

Optimal Depth-Three Circuits for Inner Product

Mohit Gurumukhani ✉ 🏠 

Cornell University, Ithaca, NY, USA

Daniel Kleber ✉

Department of Computer Science and Engineering, University of California, San Diego,
La Jolla, CA, USA

Ramamohan Paturi ✉

Department of Computer Science and Engineering, University of California, San Diego,
La Jolla, CA, USA

Christopher Rosin ✉ 🏠

Constructive Codes, USA

Navid Talebanfard ✉ 

University of Sheffield, UK

Abstract

We show that Inner Product in $2n$ variables, $\text{IP}_n(x, y) = x_1y_1 \oplus \dots \oplus x_ny_n$, can be computed by depth-3 bottom fan-in 2 circuits of size $\text{poly}(n) \cdot (9/5)^n$, matching the lower bound of Göös, Guan, and Mosnoi (Inform. Comput.'24). Our construction is obtained via the following steps.

1. We provide a general template for constructing optimal depth-3 circuits with bottom fan-in k for an arbitrary function f . We do this in two steps. First, we partition $f^{-1}(1)$ into *orbits* of its automorphism group. Second, for each orbit, we construct *one* k -CNF that (a) accepts the largest number of inputs from that orbit and (b) rejects all inputs rejected by f .
2. We instantiate the template for IP_n and $k = 2$. Guided by the intuition (which we call *modularity principle*) that optimal 2-CNFs can be constructed by taking the conjunction of variable-disjoint copies of smaller 2-CNFs, we use computer search to identify a small set of *building block* 2-CNFs over at most 4 variables.
3. We again use computer search to discover appropriate combinations (disjoint conjunctions) of building blocks to arrive at optimal 2-CNFs and analyze them using techniques from analytic combinatorics.

We believe that the approach outlined in this paper can be applied to a wide range of functions to determine their depth-3 complexity.

2012 ACM Subject Classification Theory of computation → Circuit complexity

Keywords and phrases Depth 3 circuits, Circuit lower bounds, Inner product, Analytic combinatorics

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.27

Related Version *Full Version*: <https://arxiv.org/abs/2601.04446>

Supplementary Material *Software*: <https://github.com/mjguru/Inner-Product>
archived at `swb:1:dir:22913678e86e5ba7afef126b288dbeeeb431a715`

Funding *Mohit Gurumukhani*: Supported by a Sloan Research Fellowship, NSF CAREER Award 2045576, and NSF Award CCF-2514586.

Daniel Kleber: Partially supported by NSF grant 2212136.

Ramamohan Paturi: Partially supported by NSF grant 2212136.

1 Introduction

A Σ_3 circuit is a depth-3 OR-AND-OR circuit with unbounded fan-in. Despite their simplicity, these circuits are surprisingly powerful. Any n -variate Boolean function can be computed by a Σ_3 circuit of size $O(2^{n/2})$ [4], a quadratic speed-up over the trivial construction. More



© Mohit Gurumukhani, Daniel Kleber, Ramamohan Paturi,
Christopher Rosin, and Navid Talebanfard;
licensed under Creative Commons License CC-BY 4.0

41st Computational Complexity Conference (CCC 2026).

Editor: Dana Moshkovitz; Article No. 27; pp. 27:1–27:25



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



strikingly, Valiant [19] showed that any linear-size fan-in 2 circuit of logarithmic depth can be computed by a Σ_3 circuit of size $2^{O(n/\log \log n)}$ size. Therefore, truly exponential lower bounds for depth-3 circuits imply non-linear lower bounds for log-depth circuits. However, the best known depth-3 lower bound for an explicit function is only $2^{\Omega(\sqrt{n})}$ and beating this barrier remains a major open problem in the field [9].

Even a further restriction of depth-3 circuits shows surprising power. Let Σ_3^k be the class of Σ_3 circuits in which every bottom OR gate is connected to at most k input bits. Equivalently, these circuits can be viewed as disjunctions of k -CNFs and the number of these formulas serves as a natural measure of size for the circuit. We denote the smallest number of k -CNFs needed to express a function f by $\text{Size}_3^k(f)$.¹ Valiant's depth reduction shows that linear-size log-depth circuits can in fact be computed by $\Sigma_3^{n^\epsilon}$ circuits of subexponential size where $\epsilon > 0$ can be chosen arbitrarily. More recently, Golovnev, Kulikov, and Williams [7] showed that *unrestricted* circuits of size cn for every $c < 3.9$ can be computed by Σ_3^{16} circuits of size $2^{\epsilon n}$ for some $\epsilon < 1$. Therefore, *near-maximal* lower bounds for Σ_3^{16} , i.e., lower bounds of the form $2^{n-o(n)}$, imply unrestricted circuit lower bounds beyond the state-of-the-art which is only $3.1n - o(n)$ [5, 12].

Near-maximal lower bounds are only known for Σ_3^2 and are due to Paturi, Saks, and Zane [17]. For $k \geq 3$, the state-of-the-art is due to Paturi, Pudlák, Saks, and Zane [15] who showed a Σ_3^k lower bound of $2^{cn/k}$ where $c > 1$ is a constant, building on a previous work by Paturi, Pudlák, and Zane [16] who proved a tight lower bound of $\Omega(2^{n/k})$ for Parity.

Near-maximal lower bound arguments

Paturi, Saks, and Zane [17] proved a crucial property of 2-CNFs. They showed that any 2-CNF with $2^{\Omega(n)}$ satisfying assignments, must accept a large *projection*, i.e., an affine space with dimension $\Omega(n)$ defined by equations of the form $x = 0, 1$, $x = y$, and $x = y + 1$. Therefore, any good *affine disperser*, i.e., any function that is not constant under large affine spaces, requires Σ_3^2 circuits of size $2^{n-o(n)}$. Several explicit constructions of such functions are known including with asymptotically optimal dependence on dimension - $O(\log(n))$ by Li [13]. Affine dispersers serve as a natural candidate for near-maximal lower bounds; the known explicit constructions are complicated poly-time algorithms, and the only known circuit construction has super-linear size [10]. Furthermore, state-of-the-art unrestricted circuit lower bounds hold for affine dispersers [5, 12].

In order to prove strong Σ_3^k lower bounds for affine dispersers, we cannot hope to apply the projection argument; there are 3-CNF formulas accepting exponentially many assignments which only accept projections of constant dimension [6]. However, a more general statement can be true. In particular, we do not know the answer to the following question: *is it true that every k -CNF with exponentially many satisfying assignments must accept an affine space of linear dimension?* An affirmative answer yields a near-maximal Σ_3^k lower bound for every good affine disperser. However, if the answer is negative, then there may not be a unified proof showing the hardness of *every* affine disperser. Instead, the hardness of each affine disperser may have to be demonstrated in a unique way. This suggests that by determining the complexity of *concrete* affine dispersers, we may be able to develop techniques that will ultimately allow us to prove near-maximal lower bounds.

¹ Note that this quantity is off by a factor of n^k from the smallest Σ_3^k sized circuit computing f . However, since k is small enough in this paper and the circuit size is exponential, this difference is immaterial.

Depth-3 complexity of Inner Product

The Inner Product modulo 2 function IP_n on n coordinates defined as $\text{IP}_n(x, y) = x_1y_1 \oplus \dots \oplus x_ny_n$ serves as an ideal starting point since it has a very simple and concrete definition and it is known to be a moderately good affine disperser (for dimension $n + 1$). Frankl, Gryaznov, and Talebanfard [6] pointed out that even for Σ_3^2 circuits, the exact complexity of IP_n is unclear as the projection technique fails to yield a strong bound. Golovnev, Kulikov, and Williams [7] initiated the study of depth-3 complexity of IP_n for independent reasons.

IP_n can be trivially computed by Σ_3^2 circuits of size $O(2^n)$: $\text{IP}_n = \bigvee_{I \subseteq [n], |I|=\text{odd}} (\bigwedge_{i \in I} (x_i \wedge y_i) \wedge_{i \notin I} (\neg x_i \vee \neg y_i))$. A simple lower bound of $\Omega(2^{n/2})$ follows by a reduction from Parity. Göös, Guan, and Mosnoi [8] showed, somewhat unexpectedly, that both of these bounds can be improved: $(9/5)^n \simeq 2^{0.847n} \leq \text{Size}_3^2(\text{IP}_n) \leq 2^{0.965n}$. They proved the lower bound by identifying a hard probability distribution on inputs to IP_n and characterizing the formulas that maximize the probability of accepting an input under this restriction. For the upper bound, they used the formulas in the lower bound to cover solutions with small weight, i.e., the number of coordinates on which both x and y variables are 1 is small. For the remaining solutions, they used a small number of the *matching* formulas. We will discuss these formulas later as we will also use them in our construction.

Amano [1] improved the upper bound even further to $2^{0.952n}$; this was done by observing that a depth-3 circuit for IP_n can be constructed by partitioning the coordinates into blocks of size t for a small value of t , and then combining depth-3 circuits that compute IP_t and $\neg \text{IP}_t$ on these blocks. To find such depth-3 circuits for a fixed t , the paper used computer search.

Our results

In this paper, we determine $\text{Size}_3^2(\text{IP}_n)$ up to polynomial factors by showing that the lower bound of [8] is, in fact, tight.

► **Theorem 1** (Main result). $\text{Size}_3^2(\text{IP}_n) \leq \text{poly}(n) \cdot (9/5)^n$

The paper makes two main contributions. Firstly, we give a general framework for the construction of optimal Σ_3^k circuits for a large class of *highly symmetric* functions. We show that it is sufficient to find consistent *extremal* k -CNFs to construct Σ_3^k optimal circuits for a function f . Secondly, our approach to circuit design is somewhat novel: Guided by modularity principle we use computer search to find the building blocks and search again for optimal disjoint conjunctions of building blocks to arrive at the desired constructions. Our methodology is general and flexible and we are optimistic that it can be applied to other functions.

2 A recipe for constructing depth-3 circuits

Let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be an arbitrary Boolean function and consider the task of proving a tight upper bound on $\text{Size}_3^k(f)$. In this section, we will provide a general framework for tackling such a task and will instantiate it to construct optimal circuits for IP . Our framework reduces the task of constructing Σ_3^k circuits for f to the task of constructing many different k -CNFs, each *consistent* with f , and each one maximizing the number of satisfying assignments with a certain property. We say a k -CNF F is consistent with the function f if for all $x \in \{0, 1\}^n$, it holds that $F(x) = 1 \implies f(x) = 1$, i.e., $\text{sol}(F) \subseteq f^{-1}(1)$, where $\text{sol}(F)$ is the set of satisfying assignments of F . We use the term *consistent* k -CNF to mean k -CNF consistent with a function when the function is clear from the context.

27:4 Optimal Depth-Three Circuits for Inner Product

To describe our framework, we will require some basic concepts from group theory; see Section 6 for a quick refresher. We define the notion of an automorphism group and orbits of a function as follows:

► **Definition 2** (Automorphism group and orbits). *For any $f : \{0,1\}^n \rightarrow \{0,1\}$, define the automorphism group of f , Aut_f , to be the group of permutations π such that for all $x \in \{0,1\}^n$, $f(x_{\pi(1)}, \dots, x_{\pi(n)}) = f(x_1, \dots, x_n)$. The action $\text{Aut}_f \times f^{-1}(1) \mapsto f^{-1}(1)$ of Aut_f on $f^{-1}(1)$ is defined as follows. $\pi \in \text{Aut}_f$ and $x_1 \dots x_n \in f^{-1}(1)$ maps to $x_{\pi(1)} \dots x_{\pi(n)}$. We denote the set of orbits of this action by Orb_f .*

Note that for each $S \in \text{Orb}_f$ we have $S \subseteq f^{-1}(1)$. We now define the following quantities associated with optimal consistent k -CNFs for each orbit.

► **Definition 3.** *For any $f : \{0,1\}^n \rightarrow \{0,1\}$ and $S \in \text{Orb}_f$, define*

$$\mu_{f,S,k} := \max_{F: F \text{ is an } n\text{-variate consistent } k\text{-CNF}} |\text{sol}(F) \cap S|,$$

$$\rho_{f,S,k} := |S|/\mu_{f,S,k}, \text{ and } \rho_{f,k}^* = \max_{S \in \text{Orb}_f} \rho_{f,S,k}.$$

In words, $\mu_{f,S,k}$ is the largest number of assignments S that are accepted by a consistent k -CNF, $\rho_{f,S,k}$ is a lower bound on $\text{Size}_3^k(f)$ due to S and $\rho_{f,k}^*$ is the maximum of such lower bounds over all orbits. We now present the main result of our framework which is proved in Section 2.

