Electronic Colloquium on Computational Complexity, Report No. 143 (2011)

# Jacobian hits circuits: Hitting-sets, lower bounds for depth-$D$ occur-$k$ formulas & depth-3 transcendence degree-$k$ circuits

Manindra Agrawal * Chandan Saha † Ramprasad Saptharishi ‡ Nitin Saxena §¶

**Abstract**

We present a single, common tool to strictly subsume *all* known cases of polynomial time blackbox polynomial identity testing (PIT) that have been hitherto solved using diverse tools and techniques. In particular, we show that polynomial time hitting-set generators for identity testing of the two seemingly different and well studied models - depth-3 circuits with bounded top fanin, and constant-depth constant-read multilinear formulas - can be constructed using one common algebraic-geometry theme: *Jacobian* captures algebraic independence. By exploiting the Jacobian, we design the *first* efficient hitting-set generators for broad generalizations of the above-mentioned models, namely:

- depth-3 ($\Sigma\Pi\Sigma$) circuits with constant *transcendence degree* of the polynomials computed by the product gates (*no* bounded top fanin restriction), and
- constant-depth constant-*occur* formulas (*no* multilinear restriction).

Constant-*occur* of a variable, as we define it, is a much more general concept than constant-read. Also, earlier work on the latter model assumed that the formula is multilinear. Thus, our work goes further beyond the results obtained by Saxena & Seshadhri (STOC 2011), Saraf & Volkovich (STOC 2011), Anderson et al. (CCC 2011), Beecken et al. (ICALP 2011) and Grenet et al. (FSTTCS 2011), and brings them under one unifying technique.

In addition, using the same Jacobian based approach, we prove exponential lower bounds for the immanant (which includes permanent and determinant) on the *same* depth-3 and depth-4 models for which we give efficient PIT algorithms. Our results reinforce the intimate connection between identity testing and lower bounds by exhibiting a concrete mathematical tool - the Jacobian - that is equally effective in solving both the problems on certain interesting and previously well-investigated (but not well understood) models of computation.

## 1 Introduction

A polynomial in many variables, when written down verbosely as a sum of monomials, might have a humongous expression. *Arithmetic circuits*, on the other hand, provide a succinct way to represent multivariate polynomials. An arithmetic circuit, consisting of addition (+) and multiplication ($\times$) gates, takes several variables as input and outputs a polynomial in those variables. The study of arithmetic circuits - as to which algorithmic questions on polynomials can be resolved efficiently

*Indian Institute of Technology Kanpur, `manindra@cse.iitk.ac.in`. Research done while visiting MPI-Informatik under Humboldt Forschungspreise.

†Max Planck Institute for Informatics, `csaha@mpi-inf.mpg.de`. Supported by IMPECS Fellowship.

‡Chennai Mathematical Institute, `ramprasad@cmi.ac.in`. Supported by MSR India Ph.D Fellowship.

§Hausdorff Center for Mathematics, `ns@hcm.uni-bonn.de`

¶The authors are grateful to MPI-Informatik, Saarbrücken for its generous hospitality.



ISSN 1433-8092

in this model of computation, and which polynomials do not admit any polynomial-sized circuit representation - form the foundation of algebraic complexity theory.

One particular algorithmic question, the problem of *polynomial identity testing* (PIT), occupies a pivotal position in the theory of arithmetic circuit complexity. It is the problem of deciding if the output of a given arithmetic circuit is an identically zero polynomial. Being such an elementary problem, identity testing has enjoyed its status of prime importance by appearing in several fundamental results including primality testing [AKS04], the $\mathsf{PCP}$ theorem [ALM[+]98] and the $\mathsf{IP} = \mathsf{PSPACE}$ result [LFKN90, Sha90], among many others like graph matching [Lov79, MVV87], polynomial interpolation [CDGK91], matrix completion [IKS10], polynomial solvability [KY08], factorization [SV10] and learning of arithmetic circuits [KS06]. What is more intriguing is that there is an intimate connection between identity testing and lower bounds [KI03, HS80, AvM10], especially the problem of separating the complexity classes $\mathsf{VP}$ from $\mathsf{VNP}$ (which must necessarily be shown before showing $\mathsf{P} \neq \mathsf{NP}$ [Val79, SV85]). Proving $\mathsf{VP} \neq \mathsf{VNP}$ amounts to showing that an explicit class of polynomials, like the Permanent, cannot be represented by polynomial-sized arithmetic circuits, which in turn would follow if identity testing can be derandomized using a certain kind of pseudo-random generator [Agr05, KI03]. (Note that identity testing has a simple and efficient randomized algorithm - pick a random point and evaluate the circuit at it [Sch80, Zip79, DL78].)

During the past decade, the quest for derandomization of PIT has yielded several results on restricted models of circuits. But, fortunately, the search has been made more focussed by a result [AV08, VSBR83] which states that a polynomial time *blackbox* derandomization of identity testing for depth-4 circuits (via a certain pseudo-random generator) implies a quasi-polynomial time derandomization of PIT for *poly-degree*[1] circuits. By polynomial time blackbox test, we mean:

> A polynomial time hitting-set generator, which is a boolean Turing machine that produces a set of points with *small* integer coordinates. These points are then fed (one by one) into the circuit, which internally uses *any* arithmetic, to output the evaluations of the polynomial at those points. (For small characteristic $p$, one works with a field extension, where each coordinate of a point is a small vector of integers in $\{0, .., p-1\}$.)

With depth-4 as the final frontier, the results that have been achieved so far include polynomial time hitting-set generators for the following models:

- depth-2 ($\Sigma\Pi$) circuits (equivalently, the class of *sparse* polynomials) [KS01],
- depth-3 ($\Sigma\Pi\Sigma$) circuits with constant top fanin [SS11],
- depth-4 ($\Sigma\Pi\Sigma\Pi$) multilinear circuits with constant top fanin [SV11],
- constant-depth constant-read multilinear formulas [AvMV11] (& sparse-substituted variants),
- circuits *generated* by sparse polynomials with constant transcendence degree [BMS11a].

To our knowledge, these are the only instances for which polynomial time hitting-set generators are known. The result on depth-3 bounded top fanin circuits is based upon the Chinese Remaindering technique of [KS07] and the ideal-theoretic framework studied in [SS10]. Their work followed after a sequence of developments in rank bound estimates [DS05, KS08, SS09, KS09, SS10], some using incidence geometry - although, this result [SS11] in particular is not rank based. On the other hand, the work on constant-depth multilinear formulas [AvMV11, SV11] is obtained by building upon and extending the techniques of other earlier results [KMSV10, SV09, SV08] on 'read-once' models. At a high level, this involved a study of the structure of multilinear formulas under the application of partial derivatives with respect to a carefully chosen set of variables and invoking depth-3 rank

[1]Circuits computing polynomials with degree bounded by a polynomial function in the size of the circuit.

bounds (cf. [SY10] for details). More recently, a third technique has emerged in [BMS11a] which is based upon the concept of *algebraic independence* of polynomials. They showed that for any given poly-degree circuit $C$ and sparse polynomials $f_1, \ldots, f_m$ with constant transcendence degree, a hitting-set generator for $C(f_1, \ldots, f_m)$ can be constructed in polynomial time.

**Our contribution -** With these diverse techniques floating around the study of hitting-set generators, one wonders: Could there be one single tool that is sufficiently powerful to capture all these models? Is there any unique feature underlying these seemingly different models that can lend itself to the conception of such a unifying tool? The answer to both these questions, as we show in this work, is *yes*. The key to this lies in studying the properties of the *Jacobian*, a mathematical object lying at the very core of algebraic independence. And as for the 'unique feature', notice that in the above four models some *parameter* of the circuit is *bounded* - be it bounded top fanin, bounded read of variables, or bounded transcendence degree. (Bounded depth should not be seen as an extra restriction on the circuit model because of [AV08]). At an intuitive level, it seems to us that it is this 'bounded parameter'-ness of the circuit that makes the Jacobian perform at its best.

In the process of finding a universal technique, we strengthen the earlier results significantly. We construct hitting-set generators not only for depth-3 circuits with bounded top fanin, but also for circuits of the form $C(T_1, \ldots, T_m)$, where $C$ is a poly-degree circuit and $T_1, \ldots, T_m$ are products, of linear polynomials, with bounded transcendence degree. In case of depth-3 circuits, $C(T_1, \ldots, T_m)$ is simply $T_1 + \ldots + T_m$. Further, we remove the restriction of multilinearity totally from the constant-depth constant-read model and construct the *first* hitting-set generator for this class. The condition of constant-read is also replaced by the more general notion of constant-*occur*.

At this point, one is faced with a natural question: how effective is this new tool in proving lower bounds? The intimate connection between efficient algorithms and lower bounds has recurrently appeared in various contexts [Wil11, Rag08, Uma03, PSZ00, IW97]. For arithmetic circuits, this link is provably tight [KI03, Agr05, AV08]: Derandomizing identity testing is *equivalent* to proving circuit lower bounds. Which means, one might have to look for techniques that are powerful enough to handle the dual worlds of algorithm design and lower bounds with equal effectiveness - for e.g. the *partial derivative technique* has been used to prove lower bounds and identity testing (albeit non-blackbox) on restricted models (survey [CKW11]); the $\tau$*-conjecture* is another such example [GKPS11]. In this work, we demonstrate a third tool - the Jacobian - using which we prove exponential lower bounds for the immanant (which includes determinant and permanent) on the *same* depth-3 and depth-4 models for which we give efficient PIT algorithms. In particular, this includes depth-4 constant-occur formulas, depth-4 circuits with constant transcendence degree of the underlying sparse polynomials (which significantly generalizes the lower bound result in [GKPS11]), and depth-3 circuits with constant transcendence degree of the polynomials computed by the product gates. To our knowledge, all these lower bounds are new and it is not known how to prove them using earlier techniques. (A gist of this paper is provided in Figure 1, Section 6.)

### 1.1 A tale of two PITs (& three lower bounds)

A set of polynomials $\mathbf{f} = \{f_1, \cdots, f_m\} \subset \mathbb{F}[x_1, \cdots, x_n]$ (in short, $\mathbb{F}[\mathbf{x}]$) is *algebraically independent* over $\mathbb{F}$ if there is no nonzero polynomial $H \in \mathbb{F}[y_1, \cdots, y_m]$ such that $H(f_1, \cdots, f_m)$ is identically zero. A maximal subset of $\mathbf{f}$ that is algebraically independent is a *transcendence basis* of $\mathbf{f}$ and the size of such a basis is the *transcendence degree*[2] of $\mathbf{f}$ (denoted $\mathrm{trdeg}_{\mathbb{F}}\, \mathbf{f}$). Our first theorem states:

[2]Since algebraic independence satisfies the matroid property cf. [Oxl92], transcendence degree is well-defined.

**Theorem 1.1.** *Let $C$ be a poly-degree circuit of size $s$ and each of $T_1, \ldots, T_m$ be a product of $d$ linear polynomials in $\mathbb{F}[x_1, \ldots, x_n]$ such that* $\mathrm{trdeg}_{\mathbb{F}} \{T_1, \ldots, T_m\} \leq r$*. A hitting-set for $C(T_1, \ldots, T_m)$ can be constructed in time polynomial in $n$ and $(sd)^r$, assuming* $\mathrm{char}(\mathbb{F}) = 0$ *or* $> d^r$*.*

If $C$ is a single $+$ gate, we get a hitting-set generator for depth-3 circuits with constant *transcendence degree* of the polynomials computed by the product gates (there is *no* restriction on top fanin).

Our second result uses the following generalization of *read-$k$* formulas (where every variable appears in at most $k$ leaf nodes of the formula) to *occur-$k$* formulas. Two reasons behind this generalization are: One, to accommodate the power of exponentiation - if we take the $e$-th power of a read-$k$ formula using a product gate, the 'read' of the resulting formula goes up to $ek$ - we would like to avoid this superfluous blow up in read. Two, a read-$k$ formula has size $O(kn)$, which severely hinders its power of computation - for instance, determinant and permanent cannot even be expressed in this model when $k$ is a constant [Kal85]. This calls for the following definition.

**Definition 1.** *An* occur-$k$ formula *is a rooted tree with internal gates labelled by $+$ and $\times\curlywedge$ (power-product gate). A $\times\curlywedge$ gate, on inputs $g_1, \ldots, g_m$ with incoming edges labelled $e_1, \ldots, e_m \in \mathbb{N}$, computes $g_1^{e_1} \cdots g_m^{e_m}$. At the leaves of this tree are depth-2 formulas computing sparse polynomials (*leaf nodes*), where every variable occurs in at most $k$ of these sparse polynomials.*

Size of a $\times\curlywedge$ gate is defined as the integer $(e_1 + \cdots + e_m)$ associated with its incoming edges, while size of a $+$ gate is counted as one. Size of a leaf node is the size of the corresponding depth-2 formula. With these conventions, *size* of an occur-$k$ formula is defined to be the total size of all its gates (and leaf nodes) plus the number of edges. *Depth* is defined to be the number of layers of $+$ and $\times\curlywedge$ gates plus 2 (the 'plus 2' accounts for the depth-2 formulas at the leaves). Thus, occur-$k$ is more relaxed than the traditional read-$k$ as it packs the "power of powering" (to borrow from [GKPS11]), and the leaves are sparse polynomials (at most $kn$ many) whose dependence on its variables is arbitrary. E.g. $(x_1^3 x_2 + x_1^2 x_3^2 + x_1 x_4)^e$ is *not* read-1 but is trivially depth-3 occur-1.

**Theorem 1.2.** *A hitting-set for a depth-$D$ occur-$k$ formula of size $s$ can be constructed in time polynomial in $s^R$, where $R = (2k)^{2D \cdot 2^D}$ (assuming* $\mathrm{char}(\mathbb{F}) = 0$ *or* $> s^R$*).*

A tighter analysis for depth-4 occur-$k$ formulas yields a better time complexity. Note that a depth-4 occur-$k$ formula allows unbounded top fanin. Also, it can be easily seen to subsume $\Sigma\Pi\Sigma\Pi(k)$ multilinear circuits studied by [SV11, KMSV10].

**Theorem 1.3.** *A hitting-set for a depth-4 occur-$k$ formula of size $s$ can be constructed in time polynomial in $s^{k^2}$ (assuming* $\mathrm{char}(\mathbb{F}) = 0$ *or* $> s^{2k}$*).*

For constant-depth, the above theorems not only remove the restriction of multilinearity (and relax read-$k$ to occur-$k$), but further improve upon the time complexity of [AvMV11] and [SV11]. The hitting-set generator of [AvMV11] works in time $n^{k^{O(k^2)} + O(kD)}$, which is super-exponential when $k = \Omega(s^{\varepsilon / 2D \cdot 2^D})$ for any positive $\varepsilon < 1$ and a constant $D$, whereas the generator in Theorem 1.2 runs in sub-exponential time for the same choice of parameters. The running time of [SV11] is $s^{O(k^3)}$, which is slightly worse than that of Theorem 1.3.

Since any polynomial has an exponential-sized depth-2, occur-1 formula (just the sparse representation), proving lower bounds on this model is an interesting proposition in its own right.

