What is in #P and what is not?

Christian Ikenmeyer* Igor Pak**

August 26, 2022

Abstract

For several classical nonnegative integer functions, we investigate if they are members of the counting complexity class #P or not. We prove #P membership in surprising cases, and in other cases we prove non-membership, relying on standard complexity assumptions or on oracle separations.

We initiate the study of the polynomial closure properties of #P on affine varieties, i.e., if all problem instances satisfy algebraic constraints. This is directly linked to classical combinatorial proofs of algebraic identities and inequalities. We investigate #TFNP and obtain oracle separations that prove the strict inclusion of #P in all standard syntactic subclasses of #TFNP − 1.

Keywords: Counting complexity, combinatorial proofs, TFNP, #P, GapP

ACM computing classification system:

- Mathematics of computing ∼ Discrete mathematics ∼ Combinatorics
- Theory of computation ∼ Computational complexity and cryptography ∼ Complexity classes

AMS subject classification: 05A20, 68Q15 and 68R99

*University of Liverpool, United Kingdom, christian.ikenmeyer@liverpool.ac.uk

**University of California, Los Angeles, USA, pak@math.ucla.edu

Contents

1 Introduction

1.1 Foreword

Finding a *combinatorial interpretation* is an everlasting problem in Combinatorics. Having combinatorial objects assigned to numbers brings them depth and structure, makes them alive, sheds light on them, and allows them to be studied in a way that would not be possible otherwise. Once combinatorial objects are found, they can be related to other objects via bijections, while the numbers' positivity and asymptotics can then be analyzed.

Historically, this approach was pioneered by J.J. Sylvester in his "*constructive theory of partitions*" [Syl82]. There, Sylvester was able to rederive a host of old partition identities and prove many new ones by interpreting the coefficients on both sides as the numbers of certain *Ferrers shapes* (now called *Young diagrams*), and relating two sides to each other. G.H. Hardy marveled at such proofs, calling them "striking" and "unlike any other" [Har40], see also [Pak06].

Since the 1960s, this approach became a staple in *Enumerative Combinatorics*, reaching as far as undergraduate textbooks [SW86], monographs [Loe11] and multimedia compendia [Vie16]. In *Algebraic Combinatorics*, even one combinatorial interpretation can introduce revolutionary changes. Notably, a Young tableau interpretation of the *Littlewood–Richardson* (LR-) *coefficients* $c^{\lambda}_{\mu\nu}$ was discovered in [LR34]. These numbers describe the structure constants of the *Schur functions* multiplication [Mac95, Sta12]. Over the last few decades, this result led to an avalanche of developments, culminating with a complete resolution of the *Horn problem* [Kly98] (see also [Ful98]), proof of the *saturation conjecture* [KT99], and polynomial time algorithms for the vanishing of the LR–coefficients [BI13b, MNS12, Ike16].

When a combinatorial interpretation exists it is a modern wonder, a starting point of a combinatorial investigation. *But what if none is known?* Such examples in Enumerative Combinatorics are too numerous to be listed, see e.g. [Pak18, §4]. In Algebraic Combinatorics, the following are the top three "most wanted" combinatorial interpretations, all from *Stanley's list* [Sta00]:

- *Kronecker coefficients* $g(\lambda, \mu, \nu)$ which generalize LR–coefficients and give structure constants of tensor products of $\mathfrak{S}_n$-modules. This celebrated problem goes back to Murnaghan [Mur38] and plays a crucial role in *Geometric Complexity Theory* (GCT), see [Mul09]. See [BDO15, IMW17, PP17, PPY19] and §9.4, for some recent combinatorial and complexity work on the subject.

- *plethysm coefficients* $p_\lambda(\mu, \nu)$ which describe decompositions of Schur functors of $\mathfrak{S}_n$-modules, and is the main subject of GCT7 [Mul07], see also [BIP19, FI20, IP17]. They also appear in connection to the *Foulkes conjecture* in Representation Theory, see [Bri93, CIM17, Lan15].

- *Schubert coefficients* $c(u, v, w)$ which give structure constants of the product of Schubert polynomials, defined by Lascoux and Schützenberger [LS82] in the context of *cohomology of the Grassmannian*, see [Mac91, Man01]. We refer to [Knu16, KZ20, MPP14] for examples of positive results.