► **Lemma 4.** *For all $k, n \in \mathbb{N}$ and all functions $f : \{0,1\}^n \rightarrow \{0,1\}$, we have $\rho_{f,k}^* \leq \text{Size}_3^k(f) \leq O(n \cdot |\text{Orb}_f| \cdot \rho_{f,k}^*)$.*

We note that if $|\text{Orb}_f|$ of a function f is bounded by a polynomial, then $\text{Size}_3^k(f)$ and $\rho_{f,k}^*$ are within a polynomial factor of each other. Therefore, to determine the Σ_3^k complexity of f , it is enough to find a consistent k -CNF with the maximum number of solutions in S for every orbit S . This is indeed the case for IP_n and is exactly the strategy we will use to construct optimal Σ_3^2 circuits for it.

Depth-3 upper bound – Proof of Lemma 4

We will first show that we can cover an entire orbit S by using isomorphic copies of a consistent k -CNF F where the number of copies is bounded by $O(n|S|/|\text{sol}(F) \cap S|)$.

► **Claim 5.** Let $f : \{0,1\}^n \rightarrow \{0,1\}$ be an arbitrary function and let F be a k -CNF consistent with f . Then, for any $S \in \text{Orb}_f$, there exist consistent k -CNFs $F_1, \dots, F_t$ such that $S \subseteq \text{sol}(F_1) \cup \dots \cup \text{sol}(F_t)$ where $t = O(n|S|/|\text{sol}(F) \cap S|)$.

We first show how the proof of Lemma 4 follows from the claim and present the proof of the claim following it.

Proof of Lemma 4. We prove the lower bound and the upper bound part of the inequality separately. For the lower bound, fix any Σ_3^k circuit C for f . We express $C = \bigvee_{i=1}^m F_i$ where $m \in \mathbb{N}$ and each F_i is a k -CNF. We will show that for all $S \in \text{Orb}_f$, it holds that $m \geq \rho_{f,S,k}$, which implies $m \geq \rho_{f,k}^*$. Fix any $S \in \text{Orb}_f$. We observe that each F_i must be consistent with f and so for all $i \in [m]$, we must have that $|\text{sol}(F_i) \cap S| \leq \mu_{f,S,k}$. Moreover, each assignment in S must be covered by some F_i . Indeed, for each $\alpha \in S$, since $C(\alpha) = f(\alpha) = 1$ and $C = \bigvee_{i=1}^m F_i$, there must exist some $i \in [m]$ such that $F_i(\alpha) = 1$. Hence, we must have $m \geq |S|/\mu_{f,S,k} = \rho_{f,S,k}$ as desired.

For the upper bound, we construct a circuit C for f as follows. let $S \in \text{Orb}_f$ and let F_S be the k -CNF consistent with f that maximizes $|\text{sol}(F_S) \cap S|$. By definition, $|\text{sol}(F_S) \cap S| = \mu_{f,S,k}$. From Claim 5 we get consistent k -CNFs $F_{S,1}, \dots, F_{S,t}$ which cover S where $t = t(S) = O(n|S|/|\text{sol}(F) \cap S|) = O(n\rho_{f,S,k})$. We define $C = \bigvee_{S \in \text{Orb}_f, i \in [t(S)]} F_{S,i}$. C accepts all assignments from each of the orbits S . Since each $F_{S,i}$ is consistent with f , C correctly rejects all assignments from $f^{-1}(0)$ and hence C indeed computes the function f . The total number of k -CNFs in C is $O(n \sum_{S \in \text{Orb}_f} \rho_{f,S,k}) = O(n \cdot |\text{Orb}_f| \cdot \rho_{f,k}^*)$ as claimed. $\blacktriangleleft$

Proof of Claim 5. Let $S \in \text{Orb}_f$ and F be an n -variate k -CNF consistent with f . Let $t = 2|S| \ln |S|/|\text{sol}(F) \cap S| = O(n|S|/|\text{sol}(F) \cap S|)$. For each $i \in [t]$, define $F_i = F^{\pi_i}$ to be the permutation of F under a random permutation $\pi_i \in \text{Aut}_f$. Fix any $\alpha \in S$. It follows (see Lemma 21 for the general group theoretic claim and its proof) that for a random $\pi \in \text{Aut}_f$, $\Pr[\alpha \in \text{sol}(F^\pi)] = |\text{sol}(F) \cap S|/|S|$, which implies

$$\begin{aligned} \Pr[\alpha \notin \text{sol}(F_1) \cup \dots \cup \text{sol}(F_t)] &\leq (1 - |\text{sol}(F) \cap S|/|S|)^t \\ &\leq \exp(-t \cdot |\text{sol}(F) \cap S|/|S|) \\ &< 1/|S| \end{aligned}$$

where the last inequality follows by the choice of t and the fact that $|S| \leq 2^n$. By a union bound over all $\alpha \in S$, we conclude that there exists a choice of F_i s covering all assignments in S . $\triangleleft$

3 Construction for IP

In this section, we instantiate our framework from Section 2 to construct optimal Σ_3^2 circuits for IP. We will prove the following main lemma which together with Lemma 4 proves Theorem 1. In the following, we also use IP_n^0 for IP_n and $\text{IP}_n^1 = 1 - \text{IP}_n^0$ for the complement of IP_n^0 .

► **Lemma 6.** *There exists a constant c such that for all integer $n \geq 2$, $\rho_{\text{IP}_n^1, 2}^* \leq n^c \cdot (9/5)^n$.*

To prove the lemma, we will identify the orbits of IP_n in Section 3.1 and the building block 2-CNFs in Section 3.2. In Section 3.3, we will reduce Lemma 6 to a version of itself (Lemma 14) that holds only for infinitely many even n and reduce the constructions for IP_n^1 to those of IP_n^0 . Our plan is to construct 2-CNFs for each of the orbits by taking a *disjoint* conjunction of copies of a small number of the building-block 2-CNFs. For this purpose, we divide the orbits into six regions in Section 3.4, state Lemma 16 which is a version of Lemma 14 specialized to regions. We prove Lemma 16 in Section 7.

3.1 Orbits of IP and spectra of 2-CNFs

We now describe the orbits of IP_n^1 and IP_n^0 and introduce the notion of the *spectrum* of a 2-CNF.

We observe that $\text{Aut}_{\text{IP}_n^1} = \text{Aut}_{\text{IP}_n^0}$ and they contain all permutations that permute the n coordinates, and permute x_i and y_i within each coordinate. For input $(x, y) \in (\{0, 1\}^n)^2$, define the following quantities:

1. $d_2(x, y) = |\{i \in [n] : x_i = y_i = 1\}|$.
2. $d_1(x, y) = |\{i \in [n] : x_i \neq y_i\}|$.
3. $d_0(x, y) = |\{i \in [n] : x_i = y_i = 0\}|$.

We have $d_2(x, y) + d_1(x, y) + d_0(x, y) = n$. More importantly, for $j \in \{0, 1, 2\}$, $d_j(x, y)$ are invariant under a permutation $\pi \in \text{Aut}_{\text{IP}_n^1}$, that is, $d_j(x', y') = d_j(x, y)$ where $(x', y') = \pi(x, y)$. We also observe that if $d_j(x, y) = d_j(x', y')$ for all $j \in \{0, 1, 2\}$, then there exists a permutation $\pi \in \text{Aut}_{\text{IP}_n^1}$ such that $\pi(x, y) = (x', y')$. Hence, d_j precisely determine the orbits of IP_n^1 and IP_n^0 . In particular we have $|\text{Orb}_{\text{IP}_n^1}| = |\text{Orb}_{\text{IP}_n^0}| = O(n^2)$. We parameterize the orbits of IP_n as follows:

► **Definition 7** (Orbits of IP_n^1 and IP_n^0). For $d_2, d_1, d_0 \in [n]$ such that $d_2 + d_1 + d_0 = n$, let $S(d_2, d_1, d_0) \subseteq (\{0, 1\}^n)^2$ denote the set of assignments $(x, y) \in (\{0, 1\}^n)^2$ such that $d_j(x, y) = d_j$ for all $j \in \{0, 1, 2\}$. Note that $S(d_2, d_1, d_0)$ is an orbit of the function $\text{IP}_n^{d_2 \bmod 2}$ and for any $(x, y) \in S(d_2, d_1, d_0)$, $\text{IP}_n^0(x, y) = d_2 \bmod 2$.

We record the following fact regarding size of each of the orbits:

► **Fact 8.** $|S(d_2, d_1, d_0)| = \binom{n}{d_2} \binom{n-d_2}{d_0} 2^{n-d_2-d_0} = \binom{n}{d_2} \binom{d_1+d_0}{d_1} 2^{d_1}$.

We now define the spectrum of a 2-CNF F as the generating function that encodes the number of its satisfying assignments from each of the orbits.

► **Definition 9** (Spectrum of 2-CNF). For a 2-CNF F over $2n$ variables, define the spectrum of F to be the generating function $G(x, y, z) = \sum_{p,q,r \in \mathbb{N}} C_{p,q,r} x^p y^q z^r$ where $C_{p,q,r}$ is the number of assignments $\alpha \in \{0, 1\}^{2n}$ such that $F(\alpha) = 1$, $d_2(\alpha) = p$, $d_1(\alpha) = q$, and $d_0(\alpha) = r$.

The spectrum of 2-CNF of F is always a finite homogeneous polynomial. If F is consistent with IP_n^0 , then $C_{p,q,r} = 0$ if (p, q, r) is not an orbit of IP_n^0 .

3.2 Building blocks for our constructions

In this section, we identify a small number of building block 2-CNFs which will be used to construct 2-CNFs; we will also define the notion of a disjoint conjunction of the building blocks. Since any function over 2 variables can be expressed as a 2-CNF, our building block descriptions will be expressed as AND of functions over 2 variables.

► **Definition 10** (Building block 2-CNFs and their spectra). We define the following building blocks:

- **Id₂** denotes the 2-CNF over 1 coordinate (2 variables) that accepts iff both variables equal 1; its spectrum is x .
- **Id₁** denotes the 2-CNF over 1 coordinate that accepts iff both variables do not equal each other; its spectrum is $2y$.
- **Id₀** denotes the 2-CNF over 1 coordinate that accepts iff both variables equal 0; its spectrum is z .
- **NAND** denotes the 2-CNF over 1 coordinate that accepts iff both variables do not equal 1; its spectrum is $2y + z$.
- **Matching** denotes the 2-CNF over over 2 coordinates defined as follows. On input $(x, y) \in (\{0, 1\}^2)^2$, the 2-CNF accepts iff $x_1 = x_2$, and $y_1 = y_2$; its spectrum is $x^2 + 2y^2 + z^2$.
- **2Imp** denotes the 2-CNF over 2 coordinates defined as follows. On input $(x, y) \in (\{0, 1\}^2)^2$, the 2-CNF accepts iff the following holds: $x_1 = x_2$, $x_1 \implies y_1$, and $x_2 \implies y_2$; its spectrum is $x^2 + y^2 + 2yz + z^2 = x^2 + (y + z)^2$.

► **Definition 11** (Disjoint conjunction). Let F_1 and F_2 be arbitrary 2-CNFs over n_1 and n_2 variables respectively. Then, $F_1 \wedge F_2$ is the 2-CNF over $n_1 + n_2$ variables obtained by adding all clauses from both F_1 and F_2 where we associate the first n_1 variables with variables of F_1 and the last n_2 variables with variables of F_2 .

We next record a couple of crucial observations regarding the spectrum of the disjoint conjunction of 2-CNFs and 2-CNFs consistent with either IP_n^0 or IP_n^1 .

► **Fact 12** (Spectra of disjoint conjunctions 2-CNFs). *Let F_1 and F_2 be arbitrary 2-CNFs over n_1 and n_2 variables with spectra G_1 and G_2 respectively. Then, the spectrum of $F_1 \wedge F_2$ equals $G_1 \cdot G_2$. Furthermore, if F_1 is consistent with $\text{IP}_{n_1}^{b_1}$ and that F_2 is consistent with $\text{IP}_{n_2}^{b_2}$ for some $b_1, b_2 \in \{0, 1\}$, then $F_1 \wedge F_2$ is consistent with $\text{IP}_{n_1+n_2}^{b_1 \oplus b_2}$.*

We note that each building block from Section 3.2 is consistent with IP_n^b for some $b \in \{0, 1\}$ for an appropriate n . This implies:

► **Corollary 13.** *Let F be a 2-CNF over n variables obtained by (repeatedly) applying disjoint conjunction to building blocks from Definition 10. Then, F is consistent with IP_n^b for some $b \in \{0, 1\}$.*

3.3 Reducing to infinitely often case

We now state a version of Lemma 6 that only holds for infinitely many n , each a multiple of some even constant.

► **Lemma 14.** *There exist constants c , even $m \in \mathbb{N}$ such that the following holds: For all $p, q, r, n \in \mathbb{N}$ such that $n = p + q + r$ and m divides each of p, q, r and n , there exists a 2-CNF F over n variables that is consistent with IP_n^0 and satisfies:*

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^c \cdot (9/5)^n.$$

We will now show that Lemma 6 follows from Lemma 14. Lemma 14 will be proven by reducing it to Lemma 16 in Section 3.4.

