

# Lower Bounds for Planar Arithmetic Circuits

C. Ramya\*

Pratik Shastri †

December 15, 2023

## Abstract

*Arithmetic circuits* are a natural well-studied model for computing multivariate polynomials over a field. In this paper, we study *planar arithmetic circuits*. These are circuits whose underlying graph is planar. In particular, we prove an  $\Omega(n \log n)$  lower bound on the size of planar arithmetic circuits computing explicit bilinear forms on  $2n$  variables. As a consequence, we get an  $\Omega(n \log n)$  lower bound on the size of arithmetic formulas and planar algebraic branching programs computing explicit bilinear forms. This is the first such lower bound on the formula complexity of an explicit bilinear form. In the case of read-once planar circuits, we show  $\Omega(n^2)$  size lower bounds for computing explicit bilinear forms. Furthermore, we prove fine separations between the various planar models of computations mentioned above.

In addition to this, we look at multi-output planar circuits and show  $\Omega(n^{4/3})$  size lower bound for computing an explicit linear transformation on  $n$ -variables. For a suitable definition of multi-output formulas, we extend the above result to get an  $\Omega(n^2/\log n)$  size lower bound. As a consequence, we demonstrate that there exists an  $n$ -variate polynomial computable by  $n^{1+o(1)}$ -sized formulas such that any multi-output planar circuit (resp., multi-output formula) simultaneously computing all its first-order partial derivatives requires size  $\Omega(n^{4/3})$  (resp.,  $\Omega(n^2/\log n)$ ). This shows that a statement analogous to that of Baur, Strassen[3] does not hold in the case of planar circuits and formulas.

## 1 Introduction

*Arithmetic circuits* are a natural computational model for computing multivariate polynomials over a field. These are directed acyclic graphs whose in-degree 0 vertices are labeled by variables or field constants and whose internal vertices are either addition or multiplication gates. Arithmetic circuits compute polynomials in the natural way. The two important

---

\*The Institute of Mathematical Sciences (a CI of Homi Bhabha National Institute), Chennai, India. email: ramyac@imsc.res.in

†The Institute of Mathematical Sciences (a CI of Homi Bhabha National Institute), Chennai, India. email: pratiks@imsc.res.in

parameters of a circuit are *size* which is the number of vertices in it and *depth* which is the length of a longest input to output path. One of the primary goals in Algebraic Complexity Theory is to prove lower bounds on the size of arithmetic circuits computing an explicit polynomial. Despite consistent efforts, the best known size lower bound is due to Baur and Strassen[3] who proved that any fan-in 2 circuit computing the polynomial  $x_1^n + \dots + x_n^n$  has size  $\Omega(n \log n)$ .

*Arithmetic formulas* are circuits where the underlying graph is a tree. Kalarkoti[10] has proved that a certain complexity measure based on the transcendence degree (algebraic rank) of polynomials is a lower bound on the formula size. Using this technique, he proves that the formula complexity of  $\text{Det}_{n \times n}$ , the determinant polynomial, is  $\Omega(n^3)$  (note that this is an  $\Omega(m^{3/2})$  size lower bound where  $m = n^2$  is the number variables of  $\text{Det}_{n \times n}$ ). Chatterjee et al.[7] proved that over fields of characteristic greater than  $0.1n$ , any formula computing the elementary symmetric polynomial on  $n$  variables of degree  $0.1n$  requires size  $\Omega(n^2)$ .

*Algebraic branching programs*(ABPs) are yet another model for computing polynomials. ABPs are directed acyclic graphs with designated source and sink vertices. Their edges are labeled by variables or constants. The polynomial computed by an ABP is defined to be the sum of weights of all paths from source to sink (the weight of a path is the product of edge labels in the path). Chatterjee et al. [6] show that any ABP computing the polynomial  $x_1^n + \dots + x_n^n$  has size at least  $\Omega(n^2)$ . In fact, the results by Chatterjee et al. [6] also hold for ABPs where the edges are labeled by affine forms.

In the regime of constant degree polynomials, Raz[18] showed that for any constant  $r$  and any field  $\mathbb{F}$ , there is an explicit  $n$ -variate polynomial of total-degree  $O(r)$ , with  $\{0, 1\}$  coefficients such that any depth- $r$  arithmetic circuit for it (using constants from  $\mathbb{F}$ ) has size  $n^{1+\Omega(1)}$ . As observed by Raz[18], super-quadratic lower bound of  $\Omega(n^{2+\epsilon})$  for constant-depth circuits for an explicit polynomial of constant degree implies a strong superlinear ( $\Omega(n^{1+\epsilon/2})$ ) lower bound for general circuits. Also, recently Chatterjee et al.[6] demonstrated that a superlinear lower bound on the size any ABP for a homogeneous constant-degree polynomial would imply a superlinear lower bound on its determinantal complexity.

From the preceding discussion, it is imperative to note that we do not know any explicit constant degree polynomial that has superlinear circuit complexity.

*Bilinear forms* are a special and important class of degree two polynomials. These are polynomials form  $\mathbf{y}^T M \mathbf{x}$  where  $\mathbf{y}$  and  $\mathbf{x}$  are vectors of  $n$  variables each and  $M \in \mathbb{F}^{n \times n}$  is a matrix. A natural model to compute bilinear forms is the model of *bilinear circuits*. A bilinear circuit is an arithmetic circuit in which every product gate has two inputs one of which is a linear form in  $\mathbf{x}$ -variables and the other a linear form in  $\mathbf{y}$ -variables. It is well known and easy to prove that any circuit computing a bilinear form can be converted into a bilinear circuit computing the same bilinear form with only constant blowup in size. Therefore, superlinear lower bounds for bilinear circuits imply superlinear lower bounds on the size of general circuits computing an explicit bilinear form. No such lower bounds are known.

*Bilinear formulas* are defined similarly: it is a formula in which every product gate computes a product of two linear forms, one in the  $x$  variables and one in the  $y$  variables.

Superlinear (i.e.,  $\Omega(n \log^2 n / \log \log n)$ ) lower bounds on the *bilinear* formula complexity of explicit bilinear forms follow from lower bounds on the size of depth two superconcentrators (superconcentrators are directed graphs with strong connectivity properties, see 12 for a formal definition).

Recall that a circuit is *linear* if it has no multiplication gates. We remark here that for any matrix  $M$  the depth two linear circuit complexity of the linear transformation  $M\mathbf{x}$  and the *bilinear* formula complexity of  $\mathbf{y}^T M\mathbf{x}$  are essentially the same (see [15], Equation 2). On the other hand, unlike circuits, it is *not* known whether formulas can be bilinearized efficiently, so the lower bounds on bilinear formula complexity do not imply lower bounds on general formula complexity. As noted by Nisan and Wigderson ([15], Section 1.2), no superlinear lower bounds on the formula complexity of explicit bilinear forms are known. In this paper, we prove the first such lower bounds. To get around the bilinearization obstacle for formulas, we work with planar circuits, a more general object than formulas. We observe that it is possible to bilinearize planar circuits efficiently, and that it is also possible to prove superlinear ( $\Omega(n \log n)$ ) lower bounds for planar, bilinear circuits.

## 1.1 Our results

In this article, we study *planar arithmetic circuits*, i.e., unbounded depth unbounded fan-in arithmetic circuits whose underlying graph is *planar*. Recall that a graph is said to be *planar* if it can be drawn on the plane without edge crossings. Note that every formula is a planar circuit as every tree is a planar graph. We begin by observing (in Lemma 9) that any arithmetic circuit can be converted into an equivalent planar arithmetic circuit with at most quadratic blowup in size by introducing *crossover gadgets*.

Our main result is a superlinear lower bound on the size of planar arithmetic circuits computing a class of explicit bilinear forms. In the sequel, we say a matrix  $M \in \mathbb{F}^{n \times n}$  is totally regular if all of its square minors are non-singular.

**Theorem 1.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix and  $\mathbf{x}$  and  $\mathbf{y}$  be vectors of  $n$  variables each. Then, any planar arithmetic circuit computing the bilinear form  $\mathbf{y}^T M\mathbf{x}$  has size  $\Omega(n \log n)$ .*

Over infinite fields, there exist *explicit* totally regular matrices (such as Cauchy matrices) whose entries are uniformly computable in time polynomial in the dimension of the matrix, so our bound applies to a class of explicit bilinear forms. It is important to note that our lower bound works for unbounded-depth and unbounded fanin planar arithmetic circuits. Since formulas are a subclass of planar circuits, Theorem 1 implies a superlinear lower bound on the size of formulas computing certain explicit bilinear forms. Previously, no such lower bounds were known (see [15], Section 1.2).

As a corollary, we get the following separation between circuits and planar circuits:

**Corollary 2.** *Over any infinite field  $\mathbb{F}$ , there exists an infinite family  $\{M_n\}_{n \geq 1}$  of totally regular matrices (where  $M_n \in \mathbb{F}^{n \times n}$ ) such that the bilinear form  $\mathbf{y}^T M_n \mathbf{x}$  can be computed by an arithmetic circuit of size  $O(n)$  but any planar arithmetic circuit computing it requires size  $\Omega(n \log n)$ .*

Next, we consider *read-once planar arithmetic circuits* which are a special class of planar arithmetic circuits where every variable appears as a leaf at most once. In particular, we prove quadratic lower bounds for read-once planar circuits computing bilinear forms:

**Theorem 3.** *Let  $M \in \mathbb{F}^{n \times n}$  be a totally regular matrix and  $\mathbf{x}$  and  $\mathbf{y}$  be vectors of  $n$  variables each. Then, any read-once planar arithmetic circuit computing the bilinear form  $\mathbf{y}^T M \mathbf{x}$  has size  $\Omega(n^2)$ .*

It is not hard to see that this bound is tight. Furthermore, by the planarization procedure in Lemma 9, a super-quadratic lower bound for read-once planar arithmetic circuits for an explicit family of polynomials implies a superlinear lower bound for general arithmetic circuits (for the same family), a long-standing open problem in algebraic complexity theory.

Using explicit superconcentrators of depth two, we get the following separations between circuits, read-once planar circuits and arithmetic formulas:

- Corollary 4.**
1. *Over any infinite field  $\mathbb{F}$ , there exists an explicit family  $\{f_n\}_{n \geq 1}$  of degree 4,  $n^{1+o(1)}$ -variate polynomials such that  $f_n$  is computable by a circuit of size  $n^{1+o(1)}$  but any read-once planar circuit for  $f_n$  requires size  $\Omega(n^2)$ .*
  2. *Over any infinite field  $\mathbb{F}$ , there exists an explicit family  $\{f_n\}_{n \geq 1}$  of degree 4,  $n^{1+o(1)}$ -variate polynomials such that  $f_n$  is computable by an arithmetic formula of size  $n^{1+o(1)}$  but any read-once planar circuit for  $f_n$  requires size  $\Omega(n^2)$ .*
  3. *The polynomial  $x_1^n + \dots + x_n^n$  has a planar arithmetic circuit of size  $O(n \log n)$  but any formula computing it requires size  $\Omega(n^2)$ .*

*Separations 2, 3 together imply that formula complexity and read-once circuit complexity are incomparable measures in the arithmetic setting.*

It is easy to see that any algebraic branching program can be converted into an equivalent arithmetic circuit computing the same polynomial without much blowup in size. We observe that this can be done while preserving planarity.

Thus,  $\Omega(n \log n)$  lower bound for planar circuits also extends to planar algebraic branching programs (ABPs where the underlying DAG is planar). In fact, it holds for unlayered planar ABPs:

**Theorem 5.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix and  $\mathbf{x}$  and  $\mathbf{y}$  be vectors of  $n$  variables each. Then, any (not necessarily layered) planar ABP computing the bilinear form  $\mathbf{y}^T M \mathbf{x}$  has size  $\Omega(n \log n)$ .*

All our lower bounds for bilinear forms use variants of the *planar separator theorem* (Lipton and Tarjan[13]) combined with a rank argument. One version of the planar separator theorem says that any  $n$ -vertex planar graph can be partitioned into two disconnected components of size  $\geq n/3$  by removing a small ( $\leq 2\sqrt{2}\sqrt{n}$ ) number of vertices. The rough idea is that the existence of a small separator induces algebraic dependencies among the polynomials computed in the circuit.