In all three cases, there is a widespread belief that these coefficients must have a combinatorial interpretation. A positive resolution of either problem would be a major breakthrough culminating decades long study. In the context of GCT, Mulmuley conjectured [Mul09] that both Kronecker and plethysm coefficients are in #P (see [Val79]), as a step towards proving that $\mathsf{P} \neq \mathsf{NP}$. Note that all three functions are in $\mathsf{GapP}_{\geq 0}$, suggesting commonality of the obstacles.

Now, *what if the community is wrong*, and these functions are not in #P? Such a possibility has only been raised recently [Pak19, Spe11]. According to Popper's philosophy, a belief needs to be *disprovable* in order to be *scientific* [Pop62]. Until now there has been little effort towards proving that some natural combinatorial functions are not in #P (see below). With this paper we initiate a systematic study of this problem.

We show that many natural combinatorial functions are not in #P under various complexity assumptions. In a positive direction, we prove that many functions *are* in #P, some strikingly close to those that are not.

1.2 Motivational examples of #P functions

Let $\mathsf{GapP}_{\geq 0}$ be the class of nonnegative functions in $\mathsf{GapP} := \{f_1 - f_2 \mid f_1, f_2 \in \#\mathsf{P}\}$.[1] More generally, we consider the class $\mathsf{PolynP} := \big\{\varphi(f_1, \ldots, f_k) \mid \varphi \in \mathbb{Q}[x_1, \ldots, x_k], f_i \in \#\mathsf{P}\big\}$, and study the class $\mathsf{PolynP}_{\geq 0}$ of nonnegative functions in PolynP. The place to start is to look for natural integer functions in these classes and ask if they lie in #P. For the three functions as above the problem remains open, but what is known in other cases? Consider the following motivating examples:

(1) Let $e : P \to \mathbb{N}$ be the number of linear extensions of P, where $P = (X, \prec)$ is a poset with n elements. Recall that $e(P) \geq 1$, so $e'(P) := e(P) - 1 \in \mathsf{GapP}_{\geq 0}$. Now observe that $e' \in \#\mathsf{P}$ simply because finding the lex-smallest linear extension L can be done in polynomial time (see e.g. [CW95]), so $e'(P)$ counts linear extensions of P that are different from L. Note aside that since e is #P-complete [BW91], then so is e'.

(2) Recall *Sperner's lemma* which states that for every $\{1, 2, 3\}$-coloring χ of interior vertices in a side-length n-triangle region Δ_n of the plane whose sides are colored 1, 2 and 3, respectively, there is a *rainbow* (123) triangle. We trust the reader is familiar with the setting, see e.g. [Pap94a] and [MM11, §6.7]. Here n is given in binary and χ is given by a polynomial size circuit. Denote by $t(\chi)$ the number of rainbow triangles, so that $t(\chi) - 1 \in \mathsf{GapP}_{\geq 0}$.

Since the typical proof of Sperner's lemma involves tracing down the path of non-rainbow triangles until a rainbow triangle is reached, it may come as a surprise that $t(\chi) - 1 \in \#\mathsf{P}$. Indeed, simply observe that $t(\chi) - 1 = 2t_-(\chi)$, where $t_\pm(\chi)$ denotes the number of rainbow triangles with positive/negative orientation. This follows from $t(\chi) = t_+(\chi) + t_-(\chi)$ and $t_+(\chi) - t_-(\chi) = 1$ equations, see e.g. [Pak03, §8].

(3) Let G be a simple graph with at least one edge, and let $f(G)$ be the number of proper 3-colorings of G. Then $f(G)/6$ is an integer valued function in $\mathsf{PolynP}_{\geq 0}$ by taking into account permutations of colors. Of the six possible 3-colorings corresponding to a given 3-coloring one can easily choose the lex-smallest, implying that $f(G)/6 \in \#\mathsf{P}$.[2] Such solution is not always possible in other problems, see §1.3(4), and algorithmic approaches to equivalence problems have been studied in [BG83, BG84, FG11].