Proof of Lemma 6. Fix any integer $n \geq 2$ and let c_0 and m be as given by Lemma 14. Now consider any orbit S of IP_n . It suffices to show that there exists a 2-CNF F consistent with IP_n such that $\frac{|S|}{|\text{sol}(F) \cap S|} \leq n^c \cdot (9/5)^n$ where c is a constant. As discussed in Section 3.1, S is parameterized by integers $p, q, r \in \mathbb{N}$ such that $p + q + r = n$ where p is odd and $S = S(p, q, r)$.

Let $p' \leq p, q' \leq q, r' \leq r$ be the largest multiples of m less than or equal to p, q, r respectively. Let $n' = p' + q' + r'$. n' must also be a multiple of m and $n - n' \leq 3m$. We apply Lemma 14 for p', q', r', n' to obtain 2-CNF F' over n' variables such that $\frac{|S(p', q', r')|}{|\text{sol}(F') \cap S(p', q', r')|} \leq (n')^{c_0} \cdot (9/5)^{n'}$. Let F be the 2-CNF over n variables obtained by the disjoint conjunction of $p - p'$ copies of Id_2 , $(q - q')$ copies of Id_1 , $(r - r')$ copies of Id_0 , and one copy of F' .

We know that F' is consistent with $\text{IP}_{n'}^{p' \pmod{2}}$, Id_1 and Id_0 are consistent with IP_1^0 and Id_2 is consistent with IP_1^1 . Applying Fact 12, we get that F is consistent with $\text{IP}_n^{p \pmod{2}}$ as desired. Next, using Fact 8 and the fact that $n - n' \leq 3m, p - p' \leq m, q - q' \leq m, r - r' \leq m$, we infer that there exists a constant c_1 such that $\frac{|S(p, q, r)|}{|S(p', q', r')|} \leq n^{c_1}$. We also have $|\text{sol}(F) \cap S(p, q, r)| = 2^{q-q'} |\text{sol}(F') \cap S(p', q', r')|$. We thus have

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &= \frac{|S(p, q, r)|}{|S(p', q', r')|} \cdot \frac{|S(p', q', r')|}{|\text{sol}(F) \cap S(p, q, r)|} \\ &= 2^{q'-q} \cdot \frac{|S(p, q, r)|}{|S(p', q', r')|} \cdot \frac{|S(p', q', r')|}{|\text{sol}(F') \cap S(p', q', r')|} \\ &\leq 2^m \cdot n^{c_1} \cdot n^{c_0} \cdot (9/5)^n \\ &\leq n^c \cdot (9/5)^n \end{aligned}$$

for a suitable constant c . ◀

3.4 Splitting orbits into regions

In this subsection, we divide the orbits of IP_n^1 and IP_n^0 into six regions and prove Lemma 14 for each of the six regions.

► **Definition 15.** For $i \in [6]$ and $n \in \mathbb{N}$, let $\mathcal{R}_i(n) \subseteq \mathbb{N} \times \mathbb{N} \times \mathbb{N}$ is a set of triples $(p, q, r) \in \mathbb{N}$ such that $p + q + r = n$ and

1. $\mathcal{R}_1(n) : \frac{1}{2}n - \frac{25}{32}p \geq 0, \frac{5}{4}n - \frac{25}{32}p - \frac{25}{4}r \geq 0, \text{ and } -\frac{5}{4}n + \frac{25}{16}p + \frac{25}{4}r \geq 0.$
2. $\mathcal{R}_2(n) : \frac{1}{2}n - \frac{25}{32}p \leq 0.$
3. $\mathcal{R}_3(n) : \frac{1}{2}n - \frac{25}{32}p \geq 0, \frac{5}{4}n - \frac{25}{32}p - \frac{25}{4}r \leq 0, \text{ and } \frac{1}{4}n - p \leq 0.$
4. $\mathcal{R}_4(n) : \frac{1}{2}n - \frac{25}{32}p \geq 0, -\frac{5}{4}n + \frac{25}{16}p + \frac{25}{4}r \leq 0, -\frac{1}{20}n + r \geq 0, \text{ and } \frac{1}{4}n - p \leq 0.$
5. $\mathcal{R}_5(n) : \frac{1}{2}n - \frac{25}{32}p \geq 0, \text{ and } -\frac{1}{20}n + r \leq 0.$
6. $\mathcal{R}_6(n) : \frac{1}{4}n - p \geq 0.$

From here on, we refer to (p, q, r) as the orbit corresponding to it.

For each region $\mathcal{R}_i$ and each orbit in the region, we will construct a 2-CNF such that its spectrum satisfies the following property.

► **Lemma 16.** For all $i \in [6]$, there exist constants K_i and even $m_i \in \mathbb{N}$ such that the following holds: For all $p, q, r, n \in \mathbb{N}$ such that $(p, q, r) \in \mathcal{R}_i$, $n = p + q + r$, and m_i divides each of p, q, r , there exists a 2-CNF F over $2n$ variables that is consistent with IP_n^0 and satisfies:

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{K_i} \cdot (9/5)^n.$$

We first show that Lemma 14 follows from Lemma 16. We will prove Lemma 16 in Section 7.

Proof of Lemma 14. We apply Lemma 16 and obtain even m_i and K_i . Let m be the least common multiple of m_1, m_2, m_3, m_4, m_5 , and m_6 and let c to be the maximum of $K_1, K_2, K_3, K_4, K_5, K_6$. We observe that the six regions cover the set of all triples such that $p + q + r = n$, completing our proof. ◀

4 An overview of constructions

In this section we provide an overview of our constructions that will lead to the proof of Lemma 16.

For each $i \in [6]$ and $(p, q, r) \in \mathcal{R}_i$, we construct a 2-CNF by taking a disjoint conjunction of an appropriate number of building block 2-CNFs while ensuring that F is consistent with IP_n^0 . We then argue that F captures many assignments from $S(p, q, r)$ by lower bounding the coefficient of the monomial $x^p y^q z^r$ in its spectrum. This is the common framework of construction and analysis. To illustrate how the framework works, we provide a sketch of the construction and its analysis for region 1 in the following. Our analysis of the constructions will vary in their techniques across regions, ranging from direct calculations to reducing to an optimization problem that we solve using a numerical optimization software (see Section 7.3 for an example). For a lot of regions, we will use tools from multivariate analytic combinatorics to lower bound the coefficient of the desired monomial.

A more complete description of the constructions for each region and their analysis is presented in Section 7 and Section 8.

Sketch of the construction and analysis for $\mathcal{R}_1$

Let $A = \frac{5}{4}n - \frac{25}{32}p - \frac{25}{4}r$, $B = \frac{-5}{4}n + \frac{25}{16}p + \frac{25}{4}r$, $C = \frac{1}{2}n - \frac{25}{32}p$. We select $m_1 = 8000$ (which is not necessarily optimized) and require that m_1 divides p , q , and r . A, B, C are non-negative integers since $(p, q, r) \in \mathcal{R}_1$. We take a disjoint conjunction of A copies of **Matching**, B copies of **2Imp** and $2C$ copies of **NAND** building blocks to construct our 2-CNF F . Since m_1 is even, we have that F is consistent with IP_n^0 .

The spectrum of F equals $P(x, y, z) = (x^2 + 2y^2 + z^2)^A (x^2 + (y + z)^2)^B ((2y + z)^2)^C$ and the coefficient of $x^p y^q z^r$ in $P(x, y, z)$ equals $|\text{sol}(F) \cap S(p, q, r)|$. To bound the coefficient, we define $f(u, v) = P(u^{1/2}, v, 1)$. Since $p + q + r = n$ and P is a homogeneous polynomial, the coefficient of $u^{p/2} v^q$ in $f(u, v)$ equals the coefficient of $x^p y^q z^r$ in $P(x, y, z)$. We set the gradient of $f(u, v)$ to 0 and compute the non-negative real solutions of this equation to get the unique *critical point* at $(16, 2)$. We then apply a standard result in analytic combinatorics – Lemma 29 – that states that for *well-behaved* f , the value of the coefficient of $u^{p/2} v^q$ (up to a polynomial factor) equals $\frac{f(16, 2)}{16^{p/2} 2^q} = \frac{25^{A+B+C}}{16^{p/2} 2^q} = \frac{5^n}{2^{2p+q}}$ where we used the fact that $A + B + C = n/2$. We now have $|\text{sol}(F) \cap S(p, q, r)| = \frac{5^n}{2^{2p+q}}$ up to a polynomial factor. We upper bound the number of 2-CNFs required to cover the orbit $S(p, q, r)$.

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &\leq \binom{n}{p} \binom{n-p}{r} \cdot 2^{n-p-r} \cdot \frac{2^{2p+q}}{5^n} \\ &= \binom{n}{p, q, r} \cdot 4^{-r} \cdot (4/5)^n \\ &\leq \sum_{i, j, k \in \mathbb{N}: i+j+k=n} \binom{n}{i, j, k} \cdot 1^i \cdot 1^j \cdot 4^{-k} \cdot (4/5)^n \\ &= (9/4)^n \cdot (4/5)^n = (9/5)^n \end{aligned}$$

where we used Fact 8 to establish the first inequality.

► **Remark 17.** From above calculations, we can identify that for the orbit of $(4n/9, 4n/9, n/9)$, we require $(9/5)^n$ many 2-CNFs to cover it.

5 Conclusions

The immediate question left open is to determine $\text{Size}_3^k(\text{IP}_n)$ for every $k \geq 3$. We believe that a similar methodology should work for upper bounds. However, the challenge is to find an argument that scales with k . It also remains open to find a non-trivial lower bound argument for $k \geq 3$.

Another related open question is to construct an explicit degree 2 polynomial that requires Σ_3^k circuits of size $2^{n-o(n)}$ for all constant k ; this is open even for $k = 2$. Existentially, Impagliazzo, Paturi, and Zane showed that a random degree 2 polynomial requires maximal Σ_3^k circuits [11].

How did we discover these circuits?

We here document our journey for constructing optimal circuits for IP_n . We hope this inspires future work to find optimal or close to optimal Σ_3^k circuits for various explicit functions.

Modular constructions

As a starting point, we believed that optimal circuits for IP_n must be modular, i.e., obtained by combining disjoint copies of smaller 2-CNFs. Our reasons for such a belief are twofold:

27:10 Optimal Depth-Three Circuits for Inner Product

All previous best known constructions of IP_n had this property and optimal (and conjectured to be optimal) Σ_3^k circuits for other simple functions such as Parity and Majority have this property.

Searching for better circuits for small n

Our first improvement came from examining the construction of Σ_3^2 circuits for IP_n by Amano [1]. To construct such circuits, the paper used an integer programming solver on a set of constraints to find a disjunction of 7 2-CNFs that compute IP_4^1 and another disjunction of 7 2-CNFs that compute IP_4^0 . For the general construction, a combination of these formulas is used.

We searched for better circuits using Kissat SAT solver under a similar set of constraints and identified a disjunction of 14 2-CNF formulas that computes IP_5^1 and another disjunction of 13 2-CNF formulas that computes IP_5^0 . Repeated use of these formulas as in Amano's construction gives a slightly improved construction for IP_n with size 2^{cn} where $c \approx 0.9509$. This inspired us to seek further improvements.

Search for small Pareto-optimal building blocks

Suppose A and B are 2-CNFs consistent with IP_n^b for some $b \in \{0, 1\}$ and $n \in \mathbb{N}$. Then, we say A dominates B if each coefficient in A 's spectrum is at least as large as the corresponding coefficient in B 's spectrum. We performed an exhaustive computer search for Pareto-optimal (non-dominated) 2-CNF building blocks for IP_2^0 that identified three distinct Pareto-optimal building blocks: **Matching**, **2Imp** and disjoint conjunction of 2 copies of **NAND**. In particular, **2Imp** was a new building block first identified by this exhaustive search and is *crucial* in our optimal constructions. We also identified other Pareto-optimal building blocks for small values of n that we did not end up using.

Search for compositions of building blocks using disjoint conjunctions

After we identified a suitable set of building blocks, we are left with the task of finding disjoint conjunctions of the building blocks to cover the assignments in each orbit. This task turned out to be highly nontrivial and we resorted to computer search again to gain intuition to guide us towards optimal general constructions. We searched for efficient disjoint conjunctions of building blocks for each orbit of IP_n for some large fixed values of n . For this, we enumerated all distinct compositions of the building blocks. This exercise pointed to the best composition for each orbit and helped identify the hardest orbits.