We now turn our attention to planar circuits computing multiple linear forms. In a seminal work, Lipton and Tarjan[12] applied the planar separator theorem to (among other things) obtain a quadratic lower bound on the size of planar superconcentrators. Valiant[23] had already observed that if  $M$  is totally regular then any read-once circuit computing the linear transformation  $M\mathbf{x}$  must be an  $n$ -superconcentrator. This gives a quadratic lower bound on the read-once planar circuit complexity of  $A\mathbf{x}$ . First we relax the read-once condition and show an  $\Omega(n^{4/3})$  lower bound on the size of planar circuits computing  $A\mathbf{x}$  for any totally regular  $A$ .

**Theorem 6.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix and  $\mathbf{x}$  be a vector of  $n$  variables. Then, any planar circuit that computes  $M\mathbf{x}$  requires size  $\Omega(n^{4/3})$ .*

Next, we consider multi-output formulas. A formula is said to compute polynomials  $f_1, \dots, f_t$  if there exist  $t$  gates in it that compute  $f_1, \dots, f_t$  resp.

**Theorem 7.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix and  $\mathbf{x}$  be a vector of  $n$  variables. Then, any multi-output formula for computing  $M\mathbf{x}$  requires size  $\Omega(n^2 / \log n)$ .*

Finally, we look at the implications of these lower bounds on the complexity of first order partial derivatives. In particular, we note that a result analogous to that of Baur and Strassen [3] cannot hold for formulas and planar circuits, while it does hold for read-once planar circuits. The Baur-Strassen theorem says that if a polynomial  $f$  has a (fan-in 2) circuit of size  $s$  then there is a (fan-in 2) circuit of size  $O(s)$  computing all first order partial derivatives of  $f$ .

## 1.2 Prior work on lower bounds for bilinear forms

As mentioned earlier, no superlinear lower bounds on the size of general circuits computing explicit bilinear forms are known. This motivates the study of the complexity of bilinear forms in more restricted models such as arithmetic formulas, bilinear formulas and bounded coefficient models of computation. In [15], Nisan and Wigderson defined the notion of a bilinear formula and noticed that bilinear formula complexity of the bilinear form  $\mathbf{y}^T A\mathbf{x}$  and depth two linear circuit complexity of the  $A\mathbf{x}$  are equivalent notions ([15], Equation 2). This connection means that lower bounds for depth two linear circuits apply to bilinear formulas. It remains open whether formulas can be bilinearized efficiently.

Further, Nisan and Wigderson [15] looked at the bilinear, bounded coefficient formula complexity of computing certain bilinear forms and, using spectral techniques, proved lower bounds of the form  $\Omega(n^{1+\delta})$  for this model. In a breakthrough paper, Raz [17] extended the techniques in [15] vastly and proved an  $\Omega(n^2 \log n)$  lower bound on the bounded coefficient circuit complexity of matrix multiplication.

## 1.3 Related work on planar boolean circuits

Planar circuits are well studied in the boolean setting. Lipton and Tarjan[13] initiated the study of planar boolean circuits by proving quadratic lower bounds on the size of read-once planar circuits computing multi-output boolean functions.

The read-once restriction was first relaxed by Savage [20]. He showed superlinear ( $n^{1+\delta}$  for various constants  $\delta$ ) lower bounds on the planar circuit complexity of various multi-output boolean functions.

The case of single output functions turned out to be harder and lower bounds for these were first proved by Savage [19] (in the read-once case, an  $\Omega(n^2)$  lower bound) and Turan [22] (in the general case, an  $\Omega(n \log n)$  lower bound). In [22], Turan also showed that read-once planar circuit complexity and formula complexity are incomparable measures in the boolean world.

All known lower bounds on the planar circuit complexity of boolean functions use variants of the planar separator theorem combined with “crossing sequence arguments”. These arguments make use of the fact that on any input, each wire of the circuit will either carry a zero or a one, specifically the number of possibilities is a constant. This is obviously not true in the case of arithmetic circuits over large fields, so the crossing sequence arguments do not carry over. To get around this, we use *rank* based methods. We are able to prove lower bounds for bilinear forms (which are degree two polynomials) using rank based methods whereas the boolean functions in previous works have degree (in the sense of [14]) linear in the number of variables.

## 2 Preliminaries

In this section, we formally introduce all algebraic models of computation considered in this paper and some other graph-theoretic preliminaries that are crucial to our proofs.

**Definition 1** (Arithmetic Circuits). Let  $\mathbb{F}$  be a field. An *arithmetic circuit*  $\Phi$  over  $\mathbb{F}$  is a directed acyclic graph with vertices of in-degree zero or two. A vertex of out-degree 0 is called an output gate. A vertex of in-degree zero is called an input gate and is labeled by elements from  $X \cup \mathbb{F}$ . Every other gate is labeled either by  $+$  or  $\times$ . Every gate in  $\Phi$  naturally computes a polynomial in  $\mathbb{F}[X]$  and the polynomial(s) computed by  $\Phi$  is (are) the polynomial(s) computed at the output gate(s). We allow edges (wires) to be labelled by field elements, these simply scale the polynomial. The *size* of  $\Phi$  is the number of gates in  $\Phi$  and *depth* of  $\Phi$  is the length of the longest path from an input gate to an output gate. For a polynomial  $f$ , we let  $C(f)$  denote the size of the smallest arithmetic circuit computing  $f$ .

**Definition 2** (Arithmetic Formula). An *arithmetic formula* is an arithmetic circuit where the underlying undirected graph is a tree. For a polynomial  $f$ , we let  $L(f)$  denote the size of the smallest arithmetic formula computing  $f$ . We say a formula computes polynomials  $f_1, \dots, f_k$  if there exist  $k$  nodes in the formula that compute  $f_1, \dots, f_k$ . Note that a formula that computes multiple polynomials may have gates with fan-out  $\geq 2$ . The restriction is that there should be no cycles in the underlying undirected graph. For polynomials  $f_1, \dots, f_k$ , we let  $L(f_1, \dots, f_k)$  denote the size of the smallest formula computing  $f_1, \dots, f_k$ .

**Definition 3** (Planar arithmetic circuits). A *planar arithmetic circuit* is an arithmetic circuit whose underlying graph is planar. For a polynomial  $f$ , we let  $C_p(f)$  denote the size of the smallest planar arithmetic circuit computing  $f$ .

**Definition 4** (Read-once planar arithmetic circuits). A *read-once planar arithmetic circuit* is a planar arithmetic circuit in which each variable labels at most one input gate. For a polynomial  $f$ , we let  $C_p^r(f)$  denote the size of the smallest read-once planar circuit computing  $f$ .

**Definition 5** (Algebraic Branching Programs(ABPs)). An *Algebraic Branching Program* (ABP)  $P$  is a layered directed acyclic graph with one vertex  $s$  of in-degree zero (source) in the first layer and one *sink* vertex  $t$  of out-degree zero (sink) in the last layer. Every edge in  $e$  in  $P$  is labelled by an element in  $X \cup \mathbb{F}$ . Let weight of a path be the product of its edge labels and the polynomial computed by an ABP  $P$  is the sum of weights of all  $s$  to  $t$  paths in  $P$ . The *size* of an ABP is the number of vertices in it. For a polynomial  $f$ , we let  $A(f)$  denote the size of the smallest ABP computing  $f$ .

**Definition 6** (Planar ABPs). A *Planar ABP* is an ABP whose underlying graph is planar. For a polynomial  $f$ , we let  $A_p(f)$  denote the size of the smallest planar ABP computing  $f$ .

**Definition 7** (Linear forms). By a *linear form* we mean a homogeneous degree one polynomial.

**Definition 8** (Bilinear forms). Let  $\mathbb{F}$  be a field and let  $\mathbf{x} = (x_1, \dots, x_n)^T$ ,  $\mathbf{y} = (y_1, \dots, y_n)^T$  be two vectors of variables. The *bilinear form* defined by a matrix  $M \in \mathbb{F}^{n \times n}$  is the polynomial  $f(\mathbf{x}, \mathbf{y}) = \mathbf{y}^T M \mathbf{x}$ . We say the bilinear form  $\mathbf{y}^T M \mathbf{x}$  has rank  $r$  if  $\text{rank}(M) = r$ . We say that a family of bilinear forms  $\{\mathbf{y}^T M_n \mathbf{x}\}_{n \geq 0}$  (where  $M_n \in \mathbb{F}^{n \times n}$ ) is explicit if there is an algorithm that on input  $n$  in unary outputs all entries of  $M_n$  in  $\text{poly}(n)$  time.

**Definition 9** (Bilinear circuits). A bilinear circuit with inputs  $\{x_1, \dots, x_n\}$  and  $\{y_1, \dots, y_n\}$  is an arithmetic circuit in which every product gate has exactly two children, one of which computes a linear form in the  $x$  variables and the other computes a linear form in the  $y$ 's. Every bilinear form  $f$  is clearly computed by some bilinear circuit, and we let  $C^b(f)$  denote the size of the smallest bilinear circuit computing  $f$ . It is well known and easy to show that for any bilinear form  $f$ ,  $C^b(f) = O(C(f))$ . For a bilinear form  $f$ , we let  $C_p^b(f)$  denote the size of the smallest circuit computing  $f$  that is both planar and bilinear.

**Definition 10** (Bilinear formulas). A *bilinear formula* with inputs  $\{x_1, \dots, x_n\}$  and  $\{y_1, \dots, y_n\}$  is an arithmetic formula in which every product gate has exactly two children, one of which computes a linear form in the  $x$  variables and the other computes a linear form in the  $y$ 's. Without loss of generality we can assume [15] that a bilinear formula for  $\mathbf{y}^T M \mathbf{x}$  is a sum of products of two linear forms, one in the  $x$  variables and one in the  $y$  variables, i.e, it has the following structure:

$$\mathbf{y}^T M \mathbf{x} = \sum_{i=1}^k \mathbf{y}^T (u_i v_i^T) \mathbf{x}$$

Equivalently, the bilinear formula above gives the factorization  $M = UV$  of  $M$  where  $u_i$ 's are the columns of  $U$  and  $v_i$ 's are the rows of  $V$ . The size of such a bilinear formula is the number of non-zero entries in all the vectors  $u_i, v_i$ . Every bilinear form  $f$  is clearly computed by some bilinear formula and we let  $L^b(f)$  denote the size of the smallest bilinear formula computing  $f$ . It is not known whether  $L^b(f) = O(L(f))$  holds for every bilinear form  $f$ .

**Definition 11** (Totally regular matrix). Let  $\mathbb{F}$  be a field. We say that a matrix  $M \in \mathbb{F}^{n \times n}$  is *totally regular* if every square minor of  $M$  is non-singular. We say that a family  $\{A_n\}_{n \geq 0}$  of totally regular matrices (where  $A_n \in \mathbb{F}^{n \times n}$ ) is *explicit* if there is an algorithm that takes  $n$  as input in unary and outputs the entries of  $A_n$  in  $\text{poly}(n)$  time.

**Definition 12** (Superconcentrators). An  $n$ -*superconcentrator* is a directed acyclic graph  $G = (V, E)$  with  $n$  inputs  $I_1, \dots, I_n$  and  $n$  outputs  $O_1, \dots, O_n$  such that  $\forall k \in [n]$ , for all subsets  $I' \subseteq I$  such that  $|I'| = k$  and all subsets  $O' \subseteq O$  such that  $|O'| = k$ , there exist  $k$  vertex disjoint paths from  $I'$  to  $O'$ . We say  $G$  has depth  $d$  if the longest path in  $G$  has length  $d$ . We define the size of such a graph to be the number of edges in it. We say that a family of  $n$ -superconcentrators (one for each  $n$ , the  $n$ -th member of the family must have  $n$  inputs and outputs) is *explicit* if there is an algorithm that  $\forall n$ , outputs the  $n$ -th superconcentrator in the family in  $\text{poly}(n)$  time.