(4) Let $\delta(k, G) := m_k(G)^2 - m_{k-1}(G)\, m_{k+1}(G)$, where $m_k(G)$ is the number of k-matchings in graph G. The function $\delta \in \mathsf{GapP}$ by definition. By the celebrated result of Heilmann and Lieb [HL72], the sequence $m_1(G), m_2(G), \ldots$ is *log-concave*, implying that $\delta \in \mathsf{GapP}_{\geq 0}$. This result is a starting point of many combinatorial investigations [God93], including notably the "interlacing families" series [MSS13]. While all signs point to δ being "difficult to handle", it was observed in [Pak19] that a beautiful proof in [Kra96] easily implies that $\delta \in \#\mathsf{P}$.

(5) Recall *Fermat's little theorem*: For every prime p and $a \in \mathbb{N}$, we have:

$$a^p \equiv a \pmod{p}.$$

[1] The closure $\mathsf{GapP} = \#\mathsf{P} - \#\mathsf{P}$ of #P under subtraction was introduced in [FFK94] and indep. in [Gup95].

[2] It is important to emphasize that while $f(G)$ is #P-complete, it is completely irrelevant to the conclusion. Crucially, the *lex-smallest test* is in P in both this and the previous example. In non–#P examples of this kind, the lex-smallest test is NP-hard (see below).

This is one of the most basic and most celebrated results in Number Theory, see e.g. [IR82, §3.4], and is the starting point of the *Miller–Rabin primality test*, see e.g. [MM11, §10.8.2]. The theorem can be rephrased as: for all $a \in \mathbb{N}$, we have

$$\varphi(a) \ := \ \tfrac{1}{p}(a^p - a) \in \mathbb{N}.$$

It is readily converted into a PolynP function by substituting $a \leftarrow N(\phi)$ as follows:

$$\tfrac{1}{p}\big(N(\phi)^p \ - \ N(\phi)\big) \ \in \ \varphi(\#\mathsf{P}) \ \subseteq \ \mathsf{PolynP},$$

where $N(\phi)$ is the number of satisfying assignments of a Boolean formula ϕ. It was shown by Peterson [Pet72] (see also [Gol56]) that this function is actually in $\#\mathsf{P}$ (see Proposition 7.3.1), by giving a combinatorial interpretation for $\varphi(a)$, and in this way reproving Fermat's little theorem. In other words, we have $\varphi(\#\mathsf{P}) \subseteq \#\mathsf{P}$, i.e., the class $\#\mathsf{P}$ is closed under the *Frobenius map* φ. At the heart of the proof of Proposition 7.3.1 is a polynomial-time algorithm for identifying lex-smallest elements as in §1.2(3), but here in a $\mathbb{Z}/p\mathbb{Z}$ orbit.

(6) Consider the following inequality by Grimmett [Gri76]:

$$\tau(G) \ \le \ \frac{1}{n}\left(\frac{2m}{n-1}\right)^{n-1}$$

for the *number of spanning trees* $\tau(G)$ in a simple graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges. One can turn this into a $\mathsf{GapP}_{\ge 0}$ function as follows:

$$f(G) \ := \ (2m)^{n-1} \ - \ n(n-1)^{n-1}\tau(G).$$

On the other hand, *given that the inequality holds*, the claim $f \in \#\mathsf{P}$ is trivial since $\tau \in \mathsf{FP}$. Indeed, since $f(G)$ can be computed in polynomial time by the matrix-tree theorem, we conclude that $f(G)$ counts the set of n-bit binary strings from 0 to $f(G) - 1$.[3] [4] This is why it is important in the examples above that our functions are not obviously in FP (e.g., being $\#\mathsf{P}$-hard is a good indication), since otherwise the problem becomes trivial.

1.3 Motivational non-examples

It may come as a surprise that the non-example comes from the simplest of the inequalities.