For the search, we used $n = 50$, $n = 100$, and $n = 200$. We have also enumerated restricted compositions for $n = 400$ and $n = 800$ near the conjectured hardest orbits. The search identified best compositions involving the building blocks **Matching**, **2Imp**, and **NAND** which led to the general constructions. From our best compositions for each value of n , we obtained depth-3 circuits of size $\text{poly}(n) \cdot 2^{cn}$ for IP_n where the dependence of c on n is presented below.

$n = 50$	$c = 0.8344320$
$n = 100$	$c = 0.8414042$
$n = 200$	$c = 0.8447647$
$n = 400$	$c = 0.8463819$
$n = 800$	$c = 0.8471913$

A least-squares fit to the data yields $0.8479999 - 0.6470509/n$, suggesting an eventual convergence to 0.8479999 for large n . This is extremely close to $\log(9/5) = 0.8479969$, which would match the lower bound from [8], motivating us to pursue an upper bound with $c = \log(9/5)$.

Restricted compositions to cover regions

Finally, we analytically solved for optimal constructions using saddle point methods with **Matching**, **2Imp**, and **NAND** as building blocks. Our analysis identified regions 2, 3, and 4 and that optimal constructions for these regions use at most 1 or 2 building blocks. We then performed a search for optimal compositions (using at most 2 types of building blocks) for each of the regions for $n = 200$. The goal of the search was to find a small number of distinct compositions that would cover a region so that the ratio of the size of an orbit to number of solutions in the orbit is at most $(9/5)^n$. This search successfully identified the desired compositions that are sufficient for our purpose.

6 Preliminaries

We use $\log(x)$ to denote logarithm of x with base 2 and $\ln(x)$ to denote the natural logarithm. The binary entropy function $H : [0, 1] \rightarrow [0, 1]$ is defined as $H(x) = x \log(1/x) + (1 - x) \log(1/(1 - x))$.

We will utilize the following well known approximation for binomial coefficients (see, for instance, equation 7.14 from [3]):

► **Theorem 18.** *For all $m, k \in \mathbb{N}$ such that $m \geq k$, we have that*

$$\frac{1}{m+1} 2^{mH(k/m)} \leq \binom{m}{k} \leq 2^{mH(k/m)}.$$

Group theory

► **Definition 19** (Group action and orbits). *An action of a finite group G on a finite set S is a mapping $(\cdot)$ from $G \times S$ to S such that*

- $e \cdot x = x$ for all $x \in S$ where e is the identity element of G ,
- $g \cdot (h \cdot x) = gh \cdot x$ for all $g, h \in G$ and $x \in S$.

The orbit of an element $x \in S$ is $G \cdot x = \{g \cdot x : g \in G\}$. The properties of the group show that orbits partition S into natural equivalence classes.

For $g \in G$ and $T \subseteq S$ define $g \cdot T = \{g \cdot y : y \in T\}$. For $x \in S$ let the stabilizer subgroup of x be defined as $G_x = \{g \in G : g \cdot x = x\}$.

► **Theorem 20** (Orbit-Stabilizer Theorem). *For every $x \in S$, $|G \cdot x| = |G|/|G_x|$.*

We need the following lemma which is an easily corollary of Orbit-Stabilizer Theorem.

► **Lemma 21.** *Let $x \in S$ and let $T \subseteq S$. Then $\Pr_g[x \in g \cdot T] = \frac{|G \cdot x \cap T|}{|G \cdot x|}$, where g is sampled uniformly from G .*

Proof. We have $\Pr_g[x \in g \cdot T] = \Pr_g[\exists y \in T : x = g \cdot y]$. The events $x = g \cdot y$ are disjoint and have non-zero probability only when y is in the same orbit as x . Therefore we can write the probability as $\sum_{y \in G \cdot x \cap T} \Pr_g[x = g \cdot y]$. It remains to show that $\Pr_g[x = g \cdot y] = 1/|G \cdot x|$ for every $y \in G \cdot x \cap T$.

27:12 Optimal Depth-Three Circuits for Inner Product

Observe that there is a bijection between $A := \{g : g \cdot y = x\}$ and G_y . Pick any $\sigma \in A$. Then the mapping π to $\sigma^{-1}\pi$ is a bijection. Therefore $\Pr_g[x = g \cdot y] = |A|/|G| = |G_y|/|G| = 1/|G \cdot x|$, where the last equality follows from Orbit-Stabilizer Theorem and the fact that y and x are in the same orbit. ◀

7 Constructing 2-CNFs for each region

In this section, we will prove Lemma 16 for each of the regions by constructing 2-CNFs for orbits in that region. Proofs provided in this section depend on results regarding asymptotes of coefficients of multivariate generating functions which are proved in Section 8.

7.1 2-CNFs for orbits in region 1

In this subsection, we will prove Lemma 16 for orbits in $\mathcal{R}_1$. We will rely on the following result regarding asymptotes of a coefficient of a monomial from a generating function, which we will prove in Section 8.1.

► **Lemma 22.** *There exists a constant C_0 such that the following holds. Let A, B, C, p, q, r, n be nonnegative integers such that $n \geq C_0$, $n = p + q + r$, $A = \frac{5}{4}n - \frac{25}{32}p - \frac{25}{4}r$, $B = \frac{-5}{4}n + \frac{25}{16}p + \frac{25}{4}r$, $C = \frac{1}{2}n - \frac{25}{32}p$. Then, the coefficient of the monomial $x^p y^q z^r$ in $P(x, y, z) = (x^2 + 2y^2 + z^2)^A (x^2 + (y + z)^2)^B ((2y + z)^2)^C$ is at least $\frac{1}{n^{C_0}} \frac{5^n}{2^{2p+q}}$.*

Proof of Lemma 16 for region 1. Let $m_1 = 32$. Now consider any $(p, q, r) \in \mathcal{R}_1$ such that m_1 divides each of p, q , and r and hence n .

Let $A = \frac{5}{4}n - \frac{25}{32}p - \frac{25}{4}r$, $B = \frac{-5}{4}n + \frac{25}{16}p + \frac{25}{4}r$, $C = \frac{1}{2}n - \frac{25}{32}p$. A, B, C are non-negative integers since m_1 divides each of n, p, q, r and by the definition of $\mathcal{R}_1$ we have that $A, B, C \geq 0$.

Our 2-CNF F is obtained by a disjoint conjunction of A copies of **Matching**, B copies of **2Imp** and $2C$ copies of **NAND**. It is easy to see that F is consistent with $\text{IP}_n^{p \bmod 2} = \text{IP}_n^0$.

The spectrum of F is $P(x, y, z) = (x^2 + 2y^2 + z^2)^A (x^2 + (y + z)^2)^B ((2y + z)^2)^C$. We now apply Lemma 22 to get

$$|\text{sol}(F) \cap S(p, q, r)| \geq n^{-K_1} \cdot \frac{5^n}{2^{2p+q}}$$

where K_1 is a constant. Since $|S(p, q, r)| = \binom{n}{p} \binom{n-p}{r} \cdot 2^{n-p-r}$. Using these, we get

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &\leq \binom{n}{p} \binom{n-p}{r} \cdot 2^{n-p-r} \cdot n^{-K_1} \cdot \frac{2^{2p+q}}{5^n} \\ &= n^{-K_1} \cdot \binom{n}{p, q, r} \cdot 4^{-r} \cdot (4/5)^n \end{aligned}$$

where the last equality follows from the fact that $p + q + r = n$. To prove $\frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{-K_0} \cdot (9/5)^n$ for some constant K_0 , it suffices to show that $\binom{n}{p, q, r} \cdot 4^{-r} \leq (9/4)^n$. We have

$$\begin{aligned} \binom{n}{p, q, r} \cdot 4^{-r} &= \binom{n}{p, q, r} 1^p \cdot 1^q \cdot 4^{-r} \\ &\leq \sum_{i, j, k \in \mathbb{N}: i+j+k=n} \binom{n}{i, j, k} \cdot 1^i \cdot 1^j \cdot 4^{-k} \\ &= (9/4)^n \end{aligned}$$

where the last equality follows from the binomial theorem. Hence, our construction F indeed has the desired properties. ◀

7.2 2-CNFs for orbits in region 2

In this subsection, we will prove Lemma 16 for orbits in $\mathcal{R}_2$. In the proof, we use the following general claim regarding binomial coefficients which will be proved at the end of the subsection.

▷ **Claim 23.** For all $m, k \in \mathbb{N}$ such that $m \geq k$ and m, k are even, we have that

$$\binom{m}{k} \leq (m/2 + 1)^2 \binom{m/2}{k/2}$$

Proof of Lemma 16 for region 2. Let $m_2 = 4$ and consider any $(p, q, r) \in \mathcal{R}_2$ such that m_2 divides each of p, q , and r and hence n .

Let F be the 2-CNF obtained by taking the disjoint conjunction of $n/2$ copies of **Matching**. Since **Matching** is consistent with IP_n^0 , F is also consistent with IP_n^0 as desired. The spectrum of F is given by $P(x, y, z) = (x^2 + 2y^2 + z^2)^{n/2}$ and that $|\text{sol}(F) \cap S(p, q, r)|$ equals the coefficient of the monomial $x^p y^q z^r$ in $P(x, y, z)$. Examining P , we easily deduce that this coefficient equals $\binom{n/2}{p/2} \binom{n/2-p/2}{q/2} 2^{q/2}$. Since $S(p, q, r) = \binom{n}{p} \binom{n-p}{r} 2^q$, we get

$$\frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|} = \frac{\binom{n}{p} \binom{n-p}{r} 2^q}{\binom{n/2}{p/2} \binom{n/2-p/2}{q/2} 2^{q/2}} = \frac{\binom{n}{p}}{\binom{n/2}{p/2}} \cdot \frac{\binom{n-p}{r}}{\binom{n/2-p/2}{q/2}} \cdot 2^{q/2}$$

where the last equality follows using the fact that $p + q + r = n$. We apply Claim 23 to the last expression and obtain:

$$\begin{aligned} \frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|} &\leq n^4 \cdot \binom{n/2}{p/2} \cdot \binom{(n-p)/2}{q/2} \cdot 2^{q/2} \\ &\leq n^4 \cdot \binom{n/2}{p/2} \cdot \sum_{i=0}^{(n-p)/2} \binom{(n-p)/2}{i} \cdot 2^i \\ &= n^4 \cdot \binom{n/2}{p/2} \cdot 3^{(n-p)/2} \quad (\text{by binomial theorem}) \end{aligned}$$

Since we are in $\mathcal{R}_2$, we have $p \geq p_0 = 16n/50 \geq (n/2)/2$, which implies that the term $\binom{n/2}{p/2}$ decreases as p increases. Also the term $3^{(n-p)/2}$ decreases as p increases. Hence, $n^4 \cdot \binom{n/2}{p/2} \cdot 3^{(n-p)/2} \leq n^4 \cdot \binom{n/2}{p_0/2} \cdot 3^{(n-p_0)/2}$. Therefore, we obtain

$$\frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^4 \cdot \binom{n/2}{p_0/2} \cdot 3^{(n-p_0)/2}$$

We apply Theorem 18 to the last expression to get the desired upper bound on $\frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|}$:

$$\begin{aligned} \frac{S(p, q, r)}{|\text{sol}(F) \cap S(p, q, r)|} &\leq n^4 \cdot 2^{(n/2)H(p_0/n)} \cdot 3^{(n-p_0)/2} = n^4 \cdot \left(2^{H(16/25)/2} \cdot 3^{9/50} \right)^n \\ &< n^4 \cdot (1.7)^n < n^K \cdot (9/5)^n \end{aligned}$$

where H is the binary entropy function and $K = 4$. ◀

We now prove our remaining claim, an inequality involving binomial coefficients.

27:14 Optimal Depth-Three Circuits for Inner Product

Proof of Claim 23. This will follow by couple of applications of Theorem 18. We compute:

$$\begin{aligned}
\binom{m}{k} &\leq 2^{mH(k/m)} && \text{(applying Theorem 18)} \\
&= (m/2 + 1)^2 \left(\frac{1}{(m/2 + 1)^2} \cdot 2^{(m/2)H((k/2)/(m/2))} \right)^2 \\
&\leq (m/2 + 1)^2 \binom{m/2}{k/2}^2 && \text{(applying Theorem 18)}
\end{aligned}$$

as desired. $\triangleleft$

7.3 Constructing 2-CNFs for orbits in region 3

We here prove Lemma 16 for region 3. We need the following lemma regarding the coefficients of certain generating functions which will be proved in Section 8.2.