## 2.1 Preliminary observations about planar circuits

First we note that in a planar circuit, we can assume without loss of generality that the fan-in and fan-out of every gate is at most two:

**Lemma 8.** *Let  $\Phi$  be a planar circuit of size  $s$  (i.e.,  $\Phi$  has  $s$  gates) computing  $f \in \mathbb{F}[x_1, \dots, x_n]$ . Then there exists another planar circuit  $\Phi'$  computing  $f$  such that  $\text{size}(\Phi') \leq 10s$  and every gate in  $\Phi'$  has fan-in and fan-out  $\leq 2$ .*

*Proof.* For every gate  $g$  in  $\Phi$  with fan-in  $r$  (resp., fan out) larger than 2, replace the incoming (resp., outgoing) wires with a balanced binary tree (with  $r$  leaves) with all gates of the same type as  $g$ . The number of new gates added is at most thrice the number of wires in  $\Phi$ . Since  $\Phi$  is planar, the number of wires in  $\Phi$  is at most  $3s - 6$  and so the size of the new circuit  $\Phi' \leq 10s$ .  $\square$

Lipton and Tarjan [12] observe that it is possible to *planarize* boolean circuits while incurring at most a quadratic blow-up in size. We note here that it is also possible to planarize arithmetic circuits in a similar way. We planarize a circuit  $\Phi$  by fixing an embedding of the graph of  $\Phi$  and introducing a gadget at each edge crossing:

**Lemma 9.** *Let  $\Phi$  be a fan-in 2 arithmetic circuit of size  $s$  computing a polynomial  $f \in \mathbb{F}[x_1, \dots, x_n]$ . Then there exists a fan-in 2 planar arithmetic circuit  $\Phi'$  of size  $O(s^2)$  computing  $f$ .*

*Proof.* Fix an embedding  $E$  of  $\Phi$  in the plane. At every edge crossing in  $E$ , introduce the *crossover gadget* (of small enough size) shown in Figure 1. The number of edge crossings in  $E$  is at most  $\binom{\text{number of wires in } \Phi}{2}$ . Since the circuit has fan-in 2 the number of wires in  $\Phi \leq 2s$  and hence  $\text{size}(\Phi') = O(s^2)$ .  $\square$



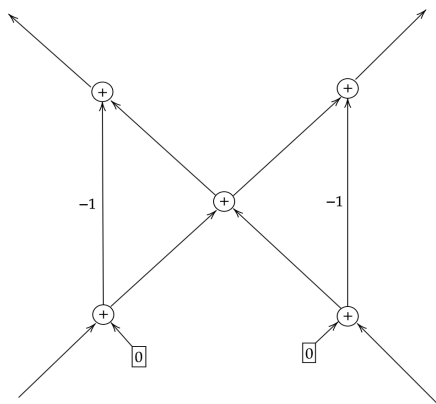


Figure 1: Crossover gadget

In the following subsection, we list some partition lemmas for graphs that will be crucial to our lower bound arguments:

## 2.2 Some partition lemmas for planar graphs

We begin with the well-known *planar separator theorem* of Lipton and Tarjan [13] which is used to prove quadratic lower bounds for read-once planar circuits:

**Theorem 10** (Lipton and Tarjan [13]). *Let  $G = (V, E)$  be a planar graph and  $w : V \rightarrow [0, 1]$  be a weight function such that  $\sum_{v \in V} w(v) \leq 1$ . Then there exists a partition  $(A, B, C)$  of  $V$  such that the sum of the weights of the vertices in  $A$  as well as the sum of the weights of the vertices in  $B$  is at most  $2/3$ ,  $|C| \leq 2\sqrt{2}\sqrt{|V|}$ , and all paths from  $A$  to  $B$  contain a vertex from  $C$ .  $\square$*

The following theorem by Turan [22] is a generalization to planar graphs of a separator lemma for trees proved by Babai et al [2], which is itself a generalization to trees of a separator lemma for sequences by Alon and Maasst [1]. It is useful for obtaining lower bounds for planar circuits:

**Theorem 11** (Turan [22]). *Let  $Z = \{z_1, \dots, z_s\}$  and  $Z' = \{z'_1, \dots, z'_s\}$  be disjoint sets and let  $G = (V, E)$  be a planar graph, some vertices of which are labelled by elements from  $Z \cup Z'$  so that each label occurs at most  $k$  times. Then there are subsets  $Z_0 \subseteq Z$ ,  $Z'_0 \subseteq Z'$  and  $V^* \subseteq V$  such that the following conditions hold:*

1.  $|Z_0| = |Z'_0| \geq s/9^k$
2.  $|V^*| \leq 450k\sqrt{|V|}$
3. After deleting  $V^*$  none of the remaining components contain labels from both  $Z_0$  and  $Z'_0$ .

For lower bounds on planar arithmetic circuits computing multiple outputs, we need the following partition lemma from [20]. It partitions a planar graph into multiple parts each having their own small separators. This is achieved by applying a version of the Lipton-Tarjan planar separator theorem multiple times.

**Theorem 12** (Savage [20]). *Let  $G = (V, E)$  be a planar graph,  $V' \subseteq V$  be a subset of its vertices and let  $1 \leq p \leq |V'|$ . Then there exists a partition  $V_1, \dots, V_p$  of  $V$  such that the following conditions hold:*

1. For all  $i \in [p]$ ,  $\frac{|V'|}{4p} \leq |V' \cap V_i| \leq \frac{4|V'|}{p}$
2. For all  $i \in [p]$  there exists a separator  $S_i$  such that  $|S_i| \leq 60\sqrt{|V|}$  and no edge joins  $V_i$  and  $V \setminus (V_i \cup S_i)$ .

If we restrict to forests instead of the more general planar graphs, we can get a statement analogous to Theorem 12 with much smaller (logarithmic) separators and we state this as our next lemma. We defer the proof of the following Lemma to Section 4.

**Lemma 13.** *Let  $F = (V, E)$  be a forest,  $V' \subseteq V$  be a subset of its vertices and let  $1 \leq p \leq |V'|$ . Then there exists a partition  $(V_1, \dots, V_p)$  of  $V$  such that the following conditions hold:*

1. For all  $i \in [p]$ ,  $\frac{|V'|}{3p} \leq |V' \cap V_i| \leq \frac{3|V'|}{p}$
2. For all  $i \in [p]$  there exists a set  $S_i$  such that  $|S_i| = O(\log(|V'|))$  and no edge joins  $V_i$  and  $V \setminus (V_i \cup S_i)$ .

### 3 Lower bounds for bilinear forms

In the following subsection, we present an  $\Omega(n \log n)$  lower bound on the planar circuit complexity (and hence, formula complexity) of bilinear forms defined by  $n \times n$  totally regular matrices. First, we note that any planar circuit computing a bilinear form can be converted into a planar bilinear circuit computing the same bilinear form with only a constant blowup in size. Next, by using the partition theorem of Turan (Theorem 11) and a careful rank argument we prove the desired lower bound.

#### 3.1 Lower Bounds for Planar Arithmetic Circuits

We begin with *bilinearization* of planar arithmetic circuits:

**Lemma 14.** *For any bilinear form  $f = \mathbf{y}^T M \mathbf{x}$ ,  $C_p^b(f) \leq 3000C_p(f)$ .*

*Proof.* Let  $\Phi$  be a planar circuit of size  $s$  for  $f$ . Assume wlog that the fan-in and fan-out of each gate in  $\Phi$  is  $\leq 2$ . Let  $E$  be a planar embedding of  $\Phi$  and let  $v$  be a gate in  $\Phi$ . The polynomial computed at  $v$  (we denote this by  $f^v$ ) can be decomposed as

$$f^v = f_x^v + f_y^v + f_{xy}^v + f_r^v$$

where  $f_x^v$  is the  $x$ -linear component of  $f^v$  (sum of all monomials in  $f^v$  of the form  $\alpha_i x_i$ ),  $f_y^v$  is the  $y$ -linear component of  $f^v$  and  $f_{xy}^v$  is the bilinear component of  $f^v$  (i.e., sum of monomials of the form  $\alpha_{ij} x_i y_j$ ).  $f_r^v$  is the rest of  $f^v$ . We now construct a bilinear circuit  $\Psi$  and planarize it to get a planar, bilinear circuit  $\Psi'$  computing  $f$ :

For every gate  $v$  in  $\Phi$  we wish to add 3 gates in  $\Psi$ , denoted by  $v_x, v_y$  and  $v_{xy}$ , that compute  $f_x^v, f_y^v$  and  $f_{xy}^v$  respectively:

- If  $v$  is a leaf with label  $x_i$  then add 3 leaves to  $\Psi$ :  $v_x$  labelled by  $x_i$  and  $v_y, v_{xy}$  labelled by 0. If  $v$  is a leaf with label  $y_i$ , again add 3 leaves to  $\Psi$ :  $v_y$  labelled by  $y_i$  and  $v_x, v_{xy}$  labelled by 0. If  $v$  is leaf with label  $\alpha \in \mathbb{F}$ , add 3 leaves  $v_x, v_y, v_{xy}$  all labelled by 0.
- If  $v$  is a sum gate in  $\Phi$  with children  $u, w$  (say  $v = \alpha u + \beta w$ ) then

$$\begin{aligned} f_x^v &= \alpha f_x^u + \beta f_x^w \\ f_y^v &= \alpha f_y^u + \beta f_y^w \\ f_{xy}^v &= \alpha f_{xy}^u + \beta f_{xy}^w \end{aligned}$$

By induction, the six gates of  $\Psi$  corresponding to  $u, w$  compute the intended polynomials. Now define  $v_x, v_y$  and  $v_{xy}$  to be sum gates with children  $(u_x, w_x), (u_y, w_y)$  and  $(u_{xy}, w_{xy})$  respectively. Label the two incoming wires to each of these gates by  $\alpha$  and  $\beta$ . By doing this, we have made the fan-outs of  $u_x, u_y, u_{xy}, w_x, w_y, w_{xy}$  all equal to 2.

- If  $v$  is a product gate in  $\Phi$  with children  $u, w$  then  $\exists \alpha, \beta \in \mathbb{F}$  such that

$$\begin{aligned} f_x^v &= \alpha \cdot f_x^u + \beta \cdot f_x^w \\ f_y^v &= \alpha \cdot f_y^u + \beta \cdot f_y^w \\ f_{xy}^v &= \alpha \cdot f_{xy}^u + \beta \cdot f_{xy}^w + f_x^u \cdot f_y^w + f_x^w \cdot f_y^u \end{aligned}$$

By induction, the six gates of  $\Psi$  corresponding to  $u, w$  compute the intended polynomials. Now define  $v_x, v_y$  to be sum gates with children  $(u_x, w_x)$  and  $(u_y, w_y)$  respectively and label the incoming wires by  $\alpha, \beta$ . Next, introduce two product gates  $p_v^1, p_v^2$  with children  $(u_x, w_y)$  and  $(u_y, w_x)$  respectively and add  $p_v^1, p_v^2$  using a sum gate  $s_v^1$ . Introduce another sum gate  $s_v^2$  computing  $\alpha u_{xy} + \beta w_{xy}$  and define  $v_{xy}$  to be a sum gate:  $v_{xy} = s_v^1 + s_v^2$ . Overall, for each product gate  $v$  in  $\Phi$ , we have added 7 gates in  $\Psi$ . Notice that by doing this we have made the fan-outs of  $u_x, w_x, u_y, w_y$  equal to 4 and the fan-outs of  $u_{xy}, w_{xy}$  equal to 2.