(1) *Cauchy–Schwartz inequality*:

$$a^2 \ + \ b^2 \ \ge \ 2ab \quad \text{where} \quad a, b \in \mathbb{R}. \tag{1.3.1}$$

Now take a, b to be counting functions. Formally, for two Boolean formulas ϕ and ψ, let

$$h(\phi, \psi) \ := \ N(\phi)^2 \ + \ N(\psi)^2 \ - \ 2N(\phi)N(\psi) \ = \ \big(N(\phi) \ - \ N(\psi)\big)^2. \tag{1.3.2}$$

By definition, the function $h \in \mathsf{GapP}_{\ge 0}$. Note, however, that if $h \in \#\mathsf{P}$, then we get a polytime witness for $N(\phi) \ne N(\psi)$. This is unlikely, as it would imply the collapse of polynomial hierarchy

[3] Combinatorialists would argue that a combinatorial interpretation should explain *why* the inequality holds in the first place. In fact, there are several schools of thought on this issue (see a discussion in [Pak18, §4]). We believe that the computational complexity approach is both the least restrictive and the most formal way to address this.

[4] In the context of GCT, motivated by the work on LR–coefficients, Mulmuley asks if Kronecker and plethysm coefficients count the number of integer points in a polytope defined by the inequalities with polynomial description [Mul09]. We do not work with this narrower notion in this paper. See, however, [KM18].

to the second level: $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$ (see Proposition 2.3.1). Colloquially, this says that under the natural complexity assumption $\mathsf{PH} \neq \Sigma_2^{\mathsf{p}}$, the Cauchy–Schwartz inequality (1.3.1) does not have a combinatorial interpretation in full generality.

(2) The *Hadamard inequality* for real $d \times d$ matrices states:

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & \cdots & a_{dd} \end{pmatrix}^2 \leq \prod_{i=1}^{d} \big(a_{i1}^2 + \ldots + a_{id}^2\big). \tag{1.3.3}$$

Geometrically, it says that the volume of a parallelepiped in $\mathbb{R}^d$ is at most the product of its basis edge lengths, with equality when these edges are orthogonal. Note that standard proofs of (1.3.3) involve the eigenvalues of $A = (a_{ij})$, see e.g. [HLP52, §2.13] and [BB61, §2.11], suggesting that translation into combinatorial language would be difficult.

Substitute all $a_{ij} \leftarrow N(\phi_{ij})$ in (1.3.3), where ϕ_{ij} are Boolean formulas. Denote by H_d the resulting counting function written in the style of (1.3.2), i.e, H_d is the difference of the right-hand side and the left-hand side of (1.3.3). It is easy to see that $H_2 \in \#\mathsf{P}$, see §2.1. For $d \geq 3$, we prove that $H_d \notin \#\mathsf{P}$ under an assumption that we call the *univariate binomial basis conjecture* (Conjecture 4.4.2). This is a general conjecture about the structure of $\#\mathsf{P}$. Formally, we show the existence of an oracle $A \subseteq \{0,1\}^*$ with $H_3(\overrightarrow{\#\mathsf{P}^A}) \not\subseteq \#\mathsf{P}^A$, see Theorem 7.2.1.

(3) For a simple graph G on n vertices, denote by $\boldsymbol{d}(G) = (d_1, \ldots, d_n)$ the degree sequence. Consider the following natural inequality:

$$\mathbb{P}[\,G \text{ is planar}\,] \leq \mathbb{P}[\,G \text{ is planar} \mid \boldsymbol{d}(G) \leqslant \boldsymbol{c}\,], \tag{1.3.4}$$

where $\boldsymbol{c} = (c_1, \ldots, c_n)$ is a given sequence, the inequality $\boldsymbol{d}(G) \leqslant \boldsymbol{c}$ is coordinate-wise: $d_i \leq c_i$ for all $1 \leq i \leq n$, and where the probability is over uniform random graphs on $[n] = \{1, \ldots, n\}$. This says that being planar correlates with having small degrees.[5]

We can convert (1.3.4) it into a $\mathsf{GapP}_{\geq 0}$ function as follows:

$$\begin{aligned} \varrho(\boldsymbol{c}) \,:=\, 2^{\binom{n}{2}} \#\big\{\text{planar graphs } G \text{ on } [n] \text{ with } \boldsymbol{d}(G) \leqslant \boldsymbol{c}\big\} \,- \\ -\, \#\big\{\text{planar graphs on } [n]\big\} \cdot \#\big\{\text{graphs } G \text{ on } [n] \text{ with } \boldsymbol{d}(G) \leqslant \boldsymbol{c}\big\}. \end{aligned}$$

This inequality is a simple special case of the *Kleitman inequality* [Kle66], which is a corollary of the *Ahlswede–Daykin inequality* [AD78] (Theorem 7.4.1). In Proposition 2.5.1, we show that the polynomial inequality implied by the Ahlswede–Daykin inequality is not in $\#\mathsf{P}$, again under the univariate binomial basis conjecture.