**Definition 2.** *[LR34] For any character $\chi : S_n \to \mathbb{C}^\times$, the* immanant *of a matrix $M = (x_{ij})_{n \times n}$ with respect to $\chi$ is defined as $Imm_\chi(M) = \sum_{\sigma \in S_n} \chi(\sigma) \prod_{i=1}^{n} x_{i\sigma(i)}$.*

Determinant & permanent are special cases of the immanant with $\chi$ as the alternating sign character & the identity character, respectively. Denote $\mathrm{Imm}_\chi(M)$ by $\mathrm{Imm}_n$ for an arbitrarily fixed $\chi$.

**Theorem 1.4.** *Any depth-4 occur-k formula that computes $Imm_n$ must have size $s = 2^{\Omega\left(n/k^2\right)}$ over any field of characteristic zero (even counting each $\times\curlywedge$ gate as size one).*

Thus, if each variable occurs in at most $n^{1/2-\varepsilon}$ $(0 < \varepsilon < 1/2)$ many underlying sparse polynomials, it takes an exponential sized depth-4 circuit to compute $\mathrm{Imm}_n$. Our next result is an exponential lower bound on the model for which hitting-set was developed in [BMS11a] (but no lower bound was shown). It is also an improvement over the result obtained in [GKPS11] which holds only for more restricted depth-4 circuits over reals.

**Theorem 1.5.** *Let $C$ be any circuit. Let $f_1, \ldots, f_m$ be sparse polynomials (of any degree) with sparsity bounded by $s$ and their* trdeg *bounded by $r$. If $C(f_1, \ldots, f_m)$ computes $Imm_n$ then $s = 2^{\Omega(n/r)}$ over any field of characteristic zero.*

Which means, any circuit involving fewer than $n^{1-\varepsilon}$ $\Sigma\Pi$-polynomials at the last levels, must have exponential size to compute $\mathrm{Imm}_n$. (The models of Theorem 1.4 & 1.5 are incomparable). The next result is on the model for which hitting-set is given by Theorem 1.1.

**Theorem 1.6.** *Let $C$ be any circuit and $T_1, \ldots, T_m$ be products of linear polynomials. If $C(T_1, \ldots, T_m)$ computes $Imm_n$ then $\mathrm{trdeg}_{\mathbb{F}}\{T_1, \ldots, T_m\} = \Omega(n)$ over any field of characteristic zero.*

Which means, any circuit involving $o(n)$ $\Pi\Sigma$-polynomials at the last levels *cannot* compute $\mathrm{Imm}_n$.

### 1.2 Our ideas

The exact reasons why our techniques work, where older ones failed, are extremely technical. However, we now give the motivating, but imprecise, ideas. To a set of products of sparse polynomials $\{T_1, \ldots, T_m\}$ we associate a polynomial – the Jacobian $J(T_1, \ldots, T_r)$. It captures the algebraic independence of $T_1, \ldots, T_r$ (assuming this to be a transcendence basis of the $T_i$'s). If we could find an $r$-variate linear map $\varphi$ that keeps $\varphi \circ J(T_1, \ldots, T_r)$ nonzero, then $\varphi(T_1), \ldots, \varphi(T_r)$ are again algebraically independent and it can be shown that for *any* $C$: $C(T_1, \ldots, T_m) = 0$ iff $C(\varphi(T_1), \ldots, \varphi(T_m)) = 0$. Since $T_i$'s are not sparse, the Jacobian is usually a difficult polynomial to work with, and so is finding $\varphi$. However, for the special models in this paper we are able to design $\varphi$ - mainly because Jacobian (being defined via partial derivatives) has a nice 'linearizing effect', on the circuit product gates, that factors itself. The $\varphi$ ultimately provides a hitting-set for $C(T_1, \ldots, T_m)$, as we reduce to a situation where $r$ is constant.

The initial idea for lower bounds is similar. Suppose $\mathrm{Imm}_n = C(T_1, \ldots, T_m)$. Then, by algebraic dependence, $J(\mathrm{Imm}_n, T_1, \ldots, T_r) = 0$. Our proofs then exploit the nature of this identity for the special models. This part requires proving several combinatorial properties of the immanant.

## 2 Preliminaries: Jacobian and faithful homomorphisms

Our contribution, in this section, is an elementary proof of Theorem 2.1, which was originally proved in [BMS11a] using Krull's *Hauptidealsatz*. Here, we state the main properties of the Jacobian and faithful homomorphisms without proofs - for details, refer to [BMS11b] (or Appendix A.1).

**Definition 3.** *The* Jacobian *of polynomials $\mathbf{f} = \{f_1, \cdots, f_m\}$ in $\mathbb{F}[x_1, \cdots, x_n]$ is the matrix $\mathcal{J}_{\mathbf{x}}(\mathbf{f}) = (\partial_{x_j} f_i)_{m\times n}$, where $\partial_{x_j} f_i = \partial f_i/\partial x_j$. Let $S \subseteq \mathbf{x} = \{x_1, \ldots, x_n\}$ and $|S| = m$. Then $J_S(\mathbf{f})$ denote the minor of $\mathcal{J}_{\mathbf{x}}(\mathbf{f})$ formed by the columns corresponding to the variables in $S$.*

**Fact 1** (Jacobian criterion)**.** *Let $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ be a finite set of polynomials of degree at most $d$, and $\mathrm{trdeg}_{\mathbb{F}}\,\mathbf{f} \le r$. If $\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) > d^r$, then $\mathrm{trdeg}_{\mathbb{F}}\,\mathbf{f} = \mathrm{rank}_{\mathbb{F}(\mathbf{x})}\mathcal{J}_{\mathbf{x}}(\mathbf{f})$.*

**Fact 2** (Chain rule). *For any finite set of polynomials* $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ *and a homomorphism* $\Phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{y}]$, *we have* $\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{f})) = \Phi\left(\mathcal{J}_{\mathbf{x}}(\mathbf{f})\right) \cdot \mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{x}))$.

**Definition 4.** *A homomorphism* $\Phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{y}]$ *(*$\mathbf{y}$ *is another set of variables) is said to be* faithful *to a finite set of polynomials* $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ *if* $\mathrm{trdeg}_{\mathbb{F}}\, \mathbf{f} = \mathrm{trdeg}_{\mathbb{F}}\, \Phi(\mathbf{f})$.

**Theorem 2.1** (Faithful is useful). *Let* $\mathbf{f} = \{f_1, \cdots, f_m\} \subset \mathbb{F}[\mathbf{x}]$ *and* $\Phi$ *be a homomorphism faithful to* $\mathbf{f}$. *For any polynomial* $C \in \mathbb{F}[y_1, \cdots, y_m]$, $C(\mathbf{f}) = 0 \Leftrightarrow C(\Phi(\mathbf{f})) = 0$.

**Lemma 2.2** (Vandermonde is faithful). *Let* $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ *be a finite set of polynomials of degree at most* $d$, $\mathrm{trdeg}_{\mathbb{F}}\, \mathbf{f} \le r$, *and* $\mathrm{char}(\mathbb{F}) = 0$ *or* $> d^r$. *Let* $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ *be a homomorphism such that* $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}\mathcal{J}_{\mathbf{x}}(\mathbf{f}) = \mathrm{rank}_{\mathbb{F}(\mathbf{z})}\Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f}))$.

*Then, the map* $\Phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}, t, y_1, \ldots, y_r]$ *that maps, for all* $i$, $x_i \mapsto \left(\sum_{j=1}^{r} y_j t^{ij}\right) + \Psi(x_i)$ *is a homomorphism faithful to* $\mathbf{f}$.

The proof of the above lemma is based upon Facts 1, 2 and an application of 'rank preserving' linear maps [GR05]. See Appendix A.1.

# 3 Hitting-set for constant transcendence degree depth-3 circuits

Let $C$ be *any* circuit and $D$ be the circuit $C(T_1, \cdots, T_m)$, where each $T_i$ is of the form $\prod_{j=1}^{d} \ell_{ij}$, every $\ell_{ij}$ is a linear polynomial in $\mathbb{F}[x_1, \ldots, x_n]$. Denote by $\mathbf{T}$ the set $\{T_1, \ldots, T_m\}$ and by $L(T_i)$ the multiset of linear polynomials that constitute $T_i$. Suppose $\mathrm{trdeg}_{\mathbb{F}}\, \mathbf{T} = k \le r$ and $\mathbf{T}_k = \{T_1, \ldots, T_k\}$ be a transcendence basis of $\mathbf{T}$. Since $\mathcal{J}_{\mathbf{x}}(\mathbf{T}_k)$ has full rank ($\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) > d^r$), without loss of generality assume that the columns corresponding to $\mathbf{x}_k = \{x_1, \cdots, x_k\}$ form a nonzero $k \times k$ minor of $\mathcal{J}_{\mathbf{x}}(\mathbf{T}_k)$. By Lemma 2.2, if we construct a $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ that keeps $J_{\mathbf{x}_k}(\mathbf{T}_k)$ nonzero then $\Psi$ can easily be extended to a homomorphism $\Phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}, t, y_1, \ldots, y_r]$ that is faithful to $\mathbf{T}$. And hence, by Theorem 2.1, it would follow that $\Phi(D) = 0$ if and only if $D = 0$.

If $T_i = \prod_{j=1}^{d} \ell_{ij}$ then $\partial_x T_i = T_i \cdot \left(\sum_{j=1}^{d} \partial_x \ell_{ij}/\ell_{ij}\right)$. By expanding, using this additive structure of $\partial_x T_i$ and the linearity of determinant wrt rows, the determinant $J_{\mathbf{x}_k}(\mathbf{T}_k)$ takes the following form,

$$J_{\mathbf{x}_k}(\mathbf{T}_k) = \sum_{\ell_1 \in L(T_1), \ldots, \ell_k \in L(T_k)} \frac{T_1 \cdots T_k}{\ell_1 \cdots \ell_k} \cdot J_{\mathbf{x}_k}(\ell_1, \cdots, \ell_k). \tag{1}$$

Call a set of linear polynomials *independent* if the correponding homogenous linear parts (i.e. the constant-free parts) are $\mathbb{F}$-linearly independent. The term $J_{\mathbf{x}_k}(\ell_1, \cdots, \ell_k)$ ensures that the above sum is only over those $\ell_1, \cdots, \ell_k$ that are independent linear polynomials (otherwise the Jacobian vanishes). The sum has the form of a depth-3 circuit, call it $H_0$, and we intend to construct a low-variate $\Psi$ such that $\Psi(H_0) \neq 0$. We show that this is achieved by a $\Psi$ that preserves the independence of a 'small' set of linear polynomials - which we call a *certificate* of $H_0$.

**Certificate of $H_0$:** We can assume that the terms $J_{\mathbf{x}_k}(\ell_1, \cdots, \ell_k)$ in equation (1) are *nonzero* field constants. Let $\mathcal{L}(H_0)$ be the set of all linear polynomials occurring in the denominator terms "$\ell_1 \cdots \ell_k$" of all the summands in sum (1). By adjusting the field constants at the numerators, we can assume that no two linear polynomials in $\mathcal{L}(H_0)$ are constant multiple of each other. This means, the depth-3 circuit $H_0$ has the form $H_0 = T \cdot \sum_L \alpha_L/\ell_1 \cdots \ell_k$, where $T := \prod_{i=1}^{k} T_k$, $\alpha_L$ is a nonzero field constant and the sum runs over *some* sets $L = \{\ell_1, \cdots, \ell_k\}$ of $k$ independent linear polynomials in $\mathcal{L}(H_0)$. Define, *content* of a depth-3 circuit $G = \sum_i P_i$, where $P_i$ is a product of linear polynomials, as $\text{cont}(G) := \gcd_i\{P_i\}$, and let the *simple part* $\text{sim}(G) := G/\text{cont}(G)$. Hence $\text{cont}(H_0) = \gcd_L\{T/\ell_1 \cdots \ell_k\}$ and

$$\text{sim}(H_0) = F_0 \cdot \sum_L \frac{\alpha_L}{\ell_1 \cdots \ell_k}, \text{ where } F_0 = \frac{T}{\text{cont}(H_0)}, \tag{2}$$

Note that, since $\ell \in \mathcal{L}(H_0)$ iff $\ell || F_0$, $F_0$ is simply the product of the linear polynomials in $\mathcal{L}(H_0)$ and so $\deg(F_0) = |\mathcal{L}(H_0)|$. For any $\ell \in \mathcal{L}(H_0)$, the terms in $\text{sim}(H_0)$ that survive modulo $\ell$ are those with $\ell$ in the denominator "$\ell_1 \cdots \ell_k$" of the above expression. Hence, $H_1 := \text{sim}(H_0) \bmod \ell_1 = F_0/\ell_1 \cdot \sum_{\ell_2,\cdots,\ell_k} \alpha_L/\ell_2 \cdots \ell_k$. We can treat $H_1$ as a depth-3 circuit in one less variable: Suppose that $\ell_1 = c_1 x_1 + \sum_{i=2}^n c_i x_i$ where $c_i$'s $\in \mathbb{F}$ and $c_1 \neq 0$, then we can replace $x_1$ by $-\sum_{i=2}^n c_i x_i / c_1$ in $\text{sim}(H_0)$, particularly in $F_0/\ell_1$ (of course, after dividing $F_0$ by $\ell_1$) as well as in each of $\ell_2, \ldots, \ell_k$ in the denominators, so that $H_1$ becomes a depth-3 circuit in $\mathbb{F}[x_2, \ldots, x_n]$. Therefore, it makes perfect sense to talk about $\text{cont}(H_1)$ and $\text{sim}(H_1)$. Observe that $\ell_2, \cdots, \ell_k$ remain independent linear polynomials modulo $\ell_1$, and so $H_1$ is a depth-3 circuit of the 'same nature' as $H_0$ but with one less linear polynomials in the denominators. Also, the linear polynomials in $\mathcal{L}(H_1)$ is a subset of the linear polynomials in $\mathcal{L}(H_0)$ modulo $\ell_1$. Extending the above argument, it is possible to define a sequence of circuits: $H_i := \text{sim}(H_{i-1}) \mod \tilde{\ell}_i$, $(1 \leq i \leq k)$ where $\tilde{\ell}_i \in \mathcal{L}(H_{i-1})$. Further, $\mathcal{L}(H_i)$ is a subset of $\mathcal{L}(H_{i-1})$ modulo $\tilde{\ell}_i$, which implies that essentially there are independent linear polynomials, say $\ell_1, \ldots, \ell_k$, in $\mathcal{L}(H_0)$ such that $\tilde{\ell}_i = \ell_i \bmod (\ell_1, \ldots, \ell_{i-1})$ and therefore $H_i = \text{sim}(H_{i-1}) \mod (\ell_1, \ldots, \ell_i)$.

**Lemma 3.1** (Certifying path)**.** *There exists independent linear polynomials $\{\ell_1, \cdots, \ell_k\} \subseteq \mathcal{L}(H_0)$ such that $H_i \neq 0 \bmod (\ell_1, \ldots, \ell_i)$, $\forall i \in [k]$, and $H_k$ is a nonzero product of linear polynomials in $\mathcal{L}(H_0)$ modulo $(\ell_1, \cdots, \ell_k)$.*
Proof: Induction on $k$, see Appendix A.2.

