Electronic Colloquium on Computational Complexity, Report No. 89 (2014)

Lower Bounds for Depth Three Arithmetic Circuits with small bottom fanin

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

Chandan Saha
Indian Institute of Science
chandan@csa.iisc.ernet.in

July 16, 2014

Abstract

Shpilka and Wigderson [SW99] had posed the problem of proving exponential lower bounds for (nonhomogeneous) depth three arithmetic circuits with bounded bottom fanin over a field $\mathbb{F}$ of characteristic zero. We resolve this problem by proving a $N^{\Omega(\frac{d}{\tau})}$ lower bound for (nonhomogeneous) depth three arithmetic circuits with bottom fanin at most τ computing an explicit N-variate polynomial of degree d over $\mathbb{F}$.

Meanwhile, Nisan and Wigderson [NW97] had posed the problem of proving superpolynomial lower bounds for homogeneous depth five arithmetic circuits. We show a lower bound of $N^{\Omega(\sqrt{d})}$ for homogeneous depth five circuits (resp. also for depth three circuits) with bottom fanin at most N^{μ}, *for any fixed* $\mu < 1$. This resolves the problem posed by Nisan and Wigderson only partially because of the added restriction on the bottom fanin (a general homogeneous depth five circuit has bottom fanin at most N).

ISSN 1433-8092

1 Introduction

The problem of proving super-polynomial lower bounds for arithmetic circuits occupies a central position in algebraic complexity theory, much like the problem of proving super-polynomial lower bounds for Boolean circuits does in Boolean complexity. The model of arithmetic circuits is an algebraic analogue of the model of Boolean circuits: an arithmetic circuit contains addition ($+$) and multiplication ($\times$) gates and it naturally computes a polynomial in the input variables over some underlying field. We typically allow the input edges to a $+$ gate to be labelled with arbitrary constants from the underlying field $\mathbb{F}$ so that a $+$ gate can in fact compute an arbitrary $\mathbb{F}$-linear combination of its inputs. As a possible stepping stone, researchers have focussed on restricted (but still nontrivial and interesting) subclasses of arithmetic circuits. In particular, circuits of low depth[1] are interesting for they correspond to computation which is highly parallel. But despite a lot of attention, proving superpolynomial lower bounds for even bounded depth arithmetic circuits remains an outstanding open problem.

Notation for low depth circuits. Bounded depth arithmetic circuits[2] consist of alternating layers of addition and multiplication gates. We will denote an arithmetic circuit of depth Δ by a sequence of Δ symbols wherein each symbol (either Σ or Π) denotes the nature of the gates at the corresponding layer and the leftmost symbol indicates the nature of the gates at the output layer. For example, a $\Sigma\Pi\Sigma$ circuit with input $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ computes a polynomial in the following manner:

$$C(\mathbf{x}) = \sum_i \prod_j \left(a_{ij0} + \sum_{k=1}^{n} a_{ijk} x_k \right), \quad \text{where each } a_{ijk} \in \mathbb{F}. \tag{1}$$

In dealing with circuits it is useful to keep track of the fanin to various gates. Towards this end, we extend the above notation and allow integer superscripts on the gate symbols (i.e. Σ or Π symbols) which denotes an upper bound on the fanin of any gate in the corresponding layer[3]. So for example a $\Sigma^{[s]}\Pi^{[e]}\Sigma^{[\tau]}$ circuit computes a polynomial of the form:

$$C(\mathbf{x}) \quad = \quad \sum_{i \le s} \prod_{j \le e} \left(\sum_{k \le \tau} a_{ijk} \cdot y_{ijk} \right) \quad \text{where each } a_{ijk} \in \mathbb{F} \text{ and } y_{ijk} \in \mathbf{x} \cup \{1\}.$$

while a $\Sigma\Pi^{[a]}\Sigma\Pi^{[b]}$ circuit computes a polynomial in the following manner:

$$C(\mathbf{x}) \quad = \quad \sum_i \prod_{j \le a} Q_{ij}(\mathbf{x}) \quad \text{where } \deg Q_{ij} \le b \text{ for all } i \text{ and } j.$$

Depth Three Circuits. Being the shallowest nontrivial subclass of arithmetic circuits, depth three arithmetic circuits, also denoted as $\Sigma\Pi\Sigma$ circuits[4] have been intensely investigated. $\Sigma\Pi\Sigma$ cir-

[1] Recall that the depth of a circuit is the maximum length of any path in the circuit.

[2] Throughout the rest of this paper, we shall deal with bounded depth circuits - indeed of depth at most 5. In this context, we will often use the words formulas and circuits interchangably, as depth-Δ circuits can be converted to depth-Δ formulas with only a polynomial blow-up in size.

[3] If there is no superscript on the symbol for a layer, then the fanin at that layer is allowed to be arbitrary.

[4] Depth three circuits with a product gate at the output, i.e. $\Pi\Sigma\Pi$-circuits, are uninteresting from the perspective of proving lower bounds for they cannot even compute irreducible polynomials of degree more than 1 (regardless of size).

cuits (more specifically tensors) arise naturally in the investigation of the complexity of polynomial multiplication and matrix multiplication[5]. Moreover, the optimal formula/circuit for some well known families of polynomials are in fact depth three circuits. In particular, the best known circuit for computing the permanent Perm_d is known as Ryser's formula [Rys63] which is a (homogeneous[6]) depth three circuit of size $O(d^2 \cdot 2^d)$. Recently it was shown [GKKS13a] that (nonhomogeneous) $\Sigma\Pi\Sigma$ circuits are surprsingly powerful - any polynomial f of *small* circuit complexity can also be computed by a (nonhomogeneous) $\Sigma\Pi\Sigma$ circuit which is *not too large*. Specifically[7], if an n-variate polynomial f of degree d can be computed by $poly(n)$-sized circuits, then it can also be computed by $n^{O(\sqrt{d})}$-sized $\Sigma\Pi\Sigma$ circuit[8].

Lower Bounds for $\Sigma\Pi\Sigma$ circuits. In a very influential piece of work, Nisan and Wigderson [NW97] showed that over any field $\mathbb{F}$, any *homogeneous* $\Sigma\Pi\Sigma$ circuit computing the determinant Det_d must be of size $2^{\Omega(d)}$. Grigoriev and Karpinski [GK98], and Grigoriev and Razborov [GR00] showed that any $\Sigma\Pi\Sigma$ arithmetic circuit over any *fixed* finite field computing Det_d must be of size at least $2^{\Omega(d)}$. This also implies that any $\Sigma\Pi\Sigma$ arithmetic circuit *over integers* computing Det_d must be of size at least $2^{\Omega(d)}$. Raz and Yehudayoff give $2^{\Omega(d)}$ lower bounds for *multilinear* $\Sigma\Pi\Sigma$ circuits[9]. But despite all this progress, even a superpolynomial lower bound for unrestricted $\Sigma\Pi\Sigma$ circuits (over an infinite field) has remained ellusive. The best known lower bound in the general $\Sigma\Pi\Sigma$ case is the quadratic lower bound due to Shpilka and Wigderson [SW99]. For more on $\Sigma\Pi\Sigma$ circuits, we refer the reader to the thesis of Shpilka [Shp01] and the references therein.

$\Sigma\Pi\Sigma$ circuits with small bottom fanin. Nisan and Wigderson noted that (nonhomogeneous) $\Sigma\Pi\Sigma$ circuits with bottom fanin just two can be exponentially more powerful than homogeneous $\Sigma\Pi\Sigma$ circuits - any homogeneous $\Sigma\Pi\Sigma$ circuit computing the elementary symmetric polynomial of degree n on $2n$ variables[10] must be of size $2^{\Omega(n)}$ but it can be computed by just $O(n^2)$-sized $\Sigma\Pi\Sigma^{[2]}$

[5] For example it can be shown that the product of two $n \times n$ matrices can be computed with $\tilde{O}(n^\omega)$ arithmetic operations if and only if the polynomial

$$M_n = \sum_{i\in[n]} \sum_{j\in[n]} \sum_{k\in[n]} x_{ij} \cdot y_{jk} \cdot z_{ki}$$

can be computed by a $\Sigma\Pi\Sigma$ circuit where the top fanin s is at most $\tilde{O}(n^\omega)$.

[6] Recall that a multivariate polynomial is said to be homogeneous if all its monomials have the same total degree. An arithmetic circuit is said to be *homogeneous* if the polynomial computed at every internal node of the circuit is a homogeneous polynomial. It is a folklore result (cf. the survey by Shpilka and Yehudayoff [SY10]) that as far as computation by polynomial-sized arithmetic circuits of unbounded depth is concerned one can assume without loss of generality that the circuit is homogeneous. Specifically, if a homogeneous polynomial f of degree d can be computed by an (unbounded depth) arithmetic circuit of size s, then it can also be computed by a *homogeneous* circuit of size $O(d^2 \cdot s)$.

[7] The quantitative version mentioned here is due to an improvement by Tavenas [Tav13].

[8] This depth reduction is only valid over fields of characteristic zero.

[9] The results of Raz and Yehudayoff are more general and extend to lower bounds for any constant depth multilinear circuit.

[10] The elementary polynomial of degree n on $2n$ formal variables is the arithmetic analog of the MAJORITY function. Formally, it is defined as

$$\text{ESYM}_n(x_1, \ldots, x_{2n}) \stackrel{\text{def}}{=} \sum_{\substack{S\subseteq[2n] \\ |S|=n}} \prod_{i\in S} x_i.$$

circuits[11]. They also noted that this contrasts sharply with the the exponential lower bounds for MAJORITY in the Boolean model and over fixed finite fields. Recently, Ramprasad Saptharishi [Sap14] pointed out to us that the depth reduction in [GKKS13a] actually yields $\Sigma\Pi\Sigma^{[O(\sqrt{d})]}$-circuits. This indicates that (nonhomogeneous) $\Sigma\Pi\Sigma^{[\tau]}$-circuits are interesting and motivates the effort to prove lower bounds for them. Indeed, Shpilka and Wigderson [SW99] had already noted this frontier in arithmetic complexity and explicitly posed the problem of proving lower bounds for (nonhomogeneous) depth three circuits with bounded bottom fanin (over fields of characteristic zero). We resolve this challenge here by proving exponential lower bounds for such circuits. Our proof techniques are based on recent developments in arithmetic circuit lower bounds.

Recent lower bound results. A series of recent works have built upon the work of Nisan and Wigderson [NW97] to prove lower bounds for *homogeneous* depth four circuits. Motivated by the depth reduction results of Agrawal and Vinay [AV08] and Koiran [Koi12] and Tavenas [Tav13] and using a complexity measure introduced in Kayal [Kay12], the work of Gupta, Kamath, Kayal and Saptharishi [GKKS13b] and Kayal, Saha and Saptharishi [KSS14] have led to lower bounds of $n^{\Omega(\sqrt{d})}$ for homogeneous depth four circuits of bottom fanin $O(\sqrt{d})$. Follow-up work by Fournier, Limaye, Malod and Srinivasan [FLMS14] showed the same lower bound for a family of polynomials in VP. Subsequently, work by Kayal, Limaye, Saha and Srinivasan [KLSS14b, KLSS14a] removed the restriction on the bottom fanin and obtained a $n^{\Omega(\sqrt{d})}$ lower bound for homogeneous depth four circuits for a family of polynomials in VNP [12]. Follow-up work by Kumar and Saraf [KS14a] showed the same lower bounds for a family of polynomials in VP[13].

Our results. Our first result is a lower bound of $N^{\Omega(\frac{d}{\tau})}$ for (nonhomogeneous) $\Sigma\Pi\Sigma^{[\tau]}$ circuits which resolves an open problem (specifically, Problem 7.5) posed by Shpilka and Wigderson in [SW99]. It also implies that the depth reduction result of [GKKS13a] is optimal *assuming that the resulting depth three circuit has bottom fanin at most* $O(\sqrt{d})$. The formal statement is as follows.

Theorem 1. Lower Bound for $\Sigma\Pi\Sigma^{[\tau]}$ circuits. *Let $\mathbb{F}$ be a field of characteristic zero. There is a family of N-variate, degree d polynomials $\{f_N\}$ in VP with $N = d^{O(1)}$ such that any $\Sigma\Pi\Sigma^{[\tau]}$ circuit over $\mathbb{F}$ computing f_N must have top fanin at least $N^{\Omega(\frac{d}{\tau})}$.*

We would like to stress here that there is no restriction of homogeneity on the $\Sigma\Pi\Sigma^{[\tau]}$ formula in the above statement. Indeed the formal degree of the $\Sigma\Pi\Sigma^{[\tau]}$ circuit can be arbitarily large (say doubly exponential) and yet we obtain the stated lower bound on the top fanin. We prove Theorem 1 by first showing a reduction from $\Sigma\Pi\Sigma^{[\tau]}$ circuits to *a subclass of homogeneous* $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ *circuits*[14] (using a result implicit in [GKKS13a]; see Lemma 5 in Secton 4). It turns out fortunately

[11] More accurately, [NW97] attribute Michael Ben-Or for an $O(n^2)$-sized $\Sigma\Pi\Sigma$ circuit for $\text{ESYM}_n(x_1, x_2, \ldots, x_{2n})$ which has the following specific form:

$$\text{ESYM}_n(\mathbf{x}) = \sum_{i=1}^{2n+1} a_i \prod_{j=1}^{2n} (x_j + i),$$

where the a_i's are appropriate field constants.