► **Lemma 24.** *There exists a constant K such that the following holds. Let $B_1, C_1, B_2, C_2, p, q, r, n$ be nonnegative integers such that $n \geq K$, $p \leq 2 \cdot \min(B_1, B_2)$, $n = p + q + r$, $r \leq n - 8$, $B_1 = 0.34n$, $C_1 = 0.16n$, $B_2 = 0.465n$, $C_2 = 0.035n$. Then, for $i \in [2]$, the coefficient of the monomial $x^p y^q z^r$ in $P_i(x, y, z) = (x^2 + (y + z)^2)^{B_i} ((2y + z)^2)^{C_i}$ is at least $\frac{1}{n^K} \frac{(2v_i+1)^{2C_i} (u_i+(v_i+1)^2)^{B_i}}{u_i^{p/2} v_i^q}$ where $v_i = \frac{-\beta_i + \sqrt{\beta_i^2 + 8rq}}{4r}$, $u_i = \frac{p(v_i+1)^2}{2B_i-p}$, and $\beta_i = 4C_i + (2B_i - p) - 3q$.*

We also need the following lemma regarding the solution of a bounded optimization problem which is solved using an optimization solver:

► **Lemma 25.** *Let $\hat{p}, \hat{q}, \hat{r}$ be arbitrary non-negative reals satisfying the following constraints:*

- $\hat{p} + \hat{q} + \hat{r} = 1$.
- $\frac{1}{2} - \frac{25}{32}\hat{p} \geq 0$.
- $\frac{5}{4} - \frac{25}{32}\hat{p} - \frac{25}{4}\hat{r} \leq 0$.
- $\frac{1}{4} - \hat{p} \leq 0$.

Let $\hat{B}_1 = 0.34, \hat{C}_1 = 0.16, \hat{B}_2 = 0.465, \hat{C}_2 = 0.035$. For $i \in [2]$, let $\hat{\beta}_i = 4\hat{C}_i + (2\hat{B}_i - \hat{p}) - 3\hat{q}$, $v_i = \frac{-\hat{\beta}_i + \sqrt{\hat{\beta}_i^2 + 8\hat{r}\hat{q}}}{4\hat{r}}$. Finally, let

$$\begin{aligned}
T_i &= -\hat{p} \log(\hat{p}) - \hat{q} \log(\hat{q}) - \hat{r} \log(\hat{r}) + \hat{q} + \\
&\quad \hat{q} \log(v_i) - 2\hat{C}_i \log(2v_i + 1) - (2\hat{B}_i - \hat{p}) \log(v_i + 1) + \\
&\quad 0.5(2\hat{B}_i - \hat{p}) \log(2\hat{B}_i - \hat{p}) + 0.5\hat{p} \log(\hat{p}) - \hat{B}_i \log(2\hat{B}_i)
\end{aligned}$$

Then, $\min(T_1, T_2) \leq 0.841$.

Proof. This is an optimization problem where we maximize $\min(T_1, T_2)$ over all $\hat{p}, \hat{q}, \hat{r}$ subject to the constraints. In particular, we substitute $\hat{q} = 1 - \hat{p} - \hat{r}$ and then maximize over all values of $\hat{p}, \hat{r}$ subject to the constraints. To solve this optimization problem, we use IbexOpt, a well known optimization tool that uses numeric interval arithmetic based library IBEX in C++ to solve global optimization problems [14]. We provide our code for this at <https://github.com/mjguru/Inner-Product/blob/main/region3.bch>. We note that the objective function we use in the code uses upper and lower bounds when evaluating logarithms to avoid numerical stability issues. These bounds only increase our objective function and so our upper bound still holds. $\blacktriangleleft$

We are now ready to prove Lemma 16 for region 3.

Proof of Lemma 16 for region 3. Let $m_3 = 2000$. Consider any $(p, q, r) \in \mathcal{R}_3$ such that m divides each of p, q, r and hence also $n = p + q + r$. Since p is even, we will construct 2-CNF F that is consistent with IP_n^0 such that $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{K_0} \cdot (9/5)^n$ for some constant K_0 .

When $r = n$, we have $|S(p, q, r) = 1|$ and in that case, we trivially construct a consistent 2-CNF covering the solution. Otherwise, since m_3 divides r and $r \leq n$, it must be that $r \leq n - m \leq n - 8$ and we use this fact when we apply Lemma 24.

Let $B_1 = 0.34n, C_1 = 0.16n, B_2 = 0.465n, C_2 = 0.035n$. We easily see that each of them is a positive even integer. For $i \in [2]$, let F_i be the 2-CNF obtained by a disjoint conjunction of B_i copies of **2Imp** and $2C_i$ copies of **NAND**. Since B_i and C_i are even, we infer that F_i is consistent with IP_n^0 . We will let F equal one of F_1, F_2 , whichever one attains larger value of $|\text{sol}(F_i) \cap S(p, q, r)|$.

The spectrum of F_i is $P_i = (x^2 + (y + z)^2)^{B_i} ((2y + z)^2)^{C_i}$. The coefficient of $x^p y^q z^r$ in P_i equals $|\text{sol}(F_i) \cap S(p, q, r)|$.

Recall that $|S(p, q, r)| = \binom{n}{p} \binom{n-p}{r} \cdot 2^q$. We apply Theorem 18 to get

$$|S(p, q, r)| \leq 2^{-p \log(p/n) - q \log(q/n) - r \log(r/n) + q}.$$

Using Lemma 24 for each P_i , we infer that

$$|\text{sol}(F_i) \cap S(p, q, r)| \geq n^{-K_1} \cdot \frac{(2v_i + 1)^{2C_i} (u_i + (v_i + 1)^2)^{B_i}}{(u_i)^{p/2} v_i^q}$$

where $v_i = \frac{-\beta_i + \sqrt{\beta_i^2 + 8rq}}{4r}$, $u_i = \frac{p(v_i + 1)^2}{2B_i - p}$, and $\beta_i = 4C_i + (2B_i - p) - 3q$. Combing the last two inequalities, we get

$$\frac{|S(p, q, r)|}{|\text{sol}(F_i) \cap S(p, q, r)|} \leq \frac{2^{-p \log(p/n) - q \log(q/n) - r \log(r/n) + q} (u_i)^{p/2} (v_i)^q}{n^{-K_1} (2v_i + 1)^{2C_i} (u_i + (v_i + 1)^2)^{B_i}}.$$

Let $\sigma_i = \frac{|S(p, q, r)|}{|\text{sol}(F_i) \cap S(p, q, r)|}$ and let $\hat{p} = \frac{p}{n}, \hat{q} = \frac{q}{n}, \hat{r} = \frac{r}{n}, \hat{B}_i = \frac{B_i}{n}, \hat{C}_i = \frac{C_i}{n}$. Then we rewrite:

$$\begin{aligned} \frac{\log(\sigma_i)}{n} - K_1 \log(n) &\leq -\hat{p} \log(\hat{p}) - \hat{q} \log(\hat{q}) - \hat{r} \log(\hat{r}) + \hat{q} + \\ &\quad 0.5\hat{p} \log(u_i) + \hat{q} \log(v_i) - 2\hat{C}_i \log(2v_i + 1) - 2\hat{B}_i \log(u_i + (v_i + 1)^2). \end{aligned}$$

We substitute $u_i = \frac{p(v_i + 1)^2}{2B_i - p}$ and simplify to obtain:

$$\begin{aligned} \frac{\log(\sigma_i)}{n} - K_1 \log(n) &\leq \hat{p} \log(\hat{p}) + \hat{q} \log(\hat{q}) + \hat{r} \log(\hat{r}) + \hat{q} + \\ &\quad \hat{q} \log(v_i) - 2\hat{C}_i \log(2v_i + 1) - (2\hat{B}_i - \hat{p}) \log(v_i + 1) + \\ &\quad 0.5(2\hat{B}_i - \hat{p}) \log(2\hat{B}_i - \hat{p}) + 0.5\hat{p} \log(\hat{p}) - \hat{B}_i \log(2\hat{B}_i) \end{aligned}$$

We now apply Lemma 25, verifying that we satisfy all the constraints, to infer that

$$\min \left(\frac{\log(\sigma_1)}{n} - K_1 \log(n), \frac{\log(\sigma_2)}{n} - K_1 \log(n) \right) \leq 0.8398 \quad (2)$$

By definition of F , we have that $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} = \min(\sigma_1, \sigma_2)$. Hence, we compute that

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &= \min(\sigma_1, \sigma_2) \\ &= 2^{n \cdot \min\left(\frac{\log(\sigma_1)}{n}, \frac{\log(\sigma_2)}{n}\right)} \\ &\leq n^{K_1} \cdot 2^{n \cdot 0.8398} \quad (\text{using Equation (2)}) \\ &\leq n^{K_1} \cdot (9/5)^n \end{aligned}$$

as desired. ◀

7.4 Constructing 2-CNFs for orbits in region 4

In this subsection, We prove Lemma 16 for region 4. We need the following lemma regarding the coefficients of certain generating functions which itself will be proved in Section 8.3

► **Lemma 26.** *There exists a constant K such that the following holds. Let $A_1, C_1, A_2, C_2, p, q, r, n$ be nonnegative integers such that $n \geq K_0$, $p \leq 2 \cdot \min(A_1, A_2)$, $n = p + q + r$, $0 < r \leq n - 8$, $A_1 = 0.34n$, $C_1 = 0.16n$, $A_2 = 0.355n$, $C_2 = 0.145n$. Also, for $i \in [2]$, define the degree 3 real polynomial $Q_i(x) = (4r)x^3 + (4A_i - 2p - 2q)x^2 + (4C_i - 2q)x - q$. Then, for $i \in [2]$, the following holds: (1) $Q_i(x)$ has a unique non-negative real root. (2) The coefficient of the monomial $x^p y^q z^r$ in $P_i(x, y, z) = (x^2 + 2y^2 + z^2)^{A_i} ((2y + z)^2)^{C_i}$ is at least $\frac{1}{n^K} \frac{(2v_i+1)^{2C_i} (u_i+2v_i^2+1)^{A_i}}{u_i^{p/2} v_i^q}$ where v_i is the unique non-negative root of the cubic polynomial $Q_i(x)$, and $u_i = \frac{p(2v_i^2+1)}{2A_i-p}$.*

We also need the solution of a bounded optimization problem which is obtained using an optimization solver.

► **Lemma 27.** *Let $\hat{p}, \hat{q}, \hat{r}$ be arbitrary non-negative reals satisfying the following constraints:*

- $\hat{p} + \hat{q} + \hat{r} = 1$.
- $\frac{1}{2} - \frac{25}{32}\hat{p} \geq 0$.
- $-\frac{5}{4} + \frac{25}{16}\hat{p} + \frac{25}{4}\hat{r} \leq 0$.
- $-\frac{1}{20} + \hat{r} \geq 0$.
- $\frac{1}{4} - \hat{p} \leq 0$.

Let $\hat{A}_1 = 0.34, \hat{C}_1 = 0.16, \hat{A}_2 = 0.355, \hat{C}_2 = 0.145$. For $i \in [2]$, let v_i be the unique non-negative root of the cubic polynomial $P_i(x) = (4\hat{r})x^3 + (4\hat{A}_i - 2\hat{p} - 2\hat{q})x^2 + (4\hat{C}_i - 2\hat{q})x - \hat{q}$, and let $u_i = \frac{p(2v_i^2+1)}{2\hat{A}_i-\hat{p}}$. Finally, let

$$\begin{aligned} T_i = & -\hat{p} \log(\hat{p}) - \hat{q} \log(\hat{q}) - \hat{r} \log(\hat{r}) + \hat{q} + \\ & -2\hat{C}_i \log(2\hat{v}_i + 1) - \hat{A}_i \log(\hat{u}_i + 2\hat{v}_i^2 + 1) + \\ & (\hat{p}/2) \log(\hat{u}_i) + \hat{q} \log(\hat{v}_i) \end{aligned}$$

Then, $\min(T_1, T_2) \leq 0.845$.

Proof. We will first assert the existence and uniqueness of the nonnegative real root of the polynomial $P_i(x)$ which follows from Lemma 26.

We now describe how we obtain the upper bound on $\min(T_1, T_2)$. This as an optimization problem where we maximize $\min(T_1, T_2)$ over all $\hat{p}, \hat{q}, \hat{r}$ subject to the constraints. In particular, we substitute $\hat{q} = 1 - \hat{p} - \hat{r}$ and then maximize over all values of $\hat{p}, \hat{r}$ subject to the constraints. To solve this optimization problem, we use IbexOpt, a well known optimization tool that uses numeric interval arithmetic based library IBEX in C++ to solve global optimization problems [14]. We provide our code for this at <https://github.com/mjguru/Inner-Product/blob/main/region4.bch>. We note that the objective function we use in the code uses upper and lower bounds when evaluating logarithms to avoid numerical stability issues. These bounds only increase our objective function and so our upper bound still holds. ◀

We are now ready to prove Lemma 16 for region 4.

Proof of Lemma 16 for region 4. We let $m_4 = 2000$. Consider any $(p, q, r) \in \mathcal{R}_4$ such that m divides each of p, q, r and hence also $n = p + q + r$. Since p is even, we will construct a 2-CNF F that is consistent with IP_n^0 such that $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{K_0} \cdot (9/5)^n$ where K_0 is a constant.