A set $\{\ell_1, \ldots, \ell_k\}$, satisfying Lemma 3.1, is called a *certifying path* of $H_0$. Fix a certifying path $\{\ell_1, \cdots, \ell_k\}$. Let $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[z_1, \ldots, z_{k+1}]$ be such that $\Psi(\ell_1), \ldots, \Psi(\ell_k)$ are independent linear polynomials in $\mathbb{F}[\mathbf{z}]$ and for every $\ell \in \cup_{i=1}^k L(T_i)$, $\ell \neq 0 \bmod (\ell_1, \ldots, \ell_k)$ iff $\Psi(\ell) \neq 0 \bmod (\Psi(\ell_1), \ldots, \Psi(\ell_k))$. We call such a $\Psi$ a *rank-$(k+1)$ preserving map* for $H_0$. It can be shown that one of the maps $\Psi_b : x_i \mapsto \sum_{j=1}^{k+1} z_j b^{ij}$, where $b$ runs over $dkn(k+1)^2$ distinct elements of $\mathbb{F}$, is a rank-$(k+1)$ preserving map for $H_0$. (It is a simple application of [GR05]. See Corollary A.2).

**Theorem 3.2** (Certificate)**.** *If $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[z_1, ..., z_{k+1}]$ is a rank-$(k+1)$ preserving map for $H_0$, then $\Psi(H_0) \neq 0$.*
Proof: Reverse induction on $k$, see Appendix A.2.

*Proof of Theorem 1.1.* As $r \geq k$, we can assume that the rank-$(k+1)$ preserving map $\Psi$ is in fact a map from $\mathbb{F}[\mathbf{x}] \to \mathbb{F}[z_1, \ldots, z_{r+1}]$. Therefore, by Lemma 2.2, $\Phi$ is a map from $\mathbb{F}[\mathbf{x}] \to \mathbb{F}[y_1, \ldots y_r, t, z_1, \ldots, z_{r+1}]$ such that: $D = 0$ iff $\Phi(D) = 0$. Since $C$ is a poly-degree circuit of size $s$, $\Phi(C(T_1, \ldots, T_m))$ is a polynomial of degree at most $ds^{O(1)}$ resp. $nrds^{O(1)}$ in the variables $\mathbf{y}, \mathbf{z}$ resp. $t$. Using [Sch80, Zip79, DL78] lemma, we can construct a hitting-set for $\Phi(D)$ in time polynomial in $n(sd)^r$. Since construction of $\Psi$ takes time $\mathsf{poly}(ndr)$, the total time taken is $\mathsf{poly}(n, (sd)^r)$. $\square$

# 4 Hitting-set for constant-depth constant-occur formulas

**Bounding the top fanin -** Let $C$ belong to the class $\mathcal{C}$ of depth-$D$ occur-$k$ formulas of size $s$. Observe that if $C(x_1, \ldots, x_n)$ is non-constant and nonzero, then there is an $i$ such that $\tilde{C} := C(x_1, \cdots, x_{i-1}, x_i + 1, x_{i+1}, \cdots, x_n) - C(x_1, \cdots, x_n) \neq 0$, assuming $\text{char}(\mathbb{F}) > s^D$ (i.e. the bound on the degree of $C$). If $C$ has a $+$ gate on top then $C(\mathbf{x}) = \sum_{i=1}^{m} T_i$, where $T_i$'s are computed by $\times\curlywedge$ gates at the next level. Since $x_i$ occurs in at most $k$ of the $T_i$'s, $\tilde{C}$ has top fanin at most $2k$. If $C$ has a $\times\curlywedge$ gate on top then $\tilde{C}$ has a $+$ gate on top with fanin 2 and depth$(\tilde{C}) = D+1$. Therefore, $\tilde{C}$ belongs to the class $\tilde{\mathcal{C}}$ of depth-$(D+1)$ occur-$2k$ formulas of size at most $(s^2+s)$, and a $+$ gate on top with fanin bounded by $2k$. Suppose $\tilde{\mathcal{H}}$ is a hitting-set for the class $\tilde{\mathcal{C}}$. Form a new set $\mathcal{H} \supset \tilde{\mathcal{H}}$ by including points $(\alpha_1+1, \alpha_2, \ldots, \alpha_n), (\alpha_1, \alpha_2+1, \ldots, \alpha_n), \ldots, (\alpha_1, \ldots, \alpha_{n-1}, \alpha_n+1)$ in $\mathcal{H}$ for every $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \tilde{\mathcal{H}}$. Observe that $\mathcal{H}$ is a hitting-set for $\mathcal{C}$ and size$(\mathcal{H}) = n \cdot$ size$(\tilde{\mathcal{H}})$. Therefore, it is sufficient if we construct $\tilde{\mathcal{H}}$. By reusing symbols, assume that $C$ is a depth-$D$ occur-$k$ formula of size $s$ with a $+$ gate on top having top fanin at most $k$.

Let $C(\mathbf{x}) = \sum_{i=1}^{k} T_i$. The goal is to construct a $\Phi$ that is faithful to $\mathbf{T} = \{T_1, \ldots, T_k\}$. Let $\mathbf{T}_r = \{T_1, \ldots, T_r\}$ be a transcendence basis of $\mathbf{T}$. Since $\mathcal{J}_{\mathbf{x}}(\mathbf{T}_r)$ has full rank (char$(\mathbb{F}) = 0$ or $> s^{Dr}$), assume that the columns corresponding to $\mathbf{x}_r = \{x_1, \ldots, x_r\}$ form a nonzero minor of $\mathcal{J}_{\mathbf{x}}(\mathbf{T}_r)$. By Lemma 2.2, it suffices to construct a $\Psi$ that keeps $J_{\mathbf{x}_r}(\mathbf{T}_r) \neq 0$.

**Proof idea -** Identify a gate with the polynomial it computes, and count *level* of a gate from the top - the gates $T_i$'s are at level 2. Suppose each $T_i$ is a $\times\curlywedge$ gate and $T_i = \prod_{\ell=1}^{d} P_{i,\ell}^{e_{i,\ell}}$, where $P_{i,\ell}$'s are gates at level 3. Since $T_i$ is also an occur-$k$ formula, $x_1, \ldots, x_r$ appear in at most $kr$ of the $P_{i,\ell}$'s, say $P_{i,1}, \ldots P_{i,kr}$. Hence, $\partial_j T_i = (\prod_{\ell=kr+1}^{d} P_{i,\ell}^{e_{i,\ell}}) \cdot (\partial_j \prod_{\ell=1}^{kr} P_{i,\ell}^{e_{i,\ell}})$ for every $1 \le i, j \le r$ and therefore, $J_{\mathbf{x}_r}(\mathbf{T}_r) = (\prod_{i=1}^{r} \prod_{\ell=kr+1}^{d} P_{i,\ell}^{e_{i,\ell}}) \cdot \det(\partial_j \prod_{\ell=1}^{kr} P_{i,\ell}^{e_{i,\ell}})$. Now notice that $\det(\partial_j \prod_{\ell=1}^{kr} P_{i,\ell}^{e_{i,\ell}})$ is a polynomial in $P_{i,\ell}$ and $\partial_j P_{i,\ell}$, for $1 \le i, j \le r$ and $1 \le \ell \le kr$. (Note the irrelevance of the exponents $e_{i,\ell}$'s.) So, if $\Psi$ is faithful to the set $\mathcal{P} := \{P_{i,\ell}, \partial_j P_{i,\ell} : 1 \le i, j \le r, 1 \le \ell \le kr\}$ and the singleton sets $\{P_{i,\ell}\}$ for $1 \le i \le r$, $kr+1 \le \ell \le d$, then $\Psi(J_{\mathbf{x}_r}(\mathbf{T}_r)) \neq 0$. Observe that the polynomials in $\mathcal{P}$ and the singleton sets are (zeroth and first order) derivatives of the gates at level 3, and further these sets involve (the derivatives of) *disjoint* groups of level-3 gates. This disjointness feature ensures that the number of such sets is at most $s$. Thus, we have reduced the problem of constructing a faithful map $\Phi$ for $\mathbf{T}$ (gates at level 2) to the problem of constructing a map $\Psi$ that is faithful to at most $s$ many sets each containing derivatives of gates at the third level. Now, the idea is to carry forward this argument recursively to deeper levels: In the next level of the recursion we reduce the problem to constructing a map that is faithful to at most $s$ sets containing (zeroth, first and second order) derivatives of disjoint groups of gates at level 4, and so on. Eventually, the recursion reaches the level of the sparse polynomials (the leaf nodes) where a faithful map can be constructed using ideas from [KS01].

Let us formalize this proof idea. For any *multiset* of variables $S$, let $\Delta_S f$ denote the partial derivative of $f$ with respect to the variables in $S$ (including repetitions, as $S$ is a multiset). Let var$(S)$ denote the set of distinct variables in $S$.

**Lemma 4.1** (Gcd trick)**.** *Let $G$ be any gate in $C$ and $S_1, \cdots, S_w$ be multisets of variables. Then there exists another occur-$k$ formula $G'$ for which, the vector of polynomials $(\Delta_{S_1} G, \cdots, \Delta_{S_w} G) = V_G \cdot (\Delta_{S_1} G', \cdots, \Delta_{S_w} G')$ such that*

1. *If $G$ is a $+$ gate then $G'$ is also a $+$ gate whose children consist of at most $k \cdot |\cup_{i=1}^{w} \mathrm{var}(S_i)|$ of the children of $G$, and $V_G = 1$.*
2. *If $G$ is a $\times\curlywedge$ gate, then $G'$ is also a $\times\curlywedge$ gate whose children consist of at most $k \cdot |\cup_{i=1}^{w} \mathrm{var}(S_i)|$ of the children of $G$, and $V_G = G/G'$.*

*Further, the gates constituting $G'$ and $V_G$ are disjoint.*
Proof: Use properties of derivation and occur-$k$, see Appendix A.3.

We say that a map is faithful to a collection of sets if it is faithful to every set in the collection. Going by the 'proof idea', suppose at the $\ell$-th level of the recursion we want to construct a $\Psi_\ell$ that is faithful to a collection of (at most) $s$ sets of polynomials, each set containing at most $r_\ell$ partial derivatives (of order up to $c_\ell$) of the gates at level $\ell$. Moreover, the sets involve derivatives of disjoint groups of gates. To begin with: $\ell = 2$ and we wish to construct a $\Psi_2$ that is faithful to just one set $\mathbf{T}$, so $r_2 \leq k$ and $c_2 = 0$. The next lemma captures the evolution of the recursion.

**Lemma 4.2** (Evolution via factoring)**.** *Let $\mathcal{U}$ be a set of $r_\ell$ derivatives (of orders up to $c_\ell$) of gates $\mathcal{G}_\mathcal{U}$ at level $\ell$, and $\mathcal{U}'$ be a transcendence basis of $\mathcal{U}$. Any $|\mathcal{U}'| \times |\mathcal{U}'|$ minor of $\mathcal{J}_\mathbf{x}(\mathcal{U}')$ is of the form $\prod_i V_i^{e_i}$, where $V_i$'s are polynomials in at most $r_{\ell+1} := (c_\ell + 1) \cdot 2^{c_\ell+1}k \cdot {r_\ell}^2$ many derivatives (of order up to $c_{\ell+1} := c_\ell + 1$) of disjoint groups of children of $\mathcal{G}_\mathcal{U}$.*
Proof: Apply the gcd trick for each $G \in \mathcal{G}_\mathcal{U}$, see Appendix A.3.

Let $\mathcal{C}_\ell$ denote the collection of sets for which we want to construct a faithful map $\Psi_\ell$ at the $\ell$-th level of the recursion. The collection $\mathcal{C}_{\ell+1}$ is formed from $\mathcal{C}_\ell$ using the above lemma: $V_i$ is a polynomial in a set of derivatives of gates at the $(\ell + 1)$-th level - denote this set of derivatives by $\mathrm{Elem}(V_i)$ - then $\mathcal{C}_{\ell+1}$ consists of the sets $\mathrm{Elem}(V_i)$ as $\mathcal{U}$ varies over all the sets in the collection $\mathcal{C}_\ell$. It follows from the lemma that the groups of gates whose derivatives form the different $\mathrm{Elem}(V_i)$'s are disjoint and therefore $|\mathcal{C}_{\ell+1}| \leq s$. Using Lemma 2.2 & 4.2, we can lift a map $\Psi_{\ell+1}$ to construct $\Psi_\ell$.

**Corollary 4.3.** *If $\Psi_{\ell+1}$ is faithful to $\mathcal{C}_{\ell+1}$ then $\Psi_\ell : x_i \mapsto \left(\sum_{j=1}^{r_\ell} y_{j,\ell} \cdot (t_\ell)^{ij}\right) + \Psi_{\ell+1}(x_i)$ is faithful to $\mathcal{C}_\ell$, where $\{y_{1,\ell}, \cdots, y_{r_\ell,\ell}, t_\ell\}$ is a fresh set of variables.*

*Proof of Theorem 1.2.* Unfolding the recursion, we eventually reach the level of the sparse polynomials at depth $D-2$ and are required to construct a map $\Psi_{D-2}$ that is faithful to a collection $\mathcal{C}_{D-2}$ of at most $s$ sets of derivatives of sparse polynomials, each set containing at most $r_{D-2}$ elements. Using the relation between $r_{\ell+1}$ and $r_\ell$ from Lemma 4.2, it is easy to bound $r_{D-2}$ by $R = (2k)^{2D \cdot 2^D}$. Let $\mathcal{U} \in \mathcal{C}_{D-2}$ with transcendence basis $\mathcal{U}'$. Any $|\mathcal{U}'| \times |\mathcal{U}'|$ minor of $\mathcal{J}_\mathbf{x}(\mathcal{U}')$ is a sparse polynomial with sparsity bounded by $s^R$ and degree bounded by $sR$. Using [KS01], the nonzeroness of this determinant is maintained by one of the maps $\Phi_p : x_i \mapsto u^{(sR+1)^i \bmod p}$ as $p$ varies from 1 to a fixed $\mathsf{poly}(s^R)$. Since $|\mathcal{C}_{D-2}| \leq s$, one of the maps $\Phi_p$, $1 \leq p \leq s \cdot \mathsf{poly}(s^R)$, preserves the rank of the Jacobian of all $\mathcal{U}$ in $\mathcal{C}_{D-2}$ - fix such a $\Phi_p$. Finally, $\Psi_{D-2} : x_i \mapsto \sum_{j=1}^{R} y_{j,D-2} t_{D-2}^{ij} + \Phi_p(x_i)$ is faithful to $\mathcal{C}_{D-2}$. Now lift this map $\Psi_{D-2}$ to $\Psi_2$ that is faithful to $\mathbf{T}$ using Corollary 4.3. The map $\Psi_2$ reduces the number of variables to $O(R)$ and hence an application of [Sch80, Zip79, DL78] lemma leads to a hitting-set generator with running time $\mathrm{poly}(s^R)$. For the Jacobian criterion to work we need $\mathrm{char}(\mathbb{F}) = 0$ or $> s^R$. $\square$