First note that  $\text{size}(\Psi) \leq 7 \cdot \text{size}(\Phi)$ . Also, note that if  $o$  is the output of  $\Phi$  computing the bilinear form  $f$ , then the gate  $o_{xy}$  in  $\Psi$  also computes  $f$ . Now from  $E$ , we produce an

embedding  $E'$  of  $\Psi$  in the plane as follows: For every gate  $v$  of  $\Phi$ , consider a  $\delta$  neighbourhood of  $E(v)$  for a value of  $\delta$  to be specified later.  $E'$  embeds the  $\leq 7$  new gates of  $\Psi$  corresponding to  $v$  inside this neighbourhood. For every wire  $w$  of  $\Phi$ , we have  $\leq 5$  corresponding wires of  $\Psi$ . Outside the  $\text{size}(\Phi)$  many disjoint  $\delta$  neighbourhoods,  $E'$  embeds these wires parallel to each other and to the embedding  $E(w)$  of the original wire  $w$ , sufficiently close together so that there are no crossings outside the neighbourhoods. Make  $\delta$  small enough so that this is possible.

It is clear that the only crossings in  $E'$  are inside the  $\delta$ -neighbourhoods. The number of crossings per neighbourhood is at most the number of wires in the neighbourhood choose two. Since each gate has fan-in and fan-out  $\leq 4$ , the number of wires in each neighbourhood is at most 28, so the number of crossings is at most 378. Eliminate each crossing using the crossover gadget in Lemma 9 to get a planar circuit  $\Psi'$  that computes  $f$ . Since each gadget introduces 7 new gates,  $\text{size}(\Psi') \leq 2646 \cdot \text{size}(\Phi) + \text{size}(\Psi) \leq 3000 \cdot \text{size}(\Phi)$ . Also, note that if  $\Phi$  is a read-once planar circuit then so is  $\Psi'$ .  $\square$

We state some simple facts about circuit  $\Psi'$  that are evident from construction in Lemma 14:

1. Each product gate in  $\Psi'$  computes a bilinear form of rank 1 i.e., it computes the product of a linear form in  $x$ -variables and a linear form in  $y$ -variables.
2. The bilinear form computed by the output of  $\Psi'$  is a linear combination of the product gates in  $\Psi'$ .

We now turn our attention to the proof of the main theorem in this paper:

**Theorem 15.** *Let  $M \in \mathbb{F}^{n \times n}$  be a totally regular matrix. Then,  $L(\mathbf{y}^T M \mathbf{x}) \geq C_p(\mathbf{y}^T M \mathbf{x}) = \Omega(n \log n)$ .*

*Proof.* Let  $\Psi$  be a planar circuit computing the bilinear form  $\mathbf{y}^T M \mathbf{x}$ . Let  $X$  denote the set of  $x$  variables  $\{x_1, \dots, x_n\}$  and  $Y$  denote the set of  $y$  variables  $Y = \{y_1, \dots, y_n\}$ . Assume for the sake of contradiction that  $\text{size}(\Psi) \leq \frac{1}{60000} n \log n$ . Then, by Lemma 14 there exists a planar, bilinear circuit  $\Phi$  computing  $\mathbf{y}^T M \mathbf{x}$  such that  $\text{size}(\Phi) \leq \frac{1}{200} n \log n$ . Let the planar graph underlying  $\Phi$  be  $G = (V, E)$  (where  $|V| \leq \frac{1}{200} n \log n$ ). It is easy to see that there exist subsets  $X_0 \subseteq X$  and  $Y_0 \subseteq Y$  such that  $|X_0| = |Y_0| \geq n/2$  and every variable in  $X_0 \cup Y_0$  appears  $\leq \frac{1}{100} \log n$  times as an input in  $\Phi$  (otherwise  $|V| > \frac{1}{200} n \log n$ ).

Applying Theorem 11 to  $G$  with  $Z = X_0$  and  $Z' = Y_0$ , we obtain  $X_1 \subseteq X_0$ ,  $Y_1 \subseteq Y_0$  and  $V^* \subseteq V$  such that the following conditions hold:

1.  $|X_1| = |Y_1| \geq \frac{n/2}{9^{\frac{1}{100} \log n}} = \Omega(n^{0.96})$
2.  $|V^*| = O(\sqrt{n}(\log n)^{3/2})$
3. Upon deleting  $V^*$  from  $G$ , no component in the resulting graph contains labels from both  $X_1$  and  $Y_1$ .

Setting all variables in  $\Phi$  outside  $X_1 \cup Y_1$  to zero, we obtain another planar, bilinear circuit  $\Phi'$  that computes the bilinear form  $\mathbf{y}'^T M' \mathbf{x}'$  where  $M'$  is the minor of  $M$  whose rows are indexed by  $Y_1$  and columns by  $X_1$ ,  $\mathbf{x}'$  is the vector of  $X_1$  variables and  $\mathbf{y}'$  is the vector of  $Y_1$  variables. Since  $M$  is totally regular,  $\text{rank}(M') = |X_1| = |Y_1| = \Omega(n^{0.96})$ . But we can show that  $\text{rank}(M') \leq 3|V^*|$ , which leads to a contradiction:

**Claim 16.**  $\text{rank}(M') \leq 3|V^*|$

*Proof of Claim 16:* As mentioned before, the output of  $\Phi'$  is a linear combination of the product gates in it. Also, each product gate  $g_i$  in  $\Phi'$  computes a product  $l_i(\mathbf{x}') \times l'_i(\mathbf{y}') = \mathbf{y}'^T A_i \mathbf{x}'$  where  $l_i(\mathbf{x}')$  and  $l'_i(\mathbf{y}')$  are linear forms in the  $X_1$  and  $Y_1$  variables respectively and  $A_i$  is a matrix with  $\text{rank} \leq 1$ . Next, we partition the product gates of  $\Phi'$  into three types, those that belong to components of  $V \setminus V^*$  that do not contain inputs labelled by  $Y_1$  variables, those that belong to components of  $V \setminus V^*$  that do contain inputs labelled by  $Y_1$  variables and those that belong to  $V^*$ . We show that the rank of any linear combination of the product gates inside a particular partition is at most  $|V^*|$ .

To that end, let  $G_1, \dots, G_k$  be the components of  $V \setminus V^*$  not containing inputs labelled by  $Y_1$ , let  $g_1 = l_1(\mathbf{x}') \times l'_1(\mathbf{y}'), \dots, g_t = l_t(\mathbf{x}') \times l'_t(\mathbf{y}')$  be the product gates of  $\Phi'$  in  $G_1 \cup \dots \cup G_k$  and let  $l_1^v(\mathbf{y}'), \dots, l_m^v(\mathbf{y}')$  be all the linear forms in the  $Y_1$  variables computed at gates (including leaves) of  $\Phi'$  that lie in  $V^*$ . Clearly,  $m \leq |V^*|$ . The following claim shows that each  $l'_i(\mathbf{y}')$  ( $1 \leq i \leq t$ ) is a linear combination of  $l_1^v(\mathbf{y}'), \dots, l_m^v(\mathbf{y}')$ .

**Sub-Claim 17.**  $\forall i \in [t], \exists \beta_{i,1}, \dots, \beta_{i,m} \in \mathbb{F}$  such that  $l'_i(\mathbf{y}') = \sum_{j=1}^m \beta_{i,j} l_j^v(\mathbf{y}')$ .

*Proof of Sub-Claim 17:* First we observe some facts about the circuit  $\Phi'$  that follow from the bilinearization procedure in Lemma 14: If a gate  $g$  in  $\Phi'$  computes a linear form in the  $Y_1$  variables then

1.  $g$  is a sum gate.
2. Every gate in the subcircuit rooted at  $g$  is also a sum gate and computes a linear form in the  $Y_1$  variables.
3. Every leaf in the subcircuit rooted at  $g$  is labelled by 0 or by  $y_i$  for some  $y_i \in Y_1$ .

Now consider a product gate  $g_i$  of  $\Phi'$  in  $G_1 \cup \dots \cup G_k$ . One of the children of  $g_i$  (say  $h$ ) computes  $l'_i(\mathbf{y}')$  which is a linear form in the  $Y_1$  variables.  $h$  either lies in  $V^*$  or it lies in the same component (say  $H \in \{G_1, \dots, G_k\}$ ) of  $G \setminus V^*$  as  $g_i$ . If  $h$  lies in  $V^*$  then the claim is trivially true for that  $i$ .

We now prove that if a gate  $h \in H$  computes a linear form in  $Y_1$  variables then  $\exists \beta_1, \dots, \beta_m \in \mathbb{F}$  such that  $h = \sum_j \beta_j l_j^v(\mathbf{y}')$ . We allow  $h$  to be a leaf. The proof is by induction on the size of the subcircuit rooted at  $h$ . For the base case, let  $h$  be a leaf. Observe that  $h$  cannot be a leaf labelled by a non-zero constant by Observation 3, it cannot be a leaf labelled by an  $X_1$  variable also by observation 3, and it cannot be a leaf labelled by a  $Y_1$  variable because  $h \in H \in \{G_1, \dots, G_k\}$ . So  $h$  is a leaf labelled by 0, and the claim is obviously true, 0 is the trivial linear combination of  $l_1^v(\mathbf{y}'), \dots, l_m^v(\mathbf{y}')$ .

For the inductive step, let  $h$  be any non-leaf gate in  $H$  computing a linear form in the  $Y_1$  variables. By observation 1,  $h$  must be a sum gate. Suppose  $h = \beta_1 h_1 + \beta_2 h_2$  with predecessors  $h_1, h_2$ . If  $h_1$  is in  $V^*$  then by observation 2,  $h_1 = l_j^v(\mathbf{y}')$  for some  $j$ . Otherwise,  $h_1 \in H$ . If it is a leaf then as before it must be labelled by 0. Again, 0 is the trivial linear combination of  $l_1^v(\mathbf{y}'), \dots, l_m^v(\mathbf{y}')$ . If  $h_1$  not a leaf, by observation 2 it satisfies the inductive hypothesis:  $\exists \gamma_1, \dots, \gamma_m \in \mathbb{F}$  such that  $h_1 = \sum_j \gamma_j l_j^v(\mathbf{y}')$ . In all cases,  $h_1$  (and symmetrically  $h_2$ ) is a linear combination of  $l_1^v(\mathbf{y}'), \dots, l_m^v(\mathbf{y}')$ , and so is  $h$ . (End of proof of Sub-Claim 17)  $\square$

Using Claim 17 we can show that a linear combination of  $g_1, \dots, g_t$  cannot have large rank:

**Sub-Claim 18.** *For any  $\alpha_1, \dots, \alpha_t \in \mathbb{F}$  let  $\mathbf{y}'^T M_1 \mathbf{x}' = \sum_{i=1}^t \alpha_i g_i$  be a linear combination of  $g_1, \dots, g_t$ . Then  $\text{rank}(M_1) \leq |V^*|$ .*

*Proof of Sub-Claim 18:* Consider the linear combination  $\mathbf{y}'^T M_1 \mathbf{x}' = \sum_{i=1}^t \alpha_i g_i$  of  $g_1, \dots, g_t$ .

$$\begin{aligned}
\mathbf{y}'^T M_1 \mathbf{x}' &= \sum_{i=1}^t \alpha_i g_i = \sum_{i=1}^t \alpha_i (l_i(\mathbf{x}') \times l_i'(\mathbf{y}')) \\
&= \sum_{i=1}^t \left( \alpha_i l_i(\mathbf{x}') \times \left( \sum_{j=1}^m \beta_{i,j} l_j^v(\mathbf{y}') \right) \right) && \text{By Claim 17} \\
&= \sum_{i=1}^t \sum_{j=1}^m \alpha_i \beta_{i,j} (l_i(\mathbf{x}') \times l_j^v(\mathbf{y}')) = \sum_{j=1}^m \sum_{i=1}^t \alpha_i \beta_{i,j} (l_i(\mathbf{x}') \times l_j^v(\mathbf{y}')) \\
&= \sum_{j=1}^m \left( \left( \sum_{i=1}^t \alpha_i \beta_{i,j} l_i(\mathbf{x}') \right) \times l_j^v(\mathbf{y}') \right)
\end{aligned}$$

This gives a decomposition of  $M_1$  as a sum of  $m$  matrices each with rank  $\leq 1$ . Therefore,  $\text{rank}(M_1) \leq m \leq |V^*|$ . (End of proof of Sub-Claim 18)  $\square$

Similarly, if  $\mathbf{y}'^T M_2 \mathbf{x}'$  is a linear combination of the product gates in the components of  $G \setminus V^*$  containing inputs labelled by  $Y_1$ , then  $\text{rank}(M_2) \leq |V^*|$ . Finally, the number of product gates in  $V^*$  is at most  $|V^*|$  so any linear combination of them will also have rank  $\leq |V^*|$ . Since the output  $\mathbf{y}'^T M' \mathbf{x}'$  of  $\Phi'$  is a linear combination of all the product gates in it, by subadditivity of rank we see that  $\text{rank}(M') \leq 3|V^*|$ . (End of proof of Claim 16)  $\square$

Claim 16 implies that  $\Omega(n^{0.96}) = |X_1| = \text{rank}(M') \leq 3|V^*| = O(\sqrt{n}(\log n)^{3/2})$ , a contradiction for large  $n$ . Hence,  $\text{size}(\Psi) \geq \frac{1}{2000} n \log n$ . (End of proof of Theorem 15)  $\square$

If the size of the underlying field  $\mathbb{F}$  is at least  $2n$ , *explicit*  $n \times n$  totally regular matrices over  $\mathbb{F}$  exist. Cauchy matrices are a standard example: Let  $(x_1, \dots, x_n)$  and  $(y_1, \dots, y_n)$  be sequences of  $2n$  distinct elements in  $\mathbb{F}$ . The  $(i, j)$ th entry of the Cauchy matrix defined by these sequences is  $1/(x_i - y_j)$ . It is not hard to see that every square minor of a Cauchy matrix is itself a Cauchy matrix, so total regularity follows from non-singularity.