Let $A_1 = 0.34n, C_1 = 0.16n, A_2 = 0.355n, C_2 = 0.145n$. Since 2000 divides n , each of these is a positive even integer. For $i \in [2]$, let F_i be the 2-CNF obtained by a disjoint conjunction of A_i copies of **Matching** and $2C_i$ copies of **NAND**. Using Corollary 13 and the fact that each of A_i and C_i are even integers, we infer that F_i is indeed consistent with IP_n^0 . Let F equal one of F_1, F_2 , whichever one attains larger value of $|\text{sol}(F_i) \cap S(p, q, r)|$.

The spectrum of F_i is $P_i = (x^2 + 2y^2 + z^2)^{A_i}((2y + z)^2)^{C_i}$. The coefficient of $x^p y^q z^r$ in P_i equals $|\text{sol}(F_i) \cap S(p, q, r)|$.

We know that $|S(p, q, r)| = \binom{n}{p} \binom{n-p}{r} \cdot 2^q$. We apply Theorem 18 and simplify using the fact $p + q + r = n$ to infer that $|S(p, q, r)| \leq 2^{-p \log(p/n) - q \log(q/n) - r \log(r/n) + q}$.

Using Lemma 26 for each P_i , we infer that

$$|\text{sol}(F_i) \cap S(p, q, r)| \geq n^{-K_1} \cdot \frac{(2v_i + 1)^{2C_i} (u_i + 2v_i^2 + 1)^{A_i}}{(u_i)^{p/2} v_i^q}$$

where v_i is the unique non-negative root of the polynomial $(4r)x^3 + (4A_i - 2p - 2q)x^2 + (4C_i - 2q)x - q$, $u_i = \frac{p(2v_i^2 + 1)}{2A_i - p}$ and K_1 a constant. We now have

$$\frac{|S(p, q, r)|}{|\text{sol}(F_i) \cap S(p, q, r)|} \leq \frac{2^{-p \log(p/n) - q \log(q/n) - r \log(r/n) + q} (u_i)^{p/2} (v_i)^q}{n^{-K_1} (2v_i + 1)^{2C_i} (u_i + 2v_i^2 + 1)^{A_i}}.$$

Let $\sigma_i = \frac{|S(p, q, r)|}{|\text{sol}(F_i) \cap S(p, q, r)|}$ and let $\hat{p} = \frac{p}{n}, \hat{q} = \frac{q}{n}, \hat{r} = \frac{r}{n}, \hat{B}_i = \frac{B_i}{n}, \hat{C}_i = \frac{C_i}{n}$. We write the inequality as

$$\begin{aligned} \frac{\log(\sigma_i)}{n} - K_1 \log(n) &\leq -\hat{p} \log(\hat{p}) - \hat{q} \log(\hat{q}) - \hat{r} \log(\hat{r}) + \hat{q} + \\ &\quad 2\hat{C}_i \log(2v_i + 1) - \hat{A}_i \log(u_i + 2v_i^2 + 1) + \\ &\quad (\hat{p}/2) \log(u_i) + \hat{q} \log(v_i) \end{aligned}$$

After verifying that we satisfy all the constraints of the lemma (since $(p, q, r) \in \mathcal{R}_5$, we must have that $r \leq 0.9n \leq n - 8$), we now apply Lemma 27, to infer that

$$\min \left(\frac{\log(\sigma_1)}{n} - K_1 \log(n), \frac{\log(\sigma_2)}{n} - K_1 \log(n) \right) \leq 0.844 \quad (2)$$

By definition of F , we have that $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} = \min(\sigma_1, \sigma_2)$. Hence, we conclude that

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &= \min(\sigma_1, \sigma_2) \\ &= 2^{n \cdot \min\left(\frac{\log(\sigma_1)}{n}, \frac{\log(\sigma_2)}{n}\right)} \\ &\leq n^{K_1} \cdot 2^{n \cdot 0.844} \quad (\text{using Equation (2)}) \\ &\leq n^{K_1} \cdot (9/5)^n \end{aligned}$$

as desired. ◀

7.5 Constructing 2-CNFs for orbits in region 5

In this subsection, we prove Lemma 16 for orbits in $\mathcal{R}_5$. We need the following result which we prove towards the end of this subsection:

▷ **Claim 28.** Let $g : (0, 0.64] \rightarrow \mathbb{R}$ be defined as $g(y) = H(y) + \varepsilon_R H(\varepsilon_R) + a - \frac{y}{2} - aH(y/2a)$ where $\varepsilon_R = \frac{1}{20}$, $a = 0.355$ and H is the binary entropy function. Then, $\max_{y \in [0, 0.64]} g(y) \leq 0.828$.

We first show how Lemma 16 follows.

Proof of Lemma 16 for region 5. Let $m_5 = 2000$ and let $(p, q, r) \in \mathcal{R}_5$ be an arbitrary orbit such that m divides each of p , q , and r and hence n . Since p is even, it suffices to construct a 2-CNF F that is consistent with IP_n^0 and $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{K_0} (9/5)^n$ where K_0 is a constant.

Let $A = 0.355n$, $C = 0.145n$. Let F be the 2-CNF obtained by disjoint conjunctions of A copies of **Matching** and $2C$ copies of **NAND** (as defined in Definition 10). Since both these building blocks are consistent with IP_n^0 , we have that F is consistent with IP_n^0 as desired.

We will now show that $\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq n^{K_0} (9/5)^n$. The spectrum of F is given by $P(x, y, z) = (x^2 + 2y^2 + z^2)^A (2y + z)^{2C}$ and $|\text{sol}(F) \cap S(p, q, r)|$ equals the coefficient of $x^p y^q z^r$ in $P(x, y, z)$. We lower bound this coefficient as follows: Choose r out of $2C$ terms of $(2y + z)$ to equal z and remaining to equal $2y$; then choose $p/2$ terms out of A_i terms to be x^2 and remaining terms to equal $2y^2$ to obtain the monomial $2^{A-p/2+2C-r} \cdot x^p y^q z^r$. Since there are $\binom{2C}{r} \cdot \binom{A}{p/2} \geq \binom{A}{p/2}$ of obtaining such a monomial, we lower bound its coefficient by $\binom{A}{p/2} \cdot 2^{A+2C-p/2-r}$. Since $p \leq 0.64n$ and $r \leq 10^{-5}n$, we have $\binom{n-p}{r} \leq \binom{n}{r}$. Since $|S(p, q, r)| = \binom{n}{p} \binom{n-p}{r} \cdot 2^q$, we get

$$\begin{aligned} \frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} &\leq \frac{\binom{n}{p} \binom{n-p}{r} 2^q}{\binom{A}{p/2} 2^{A+2C-p/2-r}} \\ &= \frac{\binom{n}{p} \binom{n-p}{r} 2^{A-p/2}}{\binom{A}{p/2}} \end{aligned}$$

where in the last line we used the fact $n = 2A + 2C = p + q + r$.

We now repeatedly apply Theorem 18 to the last expression to get the following bound:

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq (A+1) \cdot 2^{n \cdot H(p/n) + r \cdot H(r/n) + A - p/2 - A \cdot H(p/(2A))}$$

where H is the binary entropy function. Using the fact that $A+1 \leq n$ and rearranging and taking logarithm of both sides, we have

$$\frac{\log \left(\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \right)}{n} - \log(n) \leq H(\hat{p}) + \hat{r} \cdot H(\hat{r}) + \hat{A} - \hat{p}/2 - \hat{A} \cdot H(\hat{p}/(2\hat{A}))$$

where $\hat{A} = \frac{A}{n}$, $\hat{p} = \frac{p}{n}$, $\hat{r} = \frac{r}{n}$. Since $0 \leq \hat{r} \leq \frac{1}{20}$ and $\hat{r} \cdot H(\hat{r})$ is an increasing function in that range, we infer

$$\frac{\log \left(\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \right)}{n} - \log(n) \leq H(\hat{p}) + \frac{1}{20} \cdot H\left(\frac{1}{20}\right) + \hat{A} - \hat{p}/2 - \hat{A} \cdot H(\hat{p}/(2\hat{A}))$$

Let $g(y) = H(y) + \frac{1}{20}H\left(\frac{1}{20}\right) + \hat{A} - y/2 - \hat{A} \cdot H(y/2\hat{A})$ where $\hat{A} = 0.355$. Then, using Claim 28, we get that $g(\hat{p}) \leq 0.828$ and

$$\frac{\log\left(\frac{|S(p,q,r)|}{|\text{sol}(F) \cap S(p,q,r)|}\right)}{n} - \log(n) \leq 0.828$$

Rearranging and considering exponents, we conclude

$$\frac{|S(p,q,r)|}{|\text{sol}(F) \cap S(p,q,r)|} \leq n \cdot 2^{0.828n} \leq n^1 \cdot (9/5)^n$$

as desired. ◀

We now prove the claim regarding the real function g :

Proof of Claim 28. We will show that for all $y \in (0, 0.64]$, $g'(y) > 0$ which lets us conclude that $\max_{y \in [0, 0.64]} g(y) = g(0.64) \leq 0.828$ as desired.

Recall that the derivative of the binary entropy function H is $H'(x) = \log((1-x)/x)$. Using this, we compute that

$$\begin{aligned} g'(y) &= \log((1-y)/y) - \frac{1}{2} - \frac{1}{2} \log((2a-y)/y) \\ &= \log\left(\frac{1-y}{\sqrt{2y(2a-y)}}\right) \end{aligned}$$

Therefore, to show that $g'(y) > 0$ for all $y \in (0, 0.64]$, it suffices to show that for y in this domain, $1-y > \sqrt{2y(2a-y)}$. Since both sides of this inequality are positive, this is equivalent to showing that $(1-y)^2 > 2y(2a-y)$. This is equivalent to showing $3y^2 - y(4a+2) + 1 > 0$. Let $h(y) = 3y^2 - y(4a+2) + 1$. We will show that $h(y) > 0$ for all y . The discriminant of this quadratic $h(y)$ is $(4a+2)^2 - 12 = (3.42)^2 - 12 < 0$. This means $h(y)$ has no real roots. Since $h(0) = 1 > 0$, we conclude that $h(y) > 0$ for all y as desired. ◀

7.6 Constructing 2-CNFs for orbits in region 6

Proof of Lemma 16 for region 6. Let $m_6 = 2000$ and let $(p, q, r) \in \mathcal{R}_6$ be an arbitrary orbit such that m divides each of p , q and r and hence n . Since p is even, it suffices to show that we can construct a 2-CNF F that is consistent with IP_n^0 such that $\frac{|S(p,q,r)|}{|\text{sol}(F) \cap S(p,q,r)|} \leq n^{K_0}(9/5)^n$ for some constant K_0 .

Let $A = p/2$, $C = (n-p)/2$. Let F be the 2-CNF obtained by a disjoint conjunction of A copies of **Matching** and $2C$ copies of **NAND**. Since both these building blocks are consistent with IP_n^0 , we have that F is consistent with IP_n^0 as desired.

We will now show that $\frac{|S(p,q,r)|}{|\text{sol}(F) \cap S(p,q,r)|} \leq n^{K_0}(9/5)^n$. The spectrum of F is given by $P(x, y, z) = (x^2 + 2y^2 + z^2)^A (2y + z)^{2C}$ and $|\text{sol}(F) \cap S(p, q, r)|$ equals the coefficient of $x^p y^q z^r$ in $P(x, y, z)$. Since $A = p/2$, this coefficient exactly equals $\binom{2C}{q} \cdot 2^q = \binom{n-p}{q} \cdot 2^q$. Also we have that $|S(p, q, r)| = \binom{n}{p} \binom{n-p}{q} \cdot 2^q$. Therefore,

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} = \binom{n}{p}.$$

We apply Theorem 18 to get that

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq 2^{n \cdot H(p/n)}$$

27:20 Optimal Depth-Three Circuits for Inner Product

Since $0 \leq p \leq \frac{n}{4}$ and $H(\cdot)$, the binary entropy function, is increasing in this range, we obtain

$$\frac{|S(p, q, r)|}{|\text{sol}(F) \cap S(p, q, r)|} \leq 2^{n \cdot H(1/4)} \leq 2^{n \cdot (0.82)} \leq (9/5)^n$$

as desired. ◀

8 Coefficient extraction for each region

In this section, we obtain asymptotic bounds for coefficients of monomials of particular generating functions. To prove these results, we require the following lemma which bounds the coefficients of a power series:

► **Lemma 29.** *Let $p, q \in \mathbb{N}$ be such that $p^2 + q^2 \geq 4$. Let $f : \mathbb{C}^2 \rightarrow \mathbb{C}$ be analytic on a compact neighborhood $\mathcal{N}$ around $(0, 0)$. Let $f(u, v) = \sum_{r=(r_1, r_2) \in \mathbb{N}^2} C_r u^{r_1} v^{r_2}$ where for C_r are non-negative. Let $h : \mathbb{C}^2 \rightarrow \mathbb{C}$ be defined as $h(u, v) = \ln(f(u, v)) - p \ln(u) - q \ln(v)$. Assume there exist non-negative real u_0 and v_0 such that $\nabla_h(u_0, v_0) = 0$ and that $\mathcal{H}_h(u_0, v_0)$ is positive definite. Assume that $|f(u, v)|$ attains its unique global maximum on the torus $|u| = u_0, |v| = v_0$ at (u_0, v_0) . Then,*

$$C_{p,q} = \frac{f(u_0, v_0)}{2\pi u_0^{p+1} v_0^{q+1} \sqrt{\det(\mathcal{H}_h(u_0, v_0))}} \left(1 + O\left(\frac{1}{\sqrt{p^2 + q^2}}\right) \right).$$

where $\mathcal{H}_h(u_0, v_0)$ is the Hessian of h evaluated at u_0, v_0 .