[12] Meanwhile, an independent work by Kumar and Saraf [KS14b] also showed a $n^{\Omega(\log \log n)}$ lower bound for general homogeneous depth-4 circuits without the bottom fanin restriction.

[13] The result of [KS14a] is also valid over any field $\mathbb{F}$.

[14] The reduction from $\Sigma\Pi\Sigma$ formulas to homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ formulas yields a restricted class of homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ formulas wherein every product gate in the layer closest to the input layer is actually an *exponentiation*

that the proof techniques/complexity measure used in [KLSS14a, KS14a] are readily applicable to this subclass of homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits and this yields the above lower bound. Having obtained a lower bound for a subclass of homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ circuits, can our techniques be pushed further to yield lower bounds for general homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ formulas? It turns out that proving superpolynomial lower bounds for general homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ formulas was explicitly posed as an open problem by Nisan and Wigderson in [NW97]. We next give a lower bound for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ formulas with small bottom fanin. It resolves the above problem only partially because of the added restriction on the bottom fanin.

Theorem 2. Lower Bound for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits. *Let $\mathbb{F}$ be a field of characteristic zero and $\mu \in [0,1)$ be any fixed positive real number less than 1. Let $\alpha = \frac{2\mu+1}{1-\mu}$ and $\tau = O(N^{\mu})$. There is a family of N-variate, degree d polynomials $\{f_N\}$ in VNP with $N \in [d^{2+\alpha}, 2d^{2+\alpha}]$ such that any homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ formula over $\mathbb{F}$ computing f_N has size $N^{\Omega(\sqrt{d})}$.*

The family of polynomials in the above theorem is the Nisan-Wigderson design based polynomials introduced in [KSS14], and later used in [KLSS14a, KS14a], but with an altered set of parameters. The complexity measure that we use for this result is (almost) the same as the one introduced in [KLSS14a] called *the dimension of projected shifted partials under random restrictions*. An appropriate adaption of the techniques yields a lower bound for N-input homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^{\mu}]}$-circuits *for some fixed value of* $\mu < 0.1$. We felt that it would be worthwhile to push the analysis further and obtain as good a lower bound as possible while allowing the bottom fanin to be as large as possible - specifically, to allow the bottom fanin to be N^{μ} for *any* constant μ that is arbitrarily close to 1. For this, we delve deeper into the analysis of [KLSS14a] and carefully tune it at certain places, including the complexity analysis of the explicit polynomial family for which the lower bound is shown. As a corollary, we also obtain a similar lower bound for (nonhomogeneous) $\Sigma\Pi\Sigma^{[N^{\mu}]}$ circuits for *any* constant $\mu < 1$.

Corollary 3. *Let $\mathbb{F}$ be a field of characteristic zero and $\mu \in [0,1)$ be any fixed positive real number less than 1. Let $\alpha = \frac{2\mu+1}{1-\mu}$. There is a family of N-variate, degree d polynomials $\{f_N\}$ in VNP with $N \in [d^{2+\alpha}, 2d^{2+\alpha}]$ such that any $\Sigma\Pi\Sigma^{[N^{\mu}]}$ formula over $\mathbb{F}$ computing f_N has size at least $N^{\Omega(\sqrt{d})}$.*

2 Proof Overview

From depth three to homogeneous depth-5. Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a homogeneous N-variate polynomial of degree d. We first observe that if f is computed by a small (of size $N^{o(\sqrt{d})}$) $\Sigma\Pi\Sigma$ circuit $C(\mathbf{x})$ then it is also computed by a small (of size $N^{o(\sqrt{d})}$) formula $D(\mathbf{x})$ which is structurally in a sublass of *homogeneous* $\Sigma\Pi\Sigma\Pi\Sigma$ formulas (see Lemma 5 in Secton 4). Further, this reduction from depth three to homogeneous depth-5 preserves the bound on the bottom fanin of the formulas, i.e. if the bottom fanin of $C(\mathbf{x})$ is bounded by τ then same is true for $D(\mathbf{x})$. It turns out that the proof techniques/complexity measure employed in [KLSS14a, KS14a] are readily applicable to this subclass of homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits and this yields the lower bound of theorem 1. We then consider general homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits.

gate, i.e. a product gate all of whose inputs originate from the source node g, so that its output is of the form g^e for some $e \in \mathbb{Z}_{\geq 1}$. We denote such formulas as $\Sigma\Pi\Sigma\wedge\Sigma$ formulas.

Homogeneous depth five formulas. A homogeneous depth-5 formula is a representation of the form

$$D(\mathbf{x}) = \sum_i \prod_j \sum_r Q_{ijr}, \tag{2}$$

where Q_{ijr} is a product of linear forms. Also, suppose the number of variables in every linear form in Q_{ijr} (for every i, j and r) is bounded by $\tau = N^\mu$ for some fixed constant $\mu < 1$. To prove a lower bound on the size of $D(\mathbf{x})$, our overall strategy is based on the complexity measure introduced in [KLSS14a] called *the dimension of projected shifted partials under random restrictions*. As is common to many lower bounds, the proof is in two steps:

1. Upper bound the measure for any $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$-formula $D(\mathbf{x})$ as in equation (2), and
2. Lower bound the measure for an explicit (family of) polynomial(s) f.

Overall, the lower bound follows by comparing these two bounds. We will now describe the complexity measure used and then indicate why it is small for $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$-formulas.

Random restriction. The random restriction we use in this paper is quite natural and (almost) same as in [KLSS14a]. We consider the identity (2) and in that set each variable to zero independently at random with probability $(1-p)$, where $p = d^{-\beta}$ for a suitable constant $\beta > 0$ (a variable is left untouched with probability p.) For ease of exposition, it is convenient to denote a restriction in which a subset of variables $R \subseteq [N]$ is[15] set to zero (and the variables outside R are left untouched) as a homomorphism, $\sigma_R : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$. Formally, $\sigma_R : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ is a homomorphism such that $\sigma_R(f) \stackrel{\text{def}}{=} f|_{x_i=0\ \forall i \in R}$. In this notation, a random restriction can also be viewed as constructing an R by picking every variable independently at random with probability $1-p$ and then applying[16] the map σ_R to the expression given by equation (2).

The complexity measure. Let $m = x_{i_1} \cdots x_{i_k}$ be a monomial in $\mathbf{x}$. Denote $\frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} f$ by $\partial_m f$ and define

$$\boldsymbol{\partial}^{=k}_{\text{ML}} f := \{\partial_m f \mid m \text{ is a multilinear monomial of degree } k\}$$

We will refer to $\boldsymbol{\partial}^{=k}_{\text{ML}} f$ as the set of all *multilinear* k-th order partial derivatives of $f \in \mathbb{F}[\mathbf{x}]$. Let $\mathbf{x}^{=\ell}$ be the set of all multilinear monomials in $\mathbf{x}$ of degree equal to ℓ. We denote by $\mathbf{x}^{=\ell} \cdot \boldsymbol{\partial}^{=k}_{\text{ML}} f$ the set of all polynomials of the form $m \cdot g$ where $m \in \mathbf{x}^{=\ell}$ and $g \in \boldsymbol{\partial}^{=k}_{\text{ML}} f$. Define a map $\pi : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ such that when π acts on a polynomial f, it retains only and exactly the *multilinear monomials* of f. More precisely, let M_f be the set of all monomials with nonzero coefficients in f. Then, $\pi(f) := \sum_u c_u m_u$ where m_u is a multilinear monomial in M_f and coefficient of m_u in f is c_u. Naturally, π is a linear map, i.e. $\pi(af + bg) = a \cdot \pi(f) + b \cdot \pi(g)$ for every $a, b \in \mathbb{F}$ and $f, g \in \mathbb{F}[\mathbf{x}]$. The definition of π extends naturally to sets of polynomials: For $A \subseteq \mathbb{F}[\mathbf{x}]$, let $\pi(A) := \{\pi(f) \mid f \in A\}$. For integers k and ℓ, the space of projected shifted partials of f is the linear span (i.e. $\mathbb{F}$-span) of the polynomials in $\pi(\mathbf{x}^{=\ell} \cdot \boldsymbol{\partial}^{=k}_{\text{ML}} f)$. The measure we use is the dimension of this space of projected shifted partials, denoted by $\mathsf{DPSP}_{k,\ell}$ (or simply DPSP assuming parameters k and ℓ are fixed suitably):

$$\mathsf{DPSP}_{k,\ell}(f) := \dim(\pi(\mathbf{x}^{=\ell} \cdot \boldsymbol{\partial}^{=k}_{\text{ML}} f)).$$

[15] $[N]$ denotes the set of the first N positive integers, i.e. $\{1, 2, \ldots, N\}$.

[16] We will use the random restriction in two phases in Section 6 to obtain an appropriate upper bound on the measure for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ formulas.

Observe that the measure $\mathsf{DPSP}_{k,\ell}$ obeys subadditivity, i.e. $\mathsf{DPSP}_{k,\ell}(f+g) \leq \mathsf{DPSP}_{k,\ell}(f) + \mathsf{DPSP}_{k,\ell}(g)$.

From depth-5 to depth-4. Let $D(\mathbf{x})$ be a homogeneous-$\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$ formula as in equation (2) of size at most $N^{o(\sqrt{d})}$ so that in particular the total number of Q_{ijr}'s appearing in it is at most $s = N^{o(\sqrt{d})}$. We show that when a random restriction σ_R is applied on $D(\mathbf{x})$, then with high probability $\sigma_R(D(\mathbf{x}))$ can be expressed as $D_1(\mathbf{x}) + D_2(\mathbf{x})$, where $D_1(\mathbf{x})$ is computed by a homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{d}]}$ formula of top fanin at most $N^{o(\sqrt{d})}$ and $D_2(\mathbf{x})$ is a polynomial such that $\mathsf{DPSP}(D_2(\mathbf{x})) = 0$. We will argue this shortly but assuming that this happens, we can infer (via subadditivity) that

$$\begin{aligned} \mathsf{DPSP}(\sigma_R(D(\mathbf{x}))) &\leq \mathsf{DPSP}(D_1(\mathbf{x})) + \mathsf{DPSP}(D_2(\mathbf{x})) \\ &= \mathsf{DPSP}(D_1(\mathbf{x})). \end{aligned}$$

$\mathsf{DPSP}(D_1(\mathbf{x}))$ can then be upper bounded using known arguments from [KLSS14a] which in turn yields an upper bound for $\mathsf{DPSP}(\sigma_R(D(\mathbf{x})))$.

Using random restrictions to obtain a decomposition. The reason $\sigma_R(D(\mathbf{x}))$ decomposes into $D_1(\mathbf{x})$ and $D_2(\mathbf{x})$ with high probability is as follows. Let $t = \sqrt{d}$. In equation (2), suppose a Q_{ijr} has degree greater than $2t$. Such a Q_{ijr} can be expressed as $\tilde{Q}_{ijr} \cdot P_{ijr}$ with $\deg(\tilde{Q}_{ijr}) = 2t$, by simply multiplying out $2t$ linear forms in Q_{ijr}. Since bottom fanin of $D(\mathbf{x})$ is bounded by N^μ, the number of monomials in $\tilde{Q}_{ijr}$ is bounded by $N^{2\mu t}$. Monomials of $\tilde{Q}_{ijr}$ are of two kinds - those with individual degree of variables bounded by 2 (and hence have support at least t), and those with at least one variable having degree 3 or more. The probability any of the monomials in $\tilde{Q}_{ijr}$ survives under the action of the random restriction σ_R is less than $p^t \cdot N^{2\mu t}$. Running over all Q_{ijr}, with probability at least $1 - s \cdot p^t \cdot N^{2\mu t}$, we have

$$\sigma_R(D(\mathbf{x})) = \sum_i \prod_j \sum_{\substack{r \\ \deg(Q_{ijr}) \leq 2t}} \sigma_R(Q_{ijr}) + P(\mathbf{x}),$$

where every monomial in $P(\mathbf{x})$ has a variable with degree 3 or more. Now observe that for any multilinear monomial m, every monomial in $\partial_m P$ has a variable of degree 2 or more and hence $\pi(\partial_m P) = 0$, implying $\mathsf{DPSP}(P) = 0$. By taking $D_1(\mathbf{x}) = \sum_i \prod_j \sum_{r,\deg(Q_{ijr}) \leq 2t} \sigma_R(Q_{ijr})$ and $D_2(\mathbf{x}) = P(\mathbf{x})$, we come to the desired conclusion, if the "bad" probability, namely $s \cdot p^t \cdot N^{2\mu t}$, is small. Now suppose $N = d^3$ (as is the case in [KLSS14a]). Then the bad probability is $s \cdot N^{-(\frac{\beta}{3} - 2\mu)t}$ which is negligible for any constant μ less than $\beta/6$. This gives the required decomposition.