**Corollary 19.** *Over any infinite field  $\mathbb{F}$  there exists an infinite family  $\{A_n\}_{n \geq 1}$  of explicit matrices (where  $A_n \in \mathbb{F}^{n \times n}$ ) such that  $L(\mathbf{y}^T A_n \mathbf{x}) \geq C_p(\mathbf{y}^T A_n \mathbf{x}) = \Omega(n \log n)$   $\square$ .*

The following observation of Strassen is useful to get a separation between arithmetic circuits and planar arithmetic circuits:

**Lemma 20** ([24], Theorem 4.2). *Let  $\mathbb{F}$  be an infinite field. For any  $n$ -superconcentrator  $G = (V, E)$ , there exists a weight function  $w : E \rightarrow \mathbb{F}$  such that if all inputs of  $G$  are labelled by  $x_1, \dots, x_n$ , internal vertices and outputs are labelled as sum gates then the linear forms  $\ell_1, \dots, \ell_n$  computed at the output gates forms the rows of a totally regular matrix.*

It is known (Ta-Shma [21]) that there exist (explicit) superconcentrators of linear size.

**Theorem 21** (Ta-Shma [21], Corollary 1.3). *For every  $n$ ,  $\exists$  explicit  $n$ -superconcentrators of linear size and depth  $\text{poly}(\log \log n)$ .*

Combining this fact with Lemma 20 and Theorem 15 gives us the following separation:

**Corollary 22.** *Over any infinite field  $\mathbb{F}$ , there exists an infinite family  $\{M_n\}_{n \geq 1}$  of totally regular matrices (where  $M_n \in \mathbb{F}^{n \times n}$ ) such that  $C(\mathbf{y}^T M_n \mathbf{x}) = O(n)$  but  $C_p(\mathbf{y}^T M_n \mathbf{x}) = \Omega(n \log n)$ .*

It is not known if there is an explicit totally regular matrix  $M$  such that  $C(\mathbf{y}^T M \mathbf{x}) = O(n)$ . However, if it is possible to construct explicit, constant depth superconcentrators of size  $o(n \log n)$  then we can establish a separation as in Corollary 22 for explicit, constant degree polynomials. Unfortunately, we are not aware of such a construction.

### 3.2 Lower Bounds for Read-Once Planar Arithmetic Circuits

In this section, we prove lower bounds against the size of read-once planar circuits. If we do not restrict ourselves to planar circuits then the read-once restriction is not a restriction at all: one can introduce  $n$ -new input gates into a non read-once circuit and make it read-once without any asymptotic blowup in size. But as noted in [20], [22], the read-once restriction in the planar setting is quite a strong one, and we can show quadratic lower bounds for  $C_p^r(\mathbf{y}^T M \mathbf{x})$  where  $M$  is any totally regular matrix. The argument is the same as Theorem 15 except we use a different partition theorem.

**Theorem 23.** *Let  $M \in \mathbb{F}^{n \times n}$  be a totally regular matrix. Then,  $C_p^r(\mathbf{y}^T M \mathbf{x}) = \Omega(n^2)$ .*

*Proof.* Let  $X = \{x_1, \dots, x_n\}$  be the set of  $x$ -variables and  $Y = \{y_1, \dots, y_n\}$  be the set of  $y$ -variables. Let  $\Psi$  be a read-once planar circuit computing  $\mathbf{y}^T M \mathbf{x}$  and for the sake of contradiction, suppose  $\text{size}(\Psi) = o(n^2)$ . By Lemma 14, there exists a read-once, planar, bilinear circuit  $\Phi$  computing  $\mathbf{y}^T M \mathbf{x}$  such that  $\text{size}(\Phi) = o(n^2)$ . Let the graph of  $\Phi$  be  $G = (V, E)$ . Then  $|V| = o(n^2)$ . Since  $\Phi$  is read-once, it has exactly  $n$  input gates labelled by  $x_1, \dots, x_n$  resp. and exactly  $n$  input gates labelled by  $y_1, \dots, y_n$  resp. Applying Theorem 10 (ignoring directions on edges) with weight  $1/(2n)$  for each input labelled by a variable and 0 for every other gate, we obtain a partition  $(A, B, C)$  of  $V$  with the properties mentioned in Lemma 10. Since  $|V| = o(n^2)$ ,  $|C| = o(n)$ .

Note that at least one of  $A \cup C$ ,  $B \cup C$  must have  $\geq n/2$   $X$ -inputs. Without loss of generality, assume that  $A \cup C$  contains  $\geq n/2$   $X$ -inputs. Then,  $A$  contains at least  $n/2 - o(n) \geq 5n/12$   $X$ -inputs. Since the weight of  $A$  is  $\leq 2/3$ ,  $A$  contains  $\leq 4n/3 - 5n/12 = 11n/12$   $Y$ -inputs. Therefore,  $B$  contains  $\geq n/12 - o(n) \geq n/13$   $Y$ -inputs. Let  $X' \subseteq X$  and  $Y' \subseteq Y$  be such that  $|X'| = |Y'| \geq n/13$ , all the  $X'$  inputs are contained in  $A$  and all the  $Y'$  inputs are contained in  $B$ . Consider  $\Phi$  and set all inputs outside  $X' \cup Y'$  to 0 to obtain another read-once, planar, bilinear circuit  $\Phi'$  that computes the bilinear form  $\mathbf{y}'^T M' \mathbf{x}'$  where  $M'$  is the minor of  $M$  whose rows are indexed by  $Y'$  and columns by  $X'$ ,  $\mathbf{x}'$  is the vector of  $X'$  variables and  $\mathbf{y}'$  is the vector of  $Y'$  variables. Since  $M$  is totally regular,  $\text{rank}(M') = |X'| = |Y'| \geq n/13$ . However, we obtain an upper bound on the rank of  $M'$  in the following claim which gives the desired contradiction.

**Claim 24.**  $\text{rank}(M') \leq 3|C| = o(n)$ .

*Sketch.* Similar to Claim 16. Partition the product gates of  $\Phi'$  into 3 types: those appearing in  $A$ , those appearing in  $B$  and those appearing in  $C$ . We can show as in Claim 17 that there exist  $\leq |C|$  linear forms  $l_1(\mathbf{y}'), \dots, l_m(\mathbf{y}')$  such that for that every product gate  $g = \ell_g(\mathbf{x}') \cdot \ell'_g(\mathbf{y}')$  in  $A$ ,  $\ell'_g(\mathbf{y}')$  can be expressed as a linear combination of  $l_1(\mathbf{y}'), \dots, l_m(\mathbf{y}')$ . This implies that the rank of any linear combination of product gates in  $A$  (and symmetrically  $B$ ) is at most  $|C|$ . There are at most  $|C|$  product gates in  $C$  so the rank of any linear combination of them is also at most  $|C|$ . Since the output of  $\Phi'$  is a linear combination of the product gates in it, we are done by subadditivity of rank.  $\square$

Claim 24 implies that  $n/13 \leq |X'| = \text{rank}(M') \leq 3|V^*| = o(n)$ , a contradiction for large  $n$ . Hence  $\text{size}(\Psi) = \Omega(n^2)$ .  $\square$

We note that the  $\Omega(n^2)$  lower bound in Theorem 23 is tight and that a lower bound for read-once planar circuits asymptotically better than Theorem 23 implies superlinear lower bounds for general circuits. This is a direct consequence of the planarization procedure: by Lemma 9 any circuit can be planarized with at most a quadratic blowup in size.

**Observation 25.** Let  $\mathbb{F}$  be any field and let  $\{f_n\}_{n \geq 1}$  be a family of polynomials where  $f_n \in \mathbb{F}[x_1, \dots, x_n]$ . Then, for any  $\epsilon > 0$ ,  $C_p^r(f_n) = \Omega(n^{2+\epsilon})$  implies that  $C(f_n) = \Omega(n^{1+\epsilon/2})$ .

Strassen's observation in Lemma 20 combined with Theorem 23 gives us the following quadratic separation between circuit complexity and read-once planar circuit complexity:

**Corollary 26.** Over any infinite field  $\mathbb{F}$ , there exists an infinite family  $\{M_n\}_{n \geq 1}$  of totally regular matrices (where  $M_n \in \mathbb{F}^{n \times n}$ ) such that  $C(\mathbf{y}^T M_n \mathbf{x}) = O(n)$  but  $C_p^r(\mathbf{y}^T M_n \mathbf{x}) = \Omega(n^2)$ .

Here, we can use Ben-Or's trick from [15] to get an almost quadratic separation for an *explicit* family polynomials of degree 4. These polynomials are explicit in the sense that there is an algorithm which takes  $n$  as input and outputs the  $n$ th polynomial in the family in  $\text{poly}(n)$  time. In the sequel we need explicit depth-two superconcentrators which can be obtained from the following theorem of Ta-Shma:



**Theorem 27** (Ta-Shma [21], Corollary 1.2). *For every  $n \exists$  explicit  $n$ -superconcentrators of size  $n^{1+o(1)}$  and depth two.*

**Corollary 28.** *Over any infinite field  $\mathbb{F}$ , there exists an explicit family  $\{f_n\}_{n \geq 1}$  of degree 4 polynomials, where  $f_n$  is  $n^{1+o(1)}$ -variate, such that  $C(f_n) = n^{1+o(1)}$  but  $C_p^r(f_n) = \Omega(n^2)$*

*Proof.* In order to get the explicit family of polynomials, we start with an explicit depth 2 superconcentrator  $G = (V, E)$  with  $n^{1+o(1)}$  edges from Theorem 27. For every edge  $e \in E$  introduce a variable  $z_e$ . Now we construct an arithmetic circuit  $\Psi$  with inputs  $\{x_1, \dots, x_n\} \cup \{y_1, \dots, y_n\} \cup \{z_e\}_{e \in E}$  from  $G$  as follows:

1. For the  $n$  input vertices of  $G$ , add  $n$  inputs in  $\Psi$  labelled by  $x_1, \dots, x_n$  respectively.
2. For every internal vertex and output  $v$  of  $G$  add a sum gate  $g_v$  in  $\Psi$  whose inputs will be specified next.
3. For every edge  $e \in E$  going from  $u \in V$  to  $v \in V$ , introduce a product gate  $g_e$  one of whose children is  $g_u$  and the other is a newly introduced input gate labelled by  $z_e$ . Let the output of  $g_e$  feed into  $g_v$ .
4. If  $v_1, \dots, v_n$  are the outputs of  $G$ , multiply  $g_{v_1}, \dots, g_{v_n}$  by  $y_1, \dots, y_n$  respectively and add them up. This introduces  $\leq 3n$  new gates. Let the gate computing this sum be the output of  $\Psi$ .