### 4.1 Restriction to the case of depth-4

*Proof of Theorem 1.3.* Let $C = \sum_{i=1}^{k} T_i$ be a depth-4 occur-$k$ formula, where $T_i = \prod_{j=1}^{d} P_{ij}^{e_{ij}}$, $P_{ij}$'s are sparse polynomials. The discussion at the beginning of this section justifies the assumption that top fanin is $k$. Once again, assuming $\mathbf{T}_r$ to be a transcendence basis of $\mathbf{T}$, we need to design a $\Psi$ such that $\Psi(J_{\mathbf{x}_r}(\mathbf{T}_r)) \neq 0$. Let us count the number of $P_{ij}$'s that depend on the variables $\mathbf{x}_r$, the remaining $P_{ij}^{e_{ij}}$'s can be taken out common from every row of $\mathcal{J}_{\mathbf{x}_r}(\mathbf{T}_r)$ while computing its determinant - this is the first 'taking common' step. Let $c_{i\ell}$ be the number of $P_{ij}$'s present in $T_i$ that depend on $x_\ell$. The total number of sparse polynomials depending on $\mathbf{x}_r$ is therefore $\sum_{1 \leq i,\ell \leq r} c_{i\ell}$. From the condition of occur-$k$, $\sum_i c_{i\ell} \leq k$ and hence $\sum_{i,\ell} c_{i\ell} \leq rk \leq k^2$. Let $c_i := \sum_j c_{ij}$, be the number of $\mathbf{x}_r$-dependent $P_{ij}$'s present in $T_i$. For an $\mathbf{x}_r$-dependent $P_{ij}$, we can also take $P_{ij}^{e_{ij}-1}$

common from the $i$-th row of $\mathcal{J}_{\mathbf{x}_r}(\mathbf{T}_r)$ - call this the second 'taking common' step. The sparsity of every entry of the $i$-th row of the residual matrix $M$ - after the two 'taking common' steps - is bounded by $c_i s^{c_i}$, where $s$ is the size of $C$. Thus, $\det(M)$ has sparsity at most $r! \cdot \prod_{i=1}^{r} c_i s^{c_i} = s^{O(k^2)}$, which implies that $J_{\mathbf{x}_r}(\mathbf{T}_r)$ is a product of at most $s+1$ powers of sparse polynomials, each of whose sparsity is bounded by $s^{O(k^2)}$ and degree bounded by $sk$. As argued before, use [KS01] along with Lemma 2.2 to construct a hitting-set for $C$ in time $s^{O(k^2)}$ (assuming $\mathrm{char}(\mathbb{F}) = 0$ or $> s^{2k}$). $\square$

# 5 Lower bounds for the immanant

For convenience, we prove the lower bounds for $\mathrm{Det}_n$ - determinant of an $n \times n$ matrix $M = (x_{ij})$ - assuming zero characteristic. All our arguments apply to $\mathrm{Imm}_\chi(M)$ for any character $\chi$ (Appendix A.5). The following two lemmas are at the heart of our approach to proving lower bounds. Let $\mathbf{x} := \{x_{ij} : 1 \le i, j \le n\}$ and $\mathbf{T} := \{T_1, \ldots, T_m\}$, where $T_i$'s are polynomials in $\mathbb{F}[\mathbf{x}]$. (See Appendix A.4 for the proofs.)

**Lemma 5.1.** *Suppose $\mathrm{Det}_n = C(T_1, \ldots, T_m)$, where $C$ is any circuit and let $\mathbf{T}_r = \{T_1, \ldots, T_r\}$ be a transcendence basis of $\mathbf{T}$ with $r < n$. Then, there exist a set of $r+1$ variables $\mathbf{x}_{r+1} \subset \mathbf{x}$ and an equation $\sum_{i=1}^{r+1} c_i f_i \cdot M_i = 0$ such that $M_i$'s are distinct first order principal minors of $M$, $f_i$'s are distinct $r \times r$ minors of $\mathcal{J}_{\mathbf{x}_{r+1}}(\mathbf{T}_r)$, not all $f_i$'s are zero, and $c_i \in \mathbb{F}^*$.*

**Lemma 5.2.** *If $M_1, \cdots, M_t$ are distinct first order principal minors of $M$ and $\sum_{i=1}^{t} f_i \cdot M_i = 0$ (not all $f_i$'s are zero) then the total sparsity of the $f_i$'s is at least $2^{n/2-t}$.*

## 5.1 Lower bound on depth-4 occur-$k$ formulas

*Proof of Theorem 1.4.* Let $C$ be a depth-4 occur-$k$ formula of size $s$ that computes $\mathrm{Det}_n$. Since $\mathrm{Det}_n$ is irreducible we can assume a top + gate in $C$. Then $\tilde{C} := C(x_{11}+1, x_{12}, \ldots, x_{nn}) - C(\mathbf{x})$ is a depth-4 occur-$2k$ formula of size at most $2s^2$ and top fanin bounded by $2k$ (similar argument as at the beginning of Section 4). Moreover, $\tilde{C}$ computes the minor of $M$ with respect to $x_{11}$ which is essentially $\mathrm{Det}_{n-1}$. By reusing symbols, assume that $C$ is a depth-4 occur-$k$ formula with top fanin bounded by $k$, and $C$ computes $\mathrm{Det}_n$.

Let $C = \sum_{i=1}^{k} T_i = \mathrm{Det}_n$, where $T_i = \prod_{j=1}^{d} P_{ij}^{e_{ij}}$, $P_{ij}$'s are sparse polynomials. Let $\mathbf{T}_r$ be a transcendence basis of $\mathbf{T} = \{T_1, \ldots, T_k\}$. By Lemma 5.1, we have an equation $\sum_{i=1}^{r+1} c_i f_i \cdot M_i = 0$ such that $f_i$'s are distinct $r \times r$ minors of $\mathcal{J}_{\mathbf{x}_{r+1}}(\mathbf{T}_r)$ for some set of $r+1$ variables $\mathbf{x}_{r+1}$. Arguing in the same way as in the proof of Theorem 1.3 (in Section 4.1), we can throw away certain common terms from the minors $f_i$'s and get another equation $\sum_{i=1}^{r+1} g_i M_i = 0$, where the sparsity of each $g_i$ is $s^{O(k^2)}$. If we apply Lemma 5.2 on this equation, we get our desired result. $\square$

## 5.2 Lower bound on circuits generated by $\Sigma\Pi$ polynomials

*Proof of Theorem 1.5.* In Lemma 5.1, take the $T_i$'s to be sparse polynomials with sparsity bounded by $s$. Then, in the equation $\sum_{i=1}^{r+1} c_i f_i \cdot M_i = 0$, each $f_i$ has sparsity $s^{O(r)}$. Finally, apply Lemma 5.2 to obtain the desired lower bound. $\square$

## 5.3 Lower bound on circuits generated by $\Pi\Sigma$ polynomials

*Proof of Theorem 1.6.* Let $\mathbf{T} = \{T_1, \cdots, T_m\}$ be products of linear polynomials such that $C(T_1, \cdots, T_m) = \mathrm{Det}_n$ with $\mathbf{T}_k = \{T_1, \cdots, T_k\}$ being a transcendence basis. By Lemma 5.1, we get $\sum_{i=1}^{k+1} c_i f_i M_i = 0$ where the $f_i$'s are $k \times k$ minors of $\mathcal{J}_{\mathbf{x}_{k+1}}(\mathbf{T}_k)$ and wlog $f_1 \neq 0$. Like in

Section 3, we can rewrite this equation in the form $H_0 := T \cdot \sum_L \alpha_L(\mathbf{M}_{k+1})/\ell_1 \cdots \ell_k = 0$ where $\alpha_L(\mathbf{M}_{k+1}) := \sum_{i=1}^{k+1} \alpha_{L,i} M_i$ is an $\mathbb{F}$-linear combination of $\mathbf{M}_{k+1} := \{M_1, \cdots, M_{k+1}\}$. Observe that $H_0$ is a sum of products of linear polynomials, with 'coefficients' being $\mathbb{F}$-linear combinations of $\mathbf{M}_{k+1}$. And since $f_1 \neq 0$, the 'coefficient' of $M_1$ in $H_0$ is a nonzero depth-3 circuit.

The idea is to apply a similar treatment as in Section 3 to evolve $H_0$. The invariant that shall be maintained is that the coefficient of $M_1$ (modulo some linear polynomials), which is a depth-3 circuit, would stay nonzero. This would finally yield a non-trivial linear combination $\alpha_L(\mathbf{M}_{k+1}) = 0 \bmod \boldsymbol{\ell}_k$ whence we can apply the following lemma. (See Appendix A.4.)

**Lemma 5.3.** *If $M_1, \cdots, M_t$ are distinct first order principal minors of $M$ and $\sum_{i=1}^{t} \alpha_i M_i = 0 \bmod \boldsymbol{\ell}_k$ (not all $\alpha_i = 0$) for independent linear polynomials $\boldsymbol{\ell}_k$, then $t + k \geq n$.*

Formally, define the *content* of a circuit $H = T \sum_L \alpha_L(\mathbf{M}_{k+1})/\ell_1 \cdots \ell_k$ as $\mathrm{cont}(H) := \gcd_L \left\{ \frac{T}{\ell_1 \cdots \ell_k} \right\}$, and define $\mathrm{sim}(H) := H/\mathrm{cont}(H)$. Let $\mathrm{sim}(H_0)$ have the form $F_0 \sum_L \alpha_L(\mathbf{M}_{k+1})/\ell_1 \cdots \ell_k$. The coefficient of $M_1$ in the above expression is a nonzero depth-3 circuit, whose degree is $|\mathcal{L}(H_0)| - k$. Therefore by Chinese remaindering, $\exists \ell_1 \in \mathcal{L}(H_0)$ such that this coefficient is nonzero modulo $\ell_1$. Hence, we can define $H_1 := \mathrm{sim}(H_0) \bmod \ell_1$ which has the form $H_1 = F_0/\ell_1 \cdot \sum_{L \ni \ell_1} \alpha_L(\mathbf{M}_{k+1})/\ell_2 \cdots \ell_k = 0 \bmod \ell_1$. And like in Section 3, the above equation can be rewritten by replacing a variable occuring in $\ell_1$ by a suitable linear combination of the rest. Thus, we may write $H_1 = F_1 \sum_L \alpha_L(\mathbf{M}_{k+1} \bmod \ell_1)/\ell_2 \cdots \ell_k = 0$, and maintaining the invariant that the coefficient of $M_1 \bmod \ell_1$ is nonzero. Repeating this argument, we eventually obtain $H_k := F_k \cdot \alpha_L(\mathbf{M}_{k+1} \bmod \boldsymbol{\ell}_k) = 0$ while the coefficient of $M_1 \bmod \boldsymbol{\ell}_k$ is nonzero. This implies that $\alpha_L(\mathbf{M}_{k+1}) = 0 \bmod \boldsymbol{\ell}_k$ is a non-trivial equation. And Lemma 5.3 asserts that this is not possible unless $2k + 1 \geq n$ or $k \geq (n-1)/2$. ☐

# 6 Conclusion

We would like to note that the proof technique used to show the lower bound for depth-4 occur-$k$ formulas can be extended to prove an exponential lower bound for constant-depth constant-occur formulas if the following conjecture is true (see Appendix B for details).

**'Determinant of immanants' conjecture:** Let $M = (x_{ij})$ be an $n \times n$ matrix, and let $x_i$ denote the $i$-th diagonal variable $x_{ii}$. Let $M'$ be a projection of $M$ by setting $c = o(n)$ of the variables in $M$ to constants. Suppose the elements $\mathbf{x}_k := \{x_1, \cdots, x_k\}$, where $k$ is a constant independent of $n$, are partitioned into non-empty sets $\mathbf{S}_t := \{S_1, \cdots, S_t\}$. Consider $\mathcal{M}(\mathbf{S}_t)$, the set of all $t^{\text{th}}$ order principal minors of $M'$, each by choosing a $t$-tuple $B \in S_1 \times \cdots \times S_t$ as pivots. Over all possible choices of $B$, we get $m := |S_1| \cdots |S_t|$ many minors. Then for any set of diagonal variables $\mathbf{y}_m$ disjoint from $\mathbf{x}_k$, $J_{\mathbf{y}_m}(\mathcal{M}(\mathbf{S}_t)) \neq 0$.

In (Jacobian) spirit, the conjecture states that the $t^{\text{th}}$-order principal minors $N_{B_1}, \ldots, N_{B_m}$ are algebraically independent, when $n$ is sufficiently large (say, $n > c + k + m$).

Spurred by the success of Jacobian in solving the hitting-set problem for *constant-trdeg* depth-3 circuits and *constant-occur* constant-depth formulas, one is naturally inspired to investigate the strength of this approach against other 'constant parameter' models - the foremost of which is *constant top fanin* depth-4 circuits (PIT even for fanin 2?).

Another problem, which is closely related to hitting-sets and lower bounds, is reconstruction of arithmetic circuits [SY10, Chapter 5]. There is a quasi-polynomial time reconstruction algorithm [KS09], for a polynomial computed by a depth-3 constant top fanin circuit, that outputs a depth-3 circuit with quasi-polynomial top fanin. Could Jacobian be used as an effective tool to solve reconstruction problems? If yes, then it would further reinforce the versatility of this tool.

Figure 1: Comparison with the earlier efficient hitting-sets

| **Previous best** | | **This paper** | | |
|---|---|---|---|---|
| *Model* | *Running time*[1] | *Extended Model*[2] | *Running time*[1] | *$\mathrm{Imm}_n$ lower bound* |
| $\Sigma\Pi\Sigma(k)$ circuits: $T_1 + \cdots + T_k \stackrel{?}{=} 0$ | $s^k$ [SS11] | $C(T_1, \cdots, T_m) \stackrel{?}{=} 0$ poly-degree $C$ & trdeg $\{T_i\} \leq k$ | $s^k$ | trdeg $\{T_i\} = \Omega(n)$ |
| $\Sigma\Pi\Sigma\Pi(k)$ `multilinear` circuits | $s^{k^3}$ [SV11] | $\Sigma\Pi\Sigma\Pi$ occur-$k$ formulas | $s^{k^2}$ | $s = 2^{\Omega(n/k^2)}$ |
| depth-$D$, read-$k$ `multilinear` formulas | $s^R$ where $R = k^{k^2} + kD$ [AvMV11] | depth-$D$, occur-$k$ formulas | $s^R$ where $R = k^{2^D}$ | $s = 2^{\Omega(n)}$ for constant $k, D$ assuming Conjecture B.1 |
| $C(f_1, \cdots, f_m) \stackrel{?}{=} 0$ poly-degree $C$, $\Sigma\Pi$ circuits $f_i$'s & trdeg $\{f_i\} \leq k$ | $s^k$ [BMS11a] | – | – | $s = 2^{\Omega(n/k)}$ |

# References


[Agr05] Manindra Agrawal. Proving Lower Bounds Via Pseudo-random Generators. In *FSTTCS*, pages 92–105, 2005.