(4) Recall the following *Smith's theorem* [Tut46]. Let $e = (v, w)$ be an edge in a cubic graph G. Then the number $N_e(G)$ of Hamiltonian cycles in G containing e is always even. Denote $f(G, e) := N_e(G)/2$ and observe that $f \in \mathsf{PolynP}_{\geq 0}$. Is $f \in \#\mathsf{P}$? We don't know. This seems unlikely and remains out of reach with existing technology. But let us discuss the context behind this problem.

Tutte's original proof in [Tut46] uses a double counting argument. The *Price–Thomason algorithm* for finding another Hamiltonian cycle in a cubic graph [Pri77, Tho78] gives a more direct

[5] Note aside that the number of labeled planar graphs on n vertices can be computed in time polynomial in n using Tutte's generating function formulas [Tut63], see also [Noy14, Sch15]. On the other hand, the number of labeled graphs with a given upper bound on the degrees is likely not in FP, cf. [Wor18].

combinatorial proof of Smith's theorem and implies that this search problem is in PPA, the class defined by the polynomial parity argument. In fact, AnotherHamiltonianCycle is a motivational problem for PPA, while Sperner, see §1.2(2), is a motivational problem for PPAD [Pap94a][6].

Note that the Price–Thomason algorithm partitions the set of all Hamiltonian cycles into pairs, but this pairing algorithm is known to require an exponential number of steps in the worst case, see [Cam01, Kra99]. A polynomial-time algorithm instead would allow us to search for Hamiltonian cycles and only count the ones that are lexicographically smaller than their pairing partner, which would show that $N_e(G)/2 \in \#\mathsf{P}$, and $(\text{AllOtherHamiltonianCyclesThroughEdge} - 1)/2 \in \#\mathsf{P}$. Note that such a pairing algorithm (not for the symmetric group $\mathfrak{S}_2$, but for $\mathfrak{S}_3$) is the reason why $f(G)/6 \in \#\mathsf{P}$ in §1.2(3).

We study the basic search problem Leaf[7] that is used to define PPA, and that arises directly from Sperner by a parsimonious reduction from the PPAD-complete problem SourceOrSink and removing the edge directions. We show that for the corresponding counting problem we have an oracle separation that shows $\text{AllLeaves}^A/2 \notin \#\mathsf{P}^A$. In fact, for the counting version of Leaf, where we are given one leaf and count all others, we show that $\text{Leaf}^A - 1 \notin \#\mathsf{P}^A$. This has to be contrasted to Sperner, where the membership $\text{Sperner} - 1 \in \#\mathsf{P}$ relativizes, i.e., holds with respect to all oracles. The oracle instances are significantly more complicated as for the Hadamard problem, see §1.3(2).

(5) We have seen that $\text{Sperner}(\chi) - 1 = 2t_-(\chi)$, hence $(\text{Sperner} - 1)/2 \in \#\mathsf{P}$. It is easy to see that the reverse inclusion holds: The counting class $\#\mathsf{PPAD}(\text{Sperner})$ defined by the Sperner problem contains $2\#\mathsf{P} + 1$, or, in other words, $\#\mathsf{P} = (\#\mathsf{PPAD}(\text{Sperner}) - 1)/2$. For the other classes in TFNP we similarly get

$$\begin{aligned} \#\mathsf{P} &= \big(\#\mathsf{PPAD}(\text{Sperner}) - 1\big)/2 \\ &= \#\mathsf{PPADS}(\text{Sink}) - 1 \\ &= \#\mathsf{CLS}(\text{EitherSolution}(\text{Sperner}, \text{Iter})) - 1 \end{aligned}$$

and these equalities relativize. But for the more complex classes we get oracle separations: $(\#\mathsf{PPA}(\text{Leaf}) - 1)/2$, $\#\mathsf{PPP}(\text{Pigeon}) - 1$ and $\#\mathsf{PLS}(\text{Iter}) - 1$ strictly contain $\#\mathsf{P}$ with respect to an oracle.