This result seems to be a standard result in analytic combinatorics but since we could not find a ready reference to the formulation in Lemma 29, we will provide a somewhat self-contained proof in Appendix A. For now, we use Lemma 29 to obtain bounds on the coefficients of the generating functions that arise from the constructions of IP.

8.1 Bound for Region 1

Proof of Lemma 22. Let $f(u, v) = P(u^{1/2}, v, 1)$. Since the degree of x is even in $P(x, y, z)$, $2A + 2B + 2C = p + q + r = n$, and each of the three factor polynomials is homogeneous, we infer that the coefficient of $u^{p/2} v^q$ in $f(u, v)$ equals the coefficient of $x^p y^q z^r$ in $P(x, y, z)$.

Let $h : \mathbb{C}^2 \rightarrow \mathbb{C}$ as $h(u, v) = \ln(f(u, v)) - (p/2) \ln(u) - q \ln(v)$. We observe that $\nabla_h(16, 2) = 0$. We argue that the Hessian $\mathcal{H}_h(16, 2)$ of h at $u = 16$ and $v = 2$ is positive definite. To show this, let $\alpha(x, y) = h(\exp(x), \exp(y))$. We know the function α is convex since that h has nonnegative coefficients in its power series (see section 4.5 from [2]). We also observe that α is strictly convex and hence $\mathcal{H}_\alpha(16, 2)$ is positive definite. Since $\nabla_h(16, 2) = 0$, we apply the chain rule to get that $\mathcal{H}_\alpha(16, 2) = D \cdot \mathcal{H}_h(16, 2) \cdot D$ where $D = \begin{bmatrix} 16 & 0 \\ 0 & 2 \end{bmatrix}$ is a diagonal matrix. Since $\mathcal{H}_\alpha(16, 2)$ is positive definite and D is positive definite, we get that $\mathcal{H}_h(16, 2)$ is positive definite as well. We next apply triangle inequality to infer that in the (complex) torus defined by $|u| = 16, |v| = 2$, the unique global maximum occurs at $u = 16$ and $v = 2$. Lastly, we claim that at least one of p or q must have value at least $0.01n$. Indeed, if they both are smaller, then it must be that $r \geq 0.98n$, implying that $A < 0$. However, that is a contradiction since we assume A is a nonnegative integer. Let C_0 be large enough so we have $(p/2)^2 + q^2 \geq 4$ for $n \geq C_0$.

We satisfy the conditions to apply Lemma 29 with $u_0 = 16, v_0 = 2$ to extract the coefficient $u^{p/2}v^q$. We make sure that C_0 is large enough to conclude that the coefficient $u^{p/2}v^q$ equals

$$\frac{25^{A+B+C}}{C_0 4^p 2^q \text{poly}(p, q)} \left(1 + \frac{1}{\sqrt{p^2 + q^2}}\right) \geq \frac{5^n}{C_0 4^p 2^q \text{poly}(n)} \left(1 + \frac{1}{\text{poly}(n)}\right)$$

where in the last equality we used the fact that either p or q will be at least $0.01n$. $\blacktriangleleft$

8.2 Bound for region 3

In this subsection, we obtain asymptotic bounds on the coefficients of generating functions related to $\mathcal{R}_3$.

Proof of Lemma 24. For $i \in [2]$, let $f_i(u, v) = P_i(u^{1/2}, v, 1)$. Since x appears in squared form, f_i is still product of three polynomials. Furthermore, since $2B_i + 2C_i = p + q + r = n$, and both the polynomials are homogeneous, we infer that the coefficient of $u^{p/2}v^q$ in $f_i(x, y, z)$ equals the coefficient of $x^p y^q z^r$ in $P_i(x, y, z)$.

Define $h_i : \mathbb{C}^2 \rightarrow \mathbb{C}$ by $h(u, v) = \ln(f_i(u, v)) - (p/2) \ln(u) - q \ln(v)$. We will find non-negative $(u_i, v_i) \in \mathbb{R}^2$ such that $\nabla_{h_i}(u_i, v_i) = 0$. We compute that $\frac{\partial h_i}{\partial u} = -\frac{p}{2} \cdot \frac{1}{u} + \frac{B_i}{u+(v+1)^2}$. Setting this to 0 and rearranging, we infer that

$$u_i = \frac{p(v_i + 1)^2}{2B_i - p} \quad (1)$$

We next compute that $\frac{\partial h_i}{\partial v} = \frac{4C_i}{2v+1} + \frac{2B_i(v+1)}{u+(v+1)^2} - \frac{q}{v}$. Setting this to 0, rearranging, substituting in Equation (1), and using the fact that $4C_i + 2(2B_i - p) - 2q = 2r$ (since $2B_i + 2C_i = p + q + r = n$), we infer that v_i satisfies the following quadratic equation:

$$2r \cdot (v_i)^2 + \beta_i \cdot v_i - q = 0$$

where $\beta_i = 4C_i + (2B_i - p) - 3q$. This means v_i equals $\frac{\beta_i \pm \sqrt{\beta_i^2 + 8rq}}{4r}$. By the inequalities satisfied by p, q, r , it must be the case that rq is at least a constant. This makes the value inside the square root comes out to be strictly larger than β_i^2 and so one root is strictly negative and other is strictly positive. We reject the negative root since Lemma 29 requires the special point (u_i, v_i) to have non-negative real values.

With this value of u_i, v_i , we show that $\mathcal{H}_{h_i}(u_i, v_i)$ is positive definite where $\mathcal{H}_{h_i}$ is the Hessian of h_i . For this, let $\alpha_i(x, y) = h_i(\exp(x), \exp(y))$. Using the fact that h has nonnegative coefficients in its power series, we use the well known result (see section 4.5 from [2]) that the function α_i as we defined is convex. Also by picking any three non-zero terms of the power series and checking that they don't lie in single line, we see that α_i is strictly convex and so $\mathcal{H}_{\alpha_i}(u_i, v_i)$ is positive definite. Using the fact that $\nabla_{h_i}(u_i, v_i) = 0$ and chain rule, we obtain that $\mathcal{H}_{\alpha_i}(u_i, v_i) = D \cdot \mathcal{H}_{h_i}(u_i, v_i) \cdot D$ where D is the diagonal matrix with diagonal entries u_i and v_i . Since D is positive definite, we infer that $\mathcal{H}_{h_i}(u_i, v_i)$ is indeed positive definite.

We next apply the triangle inequality to infer that in the complex torus with $|u| = u_i, |v| = v_i$, the unique global maximum for f_i occurs at (u_i, v_i) (carefully checking that the values at $(\pm u_i, \pm v_i)$ are smaller). Also since $r \leq n - 8, p + q + r = n$ and $n \geq K_0$ where we will set K_0 to be large enough constant, it must be the case that $(p/2)^2 + q^2 \geq 4$. Therefore we satisfy all conditions to apply Lemma 29 with $(u_0, v_0) = (u_i, v_i)$, we infer that the coefficient of $u^{p/2}v^q$ in $f_i(u, v)$ equals

$$\frac{f(u_i, v_i)}{2\pi u_i^{p+1} v_i^{q+1} \sqrt{\mathcal{H}_{h_i}(u_i, v_i)}} \left(1 + O\left(\frac{1}{\sqrt{p^2 + q^2}}\right)\right).$$

Since $p^2 + q^2 \geq 4$, and $u_i, v_i \leq n$, and that the entries of $\mathcal{H}_{h_i}$ are bounded by $\text{poly}(u_i, v_i)$, we infer that there exists a constant K_1 such that the coefficient is at least

$$n^{-K_1} \cdot \frac{f(u_i, v_i)}{u_i^{p/2} v_i^q} = n^{-K_1} \cdot \frac{(u_i + (v_i + 1)^2)^{B_i} (2v_i + 1)^{2C_i}}{u_i^{p/2} v_i^q} \quad \blacktriangleleft$$

8.3 Bound for region 4

In this subsection, we obtain asymptotic bounds on the coefficients of generating functions related to $\mathcal{R}_4$.

Proof of Lemma 26. We first prove that Q_i has a unique real nonnegative root. First, since Q_i has degree 3, it has at least one real root. Let x_0 be any such real root. Since $Q_i(x_0) = 0$, we rearrange to infer that x_0 must satisfy:

$$\frac{4C_i x_0}{2x_0 + 1} + \frac{2x_0^2(2A_i - p)}{2x_0^2 + 1} = q$$

We see that the left side at $x = 0$ equals 0 and as $x \rightarrow +\infty$, it approaches $2C_i + 2A_i - p = q + r$ (since $n = 2A_i + 2C_i = p + q + r$). Since q and r are nonnegative and right side equals q , by intermediate value theorem, there must exist a unique value $y \geq 0$ such that left side equals q , proving our result.

We now prove the asymptotic bound. For $i \in [2]$, let $f_i(u, v) = P_i(u^{1/2}, v, 1)$. Since x appears in squared form, f_i is still product of three polynomials. Furthermore, since $2A_i + 2C_i = p + q + r = n$, and both the polynomials are homogeneous, we infer that the coefficient of $u^{p/2}v^q$ in $f_i(x, y, z)$ equals the coefficient of $x^p y^q z^r$ in $P_i(x, y, z)$.

Define $h_i : \mathbb{C}^2 \rightarrow \mathbb{C}$ by $h(u, v) = \ln(f_i(u, v)) - (p/2)\ln(u) - q\ln(v)$. We will find non-negative $(u_i, v_i) \in \mathbb{R}^2$ such that $\nabla_{h_i}(u_i, v_i) = 0$. We compute that $\frac{\partial h_i}{\partial u} = \frac{A_i}{u+2v^2+1} - \frac{p}{2u}$. Setting this to 0 and rearranging, we infer that

$$u_i = \frac{p(2v_i^2 + 1)}{2A_i - p} \quad (1)$$

We next compute that $\frac{\partial h_i}{\partial v} = \frac{4C_i}{2v+1} + \frac{4vA_i}{u+2v^2+1} - \frac{q}{v}$. Setting this to 0, rearranging, substituting in Equation (1), and using the fact that $8C_i + 4(2A_i - p) - 4q = 4r$ (since $2A_i + 2C_i = p + q + r = n$), we infer that v_i satisfies the following cubic equation:

$$(4r) \cdot v_i^3 + (4A_i - 2p - 2q) \cdot v_i^2 + (4C_i - 2q) \cdot v_i - q = 0$$

In other words, $Q_i(v_i) = 0$. From above, we know that there is a unique nonnegative root of Q_i and we let v_i be the unique nonnegative root.

With this value of u_i, v_i , we show that $\mathcal{H}_{h_i}(u_i, v_i)$ is positive definite where $\mathcal{H}_{h_i}$ is the Hessian of h_i . For this, let $\alpha_i(x, y) = h_i(\exp(x), \exp(y))$. Using the fact that h has nonnegative coefficients in its power series, we use the well known result (see section 4.5 from [2]) that the function α_i as we defined is convex. Also by picking any three non-zero terms of the power series and checking that they don't lie in single line, we see that α_i will be strictly convex and so $\mathcal{H}_{\alpha_i}(u_i, v_i)$ is positive definite. Using the fact that $\nabla_{h_i}(u_i, v_i) = 0$ and chain rule, we obtain that $\mathcal{H}_{\alpha_i}(u_i, v_i) = D \cdot \mathcal{H}_{h_i}(u_i, v_i) \cdot D$ where D is the diagonal matrix with diagonal entries u_i and v_i . Since D is positive definite, we infer that $\mathcal{H}_{h_i}(u_i, v_i)$ is indeed positive definite.