[AKS04] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Primes is in P. *Annals of Mathematics*, 160(2):781–793, 2004.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof Verification and the Hardness of Approximation Problems. *Journal of the ACM*, 45(3):501–555, 1998.

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.

[AvM10] Scott Aaronson and Dieter van Melkebeek. A note on circuit lower bounds from derandomization. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:105, 2010.

[AvMV11] Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Derandomizing Polynomial Identity Testing for Multilinear Constant-Read Formulae. In *IEEE Conference on Computational Complexity*, pages 273–282, 2011.

[BMS11a] Malte Beecken, Johannes Mittmann, and Nitin Saxena. Algebraic Independence and Blackbox Identity Testing. In *ICALP (2)*, pages 137–148, 2011.

[BMS11b] Malte Beecken, Johannes Mittmann, and Nitin Saxena. Algebraic Independence and Blackbox Identity Testing. *CoRR*, abs/1102.2789, 2011.

[CDGK91] Michael Clausen, Andreas W. M. Dress, Johannes Grabmeier, and Marek Karpinski. On Zero-Testing and Interpolation of k-Sparse Multivariate Polynomials Over Finite Fields. *Theor. Comput. Sci.*, 84(2):151–164, 1991.


[1]Estimates the bit complexity of the hitting-set generator; constant factors not stressed (also in higher exponents).
[2]We assume a zero or large characteristic.

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial Derivatives in Arithmetic Complexity (and beyond). *Foundation and Trends in Theoretical Computer Science*, 2011.

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

[DS05] Zeev Dvir and Amir Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. In *STOC*, pages 592–601, 2005.

[GKPS11] Bruno Grenet, Pascal Koiran, Natacha Portier, and Yann Strozecki. The Limited Power of Powering: Polynomial Identity Testing and a Depth-four Lower Bound for the Permanent. *Proceedings of the 31st Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, Arxiv preprint arXiv:1107.1434, 2011.

[GR05] Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. In *FOCS*, pages 407–418, 2005.

[HS80] Joos Heintz and Claus-Peter Schnorr. Testing Polynomials which Are Easy to Compute (Extended Abstract). In *STOC*, pages 262–272, 1980.

[IKS10] Gábor Ivanyos, Marek Karpinski, and Nitin Saxena. Deterministic Polynomial Time Algorithms for Matrix Completion Problems. *SIAM J. Comput.*, 39(8):3736–3751, 2010.

[IW97] Russell Impagliazzo and Avi Wigderson. P = BPP if E Requires Exponential Circuits: Derandomizing the XOR Lemma. In *STOC*, pages 220–229, 1997.

[Kal85] Kyriakos Kalorkoti. A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comput.*, 14(3):678–687, 1985.

[KI03] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *STOC*, pages 355–364, 2003.

[KMSV10] Zohar Shay Karnin, Partha Mukhopadhyay, Amir Shpilka, and Ilya Volkovich. Deterministic identity testing of depth-4 multilinear circuits with bounded top fan-in. In *STOC*, pages 649–658, 2010.

[KS01] Adam Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *STOC*, pages 216–223, 2001.

[KS06] Adam R. Klivans and Amir Shpilka. Learning Restricted Models of Arithmetic Circuits. *Theory of Computing*, 2(1):185–206, 2006.

[KS07] Neeraj Kayal and Nitin Saxena. Polynomial Identity Testing for Depth 3 Circuits. *Computational Complexity*, 16(2):115–138, 2007.

[KS08] Zohar Shay Karnin and Amir Shpilka. Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in. In *IEEE Conference on Computational Complexity*, pages 280–291, 2008.

[KS09] Zohar Shay Karnin and Amir Shpilka. Reconstruction of Generalized Depth-3 Arithmetic Circuits with Bounded Top Fan-in. In *IEEE Conference on Computational Complexity*, pages 274–285, 2009.

[KY08] Swastik Kopparty and Sergey Yekhanin. Detecting Rational Points on Hypersurfaces over Finite Fields. In *IEEE Conference on Computational Complexity*, pages 311–320, 2008.

[LFKN90] Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan. Algebraic Methods for Interactive Proof Systems. In *FOCS*, pages 2–10, 1990.

[Lov79] László Lovász. On determinants, matchings, and random algorithms. In *FCT*, pages 565–574, 1979.

[LR34] Dudley E. Littlewood and Archibald R. Richardson. Group, characters and algebras. *Philosophical Transactions of the Royal Society*, Ser. A(233):721–730, 1934.

[MVV87] Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching Is as Easy as Matrix Inversion. In *STOC*, pages 345–354, 1987.

[Oxl92] James G. Oxley. *Matroid theory*. Oxford University Press, 1992.

[PSZ00] Ramamohan Paturi, Michael E. Saks, and Francis Zane. Exponential lower bounds for depth three Boolean circuits. *Computational Complexity*, 9(1):1–15, 2000.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *STOC*, pages 245–254, 2008.

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

[Sha90] Adi Shamir. IP=PSPACE. In *FOCS*, pages 11–15, 1990.

[SS09] Nitin Saxena and C. Seshadhri. An Almost Optimal Rank Bound for Depth-3 Identities. In *IEEE Conference on Computational Complexity*, pages 137–148, 2009.

[SS10] Nitin Saxena and C. Seshadhri. From Sylvester-Gallai Configurations to Rank Bounds: Improved Black-Box Identity Test for Depth-3 Circuits. In *FOCS*, pages 21–29, 2010.

[SS11] Nitin Saxena and C. Seshadhri. Blackbox identity testing for bounded top fanin depth-3 circuits: the field doesn't matter. In *STOC*, pages 431–440, 2011.

[SV85] Sven Skyum and Leslie G. Valiant. A Complexity Theory Based on Boolean Algebra. *J. ACM*, 32(2):484–502, 1985.

[SV08] Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. In *STOC*, pages 507–516, 2008.

[SV09] Amir Shpilka and Ilya Volkovich. Improved Polynomial Identity Testing for Read-Once Formulas. In *APPROX-RANDOM*, pages 700–713, 2009.

[SV10] Amir Shpilka and Ilya Volkovich. On the relation between polynomial identity testing and finding variable disjoint factors. In *ICALP (1)*, pages 408–419, 2010.

[SV11] Shubhangi Saraf and Ilya Volkovich. Black-box identity testing of depth-4 multilinear circuits. In *STOC*, pages 421–430, 2011.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[Uma03] Christopher Umans. Pseudo-random generators for all hardnesses. *J. Comput. Syst. Sci.*, 67(2):419–440, 2003.

[Val79] Leslie G. Valiant. Completeness Classes in Algebra. In *STOC*, pages 249–261, 1979.

[VSBR83] Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM J. Comput.*, 12(4):641–644, 1983.

[Wil11] Ryan Williams. Non-uniform ACC Circuit Lower Bounds. In *IEEE Conference on Computational Complexity*, pages 115–125, 2011.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. *EUROSAM*, pages 216–226, 1979.

# A Missing Proofs

## A.1 Preliminaries: Jacobian and faithful homomorphisms

**Theorem 2.1 (restated).** *Let $\mathbf{f} = \{f_1, \cdots, f_m\} \subset \mathbb{F}[\mathbf{x}]$ and $\Phi$ be a homomorphism faithful to $\mathbf{f}$. For any polynomial $C \in \mathbb{F}[y_1, \cdots, y_m]$, $C(\mathbf{f}) = 0 \Leftrightarrow C(\Phi(\mathbf{f})) = 0$.*

*Proof.* Since $\Phi$ is faithful to $\mathbf{f}$, there is a transcendence basis (say, $f_1, \ldots, f_s$) of $\mathbf{f}$ such that $\Phi(f_1), \ldots, \Phi(f_s)$ is a transcendence basis of $\Phi(\mathbf{f})$. The function field $\mathbb{K} = \mathbb{F}(\mathbf{f})$ essentially consists of elements that are polynomials in $f_{s+1}, \ldots, f_m$ with coefficients from $\mathbb{F}(f_1, \ldots, f_s)$. Treating $C(\mathbf{f})$ as a nonzero element of $\mathbb{K}$, there is an inverse $Q \in \mathbb{K}$ such that $Q \cdot C = 1$. Since $Q$ is a polynomial in $f_{s+1}, \ldots, f_m$ with coefficients from $\mathbb{F}(f_1, \ldots, f_s)$, by clearing off the denominators of these coefficients in $Q$, we get an equation $\tilde{Q} \cdot C = P(f_1, \ldots, f_s)$, where $\tilde{Q}$ is a nonzero polynomial in $\mathbf{f}$ and $P$ is a nonzero polynomial in $f_1, \ldots, f_s$. Applying $\Phi$ to both sides of the equation, we conclude that $C(\Phi(\mathbf{f})) = \Phi(C(\mathbf{f})) \neq 0$, otherwise $P(\Phi(f_1), \ldots, \Phi(f_s)) = \Phi(P(f_1, \ldots, f_s)) = 0$ which is not possible as $\Phi(f_1), \ldots, \Phi(f_s)$ are algebraically independent and $P$ is a nontrivial polynomial. ☐

**Lemma A.1** (Vandermonde map)**.** *Let $A$ be a $r \times n$ matrix with entries in a field $\mathbb{F}$, and let $t$ be an indeterminate. Then,* $\mathrm{rank}_{\mathbb{F}(t)} \left( A \cdot (t^{ij})_{i \in [n], j \in [r]} \right) = \mathrm{rank}_{\mathbb{F}} A$.

*Proof.* Follows from Lemma 6.1 of [GR05]. ☐

**Corollary A.2.** *Let $V_1, \cdots, V_t$ be $k$-dimensional subspaces of linear polynomials in $\mathbb{F}[x_1, \cdots, x_n]$. For a constant $\alpha \in \mathbb{F}$, define the following linear homomorphism $\Psi_\alpha$ as*

$$\Psi_\alpha : x_i \mapsto \sum_{j=1}^{k} y_j \alpha^{ij}$$

*If $|\mathbb{F}| > tnk^2$, then there exists an $\alpha$ such that $\Psi_\alpha$ is an isomorphism on each of $V_1, \cdots, V_t$.*

**Lemma 2.2 (restated).** *Let $\mathbf{f} \subset \mathbb{F}[\mathbf{x}]$ be a finite set of polynomials of degree at most $d$, $\mathrm{trdeg}_{\mathbb{F}}\,\mathbf{f} \leq r$, and $\mathrm{char}(\mathbb{F}) = 0$ or $> d^r$. Let $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ be a homomorphism such that $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}\mathcal{J}_{\mathbf{x}}(\mathbf{f}) = \mathrm{rank}_{\mathbb{F}(\mathbf{z})}\Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f}))$.*

*Then, the map $\Phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}, t, y_1, \ldots, y_r]$ that maps, for all $i$, $x_i \mapsto \left(\sum_{j=1}^{r} y_j t^{ij}\right) + \Psi(x_i)$ is a homomorphism faithful to $\mathbf{f}$.*

*Proof.* Wlog let $\mathrm{trdeg}_{\mathbb{F}}\mathbf{f} = r$, which then (by Jacobian criterion) is the rank of $\mathcal{J}_{\mathbf{x}}(\mathbf{f})$. We intend to show that the matrix $\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{f}))$ is of rank $r$, which would imply (by Jacobian criterion) that $\mathrm{trdeg}_{\mathbb{F}(t,\mathbf{z})}\ \Phi(\mathbf{f}) = r$.

Consider the projection $\mathcal{J}'$ of $\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{f}))$ obtained by setting $y_1 = \cdots = y_r = 0$.

$$\begin{aligned} \mathcal{J}' = [\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{f}))]_{\mathbf{y}=\mathbf{0}} &= [\Phi(\mathcal{J}_{\mathbf{x}}(\mathbf{f})) \cdot \mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{x}))]_{\mathbf{y}=\mathbf{0}} \qquad \text{(By chain rule)} \\ &= \Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f})) \cdot \mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{x})) \end{aligned}$$

Observe that the matrix $\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{x}))$ is exactly the Vandermonde matrix that is present in Lemma A.1. Also, $\Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f}))$ has entries in $\mathbb{F}(\mathbf{z})$, and by assumption has the same rank as $\mathcal{J}_{\mathbf{x}}(\mathbf{f})$. Hence, by Lemma A.1,

$$\mathrm{rank}_{\mathbb{F}(t,\mathbf{z})}\mathcal{J}' = \mathrm{rank}_{\mathbb{F}(t,\mathbf{z})}\left(\Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f})) \cdot \mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{x}))\right) = \mathrm{rank}_{\mathbb{F}(\mathbf{z})}\Psi(\mathcal{J}_{\mathbf{x}}(\mathbf{f})) = r.$$

And since $\mathcal{J}'$ is just a projection of $\mathcal{J}_{\mathbf{y}}(\Phi(\mathbf{f}))$, the rank of the latter must also be $r$. Hence, $\Phi$ is indeed faithful. □

## A.2 Hitting-set for constant transcendence degree depth-3 circuits

**Lemma 3.1 (restated).** *There exists independent linear polynomials $\{\ell_1, \cdots, \ell_k\} \subseteq \mathcal{L}(H_0)$ such that $H_i \neq 0 \bmod (\ell_1, \ldots, \ell_i)$, $\forall i \in [k]$, and $H_k$ is a nonzero product of linear polynomials in $\mathcal{L}(H_0)$ modulo $(\ell_1, \cdots, \ell_k)$.*

*Proof.* The proof is by induction on $k$ and follows the sketch given while defining $\mathrm{sim}(\cdot)$. The degree of the nonzero polynomial $\mathrm{sim}(H_0)$ is $|\mathcal{L}(H_0)| - k$. By Chinese remaindering, there exists an $\ell_1 \in \mathcal{L}(H_0)$ such that $H_1 := \mathrm{sim}(H_0) \bmod \ell_1 \neq 0$. In the base case ($k = 1$), it is easy to see that $H_1$ is a nonzero product of linear polynomials modulo $\ell_1$. For any larger $k$, the depth-3 polynomial $H_1$ has exactly the same form as $H_0$ but with $k-1$ independent linear polynomials in the denominators. Induct on this smaller value $k-1$, keeping in mind that $\mathcal{L}(H_i) \subset \mathcal{L}(H_0)$ modulo $(\ell_1, \ldots, \ell_i)$. □

Let $\mathcal{I}_i$ and $\Psi(\mathcal{I}_i)$ denote the ideals generated by $\{\ell_1, \ldots, \ell_i\}$ and $\{\Psi(\ell_1), \ldots, \Psi(\ell_i)\}$, respectively. A rank-$(k+1)$ preserving $\Psi$ satisfies: $\ell \neq 0 \bmod \mathcal{I}_i$ iff $\Psi(\ell) \neq 0 \bmod \Psi(\mathcal{I}_i)$, for all $1 \leq i \leq k$ and $\ell \in \cup_{i=1}^{k} L(T_i)$.