But this does not give the complete picture, since non-parsimonious reductions between complete problems give different counting classes. For example if instead of leaves in a graph we count the nodes that are adjacent to leaves (which we call preleaves), then this does not change the complexity of the search problem, but it changes the counting class from $\#\mathsf{PPA}(\text{Leaf})$ to the class $\#\mathsf{PPA}(\text{Preleaf})$ (note that the functions in $\#\mathsf{PPA}(\text{Leaf})$ always attain odd values, while the functions in $\#\mathsf{PPA}(\text{Preleaf})$ do not have this restriction). The underlying argument is the *chessplayer algorithm*, see e.g. [Pap90, BCE+98], which results in non-parsimonious reductions, which then give rise to a complexity class inclusion diagram where we have an oracle with respect to which we have a strict inclusion of $\#\mathsf{P}$ in *all* the classes $\#\mathsf{PPAD} - 1$, $\#\mathsf{PPADS} - 1$, $\#\mathsf{CLS} - 1$, $\#\mathsf{PPA} - 1$, $\#\mathsf{PPP} - 1$ and $\#\mathsf{PLS} - 1$. The full class inclusion diagram of our results can be found in Figure 1. The definitions of the classes and problems can be found in §8.1.

[6] Several versions of Sperner on non-orientable manifolds are PPA-complete [Gri01, DEF+21], as well as e.g. the problems Consensus-Halving/Necklace Splitting [FRG18, FRH+20, DFHM22], and integer factoring (assuming the GRH) [Jeř16]. Main PPAD-complete problems include *Nash equilibrium* [DGP09, CD09] and *hairy ball* [GH21].

[7] Search problems are often of the type AnotherSolution, but the name does not suggest that. Leaf for example could reasonably be called AnotherLeaf. We adapt the search problem notation and drop the Another prefix and mean the corresponding problem of *counting* all but the given leaf. The problem of counting *all* leaves when we are *not* given one is called AllLeaves. Since all our problems are counting problems, we drop the customary # in front of the problem name, also to avoid having two # in the class names, see §3.2.

These problems are syntactically guaranteed to be nonnegative, but in contrast to the HADAMARD problem (for example), here the oracle separations are much more delicate, as we have to fool the Turing machine while producing instances of the correct cardinality (which is easier if the problem is a polynomial evaluated at arbitrary #P functions). To overpass these obstacles, we introduce the notion of a set-instantiator in Definition 6.1.3. We will also treat cases where we have a nonnegativity guarantee, but no further information about the reason. This requires extra care, see Propositions 2.3.3 and 7.5.5.

2 Definitions, notations and first steps

We start in §2.2 with the concept of polynomial closure properties of #P. We then prove some simple separations in §2.3 and §2.4, as a warmup before our main results in the next section.

2.1 Basic notation

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{Q}_+ = \{x \in \mathbb{Q}, x \geq 0\}$. For $i \in \mathbb{N}$ and $x \in \mathbb{R}$, we write

$$\binom{x}{i} \;=\; \frac{1}{i!}\, x(x-1)\cdots(x-i+1).$$

In particular, $\binom{i}{0} = \binom{i}{i} = 1$. We think of $\binom{x}{i}$ as a rational polynomial of degree i. Note that for $0 \leq x < i$, $x \in \mathbb{N}$, we have $\binom{x}{i} = 0$. For a vector $(a_1, \ldots, a_n) \in \mathbb{R}^n$, we use both $\vec{a}$ and $\boldsymbol{a}$ to denote this vector.

We are assuming the reader is familiar with basic complexity theory and standard complexity classes: P, NP, UP, PH, FP, #P, GapP, PPA and PPAD. We refer to [AB09, MM11, Pap94b] for the definitions and standard results, and to [Aar16, Wig19] for further background.

2.2 Closure properties

We say that a map $\varphi : \mathbb{N}^k \to \mathbb{Q}$ is *integer-valued* if it only attains integer values. Similarly, map φ is *nonnegative*, write $\varphi \geqslant 0$, if it only attains nonnegative values.