Clearly, the number of wires in  $\Psi = n^{1+o(1)}$  and so  $\text{size}(\Psi) = n^{1+o(1)}$ . The output of  $\Psi$  is a polynomial in variables  $\{x_1, \dots, x_n\} \cup \{y_1, \dots, y_n\} \cup \{z_e\}_{e \in E}$ , there are  $n^{1+o(1)}$  of them. The degree of the output polynomial is 4 since the depth of  $G$  is 2. Let  $f(\mathbf{x}, \mathbf{y}, \mathbf{z})$  be the output of  $\Psi$ . Since  $G$  was explicit and depth 2,  $f$  is also explicit and we have that  $C(f) = n^{1+o(1)}$ . Now suppose we had a read-once planar circuit  $\Phi$  computing  $f$  such that  $\text{size}(\Phi) = o(n^2)$ . Since  $G$  is a superconcentrator, there exists  $\alpha \in \mathbb{F}^{|E|}$  (by Lemma 20) such that  $f(\mathbf{x}, \mathbf{y}, \alpha) = \mathbf{y}^T \mathbf{A} \mathbf{x}$  for some totally regular matrix  $A$ . Projecting the  $z$  variables in  $\Phi$  to  $\alpha$ , we get a read-once planar circuit of size  $o(n^2)$  computing  $\mathbf{y}^T \mathbf{A} \mathbf{x}$ . This contradicts Theorem 23, and so  $\text{size}(\Phi) = \Omega(n^2)$ .  $\square$

In the boolean setting, read-once planar circuit complexity and formula complexity are known to be incomparable[22]. We show that this is the case in the arithmetic setting as well, although the separation we have is not as strong as in the boolean case. For this, we again use depth two superconcentrators. Here we note that one can build a bilinear formula for  $\mathbf{y}^T \mathbf{A} \mathbf{x}$  (for some totally regular  $A$ ) from a depth two superconcentrator. This is done by Nisan and Wigderson in [15], we give a short, different proof here that suffices for our purposes:

**Claim 29** (Nisan and Wigderson [15]). *Over any infinite field  $\mathbb{F}$ , there exists a family of totally regular matrices  $\{A_n\}_{n \geq 1}$  such that  $L(\mathbf{y}^T A_n \mathbf{x}) \leq L^b(\mathbf{y}^T A_n \mathbf{x}) = n^{1+o(1)}$*

*Proof.* By Lemma 27, there exists an explicit depth two  $n$ -superconcentrator  $G = (V = I \cup M \cup O, E)$  with  $n$  inputs  $I$ ,  $n$  outputs  $O$  and  $k = n^{1+o(1)}$  middle vertices  $M$ . By Lemma 20, we can label input vertices  $I$  of  $G$  by  $x_1, \dots, x_n$ , the middle vertices  $M$  and outputs  $O$  of  $G$  by addition gates and the edges by constants from  $\mathbb{F}$  (if  $\mathbb{F}$  is large enough) such that the outputs of the resulting circuit  $\Phi(G)$  compute linear forms  $l_1^T \mathbf{x}, \dots, l_n^T \mathbf{x}$  where the vectors  $l_1^T, \dots, l_n^T$  are the rows of a totally regular matrix  $A$ . Since  $G$  has depth 2, this gives us a factorization  $A = UV$  such that  $U \in \mathbb{F}^{n \times k}$ ,  $V \in \mathbb{F}^{k \times n}$  and  $|U| + |V| \leq |E| = O(n^{1+o(1)})$  (where  $|M|$  denotes the number of non-zero entries of a matrix  $M$ ). Let the columns of  $U$  be  $u_1, \dots, u_k$  and the rows of  $V$  be  $v_1^T, \dots, v_k^T$ . It is easy to see that  $\sum_{i=1}^k \mathbf{y}^T (u_i v_i^T) \mathbf{x}$  is a bilinear formula computing  $\mathbf{y}^T A \mathbf{x}$  and the size of the formula is at most  $|E| = n^{1+o(1)}$ , the number of edges in  $G$ .  $\square$

**Corollary 30.** *Over any infinite field  $\mathbb{F}$ , there exists an explicit family of polynomials  $\{f_n\}_{n \geq 1}$  where  $f_n$  is a polynomial in  $n^{1+o(1)}$  variables of degree 4 such that  $L(f_n) = n^{1+o(1)}$  but  $C_p^r(f_n) = \Omega(n^2)$ .*

*Proof.* The family in question is exactly the one from the previous corollary (Corollary 26).

By Theorem 23,  $C_p^r(\mathbf{y}^T A \mathbf{x}) = \Omega(n^2)$  for any totally regular matrix  $A$ . By Claim 29, there exists a totally regular  $A$  that is computed by a (bilinear) formula of size  $n^{1+o(1)}$ . This gives a separation for a (non-explicit) bilinear form.

To get an explicit polynomial for which the separation holds, repeat the procedure used in the proof of Corollary 28, i.e., introduce a new variable for every edge in the depth 2 superconcentrator. The resulting polynomial will have degree 4.  $\square$

In the other direction, we have the following easy separation:

**Lemma 31.** *There exists a family of polynomials  $\{f_n\}_{n \geq 1}$  where  $f_n$  is a polynomial in  $n$  variables of degree  $n$  such that  $C_p^r(f_n) = O(n \log n)$  but  $L(f_n) = \Omega(n^2)$ .*

*Proof.* Consider  $f(x_1, \dots, x_n) = \sum_{i=1}^n x_i^n$ . Since the degree of each  $x_i$  in  $f$  is  $n$ , in any formula computing  $f$  each  $x_i$  must appear  $n$  times and so  $L(f) \geq n^2$ . On the other hand,  $C_p^r = O(n \log n)$ , the usual circuit that raises each  $x_i$  to the  $n$ th power using  $\log n$  product gates and then adds all the products is planar.  $\square$

It would be interesting to see a multilinear polynomial which has small read-once planar circuit complexity and large formula complexity. Together, Lemma 31 and Corollary 30 show that the complexity measures  $C_p^r$  and  $L$  are incomparable.

### 3.3 Lower Bounds for Planar Algebraic Branching Programs

As mentioned in the introduction, planar algebraic branching programs can be converted into planar arithmetic circuits without blowup in size. We describe this conversion in the following lemma:

**Lemma 32.** *Let  $A$  be a planar ABP with  $v$  vertices and  $e$  edges computing  $f(x_1, \dots, x_n) \in \mathbb{F}[x_1, \dots, x_n]$ . Then there exists a planar circuit  $\Phi$  (with indegree of each gate  $\leq 2$ ) computing  $f$  with  $\text{size}(\Phi) = O(v)$ .*

*Proof.* Let  $G = (V, E)$  be the planar graph of the ABP  $A$  computing  $f$ . Let  $s$  be its source and  $t$  the sink. Construct a circuit  $\Phi$  from  $G$  as follows:

1. Remove the source  $s$  from the graph and let every  $u \in N^+(s)$  be a leaf of the circuit labelled by  $[s, u]$  ( $N^+(s)$  denotes the out neighbourhood of  $s$ ,  $[u, v]$  denotes the label of the edge  $uv$  in  $A$ ).
2. Let every  $v \in V \setminus (\{s\} \cup N^+(s))$  be a sum gate.
3. Subdivide every edge  $e = (u, v)$  of  $G$  and let the newly introduced vertex  $v_e$  be a product gate. Attach a leaf  $l_e$  to  $v_e$  labelled by  $[u, v]$ . So  $v_e$  is a product gate with predecessors  $l_e$  and  $u$ .

It is easy to see that the resulting circuit  $\Phi'$  with output  $t$  computes the polynomial  $f(x_1, \dots, x_n)$  and  $\text{size}(\Phi') \leq v - 1 + 2e = O(v + e) = O(v)$  (Note:  $G$  is planar, so  $e = O(v)$ ). Furthermore, since  $G$  is planar,  $\Phi'$  is also planar. Now convert  $\Phi'$  into a fan-in 2 circuit  $\Phi$  by replacing the incoming wires at every gate by a balanced binary tree. Again, this preserves planarity. We have introduced at most  $4e$  additional gates while going from  $\Phi'$  to  $\Phi$  and so  $\text{size}(\Phi) = O(v)$ . Clearly  $\Phi$  computes  $f$ , and we are done.  $\square$

We can now use the lower bound for planar circuits from Theorem 15 to get a lower bound for planar ABPs:

**Theorem 33.** *Let  $M \in \mathbb{F}^{n \times n}$  be a totally regular matrix. Then,  $A_p(\mathbf{y}^T M \mathbf{x}) = \Omega(n \log n)$ .*

*Proof.* Combining Lemma 32 with the lower bound for planar arithmetic circuits we get the desired result.  $\square$

By modifying Strassen's construction a bit (Lemma 20), it is possible to establish separations between ABP complexity and planar ABP complexity:

**Lemma 34.** *Over any infinite field  $\mathbb{F}$ , there exists an infinite family  $\{M_n\}_{n \geq 1}$  of totally regular matrices (where  $M_n \in \mathbb{F}^{n \times n}$ ) such that  $A(\mathbf{y}^T M_n \mathbf{x}) = O(n)$ .*

*Proof.* We start with an explicit  $n$ -superconcentrator  $G = (V, E)$  of linear size, ie,  $|V|, |E| = O(n)$  (see Theorem 21). We add two new vertices  $s, t$  to  $G$ , connect  $s$  to every input of  $G$  and  $t$  to every output. For every  $i \in [n]$  label the  $i$ th edge out of  $s$  by  $x_i$  and the  $i$ th edge into  $t$  by  $y_i$ . Now note that the internal nodes behave exactly like sum gates in a linear circuit, so we may label each edge by a constant from  $\mathbb{F}$  such that the resulting ABP computes  $\mathbf{y}^T M \mathbf{x}$  and  $M \in \mathbb{F}^{n \times n}$  is totally regular (by Lemma 20).  $\square$

Combining Lemma 34 and Theorem 33 we get the desired separation:

**Corollary 35.** *Over any infinite field  $\mathbb{F}$ , there exists an infinite family  $\{M_n\}_{n \geq 1}$  of totally regular matrices (where  $M_n \in \mathbb{F}^{n \times n}$ ) such that  $A(\mathbf{y}^T M_n \mathbf{x}) = O(n)$  but  $A_p(\mathbf{y}^T M_n \mathbf{x}) = \Omega(n \log n)$ .*

## 4 Lower Bounds for Multi-Output Planar Circuits and Partial Derivative Complexity

In the case of multi-output planar circuits, it is possible to show lower bounds better than  $\Omega(n \log n)$ . As mentioned earlier, an  $\Omega(n^2)$  lower bound on the size of any read-once planar circuit computing  $M\mathbf{x}$  (for any totally regular  $M$ ) follows from the work of Valiant [23] and Lipton and Tarjan [12]. In this section we show an  $\Omega(n^{4/3})$  lower bound on the size of (not necessarily read-once) planar circuits computing  $M\mathbf{x}$  and an  $\Omega(n^2/\log n)$  lower bound on the size of multi-output formulas computing  $M\mathbf{x}$ . We use (resp.) Theorem 12 and Lemma 13 for these lower bounds. Similar statements for multi-output planar boolean circuits are proved in [22], [20] using crossing sequence arguments.