**Theorem 3.2 (restated).** *If $\Psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[z_1, ..., z_{k+1}]$ is a rank-$(k+1)$ preserving map for $H_0$, then $\Psi(H_0) \neq 0$.*

*Proof.* Let $\{\ell_1, \ldots, \ell_k\}$ be the certifying path of $H_0$ fixed above. The proof is by reverse induction on $k$: Assuming $\Psi(H_i) \neq 0 \bmod \Psi(\mathcal{I}_i)$, we show that $\Psi(H_{i-1}) \neq 0 \bmod \Psi(\mathcal{I}_{i-1})$ for $k \geq i \geq 2$. The base case: By Lemma 3.1, $H_k$ is a *nonzero* product of linear polynomials in $\mathcal{L}(H_0)$ modulo $\mathcal{I}_k$, so by the definition of a rank-$(k+1)$ preserving map, $\Psi(H_k) \neq 0 \mod \Psi(\mathcal{I}_k)$ (ideal generated by independent linear polynomials is an integral domain). By construction, $H_{i-1} = \mathrm{cont}(H_{i-1}) \cdot \mathrm{sim}(H_{i-1}) = \mathrm{cont}(H_{i-1}) \cdot [q_i \ell_i + H_i] \mod \mathcal{I}_{i-1}$, for some polynomial $q_i$. Which means, $\Psi(H_{i-1}) = \Psi(\mathrm{cont}(H_{i-1})) \cdot [\Psi(q_i)\Psi(\ell_i) + \Psi(H_i)] \mod \Psi(\mathcal{I}_{i-1})$. If $[\Psi(q_i)\Psi(\ell_i) + \Psi(H_i)] = 0 \mod \Psi(\mathcal{I}_{i-1})$, then $\Psi(H_i) = 0 \mod \Psi(\mathcal{I}_i)$ which contradicts the induction hypothesis. Also, by Lemma 3.1, $H_{i-1} \neq 0 \mod \mathcal{I}_{i-1}$ implying that $\mathrm{cont}(H_{i-1}) \neq 0 \mod \mathcal{I}_{i-1}$. Since, $i \geq 2$, the linear polynomials in the term $\mathrm{cont}(H_{i-1})$ belong to $\mathcal{L}(H_0)$ modulo the ideal $\mathcal{I}_{i-1}$, once again by using the rank-$(k+1)$ preserving property of $\Psi$, we infer that $\Psi(\mathrm{cont}(H_{i-1})) \neq 0 \mod \Psi(\mathcal{I}_{i-1})$. Therefore, $\Psi(H_{i-1}) \neq 0 \mod \Psi(\mathcal{I}_{i-1})$. Finally, to obtain $\Psi(H_0) \neq 0$ from $\Psi(H_1) \neq 0 \mod \Psi(\mathcal{I}_1)$, use the same argument as above and that $\Psi(\ell) \neq 0$ for every $\ell \in \cup_{i=1}^{k} L(T_i)$. □

### A.3 Hitting-set for constant-depth constant-occur formulas

**Lemma 4.1 (restated).** *Let $G$ be any gate in $C$ and $S_1, \cdots, S_w$ be multisets of variables. Then there exists another occur-$k$ formula $G'$ for which, the vector of polynomials $(\Delta_{S_1} G, \cdots, \Delta_{S_w} G) = V_G \cdot (\Delta_{S_1} G', \cdots, \Delta_{S_w} G')$ such that*

1. *If $G$ is a $+$ gate then $G'$ is also a $+$ gate whose children consist of at most $k \cdot |\cup_{i=1}^{w} \mathrm{var}(S_i)|$ of the children of $G$, and $V_G = 1$.*
2. *If $G$ is a $\times\curlywedge$ gate, then $G'$ is also a $\times\curlywedge$ gate whose children consist of at most $k \cdot |\cup_{i=1}^{w} \mathrm{var}(S_i)|$ of the children of $G$, and $V_G = G/G'$.*

*Further, the gates constituting $G'$ and $V_G$ are disjoint.*

*Proof.*
1. Suppose $G = H_1 + \cdots + H_m$. Then at most $k \cdot |\cup \mathrm{var}(S_i)|$ of its children depend on the variables present in $\cup \mathrm{var}(S_i)$; let $G'$ be the sum of these children. Then, $\Delta_{S_i} G = \Delta_{S_i} G'$ as the other gates are independent of the variables in $\cup S_i$.
2. Suppose $G = H_1^{e_1} \cdots H_m^{e_m}$. Since $G$ is a gate in an occur-$k$ formula, at most $k \cdot |\cup \mathrm{var}(S_i)|$ of the $H_i$'s depend on the variables in $\cup S_i$; call these $H_1, \cdots, H_t$. Let $G' := H_1^{e_1} \cdots H_t^{e_t}$ and $V_G := G/G'$. Then, $\Delta_{S_i} G = V_G \cdot \Delta_{S_i} G'$ as claimed. □

**Lemma 4.2 (restated).** *Let $\mathcal{U}$ be a set of $r_\ell$ derivatives (of orders up to $c_\ell$) of gates $\mathcal{G}_{\mathcal{U}}$ at level $\ell$, and $\mathcal{U}'$ be a transcendence basis of $\mathcal{U}$. Any $|\mathcal{U}'| \times |\mathcal{U}'|$ minor of $\mathcal{J}_{\mathbf{x}}(\mathcal{U}')$ is of the form $\prod_i V_i^{e_i}$, where $V_i$'s are polynomials in at most $r_{\ell+1} := (c_\ell + 1) \cdot 2^{c_\ell + 1} k \cdot {r_\ell}^2$ many derivatives (of order up to $c_{\ell+1} := c_\ell + 1$) of disjoint groups of children of $\mathcal{G}_{\mathcal{U}}$.*

*Proof.* Let $G \in \mathcal{G}_{\mathcal{U}}$ be a gate at level $\ell$ and $\{U_1, \ldots, U_{e_G}\} \subset \mathcal{U}'$ be the set of all the derivatives of $G$ present in $\mathcal{U}'$. Fix any $|\mathcal{U}'| \times |\mathcal{U}'|$ sub-matrix $M$ of $\mathcal{J}_{\mathbf{x}}(\mathcal{U}')$. Consider the $e_G$ rows of $M$ that contain the derivatives of $U_1, \ldots, U_{e_G}$. These rows together contain a total of $w := e_G \cdot |\mathcal{U}'|$ elements that are up to $(c_\ell + 1)$-order derivatives of $G$; view all the elements of these $e_G$ rows as a single vector $(\Delta_{S_1} G, \cdots, \Delta_{S_w} G)$ and apply Lemma 4.1 to express it as $V_G \cdot (\Delta_{S_1} G', \cdots, \Delta_{S_w} G')$. Verify that $|\cup_{i=1}^{w} \mathrm{var}(S_i)| \leq e_G \cdot c_\ell + |\mathcal{U}'| \leq e_G \cdot c_\ell + r_\ell$. So, in $\det(M)$ we can take $V_G$ common from each of these $e_G$ rows such that the elements present inside the determinant are of the form $\Delta_{S_i} G'$, where $G'$ has at most $k(e_G c_\ell + r_\ell)$ children.

Since $|S_i| \leq c_\ell + 1$, at most $k(c_\ell + 1)$ children of $G'$ depend on $\mathrm{var}(S_i)$. If $G'$ is a $+$ gate, then $\Delta_{S_i} G'$ is the sum of the derivatives of at most $k(c_\ell + 1)$ of its children (that depend on $\mathrm{var}(S_i)$). If $G'$ is a $\times\curlywedge$ gate computing $H_1^{e_1} \cdots H_t^{e_t}$ (where $t \leq k(e_G c_\ell + r_\ell)$), then $\Delta_{S_i} G'$ is a polynomial combination of the $H_i$'s and $\{\Delta_T H_j\}_{\emptyset \neq T \subseteq S_i}$ for each $H_j$ depending on $\mathrm{var}(S_i)$. Hence in either case, $\Delta_{S_i} G'$ is a polynomial in the children of $G'$ and their at most $(2^{c_\ell+1}-1)\cdot k(c_\ell+1)$ many derivatives (of order between one and $(c_\ell+1)$).

Summing over all the $w$ elements $\Delta_{S_i} G'$, the elements of the $e_G$ rows of $M$ are polynomials in at most $k(e_G c_\ell + r_\ell) + w\cdot(2^{c_\ell+1}-1)k(c_\ell+1) = k(e_G c_\ell + r_\ell) + e_G\cdot|\mathcal{U}'|\cdot(2^{c_\ell+1}-1)k(c_\ell+1)$ derivatives of the children of $G'$. Going over all $G \in \mathcal{G}_{\mathcal{U}}$, $\det(M)$ can be expressed as a product $\prod_{G\in\mathcal{G}_{\mathcal{U}}} V_G^{e_G}$ and a polynomial $V$ in at most $k(r_\ell c_\ell + r_\ell^2) + r_\ell^2\cdot(2^{c_\ell+1}-1)k(c_\ell+1) \le (c_\ell+1)2^{c_\ell+1}k{r_\ell}^2$ derivatives (of order up to $c_\ell+1$) of a group of gates in level $\ell+1$. Further, the groups of gates whose derivatives constitute the $V_G$'s and $V$ are mutually disjoint (by Lemma 4.1). ☐

### A.4 Lower bounds for the immanant

**Lemma 5.1 (restated).** *Suppose* $\mathrm{Det}_n = C(T_1,\ldots,T_m)$*, where* $C$ *is any circuit and let* $\mathbf{T}_r = \{T_1,\ldots,T_r\}$ *be a transcendence basis of* $\mathbf{T}$ *with* $r < n$*. Then, there exist a set of* $r+1$ *variables* $\mathbf{x}_{r+1} \subset \mathbf{x}$ *and an equation* $\sum_{i=1}^{r+1} c_i f_i \cdot M_i = 0$ *such that* $M_i$*'s are distinct first order principal minors of* $M$*,* $f_i$*'s are distinct* $r\times r$ *minors of* $\mathcal{J}_{\mathbf{x}_{r+1}}(\mathbf{T}_r)$*, not all* $f_i$*'s are zero, and* $c_i \in \mathbb{F}^*$*.*

*Proof.* In a column of a Jacobian matrix $\mathcal{J}_{\mathbf{x}}(\cdot)$, all the entries are differentiated with respect to a variable $x$, we will say that the column is *indexed* by $x$. Let $\mathbf{T}_r = \{T_1,\cdots,T_r\}$ be a transcendence basis of $\mathbf{T}$. Amongst the nonzero $r\times r$ minors of $\mathcal{J}_{\mathbf{x}}(\mathbf{T}_r)$ (they exist by Jacobian criterion), pick one (call the matrix associated with the minor, $N$) that maximizes the number of diagonal variables $\{x_{ii} : 1 \le i \le n\}$ indexing the columns of $N$. Let $S$ denote the set of variables indexing the columns of $N$. Since $r < n$, there exists a diagonal variable $x_{jj} \notin S$. Consider the $(r+1)\times(r+1)$ minor of $\mathcal{J}_{\mathbf{x}}(\{\mathrm{Det}_n\}\cup\mathbf{T}_r)$ corresponding to the columns indexed by $S' := S\cup\{x_{jj}\}$ - call the associated $(r+1)\times(r+1)$ matrix $\tilde{N}$. Since, $\mathrm{Det}_n = C(\mathbf{T})$, the polynomials $\mathrm{Det}_n$ and $T_1,\ldots,T_r$ are algebraically dependent and hence $\det(\tilde{N}) = 0$. Expanding $\det(\tilde{N})$ along the first row of $\tilde{N}$, which contains signed first order minors (cofactors) of $M$, we have an equation $\sum_{i=1}^{r+1} c_i f_i M_i = 0$, where $M_i$'s are distinct minors of $M$, $f_i$'s are distinct $r\times r$ minors of $\mathcal{J}_{S'}(\mathbf{T}_r)$, and $c_i \in \mathbb{F}^*$. If $M_i$ is the principal minor of $M$ with respect to the variable $x_{jj}$ then $f_i = \det(N) \neq 0$ (by construction).

It suffices to show that if $M_i$ is a non-principal minor of $M$ then $f_i = 0$. Consider any non-principal minor $M_i$ in the above sum, say it is the minor of $M$ with respect to $x_{k\ell}$. The corresponding $f_i$ is precisely the $r\times r$ minor of $\mathcal{J}_{S'}(\mathbf{T}_r)$ with respect to the columns $S'\setminus\{x_{k\ell}\} = (S\setminus\{x_{k\ell}\})\cup\{x_{jj}\}$. Hence, by the maximality assumption on the number of diagonal elements of $M$ in $S$, $f_i = 0$. ☐

**Lemma 5.2 (restated).** *If* $M_1,\cdots,M_t$ *are distinct first order principal minors of* $M$ *and* $\sum_{i=1}^t f_i\cdot M_i = 0$ *(not all* $f_i$*'s are zero) then the total sparsity of the* $f_i$*'s is at least* $2^{n/2-t}$*.*

*Proof.* The proof is by contradiction. The idea is to start with the equation $\sum_{i=1}^t f_i M_i = 0$ and apply two steps - *sparsity reduction* and *fanin reduction* - alternatively, till we arrive at a contradiction in the form of an equation $f_j\cdot M_j = 0$, where neither $f_j$ nor $M_j$ is zero if the total sparsity of the $f_i$'s is less than $2^{n/2-t}$. With an equation of the form $\sum_{i=1}^{\tau} g_i N_i = 0$, we associate four parameters $\tau$, $\mathcal{S}$, $\eta$ and $c$. These parameters are as follows: $\tau$ is called the *fanin* of the equation, $\mathcal{S}$ is the total sparsity of the $g_i$'s (we always assume that not all the $g_i$'s are zero), every $N_i$ is a distinct first order principal minor of a symbolic $\eta\times\eta$ matrix $N = (x_{ij})$, and $c$ is the maximum number of entries of $N$ that are set as constants. To begin with, $g_i = f_i$ and $N_i = M_i$ for all $1 \le i \le t$, so $\tau = t$, $\mathcal{S} = s$ (the total sparsity of the $f_i$'s), $\eta = n$, $N = M$ and $c = 0$. In the 'sparsity reduction' step, we start with an equation $\sum_{i=1}^{\tau} g_i N_i = 0$, with parameters $\tau$, $\mathcal{S}$, $\eta$, $c$ and arrive at an equation $\sum_{i=1}^{\tau'} g'_i N'_i = 0$ with parameters $\tau'$, $\mathcal{S}'$, $\eta'$, $c'$ such that $\tau' \le \tau$, $\mathcal{S}' \le \mathcal{S}/2$, $\eta - 1 \le \eta' \le \eta$, and $c' \leq c+1$. In the 'fanin reduction' step, we start with an equation $\sum_{i=1}^{\tau} g_i N_i = 0$, with parameters $\tau$, $\mathcal{S}$, $\eta$, $c$ and arrive at an equation $\sum_{i=1}^{\tau'} g_i' N_i' = 0$ with parameters $\tau'$, $\mathcal{S}'$, $\eta'$, $c'$ such that one of the two cases happens - Case 1: $\tau' \leq \tau - 1$, $\mathcal{S}' \leq \mathcal{S}$, $\eta' = \eta - 1$, and $c' = c$; Case 2: $\tau' = 1$, $\mathcal{S}' \leq \mathcal{S}$, $\eta' = \eta$, and $c' \leq c + \tau$.