Extension for arbitrary $\mu < 1$. Combining the above decomposition argument with the lower bound available for homogeneous-$\Sigma\Pi\Sigma\Pi^{[\sqrt{d}]}$-circuits (which imposes some additional constraints on how large β can be), we get that if μ is sufficiently small (say, 0.01), any homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$ formula computing the same family of Nisan-Wigderson design based polynomials as used in [KLSS14a], has size $N^{\Omega(\sqrt{d})}$. However, in order to prove the same size lower bound for *any* constant $\mu < 1$, we delve deeper into the analysis of [KLSS14a] and carefully tune it at certain places, including the complexity analysis of the explicit polynomial family for which the lower bound is shown.

3 Preliminaries

Affine forms and linear forms. An *affine form* is simply another name for a degree one polynomial, with a (possibly) nonzero constant term. Thus an affine form $\ell(\mathbf{x})$ looks like

$$\ell(\mathbf{x}) = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n,$$

where each $a_i \in \mathbb{F}$. The *weight* of such an affine form $\ell(\mathbf{x})$ will be the number of nonzero coefficients in it, i.e.

$$\text{weight of } \ell \stackrel{\text{def}}{=} |\{i \in [0..n] : a_i \neq 0\}|$$

A homogeneous degree one polynomial (i.e. one whose constant term a_0 is zero) we will refer to as a *linear form*.

Notation for circuits with exponentiation gates. Sometimes a multiplication gate in our circuit will have the feature that all its incoming edges originate from a single gate g (thus computing g^e, if there are e wires entering the multiplication gate). We will refer to such gates as *exponentiation gates* and denote them by the symbol $\wedge$. So for example, a $\Sigma\wedge\Sigma$ circuit computes a polynomial in the following manner:

$$C(\mathbf{x}) \quad = \quad \sum_{i\in[s]} \ell_i(\mathbf{x})^{e_i} \quad \text{where each } \ell_i \in \mathbb{F}[\mathbf{x}] \text{ is an affine form.}$$

A numerical estimate. The following numerical estimate from [GKKS13b] will be useful.

Lemma 4. *Let $a(n)$, $f(n)$, $g(n)$: $\mathbb{Z}_{>0} \mapsto \mathbb{Z}$ be integer valued functions such that $(|f| + |g|) = o(a)$. Then*

$$\ln \frac{(a+f)!}{(a-g!)} = (f+g)\ln a \pm O\left(\frac{f^2+g^2}{a}\right)$$

4 Depth Three Circuits with small bottom fanin

In this section, we will first see a reduction from (nonhomogeneous) $\Sigma\Pi\Sigma^{[\tau]}$ to a subclass of homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits. It can be easily inferred from the proof of lemma V.3[17] in [GKKS13a] but we nevertheless give a proof here for completeness.

Lemma 5. *(implicit in [GKKS13a].) Let $d \geq 1$ be an integer and $\mathbb{F}$ be an infinite field of characteristic larger than d (or of zero characteristic). Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a homogeneous N-variate polynomial of degree d computed by a $\Sigma^{[s]}\Pi^{[e]}\Sigma^{[\tau]}$ circuit. Then f can also be computed by a homogeneous $\Sigma^{[s\cdot\exp(\sqrt{d})]}\Pi\Sigma^{[e]}\wedge\Sigma^{[\tau]}$ circuit.*

Proof. The premise that f can be computed by a $\Sigma^{[s]}\Pi^{[e]}\Sigma^{[\tau]}$ circuit means that there exist $s \cdot e$ affine forms ℓ_{ij}'s each of weight at most τ such that

$$f(\mathbf{x}) = \sum_{i=1}^{s}\prod_{j=1}^{e} \ell_{ij}(\mathbf{x}). \tag{3}$$

[17] Ramprasad Saptharishi [Sap14] has recently communicated to us that the consequence in the original lemma in [GKKS13a] can be slightly improved quantitatively.

Expressing f as a sum of projections of elementary symmetric polynomials. We will first ensure that each of the affine forms ℓ_{ij} has a nonzero constant term. We can do this by applying a random shift of the form $\mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$ to the above identity. That is, pick a random point $\mathbf{a} \in \mathbb{F}^n$ and replacing $\mathbf{x}$ by $\mathbf{x} + \mathbf{a}$ in the identity (3) we get

$$\begin{aligned} f(\mathbf{x}+\mathbf{a}) &= \sum_{i=1}^{s} \prod_{j=1}^{e} \ell_{ij}(\mathbf{x}+\mathbf{a}) \\ &= \sum_{i=1}^{s} \alpha_i \prod_{j=1}^{e} (1 + m_{ij}(\mathbf{x})), \quad \text{where } m_{ij}(\mathbf{x}) \stackrel{\text{def}}{=} \ell_{ij}(\mathbf{x}) - \ell_{ij}(\mathbf{0}) \text{ is a linear form of} \\ &\qquad\qquad \text{weight at most } \tau \text{ and } \alpha_i \stackrel{\text{def}}{=} \prod_{j=1}^{e} \ell_{ij}(\mathbf{a}) \end{aligned}$$

Comparing the homogeneous components of degree d on the two sides of the above identity we get

$$f(\mathbf{x}) = \sum_{i=1}^{s} \alpha_i \cdot \mathrm{ESYM}_d(m_{i1}, \ldots, m_{ie}), \tag{4}$$

where

$$\mathrm{ESYM}_d(y_1, \ldots, y_e) \stackrel{\text{def}}{=} \sum_{\substack{S \subseteq [e] \\ |S| = d}} \prod_{i \in S} y_i$$

is the elementary symmetric polynomial of degree d on the e formal variables $y_1, y_2, \ldots, y_e$.

Expressing ESYM_d in terms of the power symmetric polynomials. We now use Newton's identities to express each elementary symmetric polynomial that occurs above in terms of the power-symmetric polynomials defined as:

$$\mathrm{PSYM}_r(y_1, \ldots, y_e) \stackrel{\text{def}}{=} \sum_{j \in [e]} y_j^r.$$

We use the following implication of Newton's identities (cf. [Lit50]):

$$\mathrm{ESYM}_d = \frac{1}{d!} \cdot \begin{vmatrix} \mathrm{PSYM}_1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \mathrm{PSYM}_2 & \mathrm{PSYM}_1 & 2 & 0 & \cdots & 0 & 0 \\ \mathrm{PSYM}_3 & \mathrm{PSYM}_2 & \mathrm{PSYM}_1 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathrm{PSYM}_{d-1} & \mathrm{PSYM}_{d-2} & \mathrm{PSYM}_{d-3} & \mathrm{PSYM}_{d-4} & \cdots & \mathrm{PSYM}_1 & n-1 \\ \mathrm{PSYM}_d & \mathrm{PSYM}_{d-1} & \mathrm{PSYM}_{d-2} & \mathrm{PSYM}_{d-3} & \cdots & \mathrm{PSYM}_2 & \mathrm{PSYM}_1 \end{vmatrix}.$$

In particular, this means that ESYM_d can be expressed as a polynomial function of the PSYM_i's. Let us now count how many terms are there in such a polynomial expression. Expanding out the determinant above we see that there exist scalars $\beta_{\mathbf{a}}$'s such that

$$\mathrm{ESYM}_d(\mathbf{y}) = \sum_{\substack{\mathbf{a}=(a_1,\ldots,a_d) \in \mathbb{Z}^d_{\geq 0} \\ \sum_i i \cdot a_i = d}} \beta_{\mathbf{a}} \cdot \prod_{i \in [d]} \mathrm{PSYM}_i^{a_i}(\mathbf{y}). \tag{5}$$

The number of solutions of $\sum_{i\in[d]} i \cdot a_i = d$ is exactly the number of ways to partition the natural number d and hence is $2^{\Theta(\sqrt{d})}$ by the Hardy-Ramanujan estimate for the partition function [HR18]. Hence the number of terms in the above summation is $2^{\Theta(\sqrt{d})}$. In particular this means that $\text{ESym}_d(\mathbf{y})$ is computed by a *homogeneous* $\Sigma^{[\exp(\sqrt{d})]}\Pi\Sigma^{[e]}\wedge$-circuit.

Combining (4) and (5) to get a homogeneous $\Sigma\Pi\Sigma\wedge\Sigma$ circuit for f. If we now replace each occurence of ESym_d in equation (4) by its homogeneous $\Sigma\Pi\Sigma\wedge$ circuit given by the identity (5) , we see that $f(\mathbf{x})$ is computed by a homogeneous $\Sigma^{[s\cdot\exp(\sqrt{d})]}\Pi\Sigma^{[e]}\wedge\Sigma^{[\tau]}$ circuit. This proves the lemma.

□

We next observe that the homogeneous $\Sigma\Pi\Sigma\wedge\Sigma$-circuit in the outcome of the above lemma corresponds to a certain structured form for expressing f that we make precise below. For ease of subsequent exposition, let us introduce the following notation/terminology. Let $m = x_1^{e_1}\cdot x_2^{e_2}\cdot\ldots\cdot x_N^{e_N}$ in $\mathbb{F}[x_1, x_2, \ldots, x_N]$ be a monomial. The support of m, denoted $\text{Supp}(m)$ is the subset of variables appearing in it, i.e.

$$\text{Supp}(m) \stackrel{\text{def}}{=} \{i \ : \ e_i \geq 1\} \subseteq [N].$$

The support size of a polynomial Q, denoted $|\text{Supp}(Q)|$ is the maximum support size of any monomial appearing in Q.

Proposition 6. *Let $d \geq 1$ be an integer and $\mathbb{F}$ be an infinite field of characteristic larger than d (or of zero characteristic). Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a homogeneous N-variate polynomial of degree d computed by a $\Sigma^{[s]}\Pi^{[e]}\Sigma^{[\tau]}$ circuit. Then f admits an expression of the form*

$$f(\mathbf{x}) = \sum_{i}^{s\cdot\exp(\sqrt{d})} \prod_{j} Q_{ij}, \quad \text{Supp}(Q_{ij}) \leq \tau \tag{6}$$

Proof. The premise that f can be computed by a $\Sigma^{[s]}\Pi^{[e]}\Sigma^{[\tau]}$ circuit means that there exist $s\cdot e$ affine forms ℓ_{ij}'s each having at most τ nonzero coefficients such that

$$f(\mathbf{x}) = \sum_{i=1}^{s}\prod_{j=1}^{e} \ell_{ij}(\mathbf{x}). \tag{7}$$

First observe that if we have a linear form ℓ in which at most τ coefficients are nonzero, then for all $j \geq 1$, we have

$$\text{Supp}(\ell^j) \leq \tau.$$

In particular, this means that for all $r \geq 1$ and all $i \leq s$ we have $\text{Supp}(\text{PSym}_r(\ell_{i1}, \ell_{i2}, \ldots, \ell_{ie})) \leq \tau$. By the proof of lemma 5 we get that f can be expressed as a sum of product of the PSym_r's in a homogeneous fashion, with the expression having $s\cdot\exp(\sqrt{d})$ many terms. Hence f has a representation of the form given by equation (6). □

This means that our problem reduces to proving lower bounds for representations of the form given by the right-hand side of equation (6) which we refer to as τ-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. It turns out that such representations occur also as an intermediate step in prior work and [KLSS14a] explicitly gives an $N^{\Omega(\frac{d}{\tau})}$ lower bound for such representations.

Theorem 7. [KLSS14a]. *There exists an explicit family $\{f_N\}$ of homogeneous degree d polynomials on $N = d^3$ variables in* VNP *such that any τ-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuit computing f_N has top fanin at least $N^{\Omega(\frac{d}{\tau})}$.*

In the follow-up work of [KS14a], the class of τ-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits occurs implicitly. It follows from their work that the above lower bound is in fact valid for the family of iterated matrix multiplication polynomial which is in VP (in fact is complete for a subclass of VP called algebraic branching programs).