We next apply the triangle inequality to infer that in the complex torus with $|u| = u_i, |v| = v_i$, the unique global maximum for f_i occurs at (u_i, v_i) (carefully checking that the values at $(\pm u_i, \pm v_i)$ are smaller). Also since $r \leq n - 8$, $p + q + r = n$ and $n \geq K_0$

where we will set K_0 to be large enough constant, it must be the case that $(p/2)^2 + q^2 \geq 4$. Therefore we satisfy all conditions to apply Lemma 29 with $(u_0, v_0) = (u_i, v_i)$, we infer that the coefficient of $u^{p/2}v^q$ in $f_i(u, v)$ equals

$$\frac{f(u_i, v_i)}{2\pi u_i^{p+1} v_i^{q+1} \sqrt{\mathcal{H}_{h_i}(u_i, v_i)}} \left(1 + O \left(\frac{1}{\sqrt{p^2 + q^2}} \right) \right).$$

Since $p^2 + q^2 \geq 4$, and $u_i, v_i \leq n$, and that the entries of $\mathcal{H}_{h_i}$ are bounded by $\text{poly}(u_i, v_i)$, we infer that there exists a constant K_1 such that the coefficient is at least

$$n^{-K_1} \cdot \frac{f(u_i, v_i)}{u_i^{p/2} v_i^q} = n^{-K_1} \cdot \frac{(u_i + v_i^2 + 1)^{A_i} (2v_i + 1)^{2C_i}}{u_i^{p/2} v_i^q} \quad \blacktriangleleft$$

References

- 1 Kazuyuki Amano. Depth-three circuits for inner product and majority functions. In Satoru Iwata and Naonori Kakimura, editors, *34th International Symposium on Algorithms and Computation, ISAAC 2023, December 3-6, 2023, Kyoto, Japan*, volume 283 of *LIPIcs*, pages 7:1–7:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ISAAC.2023.7.
- 2 Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- 3 Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.
- 4 Vlado Dancík. Complexity of boolean functions over bases with unbounded fan-in gates. *Inf. Process. Lett.*, 57(1):31–34, 1996. doi:10.1016/0020-0190(95)00182-4.
- 5 Magnus Gausdal Find, Alexander Golovnev, Edward A. Hirsch, and Alexander S. Kulikov. Improving $3n$ circuit complexity lower bounds. *Comput. Complex.*, 32(2):13, 2023. doi:10.1007/S00037-023-00246-9.
- 6 Peter Frankl, Svyatoslav Gryaznov, and Navid Talebanfard. A variant of the VC-dimension with applications to depth-3 circuits. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 72:1–72:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.ITCS.2022.72.
- 7 Alexander Golovnev, Alexander S. Kulikov, and R. Ryan Williams. Circuit depth reductions. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*, volume 185 of *LIPIcs*, pages 24:1–24:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ITCS.2021.24.
- 8 Mika Göös, Ziyi Guan, and Tiberiu Mosnoi. Depth-3 circuits for inner product. *Inf. Comput.*, 300:105192, 2024. doi:10.1016/J.IC.2024.105192.
- 9 Johan Håstad, Stasys Jukna, and Pavel Pudlák. Top-down lower bounds for depth-three circuits. *Comput. Complex.*, 5(2):99–112, 1995. doi:10.1007/BF01268140.
- 10 Xuanguo Huang, Peter Ivanov, and Emanuele Viola. Affine extractors and AC0-parity. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, volume 245 of *LIPIcs*, pages 9:1–9:14, 2022. doi:10.4230/LIPIcs.APPROX/RANDOM.2022.9.
- 11 Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001. doi:10.1006/JCSS.2001.1774.
- 12 Jiayu Li and Tianqi Yang. $3.1n - o(n)$ circuit lower bounds for explicit functions. In *54th Annual ACM SIGACT Symposium on Theory of Computing, STOC*, pages 1180–1193. ACM, 2022. doi:10.1145/3519935.3519976.
- 13 Xin Li. Two source extractors for asymptotically optimal entropy, and (many) more. In *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 1271–1281. IEEE, 2023. doi:10.1109/FOCS57990.2023.00075.

- 14 Jordan Ninin. Global optimization based on contractor programming: An overview of the ibex library. In *International Conference on Mathematical Aspects of Computer and Information Sciences*, pages 555–559. Springer, 2015. doi:10.1007/978-3-319-32859-1_47.
- 15 Ramamohan Paturi, Pavel Pudlák, Michael E. Saks, and Francis Zane. An improved exponential-time algorithm for k -SAT. *J. ACM*, 52(3):337–364, 2005. doi:10.1145/1066100.1066101.
- 16 Ramamohan Paturi, Pavel Pudlák, and Francis Zane. Satisfiability coding lemma. *Chic. J. Theor. Comput. Sci.*, 1999, 1999. URL: <http://cjtcs.cs.uchicago.edu/articles/1999/11/contents.html>.
- 17 Ramamohan Paturi, Michael E. Saks, and Francis Zane. Exponential lower bounds for depth three boolean circuits. *Comput. Complex.*, 9(1):1–15, 2000. doi:10.1007/PL00001598.
- 18 Robin Pemantle, Mark C Wilson, and Stephen Melczer. *Analytic combinatorics in several variables*, volume 212. Cambridge University Press, 2024.
- 19 Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977)*, pages 162–176. Lecture Notes in Comput. Sci., Vol. 53, 1977.

A Proving coefficient asymptotics

In this section, we will prove our required lemma regarding bounding coefficients of power series of an analytic function - Lemma 29.

To help prove this, we will require the following well known result that coefficient expansion of a generating function can be expressed as a Cauchy integral (see equation 1.5 from [18]):

► **Theorem 30** (Bivariate Cauchy integral formula). *Let $f : \mathbb{C}^2 \rightarrow \mathbb{C}$ be analytic on a compact neighborhood $\mathcal{N}$ around $(0, 0)$. Let $f(u, v) = \sum_{r=(r_1, r_2) \in \mathbb{N}^2} C_r u^{r_1} v^{r_2}$ where for all $i \in \mathbb{N}^2$, $C_i \in \mathbb{R}$ and $C_i \geq 0$. Then, for all $r \in \mathbb{N}^2$, it holds that*

$$C_r = \oint_{x \in \mathcal{N}} \left(\frac{1}{2\pi i} \right)^2 f(x) \cdot x_1^{-r_1-1} \cdot x_2^{-r_2-1} dx.$$

We will utilize the following result regarding bounding Cauchy integral:

► **Theorem 31** (Theorem 5.2 from [18], specialized). *Let $A, \phi : \mathbb{C}^2 \rightarrow \mathbb{C}$ be analytic on a compact neighborhood $\mathcal{N}$ around $(0, 0)$. Furthermore assume that the real part of ϕ is nonnegative on $\mathcal{N}$ and vanishes only at $(0, 0)$, and that the Hessian matrix $\mathcal{H}_\phi$ of ϕ is such that $\mathcal{H}_\phi(0, 0)$ is non-singular. Let, $I(\lambda) = \int_{\mathcal{N}} A(z) e^{-\lambda \phi(z)}$. Then for $\lambda \leq \frac{1}{2}$, we have that:*

$$I(\lambda) = A(0, 0) \frac{2\pi}{\sqrt{\det(\mathcal{H}_\phi(0, 0))}} \cdot \frac{1}{\lambda} \left(1 + O\left(\frac{1}{\lambda}\right) \right).$$

Using these, we obtain the following estimate regarding estimating coefficient of an analytic function:

Proof of Lemma 29. We apply Theorem 30 with the contour $\mathcal{C}$ being product of circles with $|u| = u_0, |v| = v_0$ so that

$$C_{p,q} = \frac{-1}{4\pi^2} \int_{(u,v) \in \mathcal{C}} f(u, v) u^{-p-1} v^{-q-1} du dv.$$

For $\theta = (\theta_1, \theta_2) \in [-\pi, \pi]^2$, we perform change of variables to express the points in the contour as $u = u(\theta) = u_0 e^{i\theta_1}, v = v(\theta) = v_0 e^{i\theta_2}$. We see that $du = u_0 i e^{i\theta_1} d\theta_1$ and $dv = v_0 i e^{i\theta_2} d\theta_2$. We also use the fact that $\exp(h(u, v)) = \frac{f(u, v)}{u^p v^q}$ to obtain that

$$C_{p,q} = \frac{1}{4\pi^2} \int_{\theta \in [-\pi, \pi]^2} \exp(h(u(\theta), v(\theta))) d\theta.$$

Let $\lambda = \sqrt{p^2 + q^2}$. We perform one last change of variables where we let $h'(u, v) = h(u, v)/\lambda$ to express the integral as:

$$C_{p,q} = \frac{\exp(h(u_0, v_0))}{4\pi^2} \int_{\theta \in [\pi, \pi]^2} \exp(\lambda \cdot (h'(u(\theta), v(\theta)) - h'(u_0, v_0))) d\theta.$$

We will apply Theorem 31 with $A(\theta) = 1$, $\phi(\theta) = h'(u_0, v_0) - h'(u(\theta), v(\theta))$. We claim the following:

▷ Claim 32.

$$\det(\mathcal{H}_\phi(0, 0)) = \left(\frac{u_0 v_0}{\lambda}\right)^2 \det(\mathcal{H}_{h'}(u_0, v_0)).$$

We will prove this at the end and for now will just assume it.

We first show that we satisfy all of the preconditions laid out by Theorem 31. First, we have that $\phi(0, 0) = 0$ and by unique maximality of u_0, v_0 that for all $\theta \neq (0, 0)$, it follows that $\phi(\theta) > 0$. Second, we show that $\mathcal{H}_\phi(0, 0)$ is non-singular. By assumption, we know that $\mathcal{H}_h(u_0, v_0)$ is positive definite and so must be true for $\mathcal{H}_{h'}(u_0, v_0)$, making it singular. Therefore by Claim 32, it follows that $\mathcal{H}_\phi(0, 0)$ is non-singular as well. Lastly, by assumption we have that $\lambda \leq \frac{1}{2}$. Hence, we can indeed apply Theorem 31 with the same λ to infer that

$$C_{p,q} = \frac{\exp(h(u_0, v_0))}{4\pi^2} \cdot \frac{2\pi}{\sqrt{\det(\mathcal{H}_\phi(0, 0))}} \cdot \frac{1}{\lambda} \cdot \left(1 + O\left(\frac{1}{\lambda}\right)\right).$$

Using Claim 32 and the fact that $\exp(h(u_0, v_0)) = \frac{f(u_0, v_0)}{u_0^q v_0^q}$, our desired bound on $C_{p,q}$ follows.

We lastly prove our helper claim that relates $\det(\mathcal{H}_\phi(0, 0))$ and $\det(\mathcal{H}_{h'}(u_0, v_0))$.

Proof of Claim 32. Let $g(\theta) = -h'(u(\theta), v(\theta))$ so that $\phi(\theta) = h'(u_0, v_0) + g(\theta)$. Since $h'(u_0, v_0)$ is a constant, we have that $\mathcal{H}_g(0, 0) = \mathcal{H}_\phi(0, 0)$. Hence, it suffices to show that $\det(\mathcal{H}_g(0, 0)) = \left(\frac{u_0 v_0}{\lambda}\right)^2 \det(\mathcal{H}_{h'}(u_0, v_0))$. Using chain rule, we see that for $i \in [2]$:

$$\frac{\partial g}{\partial \theta_i} = \frac{\partial h'}{\partial u} \frac{\partial u}{\partial \theta_i} + \frac{\partial h'}{\partial v} \frac{\partial v}{\partial \theta_i}$$

When we evaluate this at $(0, 0)$, we will evaluate $\frac{\partial h'}{\partial u}$ and $\frac{\partial h'}{\partial v}$ at (u_0, v_0) . By assumption $\nabla_h(u_0, v_0) = 0$ and so $\nabla_{h'}(u_0, v_0) = 0$. Using above, this implies that $\nabla_g(0, 0) = 0$.

Using chain rule, the fact that $\nabla_g(0, 0) = 0$, and recalling the definition of $u(\theta)$ and $v(\theta)$, we obtain that

$$\mathcal{H}_g(0, 0) = D\mathcal{H}_{h'}(u_0, v_0)D$$

where D is the 2×2 diagonal matrix with entries iu_0, iv_0 .

With this we finally compute that

$$\begin{aligned} \det(\mathcal{H}_g(0, 0)) &= \det(D\mathcal{H}_{h'}(u_0, v_0)D) \\ &= \det(D) \cdot \det(\mathcal{H}_{h'}(u_0, v_0)) \cdot \det(D) \\ &= (u_0 v_0)^2 \det(\mathcal{H}_{h'}(u_0, v_0)) \\ &= \frac{(u_0 v_0)^2}{\lambda^2} \det(\mathcal{H}_h(u_0, v_0)) \end{aligned}$$

where for the last equality we used the fact that $h' = h/\lambda$, which implies $\mathcal{H}_{h'}(u_0, v_0) = \frac{1}{\lambda} \mathcal{H}_h(u_0, v_0)$. ◀