Naturally, starting with $\sum_{i=1}^{t} f_i M_i = 0$, the 'sparsity reduction' step can only be performed at most $\log s$ many times (since the total sparsity of the $g_i$'s reduces by at least a factor of half every time this step is executed), whereas the 'fanin reduction' step can be performed at most $t-1$ times (as the fanin goes down by at least one for every such step). Finally, when this process of alternating steps ends, we have an equation of the form $g_i \cdot N_i = 0$ (Case 2 of the fanin reduction step), where $g_i \neq 0$ and $N_i$ is a principal minor of a symbolic matrix $N$ of dimension at least $n - (\log s + t - 1)$ such that at most $(\log s + t)$ entries of $N$ are set as constants. Now, if $\log s + t \leq n - (\log s + t)$ the $N_i$ can never be zero (by Fact 3) and hence we arrive at a contradiction. Therefore, $s > 2^{n/2-t}$. Now, the details of the sparsity reduction and the fanin reduction steps.

Suppose, we have an equation $\sum_{i=1}^{\tau} g_i N_i = 0$ as mentioned above. Without loss of generality, assume that the minor $N_i$ is the minor of $N$ with respect to the $i^{th}$ diagonal element of $N$. Call all the variables $x_{ij}$ in $N$ with both $i, j > \tau$ as the *white variables*. These are the variables that are present in every minor $N_i$ in the sum $\sum_{i=1}^{\tau} g_i N_i$. The variables $x_{ij}$ where both $i, j \leq \tau$ are called the *black variables*, and the remaining are the *grey variables*. By assumption, $c$ of the variables in $N$ are set as constants.

*Sparsity reduction step* - Say $x$ is a white variable that one of the $g_i$'s depends on. Writing each $g_i$ as a polynomial in $x$, there must be two distinct powers of $x$ amongst the $g_i$'s (for if not, then $x$ can be taken common across all $g_i$'s). Let $x^\ell$ be the lowest degree and $x^h$ be the highest. Dividing the entire equation $\sum_{i=1}^{\tau} g_i N_i = 0$ by $x^\ell$, we can further assume that $\ell = 0$. Each of the $g_i$'s and $N_i$'s can be expressed as, $g_i = g_{i,0} + x \cdot g_{i,1} + \cdots + x^h \cdot g_{i,h}$ and $N_i = N_{i,0} + x \cdot N_{i,1}$, where $g_{i,j}$'s and $N_{i,j}$'s are $x$-free. Looking at the coefficients of $x^0$ and $x^{h+1}$ in the equation yields $\sum_{i=1}^{\tau} g_{i,0} \cdot N_{i,0} = 0$ and $\sum_{i=1}^{\tau} g_{i,h} \cdot N_{i,1} = 0$. Note that $N_{i,0}$'s can be thought of as principal minors of the $\eta \times \eta$ matrix $N'$ obtained by setting $x = 0$ in $N$. And each of the $N_{i,1}$'s can be thought of as minors of the $(\eta - 1) \times (\eta - 1)$ matrix $N'$ which is the matrix associated with the minor of $N$ with respect to $x$. Since the monomials in $g_{i,0}$ and $x^h g_{i,h}$ are disjoint, either the total sparsity of the $g_{i,0}$'s or the total sparsity of the $g_{i,h}$'s is $\leq \mathcal{S}/2$. Thus, one of the equations $\sum_{i=1}^{\tau} g_{i,0} \cdot N_{i,0} = 0$ or $\sum_{i=1}^{\tau} g_{i,h} \cdot N_{i,1} = 0$ yields an equation of the form $\sum_{i=1}^{\tau'} g_i' N_i' = 0$ with parameters $\tau'$, $\mathcal{S}'$, $\eta'$, $c'$ as claimed before. (In case, we choose $\sum_{i=1}^{\tau} g_{i,h} \cdot N_{i,1} = 0$ as our next equation, we also set the variables in the same columns and rows of $x$ to constants in such a way that a $g_{i,h}$ stays nonzero. This is certainly possible over a characteristic zero field [Sch80, Zip79].) The sparsity reduction step is performed whenever the starting equation $\sum_{i=1}^{\tau} g_i N_i = 0$ has a white variable among the $g_i$'s. When all the $g_i$'s are free of white variables, we perform the *fanin reduction step*.

*Fanin reduction step* - When we perform this step, all the $g_i$'s consist of black and grey variables. Pick a row $R$ from $N$ barring the first $\tau$ rows. Let $y_1, \cdots, y_\tau$ be the grey variables occuring in $R$ (these are, respectively, the variables in the first $\tau$ columns of $R$). Starting with $y_2$, divide the equation $\sum_{i=1}^{\tau} g_i N_i = 0$ by the largest power of $y_2$ common across all monomials in the $g_i$'s, and then set $y_2 = 0$. This process lets us assume that there exists at least one $g_i$ which is not zero at $y_2 = 0$. On the residual equation, repeat the same process with $y_3$ and then with $y_4$ and so on till $y_\tau$. Thus, we can assume without loss of generality that in the equation $\sum_{i=1}^{\tau} g_i N_i = 0$ there is at least one $g_i$ that is not zero when $y_2, \ldots, y_\tau$ are set to zero. Observe that if $g_1$ is the only $g_i$ that stays nonzero under the projection $y_2 = \ldots = y_\tau = 0$ then $(g_1 N_1)_{(y_2 = \ldots = y_\tau = 0)} = 0$, implying that $N_1 = 0$ under the same projection - this is Case 2 of the fanin reduction step mentioned earlier. Now, assume that there is a $g_i$ other than $g_1$ (say, $g_2$) that is nonzero under the projection $y_2 = \ldots = y_\tau = 0$. Set all the remaining variables of row $R$ to zero except $y_1$ - these are the white variables in $R$. Since the $g_i$'s are free of white variables (or else, we would have performed the 'sparsity reduction' step), none of the $g_i$'s is effected by this projection. However, $N_1$ being a minor with respect to the first diagonal element of $N$, vanishes completely after the projection. Any other $N_i$ takes the form $y_1 \cdot N_i'$, where $N_i'$ is a principal minor of a $(\eta-1)\times(\eta-1)$ matrix $N'$ which is the matrix associated with the minor of $N$ with respect to $y_1$. Therefore, after the projection, the equation $\sum_{i=1}^{\tau} g_i N_i = 0$ becomes $\sum_{i=2}^{\tau} \tilde{g}_i \cdot y_1 N_i' = 0 \Rightarrow \sum_{i=2}^{\tau} \tilde{g}_i \cdot N_i' = 0$, where $\tilde{g}_i$ is the image of $g_i$ under the above mentioned projection and further $\tilde{g_2} \neq 0$. The $\tilde{g}_i$'s might still contain variables from the first column of $N$. So, as a final step, set these variables to values so that a nonzero $\tilde{g}_i$ remains nonzero after this projection (the [Sch80, Zip79, DL78] lemma asserts that such values exist in plenty). This gives us the desired form $\sum_{i=1}^{\tau'} g_i' N_i' = 0$ with parameters $\tau'$, $\mathcal{S}'$, $\eta'$, $c'$ as claimed before (Case 1 of the fanin reduction step mentioned earlier). □

**Lemma 5.3 (restated).** *If $M_1, \cdots, M_t$ are distinct first order principal minors of $M$ and $\sum_{i=1}^{t} \alpha_i M_i = 0 \bmod \boldsymbol{\ell}_k$ (not all $\alpha_i = 0$) for independent linear polynomials $\boldsymbol{\ell}_k$, then $t + k \geq n$.*

*Proof.* Assume that $t + k < n$ (with $t \geq 1$ it means $k \leq n-2$). Since $\ell_1, \cdots, \ell_k$ are independent linear polynomials, the equation may be rewritten as $\sum_{i=1}^{t} \alpha_i M_i' = 0$ where $(M_i')$s are minors of the matrix $M'$ obtained by replacing $k$ entries of $M$ by linear polynomials in other variables. We shall call these entries as *corrupted* entries. Without loss of generality, we shall assume that $M_i'$ is the minor corresponding to the $i$-th diagonal variable and that all the $\alpha_i$'s are nonzero.

**Claim A.3.** *Each of the first $t$ rows and columns must have a corrupted entry.*

*Pf.* Suppose the first row (without loss of generality) is free of any corrupted entry. Then, setting the entire row to zero would make all $M_i' = 0$ for $i \neq 1$. But since $\sum \alpha_i M_i' = 0$, this forces $M_1'$ to become zero under the projection as well. This leads to a contradiction as $M_1'$ is a determinant of an $(n-1)\times(n-1)$ symbolic matrix under a projection, and this can not be zero unless $k \geq n-1$ (by Fact 3). □ (Claim)

Since $n - k > t$, there must exist a set of $t-1$ rows $\{R_1, \cdots, R_{t-1}\}$ of $M$ that are free of any corrupted entries. For each of these rows, set the $i$-th variable of row $R_i$ to 1, and every other variable in $R_1, \cdots, R_{t-1}$ to zero. These projections make $M_i' = 0$ for all $i \neq t$ (as in these minors an entire row vanishes). And since $\sum_{i=1}^{t} \alpha_i M_i' = 0$, this forces $M_t'$ to become zero under this projection as well. But under this projection, $M_t'$ just reduces (up to a sign) to the minor obtained from $M'$ by removing the columns $\{1, \cdots, t\}$ and rows $\{R_1, \cdots, R_{t-1}\} \cup \{t\}$. This is a determinant of an $(n-t)\times(n-t)$ symbolic matrix, containing at most $k-t$ corrupted entries, thus $k - t \geq n - t$ (by Fact 3). But then $k \geq n$, which contradicts our initial assumption. □

### A.5 Extensions to immanants

All the lower bound proofs use some very basic properties of $\mathrm{Det}_n$. These properties are general enough that they apply to any *immanant*. For any character $\chi : S_n \to \mathbb{C}^\times$, recall the definition of the immanant of an $n \times n$ matrix $M = (x_{ij})$:

$$\mathrm{Imm}_\chi(M) \quad = \quad \sum_{\sigma \in S_n} \chi(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)}$$

Since $\chi$ is a character, this in particular means that $\chi(\sigma) \neq 0$ for any $\sigma \in S_n$

**Definition 5** (Immanant minor)**.** *The* minor *of $Imm_\chi(M)$ with respect to the $(i,j)$-th entry is defined as*

$$(Imm_\chi(M))_{i,j} \quad = \quad \sum_{\substack{\sigma \in S_n \\ \sigma(i)=j}} \chi(\sigma) \prod_{k \neq i} x_{k,\sigma(k)}$$

This may also be rewritten as a scalar multiple of $\mathrm{Imm}_{\chi'}(M_{ij})$ for a suitable character $\chi' : S_{n-1} \to \mathbb{C}^\times$, where $M_{ij}$ is the submatrix of $M$ after removing the $i$-th row and $j$-th column. From the definition, it follows directly that the partial derivative of $\mathrm{Imm}_\chi(M)$ with respect to $x_{ij}$ is precisely the minor with respect to $(i,j)$.

The only crucial fact of determinants that is used in all the proofs is that a symbolic $n \times n$ determinant cannot be zero when less than $n$ of its entries are altered.

**Fact 3.** *Let $M'$ be the matrix obtained by setting $c < n$ entries of $M$ to arbitrary polynomials in $\mathbb{F}[\mathbf{x}]$. Then for any character $\chi : S_n \to \mathbb{C}^\times$, we have $Imm_\chi(M') \neq 0$.*

*Proof.* We shall say an entry of $M'$ is *corrupted* if it is one of the $c$ entries of $M$ that has been replaced by a polynomial. We shall prove this by carefully rearranging the rows and columns so that all the corrupted entries are above the diagonal. Then, since all entries below the diagonal are free, we may set all of them to zero and the immanant reduces to a single nonzero monomial.

Since less than $n$ entries of $M'$ have been altered, there exists a column that is free of any corrupted entries. By relabelling the columns if necessary, let the first column be free of any corrupted entry. Pick any row $R$ that contains a corruption and relabel the rows to make this the first row. This ensures that the first column is free of any corrupted entry, and the $(n-1) \times (n-1)$ matrix defined by rows and columns, 2 through $n$, contain less than $c-1$ corruptions. By induction, the $c-1$ corruptions may be moved above the diagonal by suitable row/column relabelling. And since the first column is untouched during the process, we now have all $c$ corruptions above the diagonal. Now setting all entries below the diagonal to zeroes reduces the immanant to a single nonzero monomial. ☐

With this fact, all our lower bound proofs of the determinant can be rewritten for any immanant.

# B Conditional immanant lower bounds for depth-$D$ occur-$k$ formulas

In this section, we present a lower bound for depth-$D$ occur-$k$ formulas similar in spirit to Theorem 1.4 by assuming the following conjecture about determinant minors.

**Conjecture B.1.** *Let $M = (x_{ij})$ be an $n \times n$ matrix, and let $x_i$ denote the $i$-th diagonal variable $x_{ii}$. Let $M'$ be a projection of $M$ by setting $c = o(n)$ of the variables in $M$ to constants. Suppose the elements $x_1, \cdots, x_k$, where $k$ is a constant independent of $n$, are partitioned into non-empty sets $S_1, \cdots, S_t$. Consider $\mathcal{M}(\mathbf{S}_t)$, the set of $t^{th}$ order principal minors of $M'$, each by choosing a $t$-tuple $B \in S_1 \times \cdots \times S_t$ as pivots. Over all possible choices of $B$, we get $m := |S_1| \cdots |S_t|$ many minors. Then for any set of diagonal variables $\mathbf{y}_m$ disjoint from $\mathbf{x}_k$, $J_{\mathbf{y}_m}(\mathcal{M}(\mathbf{S}_t)) \neq 0$.*

The conjecture roughly states that the different $t^{\text{th}}$-order principal minors are algebraically independent. We will need a generalization of Lemma 5.1 for the purposes of this section.

**Lemma B.2.** *Suppose $\{f_1, \cdots, f_s, g_1, \cdots, g_t\}$ are algebraically dependent polynomials such that* $\operatorname{trdeg}\{\mathbf{g}_t\} = t$. *Let $S \subseteq \mathbf{x}$ be a fixed set of variables of size at least $s+t$. Then there exists a set of $s+t$ variables $\mathbf{x}_{s+t} \subset \mathbf{x}$ and an equation of the form*

$$\sum_{i=1}^{r} c_i \cdot F_i \cdot G_i = 0 \qquad \text{where } r \leq \binom{s+t}{t}$$

*such that each $c_i \in \mathbb{F}^*$, each $F_i$ is a distinct $s \times s$ minor of $\mathcal{J}_{\mathbf{x}_{s+t} \cap S}(\mathbf{f}_s)$, each $G_i$ is a distinct $t \times t$ minor of $\mathcal{J}_{\mathbf{x}_{s+t}}(\mathbf{g}_t)$, and not all $G_i$'s are zero.*

Note that we are not asserting the nonzeroness of $F_i$'s. Also, Lemma 5.1 may be obtained from the above lemma by taking $f_1 = \mathrm{Det}_n$, $s = 1$ and $S$ to be the set of diagonal variables.