Theorem 8. [KS14a]. *There exists an explicit family $\{f_N\}$ of homogeneous degree d polynomials on $N = d^{O(1)}$ variables in* VP *such that any τ-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuit computing f_N has top fanin at least $N^{\Omega(\frac{d}{\tau})}$.*

Combining Proposition 6 with the above theorem immediately yields theorem 1. In the next section we move on investigating homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ circuits.

5 The lower bound for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$ formulas

Here we follow the outline given in section 2 and derive a lower bound for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$-formulas.

Step 1: an upper bound for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$-formulas. Let $0 \leq \mu < 1$ be a fixed constant. Consider a homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^\mu]}$ formula of size s as in equation (2) computing a homogeneous N-variate polynomial of degree d. We pick a random set $R \subseteq [N]$ by picking each variable independently at random with probability $1-p$, where $p = d^{-\beta}$ (for a suitable constant $\beta > 0$), and upper bound the DPSP-complexity of $\sigma_R(D(\mathbf{x}))$.

Lemma 9. *Let $t = \sqrt{d}$, $\alpha = \frac{2\mu+1}{1-\mu}$ and $d^{2+\alpha} \leq N \leq 2d^{2+\alpha}$ be an integer. If $s \leq N^{\frac{0.03}{2+\alpha}\cdot\sqrt{d}}$ then there exists a constant $0 < \beta < \alpha$ such that with probability at least $1 - \frac{1}{N^{\Omega(\sqrt{d})}}$, a random restriction σ_R satisfies:*

$$\mathsf{DPSP}_{k,\ell}(\sigma_R(D(\mathbf{x}))) \leq s \cdot \binom{\frac{d}{t}+1}{k} \cdot \binom{N}{\ell+2kt} \quad \text{for all } k,\ell \geq 0 \quad \text{satisfying } \ell + 2kt \leq \frac{N}{2}. \tag{8}$$

We defer the proof of this lemma to section 6.

Step 2.1: constructing a suitable family of polynomials. The explicit family of polynomials for which we prove the lower bound is a variant of the Nisan-Wigderson design based polynomials used in [KSS14, KLSS14a, KS14a]. The choice of this family depends on the bottom fanin of the depth 5 formulas. When the bottom fanin is $\tau = N^\mu$, for some fixed $0 \leq \mu < 1$, the family is defined as follows. For an integer d and $\alpha = \frac{2\mu+1}{1-\mu}$, let q be the smallest prime number between $d^{1+\alpha}$ and $2d^{1+\alpha}$ (such a prime is guaranteed to exist by the Bertrand-Chebyshev theorem [Erd32])[18]. We define a family of Nisan-Wigderson polynomials of degree d on $N = d \cdot q$ variables, parametrized by a number r (to be fixed later in the analysis).

$$\mathsf{NW}_r(x_{1,1}, x_{1,2}, \ldots, x_{d,q}) := \sum_{\substack{h(z)\in\mathbb{F}_q[z] \\ \deg(h)\leq r}} \prod_{i\in[d]} x_{i,h(i)},$$

[18] We are avoiding ceil/floor notations for simplicity of exposition

where $\mathbb{F}_q$ is the finite field with q elements.

Step 2.2: lower bounding the DPSP-complexity of our polynomial family. For appropriate choices of integers r, k, ℓ and a random restriction σ_R, we show that $\mathsf{DPSP}_{k,\ell}(\sigma_R(\mathsf{NW}_r))$ is large with high probability.

Lemma 10. The main technical lemma. *Let NW_r be the Nisan-Wigderson design based polynomial defined above. Suppose R is a set formed by picking each variable independently at random with probability $1-p$, where $p = d^{-\beta}$ and $\beta > 0$ is a constant less than α. Over any field $\mathbb{F}$ of characteristic zero, for $r = \frac{\alpha+\beta}{2(1+\alpha)} \cdot d - 1$, $k = \delta \cdot \sqrt{d}$ (for a small constant $\delta > 0$) and $\ell = \frac{N}{2}(1 - \frac{k \ln d}{d})$, we have*

$$\mathsf{DPSP}_{k,\ell}(\sigma_R(\mathsf{NW}_r)) \geq \frac{1}{d^{O(1)}} \min\left(\frac{p^k}{4^k} \cdot \binom{N}{k} \cdot \binom{N}{\ell}, \binom{N}{\ell + d - k}\right), \tag{9}$$

with probability at least $1 - \frac{1}{d^{\Theta(1)}}$.

We will prove this lemma in Section A of the appendix.

Final Step: comparing the two bounds. Comparing the probabilities with which equations (8) and (9) are satisfed, we see that there exists a set R such that both of them are simultaneously satisfied, implying:

$$\begin{aligned} s &\geq \frac{\mathsf{DPSP}_{k,\ell}(\sigma_R(\mathsf{NW}_r))}{\binom{\frac{d}{t}+1}{k} \cdot \binom{N}{\ell+2kt}} \\ &= N^{\Omega(\sqrt{d})} \qquad \text{(for small enough constant } \delta\text{)} \end{aligned}$$

The above implication can be worked out using the numerical estimates given in lemma 4. This proves the lower bound of theorem 2.

6 Upper bounding the measure for homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ formulas

Let $D(\mathbf{x})$ be a homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[\tau]}$ formula with bottom fanin bounded by $\tau = N^{\mu}$ where $\mu \in [0, 1)$ is a fixed constant.

$$D(\mathbf{x}) = \sum_i \prod_j \sum_r Q_{ijr}, \tag{10}$$

where Q_{ijr} is a product of linear forms. As before, let $\alpha = \frac{2\mu+1}{1-\mu}$. In this section we give a proof of lemma 9. We first show that when we apply a random restriction to a small homogeneous $\Sigma\Pi\Sigma\Pi\Sigma^{[N^{\mu}]}$ formula, then with high probability it decomposes into two pieces which are individually much easier to deal with.

Lemma 11. Decompostion under random restrictions. *Suppose that $D(\mathbf{x})$ has size $s \leq N^{\frac{0.03}{2+\alpha} \cdot \sqrt{d}}$. Then, it is possible to fix a constant $0 < \beta < \alpha$ and[19] form a set R by picking each*

[19] The requirement of $\beta < \alpha$ in the statement of lemma 11 comes from Lemma 10.

variable independently at random with probability $1-p$*, where* $p = d^{-\beta}$*, such that with probability at least* $1 - \frac{1}{N^{\Omega(\sqrt{d})}}$ *the following is true:*

$$\sigma_R(D(\mathbf{x})) = D_1(\mathbf{x}) + D_2(\mathbf{x}),$$

where $D_1(\mathbf{x})$ *is a homogeneous* $\Sigma\Pi\Sigma\Pi^{[2\sqrt{d}]}$ *formula having top fanin same as that of* $D(\mathbf{x})$*, and* $\mathsf{DPSP}_{k,\ell}(D_2(\mathbf{x})) = 0$ *for any choice of* k *and* ℓ*.*

Before proving this, let us see why it implies the required upper bound of lemma 9.

Proof of lemma 9. Using the decomposition lemma 11, with probability at least $1 - \frac{1}{N^{\Omega(\sqrt{d})}}$ we have:

$$\mathsf{DPSP}_{k,\ell}(\sigma_R(D(\mathbf{x}))) \leq \mathsf{DPSP}_{k,\ell}(D_1(\mathbf{x})).$$

Let $t = \sqrt{d}$ and k, ℓ be arbitary integers satisfying $\ell + 2kt \leq \frac{N}{2}$. Then the dimension of the projected shifted partials of $D_1(\mathbf{x})$ is upper bounded as in [KLSS14a],

$$\mathsf{DPSP}_{k,\ell}(\sigma_R(D(\mathbf{x}))) \leq s \cdot \binom{\frac{d}{t}+1}{k} \cdot \binom{N}{\ell + 2kt}. \tag{11}$$

This proves lemma 9. □

6.1 Proof of the decomposition lemma.

We will prove lemma 11 here by considering two cases separately: $0 \leq \mu \leq \frac{1}{5}$ and $\frac{1}{5} < \mu < 1$. Let $t = \sqrt{d}$.

Case 1. Suppose $0 \leq \mu \leq \frac{1}{5}$. In this case the analysis is similar to the one outlined in Section 2. Let Q_{ijr} be a product of linear forms as in equation (10) and $\deg(Q_{ijr}) > 2t$. Then Q_{ijr} can be expressed as $Q_{ijr} = \tilde{Q}_{ijr} \cdot P_{ijr}$ such that $\deg(\tilde{Q}_{ijr}) = 2t$, by simply multiplying out $2t$ linear forms in Q_{ijr}. Since the support of every linear form in Q_{ijr} is bounded by $\tau = N^{\mu}$, the number of monomials in $\tilde{Q}_{ijr}$ is bounded by $\tau^{2t} = (N^{\mu})^{2t}$. The monomials of $\tilde{Q}_{ijr}$ are of two types - those with individual degree of every variable bounded by 2 (and hence has support at least t), and those with at least one variable of degree 3 or more.

Let R be a set formed by picking every variable independently at random with probability $1-p$, where $p = d^{-\beta}$ for an appropriate choice of β (to be fixed shortly). The probability that any monomial of support at least t in $\tilde{Q}_{ijr}$ survives under the random restriction σ_R is bounded by $p^t \cdot (N^{\mu})^{2t}$. Running over all Q_{ijr} in equation (10), with probability at least $1 - s \cdot p^t \cdot (N^{\mu})^{2t}$,

$$\sigma_R(D(\mathbf{x})) = \sum_i \prod_j \sum_{\substack{r \\ \deg(Q_{ijr}) \leq 2t}} \sigma_R(Q_{ijr}) + P,$$

where every monomial in P has a variable of degree 3 or more. Naturally, $\mathsf{DPSP}_{k,\ell}(P) = 0$ for any choice of k and ℓ. Since $s \leq N^{\frac{0.03}{2+\alpha} \cdot \sqrt{d}}$, $p = d^{-\beta}$, $\alpha = \frac{2\mu+1}{1-\mu}$ and $t = \sqrt{d}$, the "bad" probability is

$$\begin{aligned} s \cdot p^t \cdot (N^{\mu})^{2t} &\leq (N^{\frac{0.03}{2+\alpha}} \cdot d^{-\beta} \cdot N^{2\mu})^t \\ &\leq (N^{\frac{0.03}{2+\alpha}} \cdot N^{-\frac{\beta}{2+\alpha}} \cdot 2^{\frac{\beta}{2+\alpha}} \cdot N^{2\mu})^t, \qquad \text{as } \left(\frac{N}{2}\right)^{\frac{1}{2+\alpha}} \leq d \leq N^{\frac{1}{2+\alpha}} \end{aligned}$$

The above quantity is at most $\frac{1}{N^{\Omega(\sqrt{d})}}$ if

1. $2\mu + \frac{0.03}{2+\alpha} < \frac{\beta}{2+\alpha}$, and
2. $0 < \beta < \alpha$.

It is easy to verify that these two conditions are satisfied if $\beta = \frac{6.5\mu+0.03}{1-\mu}$ and considering $\mu \leq \frac{1}{5}$.

Case 2. Suppose $\frac{1}{5} < \mu < 1$. In this case we apply the random restriction in two phases.

Phase 1: Pick each variable independently at random with probability $1-p_1$, where $p_1 = d^{-\beta_1}$, and form a set R_1. (β_1 will be fixed shortly.) Let g be a linear form in a product Q_{ijr}. Assume without loss of generality that the support of g is exactly $\tau = N^\mu$ (if not, simply fill in g with variables having zero coefficients). Then, the expected value of the support size of $\sigma_{R_1}(g)$ is

$$\gamma := \mathcal{E}[\text{support size of } g] = d^{-\beta_1} \cdot N^\mu.$$

By Chernoff bound,

$$\Pr\{\text{bottom fanin of } \sigma_{R_1}(D(\mathbf{x})) \geq (1+\sqrt{3}) \cdot \gamma\} \leq s \cdot e^{-\gamma}.$$

One can verify that the above probability is less than $\frac{1}{N^{\Omega(\sqrt{d})}}$ if

$$\mu \cdot (2+\alpha) > \beta_1 + \frac{1}{2} \tag{12}$$

We will set β_1 shortly to satisfy the above condition.

Phase 2: Pick each variable independently at random (and independent of Phase 1) with probability $1-p_2$, where $p_2 = d^{-\beta_2}$, and form a set R_2. (β_2 will be set to an appropriate value shortly.) We wish to study the formula $\sigma_{R_2}(\sigma_{R_1}(D(\mathbf{x}))) = \sigma_{R_1 \cup R_2}(D(\mathbf{x}))$.