We say that φ is a *closure property* of #P, if for all $f_1, \ldots, f_k \in \#\mathsf{P}$ we have $\varphi(f_1, \ldots, f_k) \in \#\mathsf{P}$. More concisely, we also write:

$$\varphi\big(\overrightarrow{\#\mathsf{P}}\big) \;\subseteq\; \#\mathsf{P}.$$

This is a generalization of the notation $\mathsf{GapP} = \#\mathsf{P} - \#\mathsf{P}$ from [FFK94].[8] Let $S \subseteq \mathbb{N}^k$ be a fixed subset. We say that φ is a *closure property* of #P *restricted* to S (or *on* S), if for all $f_1, \ldots, f_k \in \#\mathsf{P}$ which satisfy $\big(f_1(w), \ldots, f_k(w)\big) \in S$ for all $w \in \{0,1\}^*$, we have $\varphi(f_1, \ldots, f_k) \in \#\mathsf{P}$.

Note that we evaluate these #P functions on the same input. For example, in the notation of §1.2(2), the map $\varphi(t_-, t_+) := t(\chi) = t_- + t_-$ is restricted to $S = \{(t_-, t_+) \mid t_+ - t_- = 1\}$. Similarly, in the notation of §1.2(3), we have $S = 6\mathbb{N}$.

We write the restriction to S as a subscript, usually denoted $\overrightarrow{\#\mathsf{P}}_{\in S}$, but the property "$\in S$" is sometimes notationally replaced by other properties such as "≥ 1" (in which case $S = \mathbb{N}_{\geq 1}$)

[8] When defining $\overrightarrow{\#\mathsf{P}}$, two different definitions are equivalent (in the same way as for GapP). First, one can define $\overrightarrow{\#\mathsf{P}}$ via k many nondeterministic polynomial time Turing machines and consider the k-vector of their number of accepting paths as the output. Alternatively, one can define it via one nondeterministic polynomial time Turing machine that has k many different states of acceptance and one reject state (these states of acceptance are usually labeled with $+1$ and -1 in GapP). This is a complexity class of multi-output functions, as, for example, considered in [Val76].

or "even" (in which case $S = 2\mathbb{N}$). For example, in notation of §1.2(1), we have $e(P) \in \#\mathsf{P}_{\geq 1}$. Similarly, in the notation of §1.3(4), we have $N_e(G) \in \#\mathsf{P}_{\text{even}}$. This allows us to write statements such as

$$\#\mathsf{P}_{\geq 1} + 1 \subseteq \#\mathsf{P}, \quad \text{and} \quad \#\mathsf{P}^A_{\text{even}}/2 \not\subseteq \#\mathsf{P}^A$$

for the oracle A separation. More generally, in the multivariate case we write

$$\varphi\big(\overrightarrow{\#\mathsf{P}}_{\in S}\big) \subseteq \#\mathsf{P}$$

for the closure property of $\#\mathsf{P}$ restricted to S. [HR00] study the univariate case and call such a restriction a counting property. These univariate restrictions also play a role in [CGH+89] and are the main focus of [GW87]. The most famous example is probably $\mathsf{UP} = \#\mathsf{P}_{\in\{0,1\}}$ (if one identifies languages with their characteristic functions, which we do), see [Val76, GS88, Ko85, HT03]. In some contexts it is natural to consider a promise version of UP, see [VV85], but that is *different* from what we consider here. To make connections to TFNP more visible, we define $\#\mathsf{TFNP} := \#\mathsf{P}_{\geq 1}$.

Let $\varphi, \psi \in \mathbb{Q}[x_1, \ldots, x_k]$ be rational polynomials. We write $\varphi \geqslant_{\#} \psi$ if

$$\varphi(f_1, \ldots, f_k) - \psi(f_1, \ldots, f_k) \in \#\mathsf{P} \quad \text{for all} \quad f_1, \ldots, f_k \in \#\mathsf{P},$$

or, equivalently, $(\varphi - \psi)(\overrightarrow{\#\mathsf{P}}) \subseteq \#\mathsf{P}$. For example, $x^2 + 3x \geqslant_{\#} 0$. Less obviously, $x^2 \geqslant_{\#} x$, since $x^2 - x = 2\binom{x}{2}$ which counts unordered pairs (i, j), where $1 \leq i < j \leq x$. For the Hadamard inequality (1.3.3), we easily have $H_2 \geqslant_{\#} 0$, since

$$\begin{aligned} \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 &= (ad - bc)^2 = a^2d^2 - 2abcd + b^2c^2 \\ &\leqslant_{\#} a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 = (a^2 + b^2)(c^2 + d^2). \end{aligned}$$