### 4.1 Lower Bounds for Multi-Output Circuits

**Theorem 36.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix. Then  $C_p(M\mathbf{x}) = \Omega(n^{4/3})$ .*

*Proof.* Let  $\Phi$  be a planar circuit computing  $M\mathbf{x}$  with outputs gates  $O_1, \dots, O_n$  and let  $L_1 \sqcup \dots \sqcup L_n$  be its set of leaves labelled by variables, where  $L_i$  contains all the leaves labelled by  $x_i$ . Because  $M$  is totally regular the following statement is true:

**Claim 37.**  $\forall k \in [n], \forall$  subsets  $\{O_{i_1}, \dots, O_{i_k}\}$  of  $k$  outputs of  $\Phi$ ,  $\forall k$  sets  $L_{j_1}, \dots, L_{j_k}$ ,  $\exists$  a permutation  $\pi : [k] \rightarrow [k]$  such that  $\forall t \in [k], \exists$  a path  $P_t$  from an input in  $L_{j_t}$  to  $O_{\pi(i_t)}$  and the paths  $P_1, \dots, P_k$  are vertex disjoint.

*Proof.* This is an immediate generalization of Valiant's observation that the graph of a (read-once) circuit computing  $A\mathbf{x}$  is an  $n$ -superconcentrator [23] (for any totally regular  $A$ ). The idea is that if the maximum number of vertex disjoint paths from  $L_{j_1}, \dots, L_{j_k}$  to  $O_{i_1}, \dots, O_{i_k}$  was less than  $k$ , then by Menger's theorem there would exist a  $(\{L_{j_1} \cup \dots \cup L_{j_k}\}, \{O_{i_1} \cup \dots \cup O_{i_k}\})$  cut of size strictly less than  $k$ , which would in turn imply that the minor of  $M$  whose rows are indexed by  $O_{i_1}, \dots, O_{i_k}$  and columns by  $x_{j_1}, \dots, x_{j_k}$  cannot be full rank.  $\square$

Now we can apply Theorem 12 to get the desired lower bound:

Let  $G = (V, E)$  be the underlying undirected graph of  $\Phi$ . Now suppose  $\text{size}(\Phi) = |V| = o(n^{4/3})$  (otherwise we're done). Apply Theorem 12 to  $G$  with  $V' = \{O_1, \dots, O_n\}$  and  $p = (|V| + n)/n = o(n^{1/3}) < n = |V'|$ . Let  $V_1, \dots, V_p$  be the partition so obtained and let  $S_1, \dots, S_p$  be the corresponding separators. Note that the number of input gates of  $\Phi$  is at most  $(|V| + n)/2$ : If  $\Phi$  has  $t$  inputs each with indegree 0 and outdegree  $\geq 1$ , and  $k$  non-input gates each with indegree 2 out of which  $\leq n$  (the outputs) have outdegree 0 and the rest have outdegree  $\geq 1$  (so that  $t + k = |V|$ ) then  $|E| = \sum_v d^-(v) = 2k$  and  $|E| = \sum_v d^+(v) \geq |V| - n$ . So  $k \geq (|V| - n)/2 \implies t \leq (|V| + n)/2$ . Here  $d^-(v)$  and  $d^+(v)$  denote the indegree and outdegree of  $v$  resp.

So there must exist a  $V_i \in \{V_1, \dots, V_p\}$  that contains at most  $n/2$  input gates for otherwise the number of inputs of  $\Phi$  would be  $> (|V| + n)/2$ . Let  $X_i \subseteq \{x_1, \dots, x_n\}$  be a set of  $n/2$  variables not appearing in  $V_i$ . Let  $X'_i \subseteq X_i$  be such that  $|X'_i| = \Omega(n)$  and all leaves labelled by  $X'_i$  appear in  $V \setminus (V_i \cup S_i)$ . Such a set exists since  $|S_i| \leq 60\sqrt{|V|} = o(n^{2/3})$ . Let  $A_i$  be the set of outputs in  $V_i$ . Then  $|A_i| \geq \frac{n}{4p} = \omega(n^{2/3})$ . By Claim 37, there must be at least  $|A_i| = \omega(n^{2/3})$  vertex disjoint paths from the sets  $(L_j$ 's) of leaves corresponding to variables in  $X'_i$  to outputs in  $A_i$ . All these paths must go through  $S_i$ . But as we saw before,  $|S_i| = o(n^{2/3})$ . This is a contradiction, and so  $\text{Size}(\Phi) = \Omega(n^{4/3})$ .  $\square$

For proving better lower bounds for multi-output formulas we first prove the improved partition lemma (Lemma 13). For this purpose, it will be convenient to work with partition trees. We identify the vertices of a binary tree with binary strings in the natural way: the root is identified with  $\epsilon$ , the empty string, and the left and right children of a vertex  $\alpha$  are identified with  $\alpha 0, \alpha 1$  respectively. A partition tree  $T$  for a set  $V$  is a binary tree each of whose vertices are labelled by subsets of  $V$  (we say that a vertex  $\alpha$  of  $T$  is labelled by the subset  $V_\alpha$ ).  $V_\epsilon$  is labelled with  $V$  and for every non-leaf vertex  $\alpha$  of  $T$ ,  $(V_{\alpha 0}, V_{\alpha 1})$  forms a partition of  $V_\alpha$ . It is easy to see that the subsets labelling the leaves of  $T$  form a partition of  $V$ . We restate and prove lemma 13 in the language of partition trees:

**Lemma 38.** *(Equivalent to lemma 13) Let  $F = (V, E)$  be a forest,  $V' \subseteq V$  be a subset of its vertices and let  $1 \leq p \leq |V'|$ . Then there exists a partition tree  $T_p$  of  $V$  with  $p$  leaves such that the following conditions hold:*

1. For all leaves  $\alpha$  of  $T_p$ ,  $\frac{|V'|}{3p} \leq |V' \cap V_\alpha| \leq \frac{3|V'|}{p}$
2. For all vertices  $\alpha$  of  $T_p$  there exists a set  $S_\alpha$  such that  $|S_\alpha| \leq O(\log(|V'|))$  and no edge joins  $V_\alpha$  and  $V \setminus (V_\alpha \cup S_\alpha)$ .  $\square$

*Proof.* We need the following well known lemma (see [5], [4])

**Lemma 39.** *Let  $F = (V, E)$  be a forest and let  $V' \subseteq V$ . Then there exists a constant  $k$  and a partition  $(A, B, C)$  of  $V$  such that  $|A| \leq 2|V|/3$ ,  $|B| \leq 2|V|/3$ ,  $|C| \leq k$  and  $|V' \cap A| \leq 2|V'|/3$ ,  $|V' \cap B| \leq 2|V'|/3$ . Furthermore, all paths from  $A$  to  $B$  contain a vertex from  $C$ .*

Now, let  $F = (V, E)$  be a forest and let  $V' \subseteq V$ . The only difference between the proof of Theorem 12 in [20] and Lemma 38 is the use of Lemma 39 (see [22], Lemma 4 for an exposition of a similar result using the partition tree terminology we use here). We construct, by induction on  $p$ , a partition tree  $T_p$  with  $p$  leaves such that partition induced by it's leaves satisfies the required properties. For each vertex  $\alpha$  of  $T_p$ , in addition to  $V_\alpha$  we maintain also a partition  $(W_\alpha, U_\alpha)$  of  $V_\alpha$ . The case when  $p = 1$  is trivial, here  $V_\epsilon = V$ ,  $W_\epsilon = V$ ,  $U_\epsilon = \emptyset$  and  $S_\epsilon = \emptyset$ . Now suppose  $2 \leq p \leq |V'|$  and we have constructed  $T_{p-1}$ . Pick a leaf  $\alpha$  of  $T_{p-1}$  that such that  $|V' \cap V_\alpha|$  is maximized. To get  $T_p$  from  $T_{p-1}$  we attach two leaves  $\alpha 0, \alpha 1$  to  $\alpha$ . To get a complete description of  $T_p$  we must define  $V_{\alpha 0} = W_{\alpha 0} \sqcup U_{\alpha 0}$  and  $V_{\alpha 1} = W_{\alpha 1} \sqcup U_{\alpha 1}$ . Applying Lemma 39 to the subforest of  $F$  induced by  $W_\alpha$ , we get a

partition  $(A_\alpha, B_\alpha, C_\alpha)$ . Define  $W_{\alpha 0} = A_\alpha$ ,  $W_{\alpha 1} = B_\alpha$  and define  $(U_{\alpha 0}, U_{\alpha 1})$  to be a partition of  $U_\alpha \cup C_\alpha$  such that

$$|V' \cap (W_{\alpha 0} \cup U_{\alpha 0})| \leq \frac{2}{3}|V' \cap V_\alpha| \text{ and}$$

$$|V' \cap (W_{\alpha 1} \cup U_{\alpha 1})| \leq \frac{2}{3}|V' \cap V_\alpha|$$

These definitions imply that the following hold:

1. For any two leaves  $\beta, \gamma$  of  $T_p$ ,  $|V' \cap V_\beta| \leq 3|V' \cap V_\gamma|$ , this follows by induction on  $p$  and noticing that  $|V' \cap V_\alpha|$  is maximal for the leaf  $\alpha$  of  $T_{p-1}$ . This observation implies condition 1 of lemma 13.
2. For every vertex  $\alpha$  in  $T_p$ ,  $W_\alpha \leq (2/3)^{|\alpha|}|V|$ . This follows by induction on  $|\alpha|$  and condition 1 of lemma 39.
3. For every inner vertex  $\alpha$  of  $T_p$ ,  $|C_\alpha| \leq k$ . This follows from condition 3 of lemma 39.
4. For every vertex  $\alpha$  of  $T_p$ , let  $S_\alpha = \bigcup_{i=1}^{|\alpha|-1} C_{\alpha^i}$  where  $\alpha^i$  is the prefix of  $\alpha$  of length  $i$ . Then  $S_\alpha$  is a  $V_\alpha$ ,  $V \setminus (V_\alpha \cup S_\alpha)$  separator. This follows by induction on  $|\alpha|$ .
5. For every vertex  $\alpha$ , we have that  $|S_\alpha| = O(\log |V'|)$ . This follows from the following facts:
  - The maximal length of any root to leaf path in  $T_p$  must be  $O(\log |V'|)$ , since at each step in the induction we pick a leaf  $\alpha$  that maximizes  $|V' \cap V_\alpha|$  and attach leaves  $\alpha 0, \alpha 1$  such that  $|V' \cap V_{\alpha 0}|, |V' \cap V_{\alpha 1}| \leq \frac{2}{3}|V' \cap V_\alpha|$ ;
  - $|C_\beta| \leq k$  for every inner vertex  $\beta$  of  $T_p$ ; and
  - The union bound

This implies condition 2 of lemma 13. □

We can now prove, as promised, the improved lower bound for multi output formulas:

**Theorem 40.** *Let  $M \in \mathbb{F}^{n \times n}$  be any totally regular matrix. Then  $L(M\mathbf{x}) = \Omega(n^2 / \log n)$ .*

*Proof.* Let  $\Phi$  be a multi-output formula computing  $M\mathbf{x}$  with output gates  $O_1, \dots, O_n$  and let  $L_1 \sqcup \dots \sqcup L_n$  be its set of leaves labelled by variables, where  $L_i$  contains all the leaves labelled by  $x_i$ . Observe that claim 37 continues to hold.

Let  $F = (V, E)$  be the underlying undirected forest of  $\Phi$ . Now suppose  $\text{size}(\Phi) = |V| = o(n^2 / \log n)$  (otherwise we're done). Apply lemma 13 to  $F$  with  $V' = \{O_1, \dots, O_n\}$  and  $p = (|V| + n)/n = o(n / \log n) < n = |V'|$ . Let  $V_1, \dots, V_p$  be the partition so obtained and let  $S_1, \dots, S_p$  be the corresponding separators. Note that the number of input gates of  $\Phi$  is at most  $(|V| + n)/2$ : If  $\Phi$  has  $t$  inputs each with indegree 0 and outdegree  $\geq 1$ , and  $k$  non-input gates each with indegree 2 out of which  $\leq n$  (the outputs) have outdegree 0 and the rest have

outdegree = 1 (so that  $t+k = |V|$ ) then  $|E| = \sum_v d^-(v) = 2k$  and  $|E| = \sum_v d^+(v) \geq |V| - n$ . So  $k \geq (|V| - n)/2 \implies t \leq (|V| + n)/2$ . Here  $d^-(v)$  and  $d^+(v)$  denote the indegree and outdegree of  $v$  respectively.