If we set β_1 satisfying equation (12) then with high probability the bottom fanin of $\sigma_{R_1}(D(\mathbf{x}))$ is less than $(1+\sqrt{3}) \cdot \gamma$ — assume that this happens after Phase 1. The argument from here on is similar to that in Case 1. Let

$$\sigma_{R_1}(D(\mathbf{x})) = \sum_i \prod_j \sum_r Q'_{ijr},$$

where each linear form in every Q'_{ijr} has support size bounded by $(1+\sqrt{3}) \cdot \gamma$. If $\deg(Q'_{ijr}) \geq 2t$ then $Q'_{ijr} = \tilde{Q}'_{ijr} \cdot P'_{ijr}$ where $\deg(\tilde{Q}'_{ijr}) = 2t$ and number of monomial in $\tilde{Q}'_{ijr}$ is bounded by $(1+\sqrt{3})^{2t} \cdot \gamma^{2t}$. Once again, focus on those monomials in $\tilde{Q}'_{ijr}$ that have support at least t. (Each of the remaining monomials in $\tilde{Q}'_{ijr}$ has a variable of degree 3 or more.) The probability that any of those monomials in $\tilde{Q}'_{ijr}$ survives after the random restriction σ_{R_2} is applied is bounded by $p_2^t \cdot (1+\sqrt{3})^{2t} \cdot \gamma^{2t}$. Hence with probability at least $1 - s \cdot p_2^t \cdot (1+\sqrt{3})^{2t} \cdot \gamma^{2t}$,

$$\sigma_{R_1 \cup R_2}(D(\mathbf{x})) = \sigma_{R_2}(\sigma_{R_1}(D(\mathbf{x}))) = \sum_i \prod_j \sum_{r \atop \deg(Q'_{ijr}) \leq 2t} \sigma_{R_2}(Q'_{ijr}) + P',$$

where $\mathsf{DPSP}_{k,\ell}(P') = 0$ for any k, ℓ. Let us calculate the bad probability a bit more closely.

$$\begin{aligned} s \cdot p_2^t \cdot (1+\sqrt{3})^{2t} \cdot \gamma^{2t} &\leq [N^{\frac{0.03}{2+\alpha}} \cdot p_2 \cdot (1+\sqrt{3})^2 \cdot \gamma^2]^t \\ &= [N^{\frac{0.03}{2+\alpha}} \cdot d^{-\beta_2} \cdot (1+\sqrt{3})^2 \cdot d^{-2\beta_1} \cdot N^{2\mu}]^t. \end{aligned}$$

The above quantity is less than $\frac{1}{N^{\Omega(\sqrt{d})}}$ if

$$2\mu \cdot (2+\alpha) + 0.03 < \beta_2 + 2\beta_1, \qquad \text{and} \tag{13}$$

$$\beta_1 + \beta_2 < \alpha \quad \& \quad \beta_1, \beta_2 > 0 \tag{14}$$

The requirement stated in equation (14) comes from Lemma 10, as Phase 1 and 2 together amounts to setting each variable zero independently with probability $1 - p_1 p_2 = 1 - d^{-(\beta_1+\beta_2)}$. It is easy to verify that the conditions stated by equations (12), (13) and (14) are satisfied by choosing

$$\begin{aligned} \beta_1 &= \mu \cdot (2+\alpha) - 0.51 \\ \beta_2 &= 1.06, \end{aligned}$$

and keeping in mind that $\mu > \frac{1}{5}$. This completes the proof of the decomposition lemma.

7 Summary and discussion

A recent line of research on arithmetic circuit lower bounds uses the dimension of the space of shifted partials and its variant the projected shifted partials under random restriction as a complexity measure to make progress on proving lower bounds for certain interesting classes of arithmetic circuits, namely regular formulas and homogeneous depth four formulas. (The dimension of the space of shifted partials measure is in turn based on the classical measure of the dimension of the space of partial derivatives.) The formal degree of a homogeneous depth four formula (or a regular formula) is bounded by the degree (or the order of the degree) of the polynomial that it computes. At this point it was not clear if the present techniques are applicable to models where the formal degree is much higher than the degree of the computed polynomial. One very interesting (and arguably the simplest nontrivial) example of such an unrestricted formal degree model is (nonhomogeneous) depth three circuits over fields of characteresic zero - its power being exhibited by the recent work of [GKKS13a].

Our work takes a step forward in this direction by showing an exponential lower bound for (nonhomogeneous) depth three circuits with small bottom fanin over fields of characteristic zero. Along the way we also show an exponential lower bound for homogeneous depth five formulas with small bottom fanin. The second result is for an explicit polynomial in VNP. An immediate question is whether the combinatorial argument from [KS14a] can be suitably adapted so that the lower bound of theorem 2 holds for iterated matrix multiplication as well. Both these results are obtained by building upon the current techniques on shifted patials based measures. It would be very interesting to prove analogous lower bounds for less restrictive subclasses of arithmetic circuits.

- Can we drop the restriction of 'small bottom fanin' from both the models - (nonhomogeneous) depth three circuits and homogeneous depth five circuits - and still show an exponential lower bound?

A few other intriguing problems on arithmetic circuit lower bounds are worth mentioning here:

- Show a super-polynomial lower bound for homogeneous bounded depth arithmetic circuits.
- Show a super-polynomial lower bound for homogeneous arithmetic formulas.

- Show a super-polynomial separation between homogeneous product-depth-Δ formulas and homogeneous product-depth-$(\Delta - 1)$ formulas.
- Solve the above problems without the assumption of homogeneity.

Solutions to these problems, using present or new techniques, would give a significant boost to our understanding of arithmetic circuit lower bounds.

Acknowledgements

The authors would like to thank Amit Chakrabarti, Mrinal Kumar, Satya Lokam and Ramprasad Saptharishi for helpful discussions. In particular, Ramprasad pointed out to us that a lemma in [GKKS13a] can be improved quantitatively and that the $\Sigma\Pi\Sigma$ circuits which come out of the depth reduction in [GKKS13a] in fact have small bottom fanin.

References

[Alo09] Noga Alon. Perturbed Identity Matrices Have High Rank: Proof and Applications. *Combinatorics, Probability & Computing*, 18(1-2):3–15, 2009.

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

[Erd32] Paul Erdös. Beweis eines Satzes von Tschebyschef. *Acta Sci. Math. (Szeged)*, 5:194–198, 1930-1932.

[FLMS14] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *STOC*, pages 128–135, 2014.

[GK98] Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *STOC*, pages 577–582, 1998.

[GKKS13a] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Foundations of Computer Science (FOCS)*, pages 578–587, 2013.

[GKKS13b] Ankit Gupta, Neeraj Kayal, Pritish Kamath, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, 2013.

[GR00] Dima Grigoriev and Alexander A. Razborov. Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000.

[HR18] G. H. Hardy and S. Ramanujan. Asymptotic formula in combinatory analysis. *Proceedings of the London Mathematical Society*, s2-17(1):75–115, 1918.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. Technical report, Electronic Colloquium on Computational Complexity (ECCC), 2012.

[KLSS14a] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. *To appear in FOCS*, 2014.

[KLSS14b] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas. In *STOC*, pages 119–127, 2014.

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012.

[KS14a] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. *To appear in FOCS*, 2014.

[KS14b] Mrinal Kumar and Shubhangi Saraf. Superpolynomial lower bounds for general homogeneous depth 4 arithmetic circuits. In *ICALP (1)*, pages 751–762, 2014.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *STOC*, pages 146–153, 2014.

[Lit50] D.E. Littlewood. *The Theory of Group Characters and Matrix Representations of Groups*. Ams Chelsea Publishing. AMS Chelsea Pub., 2nd edition, 1950.

[NW97] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. Available at http://www.math.ias.edu/~avi/PUBLICATIONS/MYPAPERS/NW96/final.pdf.

[Rys63] H. J. Ryser. Combinatorial mathematics. *Math. Assoc. of America*, 14, 1963.

[Sap14] Ramprasad Saptharishi. Personal communication, 2014.

[Shp01] Amir Shpilka. *Lower Bounds for Small Depth Arithmetic and Boolean Circuits*. PhD thesis, The Hebrew University, 2001.

[SW99] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. In *IEEE Conference on Computational Complexity*, pages 87–, 1999. Available at http://eccc.hpi-web.de/report/1999/023/.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5:207–388, March 2010.

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.

A Proof of Lemma 10

In this section we prove lemma 10, i.e. we show that the dimension of projected shifted partial derivatives of a randomly restricted Nisan-Wigderson design based polynomial is within a 'small' factor of the maximum possible with high probability. Our proof is very similar to the proof of Lemma 13 in [KLSS14a] - in fact, we reuse quite a bit of the argument from there but carefully tune it at places to achieve the required setting of parameters. Proofs of some of the propositions in this section are collected in Section B. Let $e \stackrel{\text{def}}{=} (d-k)$ throughout the rest of this section.

Preliminaries. Note that in the construction in Section 5 of NW_r, there is a 1-1 correspondence between the variable indices in $[N]$ and points in $[d] \times [q]$. Being homogeneous and multilinear of degree d, the monomials of NW_r are in 1-1 correspondence with sets in $\binom{[N]}{d} \equiv \binom{[d]\times[q]}{d}$. Indeed, from the construction it is clear that the coefficient of any monomial in NW_r is either 0 or 1 and that there is a 1-1 correspondence between monomials in the support of NW_r and univariate polynomials of degree at most r in $\mathbb{F}_q[z]$. Now since two distinct polynomials of degree r over a field have at most r common roots we get:

Proposition 12. [A basic property of our construction.] *For any two distinct sets $D_1, D_2 \in \binom{[d]\times[q]}{d}$ in the support of NW_r, we have*

$$|D_1 \cap D_2| \quad \leq \quad r.$$

Let R be a set formed by picking each variable independently at random with probability $1-p$, where $p = d^{-\beta}$ for $0 < \beta < \alpha$. Our goal for the remainder of this section is to lower bound $\mathsf{DPSP}_{k,\ell}(\sigma_R(NW_r))$.

Reformulating our goal in terms of the rank of an explicit matrix. Let f be any homogeneous multilinear polynomial of degree d on N variables. Then we have

$$\boldsymbol{\partial}^{=k}_{\mathrm{ML}} f = \left\{ \partial^C f \ : \ C \in \binom{[N]}{k} \right\}.$$

Note that every k-th order derivative of f is homogeneous and multilinear of degree $(d-k)$. Hence

$$\pi(\mathbf{x}^{=\ell} \cdot \boldsymbol{\partial}^{=k}_{\mathrm{ML}} f) = \left\{ \mathbf{x}_A \cdot \sigma_A \left(\partial^C f\right) \ : \ A \in \binom{[N]}{\ell}, \quad C \in \binom{[N]}{k} \right\}.$$

Thus we have

Proposition 13. *For any homogeneous multilinear polynomial f of degree d on N variables and for all integers k and ℓ:*

$$\mathsf{DPSP}_{k,\ell}(f) = \dim\left(\left\{ \mathbf{x}_A \cdot \sigma_A \left(\partial^C f\right) \ : \ A \in \binom{[N]}{\ell}, \quad C \in \binom{[N]}{k} \right\} \right).$$

Now the $\mathbb{F}$-linear dimension of any set of polynomials is the same as the rank of the matrix corresponding to our set of polynomials in the natural way. In fact, we will focus our attention on a subset of rows of this matrix and prove a lower bound on the rank of the matrix defined by this subset of rows. Specifically,

Proposition 14. *Let f be a homogeneous multilinear polynomial of degree d on N variables. Let k, ℓ be integers. Define a matrix $M(f)$ as follows. The rows of $M(f)$ are labelled by pairs of subsets $(A, C) \in \binom{[N]}{\ell} \times \binom{[N]}{k}$ such that $A \cap C = \Phi$ (null set) and columns are indexed by subsets $S \in \binom{[N]}{\ell+e}$. Each row (A, C) corresponds to the polynomial*

$$f_{A,C} \stackrel{\text{def}}{=} \mathbf{x}_A \cdot \sigma_A \left(\partial^C f\right)$$

in the following way. The S-th entry of the row (A, C) is the coefficient of $\mathbf{x}_S$ in the polynomial $f_{A,C}$. Then,

$$\mathsf{DPSP}_{k,\ell}(f) \geq \mathit{rank}(M(f)).$$

So our problem is equivalent to lower bounding the rank of the matrix $M(f)$ for our constructed polynomial f. Now note that the entries of $M(f)$ are coefficients of appropriate monomials of f and it will be helpful to us in what follows to keep track of this information. We will do it by assigning a label to each cell of $M(f)$ as follows. We will think of every location in the matrix $M(f)$ being labelled with either a set $D \in \binom{[N]}{d}$ or the label `InvalidSet` depending on whether that entry contains the coefficient of the monomial $\mathbf{x}_D$ of f or it would have been zero regardless of the actual coefficients of f. Specifically, let us introduce the following notation. For sets A, B define:

1.
$$A \setminus\setminus B = \begin{cases} A \setminus B & \text{if} B \subseteq A \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

2.
$$A \uplus B = \begin{cases} A \cup B & \text{if} B \cap A = \emptyset \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

Then the label of the $((A, C), S)$-th cell of $M(f)$ is defined to be the set $(S \setminus\setminus A) \uplus C$. Equivalently, if the label of a cell of the (A, C)-th row of M is a set D then the column must be the one corresponding to $S = (D \setminus\setminus C) \uplus A$ (if C is not a subset of D or if D and A are not disjoint then D cannot occur in the row indexed by (A, C)). For the rest of this section, we will refer to $M(\sigma_R(\mathsf{NW}_r))$ simply as the matrix M. Our goal then is to show that the rank of this matrix M is reasonably close to the trivial upper bound, viz. the minimum of the number of rows and the number of columns of M with high probability. It turns out that our matrix M is a relatively sparse matrix and we will exploit this fact by using a relevant lemma from real matrix analysis to obtain a lower bound on its rank.

