials. This is done by allowing the edge labels to be arbitrary linear polynomials (including constants) and dropping the requirement that the graph be leveled. Arguing exactly as in the proof of Theorem 3.2, we have that for any homogeneous f of degree d,

$$\hat{B}_{\mathcal{A}}(f) = \inf_{\hat{z}(\cdot) \equiv 0} \hat{B}(f + \hat{z}), \tag{2}$$

where $\hat{B}$ denotes the smallest general ABP that computes a particular function and the infimum is over the zero polynomial and all identities $\hat{z}$ (not necessarily homogeneous) of $\mathcal{A}$. However, recall from Section 2.1 that for any such $\hat{z}$, its homogeneous components are also identities; thus, in particular, the degree-d homogeneous component of $f + \hat{z}$ is of the form $f + z$, where z is homogeneous of degree d.

Now, given a (general) ABP of size s and depth ℓ for a function g, it is not hard to construct (see [21]) a (standard) ABP of size $O(s\ell)$ that computes each of the degree-d homogeneous components of g. Thus there is an ABP of size $O(s\ell) = O(s^2)$ that computes $f + z$. Combining this with (2) gives

$$\hat{B}_{\mathcal{A}}(f)^2 \geq C \inf_{z(\cdot) \equiv 0} B(f + z)$$

for a universal constant C, where the infimum is over the zero polynomial and all homogeneous identities z of degree d. This is what we claimed.