So there must exist a  $V_i \in \{V_1, \dots, V_p\}$  that contains at most  $n/2$  input gates for otherwise the number of inputs of  $\Phi$  would be  $> (|V| + n)/2$ . Let  $X_i \subseteq \{x_1, \dots, x_n\}$  be a set of  $n/2$  variables not appearing in  $V_i$ . Let  $X'_i \subseteq X_i$  be such that  $|X'_i| = \Omega(n)$  and all leaves labelled by  $X'_i$  appear in  $V \setminus (V_i \cup S_i)$ . Such a set exists since  $|S_i| = O(\log(|V'|)) = O(\log n)$ . Let  $A_i$  be the set of outputs in  $V_i$ . Then  $|A_i| \geq \frac{n}{4p} = \omega(\log n)$ . By Claim 37, there must be at least  $|A_i| = \omega(\log n)$  vertex disjoint paths from the sets ( $L_j$ 's) of leaves corresponding to variables in  $X'_i$  to the outputs in  $A_i$ . All these paths must go through  $S_i$ . But as we saw before,  $|S_i| = O(\log n)$ . This is a contradiction, and so  $\text{size}(\Phi) = \Omega(n^2/\log n)$ .  $\square$

Note that this bound is optimal up to the  $\log n$  factor: any linear transformation can be computed by an  $O(n^2)$  size multi-output formula (by constructing disjoint formulas for each linear form).

## 4.2 Complexity of Partial Derivatives

We now turn our attention to partial derivative complexity. Baur and Strassen[3] proved that for any polynomial  $f \in \mathbb{F}[x_1, \dots, x_n]$ ,  $C(\partial_{x_1}(f), \dots, \partial_{x_n}(f)) = O(C(f))$ . That is, if there exists a (fan-in 2) circuit of size  $s$  computing  $f$  then there exists another multi-output (fan-in 2) circuit of size  $O(s)$  that simultaneously computes all the first order partial derivatives of  $f$ . We observe that an analogous result *cannot* hold for formulas and planar circuits, while it does hold for read-once planar circuits. Theorem 40 combined with the construction in Claim 29 immediately gives the following lower bound on the partial derivative complexity of multi-output formulas:

**Corollary 41.** *Over any infinite field  $\mathbb{F}$  there exists a family  $\{f_n\}_{n \geq 1}$  of polynomials where  $f_n \in \mathbb{F}[x_1, \dots, x_n]$  such that  $L(f_n) = n^{1+o(1)}$  but  $L(\partial_{x_1}(f_n), \dots, \partial_{x_n}(f_n)) = \Omega(n^2/\log n)$ .*

Similarly, Theorem 36 combined with the construction in Claim 29 gives us a lower bound on the partial derivative complexity of planar circuits:

**Corollary 42.** *Over any infinite field  $\mathbb{F}$  there exists a family  $\{f_n\}_{n \geq 1}$  of polynomials where  $f_n \in \mathbb{F}[x_1, \dots, x_n]$  such that  $C_p(f_n) = n^{1+o(1)}$  but  $C_p(\partial_{x_1}(f_n), \dots, \partial_{x_n}(f_n)) = \Omega(n^{4/3})$ .*

In the case of read-once planar circuits, we observe that the proof of Baur-Strassen theorem (as demonstrated in [8], this proof is due to Kaltofen and Singer [11]) extends easily to the case of read-once planar circuits. Hence, we have the following corollary:

**Corollary 43.** *For any polynomial  $f \in \mathbb{F}[x_1, \dots, x_n]$ ,  $C_p^r(\partial_{x_1}(f), \dots, \partial_{x_n}(f)) = O(C_p^r(f))$ .*

The above corollary holds as the circuit for simultaneously computing  $\partial_{x_1}(f), \dots, \partial_{x_n}(f)$  of a polynomial  $f$  can be constructed using a copy of the circuit for  $f$  and a copy of its mirror image (see [8], Theorem 9.10) and in the case of read-once planar circuits, it is easy to check that this construction introduces only a linear number of edge crossings each of which can be eliminated by introducing the crossover gadget in Lemma 9.

## 5 Discussion and Open Problems

- A long-standing open problem in algebraic complexity theory is to construct an explicit constant degree polynomial that has superlinear circuit complexity. In this regard, it is challenging to prove  $\Omega(n^{1+\delta})$  lower bound (for some  $\delta > 0$ ) for planar arithmetic circuits for any explicit  $n$ -variate polynomial, not necessarily of constant degree. Note that in the case of boolean planar circuits, the best known lower bound is  $\Omega(n \log^2 n)$  ([9]). In fact, obtaining an  $\Omega(n^{2+\delta})$  lower bound (for some  $\delta > 0$ ) for read-once planar arithmetic circuits for any explicit  $n$ -variate polynomial would also imply general circuit lower bounds.
- Although the theorem gives an  $\Omega(n \log n)$  lower bound for planar arithmetic circuits computing any bilinear form  $\mathbf{y}^T A \mathbf{x}$  where  $A$  is totally regular, the best lower bound that we can hope to get is  $\Omega\left(\frac{n \log^2 n}{\log \log n}\right)$  because depth 2 superconcentrators of size  $O\left(\frac{n \log^2 n}{\log \log n}\right)$  exist (see [16] for details). It would be interesting to see if Theorem 15 is tight for some totally regular matrix.

## Acknowledgements

We thank Meena Mahajan for insightful discussions on planar arithmetic circuits.

## References

- [1] Noga Alon and Wolfgang Maass. Meanders and their applications in lower bounds arguments. *Journal of Computer and System Sciences*, 37(2):118–129, 1988. URL: <https://www.sciencedirect.com/science/article/pii/0022000088900025>, doi:[https://doi.org/10.1016/0022-0000\(88\)90002-5](https://doi.org/10.1016/0022-0000(88)90002-5).
- [2] L. Babai, P. Pudlák, V. Rödl, and E. Szemerédi. Lower bounds to the complexity of symmetric boolean functions. *Theoretical Computer Science*, 74(3):313–323, 1990. URL: <https://www.sciencedirect.com/science/article/pii/0304397590900802>, doi:[https://doi.org/10.1016/0304-3975\(90\)90080-2](https://doi.org/10.1016/0304-3975(90)90080-2).
- [3] Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22(3):317–330, 1983. URL: <https://www>.



sciencedirect.com/science/article/pii/030439758390110X, doi:[https://doi.org/10.1016/0304-3975\(83\)90110-X](https://doi.org/10.1016/0304-3975(83)90110-X).

- [4] Sandeep Bhatt, Fan Chung, Tom Leighton, and Arnold Rosenberg. Optimal simulations of tree machines. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*, pages 274–282, 1986. doi:[10.1109/SFCS.1986.38](https://doi.org/10.1109/SFCS.1986.38).
- [5] Sandeep N. Bhatt and Charles E. Leiserson. How to assemble tree machines (extended abstract). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing, STOC '82*, page 77–84, New York, NY, USA, 1982. Association for Computing Machinery. doi:[10.1145/800070.802179](https://doi.org/10.1145/800070.802179).
- [6] Abhranil Chatterjee, Mrinal Kumar, and Ben Lee Volk. Determinants vs. algebraic branching programs. *Electron. Colloquium Comput. Complex.*, TR23-115, 2023. URL: <https://eccc.weizmann.ac.il/report/2023/115>, arXiv:TR23-115.
- [7] Prerona Chatterjee, Mrinal Kumar, Adrian She, and Ben Lee Volk. Quadratic lower bounds for algebraic branching programs and formulas. *computational complexity*, 31(2):8, Jul 2022. doi:[10.1007/s00037-022-00223-8](https://doi.org/10.1007/s00037-022-00223-8).
- [8] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial derivatives in arithmetic complexity and beyond. *Foundations and Trends in Theoretical Computer Science*, January 2010. URL: <https://www.microsoft.com/en-us/research/publication/partial-derivatives-in-arithmetic-complexity-and-beyond/>.
- [9] Hans Dietmar Gröger. A new partition lemma for planar graphs and its application to circuit complexity. In Lothar Budach, editor, *Fundamentals of Computation Theory, 8th International Symposium, FCT '91, Gosen, Germany, September 9-13, 1991, Proceedings*, volume 529 of *Lecture Notes in Computer Science*, pages 220–229. Springer, 1991. doi:[10.1007/3-540-54458-5\\_66](https://doi.org/10.1007/3-540-54458-5_66).
- [10] K. A. Kalorkoti. A lower bound for the formula size of rational functions. *SIAM Journal on Computing*, 14(3):678–687, 1985. arXiv:<https://doi.org/10.1137/0214050>, doi:[10.1137/0214050](https://doi.org/10.1137/0214050).
- [11] Erich Kaltofen and Michael F. Singer. Size efficient parallel algebraic circuits for partial derivatives. In *IV International Conference on Computer Algebra in Physical Research*, 1990. URL: <https://users.cs.duke.edu/~elk27/bibliography/91/KaSi91.pdf>.
- [12] Richard J. Lipton and Robert Endre Tarjan. Applications of a planar separator theorem. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 162–170, 1977. doi:[10.1109/SFCS.1977.6](https://doi.org/10.1109/SFCS.1977.6).
- [13] Richard J. Lipton and Robert Endre Tarjan. A separator theorem for planar graphs. *SIAM Journal on Applied Mathematics*, 36(2):177–189, 1979. arXiv:<https://doi.org/10.1137/0136016>, doi:[10.1137/0136016](https://doi.org/10.1137/0136016).

- [14] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *computational complexity*, 4(4):301–313, Dec 1994. doi:10.1007/BF01263419.
- [15] Noam Nisan and Avi Wigderson. On the complexity of bilinear forms: Dedicated to the memory of jacques morgenstern. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '95, page 723–732, New York, NY, USA, 1995. Association for Computing Machinery. doi:10.1145/225058.225290.
- [16] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000. arXiv:<https://doi.org/10.1137/S0895480197329508>, doi:10.1137/S0895480197329508.
- [17] Ran Raz. On the complexity of matrix product. In *Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing*, STOC '02, page 144–151, New York, NY, USA, 2002. Association for Computing Machinery. doi:10.1145/509907.509932.
- [18] Ran Raz. Elusive functions and lower bounds for arithmetic circuits. *Theory of Computing*, 6(7):135–177, 2010. URL: <https://theoryofcomputing.org/articles/v006a007>, doi:10.4086/toc.2010.v006a007.
- [19] John E. Savage. *Planar Circuit Complexity and The Performance of VLSI Algorithms +*, pages 61–68. Springer Berlin Heidelberg, Berlin, Heidelberg, 1981. doi:10.1007/978-3-642-68402-9\_8.
- [20] John E. Savage. The performance of multilective vlsi algorithms. *Journal of Computer and System Sciences*, 29(2):243–273, 1984. URL: <https://www.sciencedirect.com/science/article/pii/0022000084900333>, doi:[https://doi.org/10.1016/0022-0000\(84\)90033-3](https://doi.org/10.1016/0022-0000(84)90033-3).
- [21] Amnon Ta-Shma. On extracting randomness from weak random sources (extended abstract). In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, page 276–285, New York, NY, USA, 1996. Association for Computing Machinery. doi:10.1145/237814.237877.
- [22] György Turán. On the complexity of planar boolean circuits. *computational complexity*, 5(1):24–42, Mar 1995. doi:10.1007/BF01277954.
- [23] Leslie G. Valiant. On non-linear lower bounds in computational complexity. In *Proceedings of the Seventh Annual ACM Symposium on Theory of Computing*, STOC '75, page 45–53, New York, NY, USA, 1975. Association for Computing Machinery. doi:10.1145/800116.803752.
- [24] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In Jozef Gruska, editor, *Mathematical Foundations of Computer Science 1977*, pages 162–176, Berlin, Heidelberg, 1977. Springer Berlin Heidelberg.