The Surrogate Rank. Consider the matrix $B \stackrel{\text{def}}{=} M^T \cdot M$. Then B is a real symmetric, positive semidefinite matrix. From the definition of B it is easy to show that:

Proposition 15. *Over any field $\mathbb{F}$ we have*

$$\mathit{rank}(B) \leq \mathit{rank}(M).$$

Over the field $\mathbb{R}$ of real numbers we have

$$\mathit{rank}(B) = \mathit{rank}(M).$$

So it suffices to lower bound the rank of B. By an application of Cauchy-Schwarz on the vector of nonzero eigenvalues of B, one obtains:

Lemma 16. *[Alo09] Over the field of real numbers $\mathbb{R}$ we have:*

$$\mathit{rank}(B) \geq \frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)}.$$

Let us call the quantity $\frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)}$ as the surrogate rank of B, denoted $\mathrm{SurRank}(B)$. It then suffices to show that this quantity is within a 'small' factor of $U = \min(\binom{N}{\ell+e}, \binom{N}{\ell} \cdot \binom{N}{k})$ with high probability. In the rest of this section, we will first derive an exact expression for $\mathrm{SurRank}(B)$ and then show that it is close to U (again, with high probability). In the following discussion we would need an estimate of a quantity $R_d(w, r)$ that denotes the number of univariate polynomials in $\mathbb{F}_q[z]$ of degree at most r having exactly w distinct roots in $[d]$.

An estimate for $R_d(w, r)$. First note that any polynomial $h(z) \in \mathbb{F}_q[z]$ of degree at most r that has w roots in $[d]$ must be of the form

$$h(z) = (z - \alpha_1) \cdot (z - \alpha_2) \cdot \ldots \cdot (z - \alpha_w) \cdot \hat{h}(z),$$

where each α_i is in $[d]$ and $\hat{h}(z) \in \mathbb{F}_q[z]$ is of degree at most $(r-w)$. Thus we have

$$R_d(w,r) \leq q^{r-w+1} \cdot \binom{d}{w} \leq q^{r+1} \cdot \left(\frac{d}{q}\right)^w \cdot \frac{1}{w!} \tag{15}$$

A.1 Deriving an exact expression for $\mathrm{SurRank}(B)$.

We will now calculate an exact expression for $\mathrm{SurRank}(B)$, or equivalently an exact expression for $\mathrm{Tr}(B)$ and $\mathrm{Tr}(B^2)$.

Calculating $\mathrm{Tr}(B)$. Calculating $\mathrm{Tr}(B)$ is fairly straightforward. From the definition of the matrix B we have:

Proposition 17. *For any $0, \pm 1$ matrix M (i.e. a matrix all of whose entries are either 0, or $+1$ or -1) we have*

$$\mathrm{Tr}(B) = \mathrm{Tr}(M^T \cdot M) = \textit{number of nonzero entries in } M.$$

Now we can calculate the number of nonzero entries in M by going over all sets $D \in \binom{[N]}{d} \cap \mathrm{Supp}(\sigma_R(\mathsf{NW}_r))$, calculating the number of cells of M labelled with D and adding these up. Clearly

$$\sigma_R(\mathsf{NW}_r) = \sum_{D \in \mathrm{Supp}(\mathsf{NW}_r)} e_D \cdot \mathbf{x}_D,$$

where e_D is an indicator variable such that $e_D = 1$ if $\sigma_R(\mathbf{x}_D) \neq 0$, and $e_D = 0$ otherwise. Hereafter, we will refer to $\sigma_R(\mathsf{NW}_r)$ as g at some places, and the number of monomials in $\sigma_R(\mathsf{NW}_r)$ as $\mu(g)$.

$$\begin{aligned} \mu(g) &= \sum_{D \in \mathrm{Supp}(\mathsf{NW}_r)} e_D \\ \Rightarrow \mathcal{E}\left[\mu(g)\right] &= p^d \cdot q^{r+1} = \gamma \text{ (say)} \\ \Rightarrow \mathcal{E}\left[\mathrm{Tr}(B)\right] &= \gamma \cdot \binom{d}{k} \cdot \binom{N-d}{\ell}. \end{aligned}$$

Proposition 18. $\Pr\left[\mathrm{Tr}(B) \leq \frac{1}{2} \cdot \gamma \cdot \binom{d}{k} \cdot \binom{N-d}{\ell}\right] \leq \frac{10}{pd^\alpha}$. *(Proof in Section B)*

Calculating $\mathrm{Tr}(B^2)$. From the definition of $B = M^T \cdot M$ and expanding out the relevant summations we get:

Proposition 19.

$$\mathrm{Tr}(B^2) = \sum_{(A_1,C_1),(A_2,C_2) \in \left(\binom{[N]}{\ell} \times \binom{[N]}{k}\right)^2} \; \sum_{S_1,S_2 \in \binom{[N]}{\ell+e}^2} M_{(A_1,C_1),S_1} \cdot M_{(A_1,C_1),S_2} \cdot M_{(A_2,C_2),S_1} \cdot M_{(A_2,C_2),S_2}.$$

We will use the following notation in doing this calculation. For a pair of row indices $((A_1,C_1),(A_2,C_2)) \in \left(\binom{[N]}{\ell} \times \binom{[N]}{k}\right)^2$ and a pair of column indices $S_1, S_2 \in \left(\binom{[N]}{\ell+e}\right)^2$, the box $\mathbf{b}$ defined by them, denoted $\mathbf{b} = 2-\mathrm{box}((A_1,C_1),(A_2,C_2),S_1,S_2)$ is the four-tuple of cells

$$(((A_1,C_1),S_1),((A_1,C_1),S_2),((A_2,C_2),S_1),((A_2,C_2),S_2)).$$

Since all the entries of our matrix M are either 0 or 1 we have:

Proposition 20.

$$\mathrm{Tr}(B^2) = \textit{Number of boxes } \mathbf{b} \textit{ with all four entries nonzero.}$$

For a box $\mathbf{b} = 2-\mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$, its tuple of labels, denoted $\mathrm{labels}(\mathbf{b})$ is the tuple of labels of the cells $((A_1, C_1), S_1), ((A_1, C_1), S_2), ((A_2, C_2), S_1), ((A_2, C_2), S_2))$ in that order. In other words,

$$\mathrm{labels}(\mathbf{b}) = ((S_1 \setminus\setminus A_1) \uplus C_1, (S_2 \setminus\setminus A_1) \uplus C_1, (S_1 \setminus\setminus A_2) \uplus C_2, (S_2 \setminus\setminus A_2) \uplus C_2).$$

We then have

Proposition 21. $\mathrm{Tr}(B^2)$ *equals the number of boxes*

$$\mathbf{b} = 2 - \textit{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

such that all the four labels in $\mathrm{labels}(\mathbf{b})$ *are valid sets in the support of our design polynomial* $\sigma_R(\mathsf{NW}_r)$.

So our problem boils down to counting the number of boxes in which all the four labels are valid sets in the support of our polynomial $\sigma_R(\mathsf{NW}_r)$. Let us analyze the box

$$\mathbf{b} = 2 - \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

a bit closely. Suppose $\mathrm{labels}(\mathbf{b}) = (D_1, D_2, D_3, D_4)$ as shown in the table below where D_1, D_2, D_3, D_4 are valid sets in $\binom{[N]}{d}$.

	S_1	S_2
(A_1,C_1)	D_1	D_2
(A_2,C_2)	D_3	D_4

Define the following sets:

$$\begin{array}{rcll} E_1 & := & A_1 \backslash (A_1 \cap A_2) & E_2 := A_2 \backslash (A_1 \cap A_2) \\ E_3 & := & C_1 & E_4 := C_2 \\ E_5 & := & D_1 \backslash (E_2 \uplus E_3) & E_6 := D_2 \backslash (E_2 \uplus E_3) \\ & = & D_3 \backslash (E_1 \uplus E_4) & = D_4 \backslash (E_1 \uplus E_4) \end{array}$$

Note that $E_2 \uplus E_3$ must be a subset of both D_1 and D_2, similarly $E_1 \uplus E_4$ must be a subset of both D_3 and D_4. Also, $D_1 \backslash (E_2 \uplus E_3) = D_3 \backslash (E_1 \uplus E_4)$ as $(D_1 \setminus\setminus C_1) \uplus A_1 = (D_3 \setminus\setminus C_2) \uplus A_2 = S_1$. Similarly, $D_2 \backslash (E_2 \uplus E_3) = D_4 \backslash (E_1 \uplus E_4)$. Verify that D_1, D_2, D_3 and D_4 can be expressed as:

$$\begin{array}{rcll} D_1 & = & E_2 \uplus E_5 \uplus E_3 & D_2 = E_2 \uplus E_6 \uplus E_3 \\ D_3 & = & E_1 \uplus E_5 \uplus E_4 & D_4 = E_1 \uplus E_6 \uplus E_4 \end{array} \tag{16}$$

From the above definitions, if $|A_1 \cap A_2| = v$ then

$$\begin{array}{rcl} |E_1| & = & |E_2| = \ell - v \\ |E_3| & = & |E_4| = k \\ |E_5| & = & |E_6| = d - (\ell - v + k) \end{array} \tag{17}$$

Proposition 22. *Unless D_1, D_2, D_3, D_4 are all distinct sets,* labels($\mathbf{b}$) *contains at most two distinct sets. Furthermore, if D_1, D_2, D_3 are distinct then $\ell - v + k \leq r$ and $d - (\ell - v + k) \leq r$.*

Proof. We show that if D_1 equals any of D_2, D_3 or D_4 then labels($\mathbf{b}$) has at most two distinct sets. The argument is similar for other cases. Suppose $D_1 = D_2$ then by Equation 16 $E_5 = E_6$, implying $D_3 = D_4$. If $D_1 = D_3$ then again by Equation 16, $E_2 \uplus E_3 = E_1 \uplus E_4$ implying $D_2 = D_4$. Now suppose $D_1 = D_4$, then by Equation 16, $E_6 \subseteq D_1$. But $E_6 \subseteq D_2$, which means $D_2 \subseteq D_1$ as $E_2 \uplus E_3 \subseteq D_1$. Since $|D_2| = |D_1| = d$, $D_1 = D_2$ and hence $D_1 = D_2 = D_3 = D_4$.

To prove the second statement of the lemma, observe that $|D_1 \cap D_2| \geq |E_2 \uplus E_3| = \ell - v + k$. So, if $\ell - v + k \geq r + 1$ then $D_1 = D_2$. Similarly, $|D_1 \cap D_3| \geq |E_5| = d - (\ell - v + k)$. If $d - (\ell - v + k) \geq r + 1$ then $D_1 = D_3$. □

This means that any box $\mathbf{b}$ that contributes to $\mathrm{Tr}(B^2)$ must have the property that its label set labels($\mathbf{b}$) contains at most two distinct sets in the support of $\sigma_R(\mathsf{NW}_r)$, or four distinct sets in the support of $\sigma_R(\mathsf{NW}_r)$. A set D is in the support of $\sigma_R(\mathsf{NW}_r)$ if D is in the support of NW_r and $\sigma_R(\mathbf{x}_D) \neq 0$. (Recall that e_D is an indicator variable which is 1 if $\sigma_R(\mathbf{x}_D) \neq 0$, and zero otherwise.)