*Proof.* The proof is along the lines of Lemma 5.1. Amongst the nonzero $t \times t$ minors of $\mathcal{J}_{\mathbf{x}}(\mathbf{g}_t)$, pick one (call the matrix associated with the minor, $N$) that maximizes the number of variables in $S$ indexing the columns of $N$. Without loss of generality, let $\mathbf{x}_t$ be the set of variables indexing the columns of $N$. Since $|S| \geq s+t$, there exists $s$ other variables in $S$, say $\{x_{1+t}, \cdots, x_{s+t}\}$. Consider the $(s+t) \times (s+t)$ minor of $\mathcal{J}_{\mathbf{x}}(\mathbf{f}_s \cup \mathbf{g}_t)$ corresponding to the columns indexed by $\mathbf{x}_{s+t}$ - call the associated $(s+t) \times (s+t)$ matrix $\tilde{N}$.

Since $\mathbf{f}_s, \mathbf{g}_t$ are algebraically dependent, we have that $\det(\tilde{N}) = 0$. Expanding $\det(\tilde{N})$ over all possible $s \times s$ minors in the first $s$ rows, we have an equation

$$\sum_{U \subseteq \mathbf{x}_{s+t}, |U| = s} c_i \cdot F_U \cdot G_U = 0$$

where each $F_U$ is a distinct $s \times s$ minor of $\mathcal{J}_{\mathbf{x}_{s+t}}(\mathbf{f}_s)$, each $G_U$ is a distinct $t \times t$ minor of $\mathcal{J}_{\mathbf{x}_{s+t}}(\mathbf{g}_t)$, and $c_i \in \mathbb{F}^*$. If $G_U$ is the minor with respect to variables $\mathbf{x}_t$, then $G_U = \det(N) \neq 0$ (by construction). It suffices to show that if $F_U$ is a minor indexed by variables outside $S$, then $G_U = 0$. This follows, just like in Lemma 5.1, by the maximality assumption on choice of $\mathbf{x}_t$. ☐

The rest of this section shall be devoted to the proof of the following theorem.

**Theorem B.3.** *Assuming Conjecture B.1, any depth-$D$ occur-$k$ formula that computes $\mathrm{Det}_n$ must have size $s = 2^{\Omega(n)}$ over any field of characteristic zero.*

**Proof idea:** The proof proceeds on the same lines as Theorem 1.4. If $T_1, \cdots, T_k$ is a transcendence basis of gates at level 2 computing the determinant, then $\mathcal{J}_{\mathbf{x}}(\mathrm{Det}_n, T_1, \cdots, T_k)$ is a matrix of rank $k$. This yields a non-trivial equation of the form $\sum N_i^{(1)} \cdot G_i^{(1)} = 0$ where each of the $N_i^{(1)}$'s are principal minors of $M = (x_{ij})$ and $G_i^{(1)}$'s are $k \times k$ minors of $\mathcal{J}_{\mathbf{x}}(T_1, \cdots, T_k)$. Here is where we may use Lemma 4.2 to remove common factors and obtain an equation of the form $\sum N_i^{(1)} \cdot \tilde{G_i}^{(1)} = 0$ where $\tilde{G_i}^{(1)}$ is a polynomial of constantly many derivatives of polynomials computed at the next level. The above equation may be thought of as a polynomial relation amongst $\left\{N_i^{(1)}\right\} \cup \left\{\mathrm{Elem}(\tilde{G_i}^{(1)})\right\}$. Applying Lemma B.2 (with a suitable choice of $S_2$), we get an equation of the form $\sum N_i^{(2)} \cdot G_i^{(2)} = 0$ where each $N_i^{(2)}$ is a minor of $\mathcal{J}_{S_2}\left(\left\{N_i^{(1)}\right\}\right)$, and $G_i^{(2)}$s are Jacobians minors of $\bigcup \mathrm{Elem}(\tilde{G_i}^{(1)})$.

Again after removing common factors, this equation may be intepreted as a polynomial relation amongst the entries of $N_i^{(2)}$ (which are minors of order 2) and $\mathrm{Elem}(\tilde{G}_i^{(2)})$.

Repeating this argument, we finally reach the level of sparse polynomials and obtain a non-trivial equation $\sum N_i^{(D-2)} \cdot \tilde{G}_i^{(D-2)} = 0$, where each $N_i^{(D-2)}$ is a Jacobian minor of $(D-3)$-order minors, and each $\tilde{G}_i^{(D-2)}$ is a sparse polynomial. With a slightly more careful choice of the sets $S_i$ in Lemma B.2, each of the minors $N_i^{(D-2)}$ would be a minor of $\mathcal{J}_{S_{D-2}}(\mathcal{M}(S_1, \cdots, S_{D-3}))$. Assuming Conjecture B.1, we can show that such an equation is not possible unless the sparsity of the $f_i$'s is *large*, using a similar argument as in Lemma 5.2.

**Lemma B.4.** *Suppose $Det_n$ is computed by a depth-$D$ occur-$k$ formula of size $s$. Then there exist variables $x_1, \cdots, x_R$ where $R = R(k, D)$, a partition of $\mathbf{x}_R$ into non-empty sets $S_1, \cdots, S_{D'}$, $(D' \leq (D-2))$ polynomials $f_1, \cdots, f_m$ (not all zero) where $m = |\mathcal{M}(\mathbf{S}_D)|^{O(R)}$ and each $f_i$ has sparsity at most $s^R$, such that*

$$\sum_{i=1}^{m} f_i \cdot N_i = 0$$

*where each $N_i$ is a minor of $\mathcal{J}_{\mathbf{x}}(\mathcal{M}(\mathbf{S}_{D'}))$ indexed by diagonal variables.*

*Proof.* To begin with, $\mathrm{Det}_n = C(T_1, \cdots, T_m)$ where $T_1, \cdots, T_m$ are polynomials computed at the first level. So Lemma 5.1 gives a starting equation, though we do not really have a sparsity bound on the $f_i$'s. The proof shall proceed by transforming this equation into another, involving lower level polynomials, till we get a sparsity bound.

In general, we shall have an equation of the form $C_\ell(\mathcal{M}(S_1, \cdots, S_{\ell-1}), T_1^{(\ell)}, \cdots, T_{r_\ell}^{(\ell)}) = 0$, where each $T_i^{(\ell)}$ is a derivative (of order at most $\ell$) of a polynomial computed at level $\ell$ of the circuit. Without loss of generality, we may assume that $\left\{T_1^{(\ell)}, \cdots, T_{r_\ell}^{(\ell)}\right\}$ are algebraically independent. Let $m_\ell := |S_1| \cdots |S_{\ell-1}|$. Choose a set of diagonal elements $S_\ell$ of size $|\mathcal{M}(S_1, \cdots, S_{\ell-1})| + r_\ell$ that is disjoint from $S_1, \cdots, S_{\ell-1}$. Applying Lemma B.2 with $S_\ell$, we get an equation of the form

$$\sum_i c_i^{(\ell)} N_i^{(\ell)} \cdot G_i^{(\ell)} = 0$$

where $N_i^{(\ell)}$ is an $m_\ell \times m_\ell$ minor of $\mathcal{J}_{S_\ell}(\mathcal{M}(\mathbf{S}_{\ell-1}))$ indexed by diagonal variables, each $G_i^{(\ell)}$ is an $r_\ell \times r_\ell$ minor of $\mathcal{J}_{S_\ell}(\mathbf{T}_{r_\ell}^{(\ell)})$. Since $C$ is an occur-$k$ circuit, using an argument similar to Lemma 4.2, the above equation may be rewritten as

$$V_\ell \cdot \sum_i c_i^{(\ell)} N_i^{(\ell)} \cdot \tilde{G}_i^{(\ell)} = 0$$

where each $\tilde{G}_i^{(\ell)}$'s is a polynomial function of at most $r_{\ell+1} := (\ell+1)2^{\ell+1} \cdot k(r_\ell + m_\ell)r_\ell$ many derivatives of polynomials computed at level $\ell+1$. Note that $V_\ell$ cannot be zero as at least one $G_i^{(\ell)}$ was guaranteed to be nonzero by Lemma B.2. Therefore, $\sum_i c_i^{(\ell)} N_i^{(\ell)} \cdot \tilde{G}_i^{(\ell)} = 0$. Since each $\tilde{G}_i^{(\ell)}$ is a polynomial function of $r_{\ell+1}$ derivatives at the next level, we now have $C_{\ell+1}(\mathcal{M}(S_1, \cdots, S_\ell), T_1^{(\ell+1)}, \cdots, T_{r_{\ell+1}}^{(\ell+1)}) = 0$.

Unfolding this recursion, we finally reach the level of sparse polynomials, at which point we have an equation of the form

$$\sum_i c_i^{(D-2)} N_i^{(D-2)} \cdot \tilde{G}_i^{(D-2)} = 0$$

and each $\tilde{G}_i^{(D-2)}$ is a $r_{D-2} \times r_{D-2}$ Jacobian minor of sparse polynomials. Hence, each $\tilde{G}_i^{(D-2)}$ is itself a polynomial of sparsity bounded by $s^{r_{D-2}}$ as claimed. $\square$

We now have to show that an equation of the form $\sum f_i \cdot N_i = 0$ is not possible unless one of the $f_i$'s has exponential sparsity. The above lemma guarantees that at least one of the $f_i$'s are nonzero in this equation, but it could be the case that some of the $N_i$'s are zero. This was not the case in the depth-4 lower bound as each $N_i$ was just a determinant minor. However, in this case they are jacobians of minors. Conjecture B.1 asserts that the $N_i$'s are nonzero, even if *few* variables are set to zero. This assumption would be enough to get the required lower bound.

**Lemma B.5.** *Let $|\mathcal{M}(S_1, \cdots, S_D)| =: m$ be a constant and let $\{N_i\}_{i \leq t}$ be distinct $m \times m$ minors of $\mathcal{J}_{\mathbf{x}}(\mathcal{M}(\mathbf{S}_D))$ where the columns of $N_i$ are indexed by a set $T_i$ of diagonal variables of $M$ disjoint from $\bigcup_{j=1}^{D} S_j$. Suppose $f_1, \cdots, f_t$ are polynomials such that $\sum_{i=1}^{t} f_i \cdot N_i = 0$ (not all $f_i$'s are zero). Then, assuming Conjecture B.1 is true, the total sparsity of the $f_i$'s is $2^{\Omega(n)}$.*

*Proof.* The proof is along the lines of the proof of Lemma 5.2 and shall proceed by a similar series of *sparsity reduction* and *fanin reduction* steps to arrive at a contradiction. Throughout the proof, Conjecture B.1 shall assert that $N_i$'s stay nonzero (even when few variables are set to constants). We briefly describe the *sparsity reduction* and the *fanin reduction* steps and the rest of the proof would follow in essentially an identical fashion as the proof of Lemma 5.2.

Without loss of generality, assume that $\{x_1, \cdots, x_r\}$ is the union of the sets $S_i$'s and $T_i$'s. Let $N$ refer to the matrix of indeterminates that the $N_i$'s are derived from. In our case, $N$ would be obtained by (possibly) setting few variables to constants in $M = (x_{ij})$. We'll refer to all the variables $x_{ij}$ where both $i, j > r$ as *white* variables; these are present in every entry of each $N_i$. The variables $x_{ij}$ where both $i, j \leq r$ shall be called *black* variables, and the rest called *grey* variables. Here again, the *sparsity reduction* step shall be applied whenever one of the $f_i$'s depends on a *white* variable, otherwise the *fanin reduction* steps shall be applied.

*Sparsity-reduction* step - Suppose one of the $f_i$'s depend on a white variable $x$. Then each $N_i$ can be written as $N_i = N_{i,0} + \cdots + x^m N_{i,m}$, and $f_i = f_{i,0} + \cdots + f_{i,h} x^h$. One of the two equations corresponding to the coefficient of $x^0$ and $x^{h+m}$ yields a similar equation with sparsity reduced by a factor of $1/2$. Observe that $N_{i,0}$ is just $N_i \mid_{x=0}$, and hence the polynomials $\{N_{i,0}\}$ may be thought of as corresponding Jacobian minors of $N'$ obtained by setting $x = 0$ in $N'$. Also, $N_{i,m}$ is obtained by replacing every entry of the matrix corresponding to $N_i$ by its minor with respect $x$. And hence, $N_{i,m}$ can be thought of as a corresponding Jacobian minor of $N_x$ obtained by taking the minor of $N$ with respect to $x$. Thus the two equations corresponding to the coefficient of $x^0$ and $x^{h+m}$ are indeed of the same form as $\sum f_i N_i = 0$. (In the case of the coefficient of $x^{h+m}$, we need to set other variables in the row/column containing $x$ as in the proof of Lemma 5.2)

*Fanin reduction* step - Without loss of generality, let $x_1 \in T_1 \setminus T_2$. Pick a row $R$ of $N$ barring the first $r$ rows, and let $y_1, \cdots, y_r$ be the grey variables in $R$ (where $y_1$ is in the same column as $x_1$). By a similar process as in the proof of Lemma 5.2, we can assume that at least one $f_i$ is nonzero when $y_2, \cdots, y_r$ are set to zero.

If one of the $f_i$'s become zero when $y_2, \cdots, y_r = 0$, then pick any white variable $y$ in row $R$ and set every variable in row $R$ to zero besides $y$. This would ensure that the fanin of the equation reduces and each $N_i$ is now $y^m \cdot N_i'$. Each $N_i'$ may be thought of as being obtained from $N_y$, the minor of $N$ with respect to $y$. The other variables in the column of $y$ can be set to values to ensure that the $f_i$'s stay nonzero to obtain an equation of the form $\sum f_i' N_i' = 0$ of reduced fanin.

If none of the $f_i$'s become zero when $y_2, \cdots, y_r = 0$, then set every variable in row $R$ other than $y_1$ to zero. This ensures an entire column of the matrix corresponding to $N_1$ becomes zero (as $x_1$ indexes one of the columns of $N_1$), and hence $N_1$ becomes zero. On the other hand, $N_2$ remains nonzero and each surviving $N_i$ can be written as $y_1^m \cdot N_i'$, where $N_i'$ is the corresponding Jacobian minor of $N_{y_1}$. Again, the other variables in the column of $y_1$ can be set to values to ensure that $f_i$'s stay nonzero and we obtain an equation $\sum f_i' N_i' = 0$ of reduced fanin.

As in the proof of Lemma 5.2, we eventually obtain an equation of the form $f_1 N_1 = 0$ where $f_1 \neq 0$ thus implying that $N_1 = 0$. The number of variables that have been set to constants is bounded by $t + \log S$ where $S$ is the initial total sparsity of the $f_i$'s, and $N_1$ is a Jacobian minor of a symbolic matrix of dimension $n - (\log S + t - 1)$. Conjecture B.1 asserts that $N_1$ would be nonzero unless $\log S + t = \Omega(n - (\log S + t - 1))$, or $S = 2^{\Omega(n)}$. $\square$

That concludes the proof of Theorem B.3 as well.



ECCC ISSN 1433-8092

http://eccc.hpi-web.de