Corollary 23. *For any four distinct sets $D_1, D_2, D_3, D_4 \in \binom{[N]}{d}$ define*

$$\begin{aligned}
\mu_0(D_1) &\stackrel{\text{def}}{=} \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_1, D_1, D_1)\} \\
\mu_1(D_1, D_2) &\stackrel{\text{def}}{=} \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_2, D_1, D_2)\} \\
\mu_2(D_1, D_2) &\stackrel{\text{def}}{=} \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_1, D_2, D_2)\} \\
\mu_3(D_1, D_2, D_3, D_4) &\stackrel{\text{def}}{=} \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_2, D_3, D_4)\}
\end{aligned}$$

Let the support of NW_r, denoted $\mathrm{Supp}(\mathsf{NW}_r) \subset \binom{[N]}{d}$*, be the set of all sets $D \in \binom{[N]}{d}$ such that the coefficient of the monomial $\mathbf{x}_D$ in NW_r is nonzero. Define T_0, T_1, T_2, T_3 as follows:*

$$\begin{aligned}
T_0 &= \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot |\mu_0(D_1)| \\
T_1 &= \sum_{D_1 \neq D_2 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot e_{D_2} \cdot |\mu_1(D_1, D_2)| \\
T_2 &= \sum_{D_1 \neq D_2 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot e_{D_2} \cdot |\mu_2(D_1, D_2)| \\
T_3 &= \sum_{D_1 \neq D_2 \neq D_3 \neq D_4 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot e_{D_2} \cdot e_{D_3} \cdot e_{D_4} \cdot |\mu_3(D_1, D_2, D_3, D_4)| \qquad (18)
\end{aligned}$$

Then

$$\mathrm{Tr}(B^2) = T_0 + T_1 + T_2 + T_3.$$

We are using the notation $D_1 \neq D_2 \neq D_3 \neq D_4$ to mean that the four sets are distinct. The proof of Proposition 22 rules out the existence of any box $\mathbf{b}$ having labels($\mathbf{b}$) $= (D_1, D_2, D_2, D_1)$ with distinct $D_1, D_2 \in \mathrm{Supp}(\mathsf{NW}_r)$ and that is why there is no term in $\mathrm{Tr}(B^2)$ corresponding to such boxes.

Proposition 18 shows that $\mathrm{Tr}(B)$ is large with high probability. In order to lower bound $\frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)}$, we will show that $\mathrm{Tr}(B^2)$ is less than an upper bound with high probability. This is achieved by upper bounding the expected values of T_0, T_1, T_2 and T_3 and then applying Markov's inequality.

A.2 Upper bound for $\mathcal{E}[T_3]$

Let $\rho(D_1, D_2, D_3)$ be the number of pairs of rows $((A_1, C_1), (A_2, C_2))$ in which D_1, D_2, D_3 (all distinct) can possibly occur as labels (as depicted in the table before). For a fixed D_1, D_2, D_3 we upper bound $\rho(D_1, D_2, D_3)$ with the help of Equation 16. Notice that for a fixed D_1, D_2, D_3, if we specify E_2, E_3, E_4 and $A_1 \cap A_2$ then the sets A_1, C_1, A_2, C_2 are determined. Let us count the number of ways we can pick E_2, E_3, E_4 and $A_1 \cap A_2$ for a given D_1, D_2, D_3. Taking the size bounds on the sets into account from Equation 17, this quantity is upper bounded by,

$$\binom{d}{\ell - v} \cdot \binom{d - (\ell - v)}{k} \cdot \binom{\ell - v + k}{k} \cdot \binom{N - d}{v}.$$

The quantity $\binom{N-d}{v}$ is an upper bound on the number of ways we can pick $A_1 \cap A_2$ as A_1 must be disjoint from D_1. By Proposition 22, $\ell - v + k \leq r < d$, (also, $v \leq \ell < \frac{N-d}{2}$) implying

$$\rho(D_1, D_2, D_3) \leq 2^d \cdot \binom{d}{k}^2 \cdot \binom{N - d}{\ell} = \rho \quad \text{(say)}. \tag{19}$$

Hence,

$$T_3 \leq \rho \cdot \sum_{D_1 \neq D_2 \neq D_3 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot e_{D_2} \cdot e_{D_3} \tag{20}$$

Now we upper bound the expected value of the quantity $\sum_{D_1 \neq D_2 \neq D_3 \in \mathrm{Supp}(\mathsf{NW}_r)} e_{D_1} \cdot e_{D_2} \cdot e_{D_3} = \eta$ (say) in the following proposition.

Proposition 24. *$\mathcal{E}[\eta] \leq 4 \cdot \gamma^2 \cdot q^{(r+1)} \cdot \left(\frac{d}{q}\right)^d$, where γ is as in Proposition 18. This implies*

$$\mathcal{E}[T_3] \leq 4 \cdot \left(\frac{2}{d^{\frac{\alpha-\beta}{2}}}\right)^d \cdot \gamma^2 \cdot \binom{d}{k}^2 \cdot \binom{N - d}{\ell}.$$

Proof of the above proposition can be found in Section B. We show in the later sections that $\mathcal{E}[T_3]$ is negligible compared to $\mathcal{E}[T_0 + T_1 + T_2]$ and hence does not contribute much to the expected value of $\mathrm{Tr}(B^2)$.

In what follows we will derive expressions for $|\mu_0(D_1)|, |\mu_1(D_1, D_2)|$ and $|\mu_2(D_1, D_2)|$ and compute expected values of T_0, T_1 and T_2 by summing these up over $D_1, D_2 \in \mathrm{Supp}(\sigma_R(\mathsf{NW}_r))$. We first observe:

Proposition 25. *For any set $D_1 \in \binom{[N]}{d}$ and any row (A, C) of M, there can be at most one cell in that row labelled with the set D_1.*

This means that any box $\mathbf{b} = 2 - \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$ contributing to either $\mu_0(D_1)$ or $\mu_2(D_1, D_2)$, the columns S_1 and S_2 must be the same.

A.3 Calculating $\mu_0(D_1)$ and $\mathcal{E}[T_0]$.

Every box $\mathbf{b} \in \mu_0(D_1)$ is of the form $\mathbf{b} = 2 - \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_1)$ where both the entries $((A_1, C_1), S_1)$ and $((A_2, C_2), S_1)$ are both labelled by D_1. This implies $A_1 = A_2$ and $C_1 = C_2$: By Equation 16, $E_1 \subseteq D_3 = D_1$, but A_1 is disjoint from D_1 and $E_1 \subseteq A_1$. Hence, E_1 is an empty set and similarly E_2 is also an empty set. This also implies $E_3 = E_4$ from Equation 16 as $D_3 = D_1$. Analyzing this situation gives

Proposition 26.

$$|\mu_0(D_1)| = \binom{N-d}{\ell} \cdot \binom{d}{k} \quad \text{and} \quad \mathcal{E}[T_0] = \gamma \cdot \binom{N-d}{\ell} \cdot \binom{d}{k}$$

Proof. For a fixed D_1, we can choose C_1 in $\binom{d}{k}$ ways and A_1 in $\binom{N-d}{\ell}$ ways. (Recall A_1 must be disjoint from D_1.) The expression for $\mathcal{E}[T_0]$ follows immediately from Equation 18. □

A.4 Calculating $\mu_1(D_1, D_2)$ and $\mathcal{E}[T_1]$.

Let $D_1, D_2 \in \binom{[N]}{d}$ be two distinct subsets in the support of NW_r. We consider a box $\mathbf{b} = 2 - \text{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$ in $\mu_1(D_1, D_2)$. Observe that even in this case it must be that $A_1 = A_2$ and $C_1 = C_2$: By the same reason as before since D_3 equals D_1 in Equation 16. Analyzing this situation gives

Proposition 27. *If* $|D_1 \cap D_2| = w$ *then*

$$|\mu_1(D_1, D_2)| = \binom{N-2d+w}{\ell} \cdot \binom{w}{k} \quad \text{and hence} \quad \mathcal{E}[T_1] \leq d \cdot \frac{\gamma^2}{d^{(\alpha-\beta)k} \cdot k!} \cdot \binom{N-2d+k}{\ell}.$$

Proof of the above proposition is given in Section B.

A.5 Calculating $\mu_2(D_1, D_2)$ and $\mathcal{E}[T_2]$.

Let $D_1, D_2 \in \binom{[N]}{d}$ be two distinct subsets in the support of NW_r. We consider a box $\mathbf{b} = 2 - \text{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$ in $\mu_2(D_1, D_2)$. As we observed before this can happen only if $S_1 = S_2 = S$ (say). Let $|C_1 \cap C_2| = u$. Analyzing this situation gives

Proposition 28. *If* $|D_1 \cap D_2| = w$ *then*

$$\begin{aligned} |\mu_2(D_1, D_2)| &= \sum_{0 \leq u \leq k} \binom{N-2d+w}{\ell-d+k+w-u} \cdot \binom{d-w}{k-u} \cdot \binom{d-w}{k-u} \cdot \binom{w}{u}, \text{ and hence} \\ \mathcal{E}[T_2] &\leq dk \cdot \gamma^2 \cdot \binom{N-2d}{\ell-d+k} \cdot \binom{d}{k}^2. \end{aligned}$$

Proof. The expectation calculation is similar to the one in the proof of Proposition 27 - the maxima of the relevant expression is touched at $w = u = 0$. □

A.6 Lower bound on $\text{SurRank}(B)$

A comparison between the binomial coefficients $\binom{N-2d}{\ell-d+k}$ and $\binom{N-d}{\ell}$ shows that

$$\binom{N-2d}{\ell-d+k} \geq \frac{1}{3^d} \cdot \binom{N-d}{\ell}.$$

Thus, from Proposition 26, 28 and 24, the upper bound on $\mathcal{E}[T_2]$ dominates the upper bounds on $\mathcal{E}[T_0]$ and $\mathcal{E}[T_3]$. Applying Markov's inequality,

$$\text{Tr}(B^2) \leq d^2 \cdot \frac{\gamma^2}{d^{(\alpha-\beta)k} \cdot k!} \cdot \binom{N-2d+k}{\ell} + 3d^2k \cdot \gamma^2 \cdot \binom{N-2d}{\ell-d+k} \cdot \binom{d}{k}^2$$

with probability at least $1 - \frac{1}{d}$. Coupled with Proposition 18,

$$\text{SurRank}(B) \geq \min\left(\frac{\frac{1}{4}\cdot\gamma^2\cdot\binom{d}{k}^2\cdot\binom{N-d}{\ell}^2}{2d^2\cdot\frac{\gamma^2}{d^{(\alpha-\beta)k}\cdot k!}\cdot\binom{N-2d+k}{\ell}}, \frac{\frac{1}{4}\cdot\gamma^2\cdot\binom{d}{k}^2\cdot\binom{N-d}{\ell}^2}{6d^2k\cdot\gamma^2\cdot\binom{N-2d}{\ell-d+k}\cdot\binom{d}{k}^2}\right),$$

with probability at least $1 - \frac{1}{d^{\Omega(1)}}$. The first ratio is at least $\frac{p^k}{d^{O(1)}}\cdot\frac{1}{4^k}\cdot\binom{N}{k}\cdot\binom{N}{\ell}$ as

$$\frac{\binom{N-d}{\ell}^2}{\binom{N-2d+k}{\ell}} \geq \frac{1}{2^k d^{O(1)}}\cdot\binom{N}{\ell} \text{ and } d^{\alpha k}\cdot k!\cdot\binom{d}{k}^2 \geq \frac{1}{2^k d^{O(1)}}\cdot\binom{N}{k}.$$

The second ratio is at least $\frac{1}{d^{O(1)}}\cdot\binom{N}{\ell+d-k}$ as,

$$\frac{\binom{N-d}{\ell}^2}{\binom{N-2d}{\ell-d+k}} \geq \frac{1}{d^{O(1)}}\cdot\binom{N}{\ell+d-k}.$$

Therefore,

$$\text{SurRank}(B) \geq \frac{1}{d^{O(1)}}\min\left(\frac{p^k}{4^k}\cdot\binom{N}{k}\cdot\binom{N}{\ell}, \binom{N}{\ell+d-k}\right).$$

B Proofs of certain propositions

Proposition 18. $\Pr\left[\text{Tr}(B) \leq \frac{1}{2}\cdot\gamma\cdot\binom{d}{k}\cdot\binom{N-d}{\ell}\right] \leq \frac{10}{pd^\alpha}$.

Proof. As in Proposition 17, $\text{Tr}(B) = \text{Tr}(M^T\cdot M) =$ number of nonzero entries in M.

$$\begin{aligned}\text{Tr}(B) &= \mu(g)\cdot\binom{d}{k}\cdot\binom{N-d}{\ell}\\ \Rightarrow \mathcal{E}\left[\text{Tr}(B)\right] &= \gamma\cdot\binom{d}{k}\cdot\binom{N-d}{\ell}\end{aligned}$$

Hence,

$$\Pr\left[\text{Tr}(B) \leq \frac{1}{2}\cdot\gamma\cdot\binom{d}{k}\cdot\binom{N-d}{\ell}\right] = \Pr\left[\mu(g) \leq \frac{1}{2}\cdot\gamma\right].$$

It turns out that the variance of $\mu(g)$, denoted by $\mathsf{Var}(\mu(g))$, can be upper bounded as follows.

$$\begin{aligned}\mathsf{Var}(\mu(g)) &\leq \gamma\cdot(1-p^d) + \gamma^2\cdot\frac{2}{pd^\alpha}\\ \Rightarrow \Pr\left[\mu(g) \leq \frac{1}{2}\cdot\gamma\right] &\leq \frac{10}{pd^\alpha} \quad \text{(by Chebyshev's inequality)}\end{aligned}$$

The last inequality also uses the fact that $\gamma > 2pd^\alpha$ which is true since $r = \frac{\alpha+\beta}{2(1+\alpha)}\cdot d - 1$ and hence $\gamma = d^{\Omega(d)}$. Now, let us bound the variance of $\mu(g)$. In the summations below, D, D_1, D_2 run over

all elements in Supp(NW_r).

$$
\begin{aligned}
\mathsf{Var}(\mu(g)) &= \mathcal{E}[\mu(g)^2] - \mathcal{E}[\mu(g)]^2 \\
&= \mathcal{E}\left[\left(\sum_D e_D\right)^2\right] - \mathcal{E}\left[\sum_D e_D\right]^2 \\
&= \mathcal{E}\left[\sum_D e_D^2 + \sum_{\substack{D_1,D_2 \\ D_1 \neq D_2}} e_{D_1} \cdot e_{D_2}\right] - \left[\sum_D \mathcal{E}[e_D]\right]^2 \quad \text{(by linearity of expectation)} \\
&= \mathcal{E}\left[\sum_D e_D + \sum_{\substack{D_1,D_2 \\ D_1 \neq D_2}} e_{D_1} \cdot e_{D_2}\right] - \left[\sum_D \mathcal{E}[e_D]^2 + \sum_{\substack{D_1,D_2 \\ D_1 \neq D_2}} \mathcal{E}[e_{D_1}] \cdot \mathcal{E}[e_{D_2}]\right] \quad (\text{as } e_D^2 = e_D) \\
&= \mathcal{E}\left[\sum_D e_D\right] - \sum_D \mathcal{E}[e_D]^2 + \mathcal{E}\left[\sum_{\substack{D_1,D_2 \\ D_1 \neq D_2}} e_{D_1} \cdot e_{D_2}\right] - \sum_{\substack{D_1,D_2 \\ D_1 \neq D_2}} \mathcal{E}[e_{D_1}] \cdot \mathcal{E}[e_{D_2}] \\
&= p^d \cdot q^{r+1} - p^{2d} \cdot q^{r+1} + \sum_{w=0}^{r} \left[\mathcal{E}\left[\sum_{\substack{D_1,D_2 \\ D_1 \neq D_2, |D_1 \cap D_2| = w}} e_{D_1} \cdot e_{D_2}\right] - \sum_{\substack{D_1,D_2 \\ D_1 \neq D_2, |D_1 \cap D_2| = w}} \mathcal{E}[e_{D_1}] \cdot \mathcal{E}[e_{D_2}]\right] \\
&= \gamma \cdot (1 - p^d) + \sum_{w=0}^{r} \left[\sum_{\substack{D_1,D_2 \\ D_1 \neq D_2, |D_1 \cap D_2| = w}} (\mathcal{E}[e_{D_1} \cdot e_{D_2}] - \mathcal{E}[e_{D_1}] \cdot \mathcal{E}[e_{D_2}])\right] \\
&\qquad \text{(by linearity of expectation)} \\
&= \gamma \cdot (1 - p^d) + \sum_{w=0}^{r} \left[\sum_{\substack{D_1,D_2 \\ D_1 \neq D_2, |D_1 \cap D_2| = w}} \left(p^d \cdot p^{d-w} - p^d \cdot p^d\right)\right] \\
&\qquad (\text{as } \mathcal{E}[e_{D_2} | e_{D_2} = 1] = p^{d-w} \text{ if } |D_1 \cap D_2| = w) \\
&= \gamma \cdot (1 - p^d) + \sum_{w=1}^{r} \left[\sum_{D_1} \sum_{\substack{D_2 \\ D_1 \neq D_2, |D_1 \cap D_2| = w}} \left(p^{2d-w} - p^{2d}\right)\right] \\
&= \gamma \cdot (1 - p^d) + \sum_{w=1}^{r} \left[\sum_{D_1} R_d(w,r) \cdot p^{2d} \left(p^{-w} - 1\right)\right] \quad (\text{recall } R_d(w,r) \text{ from Equation 15}) \\
&\leq \gamma \cdot (1 - p^d) + p^{2d} \cdot \sum_{w=1}^{r} \left[q^{r+1} \cdot R_d(w,r) \cdot p^{-w}\right] \\
&\leq \gamma \cdot (1 - p^d) + \gamma^2 \cdot \sum_{w=1}^{r} \frac{1}{(pd^\alpha)^w} \quad \left(\text{ since } R_d(w,r) \leq q^{r+1} \cdot \left(\frac{d}{q}\right)^w \cdot \frac{1}{w!}\right) \\
&\leq \gamma \cdot (1 - p^d) + \gamma^2 \cdot \frac{2}{pd^\alpha}
\end{aligned}
$$

The last inequality is true as without loss of generality $pd^{\alpha} = d^{\alpha-\beta} > 2$.

□

Proposition 24. *$\mathcal{E}[\eta] \leq 4 \cdot \gamma^2 \cdot q^{(r+1)} \cdot \left(\frac{d}{q}\right)^d$, where γ is as in Proposition 18. This implies*

$$\mathcal{E}[T_3] \leq 4 \cdot \left(\frac{2}{d^{\frac{\alpha-\beta}{2}}}\right)^d \cdot \gamma^2 \cdot \binom{d}{k}^2 \cdot \binom{N-d}{\ell}.$$

Proof. Observe that

$$\begin{aligned} w &:= |D_1 \cap D_2| \geq |E_2 \uplus E_3| = \ell - v + k \\ w' &:= |(D_3 \cap D_1) \cup (D_3 \cap D_2)| \geq |D_3 \cap D_1| \geq |E_5| = d - (\ell - v + k) \end{aligned}$$

Hence,

$$\begin{aligned} \eta &\leq \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_1 \cap D_2| = w}} \sum_{w' \geq d - (\ell - v + k)} \sum_{\substack{D_3 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_3 \neq D_2 \neq D_1, |(D_3 \cap D_1) \cup (D_3 \cap D_2)| = w'}} e_{D_1} \cdot e_{D_2} \cdot e_{D_3} \\ \mathcal{E}[\eta] &\leq \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_1 \cap D_2| = w}} \sum_{w' \geq d - (\ell - v + k)} \sum_{\substack{D_3 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_3 \neq D_2 \neq D_1, |(D_3 \cap D_1) \cup (D_3 \cap D_2)| = w'}} p^d \cdot p^{d-w} \cdot p^{d-w'} \\ &\leq \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_1 \cap D_2| = w}} \sum_{w' \geq d - (\ell - v + k)} p^{3d-w-w'} \cdot \binom{d}{w'} \cdot q^{(r+1)-w'}, \end{aligned}$$

as the number of D_3 with $|(D_3 \cap D_1) \cup (D_3 \cap D_2)| = w'$ for a fixed D_1, D_2 is bounded by $\binom{d}{w'} \cdot q^{(r+1)-w'}$. This implies,

$$\begin{aligned} \mathcal{E}[\eta] &\leq \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_1 \cap D_2| = w}} \sum_{w' \geq d - (\ell - v + k)} p^{3d-w-w'} \cdot d^{w'} \cdot q^{(r+1)-w'} \\ &\leq \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_1 \cap D_2| = w}} p^{3d-w} \cdot q^{(r+1)} \cdot \left(\frac{d}{pq}\right)^{d-(\ell-v+k)} \cdot 2 \\ &\qquad \text{(assuming } pq > 2d \text{ as } q \geq d^{1+\alpha}) \\ &\leq 2 \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} p^{3d-w} \cdot q^{(r+1)} \cdot \left(\frac{d}{pq}\right)^{d-(\ell-v+k)} \cdot R_d(w, r) \\ &\qquad \text{(recall } R_d(w,r) \text{ from Equation 15)} \\ &\leq 2 \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq \ell - v + k} p^{3d-w} \cdot q^{(r+1)} \cdot \left(\frac{d}{pq}\right)^{d-(\ell-v+k)} \cdot q^{(r+1)} \cdot \left(\frac{d}{q}\right)^{w} \\ &\leq 4 \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} p^{3d} \cdot q^{2(r+1)} \cdot \left(\frac{d}{pq}\right)^{d} \\ &\leq 4 \cdot p^{2d} \cdot q^{3(r+1)} \cdot \left(\frac{d}{q}\right)^{d} = 4 \cdot \gamma^2 \cdot q^{(r+1)} \cdot \left(\frac{d}{q}\right)^{d} \end{aligned}$$

Therefore,

$$\begin{aligned}\mathcal{E}[T_3] &\leq \rho \cdot \mathcal{E}[\eta] \\ &\leq 2^d \cdot \binom{d}{k}^2 \cdot \binom{N-d}{\ell} \cdot 4 \cdot \gamma^2 \cdot q^{(r+1)} \cdot \left(\frac{d}{q}\right)^d\end{aligned}$$

Since $r+1 = \frac{\alpha+\beta}{2(1+\alpha)} \cdot d$ and $q \geq d^{1+\alpha}$,

$$\mathcal{E}[T_3] \leq 4 \cdot \left(\frac{2}{d^{\frac{\alpha-\beta}{2}}}\right)^d \cdot \gamma^2 \cdot \binom{d}{k}^2 \cdot \binom{N-d}{\ell}.$$

□

Proposition 27. If $|D_1 \cap D_2| = w$ then

$$|\mu_1(D_1, D_2)| = \binom{N-2d+w}{\ell} \cdot \binom{w}{k} \quad \text{and hence} \quad \mathcal{E}[T_1] \leq d \cdot \frac{\gamma^2}{d^{(\alpha-\beta)k} \cdot k!} \cdot \binom{N-2d+k}{\ell}.$$

Proof. For a given D_1, D_2, let us count the number of rows (A, C) in which D_1 and D_2 can occur as labels. Since $C \subset D_1 \cap D_2$ and $|D_1 \cap D_2| = w$, we can pick C in $\binom{w}{k}$ ways. For every choice of C, we can pick A in $\binom{N-2d+w}{\ell}$ ways as A must be disjoint from $D_1 \cup D_2$ and $|D_1 \cup D_2| = 2d - w$. By Equation 18,

$$\begin{aligned}T_1 &= \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_2 \cap D_1| = w}} e_{D_1} \cdot e_{D_2} \cdot |\mu_1(D_1, D_2)| \\ \Rightarrow \mathcal{E}[T_1] &= \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq k} \sum_{\substack{D_2 \in \mathrm{Supp}(\mathsf{NW}_r) \\ D_2 \neq D_1, |D_2 \cap D_1| = w}} p^d \cdot p^{d-w} \cdot \binom{N-2d+w}{\ell} \cdot \binom{w}{k} \\ &\leq p^{2d} \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq k} R_d(w, r) \cdot p^{-w} \cdot \binom{N-2d+w}{\ell} \cdot \binom{w}{k} \\ &\leq p^{2d} \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq k} q^{r+1} \cdot \left(\frac{d}{pq}\right)^w \cdot \frac{1}{w!} \cdot \binom{N-2d+w}{\ell} \cdot \binom{w}{k} \\ &\leq p^{2d} \cdot q^{r+1} \cdot \sum_{D_1 \in \mathrm{Supp}(\mathsf{NW}_r)} \sum_{w \geq k} \left(\frac{1}{d^{\alpha-\beta}}\right)^w \cdot \frac{1}{w!} \cdot \binom{N-2d+w}{\ell} \cdot \binom{w}{k}\end{aligned}$$

The term $\left(\frac{1}{d^{\alpha-\beta}}\right)^w \cdot \frac{1}{w!} \cdot \binom{N-2d+w}{\ell} \cdot \binom{w}{k}$ is maximized at $w = k$ as $\beta < \alpha$. So,

$$\mathcal{E}[T_1] \leq d \cdot \frac{\gamma^2}{d^{(\alpha-\beta)k} \cdot k!} \cdot \binom{N-2d+k}{\ell}.$$

□

ECCC ISSN 1433-8092

http://eccc.hpi-web.de