We emphasize again that over the reals this is *not* a valid proof of the Hadamard inequality for 2×2 matrices since the $2abcd$ term can be negative. The inequality $H_2(a, b, c, d) \geq 0$ over the reals follows from the Cauchy–Schwartz inequality in this case.

2.3 Complete squares

As in the introduction, we have $\mathsf{GapP} = \#\mathsf{P} - \#\mathsf{P} = \{f_1 - f_2 \mid f_1, f_2 \in \#\mathsf{P}\}$. We use the notation $[\mathsf{C} = 0]$ to denote the class of languages $L \subseteq \{0,1\}^*$ for which there exists a function $f \in \mathsf{C}$ with: $w \in L$ if and only if $f(w) = 0$. For example, $[\#\mathsf{P} = 0] = \mathsf{coNP}$ and $[\mathsf{GapP} = 0] = \mathsf{C}_{=}\mathsf{P}$. The following proposition about k-th powers of GapP functions is well known:

2.3.1 Proposition. *If* $\mathsf{GapP}^k \subseteq \#\mathsf{P}$ *for some even* k*, then* $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.

Proof. Recall that $\mathsf{PH} \subseteq \mathsf{NP}^{\mathsf{C}_{=}\mathsf{P}}$, which can be found for example in [Tar91], [Gre93] and [Cur16], which follows from Toda's $\mathsf{PH} \subseteq \mathsf{P}^{\#\mathsf{P}}$ theorem (see [Toda91, KVVY93, For97, For09]) as follows: $\mathsf{PH} = \mathsf{NP}^{\mathsf{PH}} \subseteq \mathsf{NP}^{\mathsf{P}^{\#\mathsf{P}}} = \mathsf{NP}^{\#\mathsf{P}} = \mathsf{NP}^{\mathsf{GapP}} =^{9} \mathsf{NP}^{\mathsf{C}_{=}\mathsf{P}}$. We now observe:
$\mathsf{PH} \subseteq \mathsf{NP}^{\mathsf{C}_{=}\mathsf{P}} = \mathsf{NP}^{[\mathsf{GapP}=0]} = \mathsf{NP}^{[\mathsf{GapP}^k=0]} \subseteq \mathsf{NP}^{[\#\mathsf{P}=0]} = \mathsf{NP}^{\mathsf{coNP}} = \Sigma_2^{\mathsf{p}}$. □

2.3.2 Corollary (Cauchy–Schwartz inequality)**.** $a^2 + b^2 \not\geqslant_{\#} 2ab$ *unless* $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.

[9] $\mathsf{NP}^{\mathsf{GapP}} \subseteq \mathsf{NP}^{\mathsf{C}_{=}\mathsf{P}}$ holds because instead of calling the oracle for a function $g \in \mathsf{GapP}$ we can nondeterministically guess its return value $i = g(w)$ and call the $\mathsf{C}_{=}\mathsf{P}$ oracle $[g - i = 0]$ on the input w to check for correctness (continue the computation if the guess was correct; reject the computation if the guess was wrong).

This innocent looking corollary has immediate negative consequences on the existence of combinatorial proofs for inequalities (in the sense of combinatorial interpretations of the difference of both sides of the inequality), for example the Cauchy inequality or the Alexandrov–Fenchel inequality, see §7.1 for the details.

Given the success in our *matching polynomial* example §1.2(4), one can ask if this example is generalizable to other log-concave properties. Formally, is it true that $g^2 \geqslant_{\#} fh$ when functions (f, g, h) are restricted to $S = \{(f, g, h) \in \mathbb{N}^3 \mid g^2 - fh \geq 0\}$? We give a negative answer to this question, suggesting that many log-concavity results and open problems (see §9.1) are unlikely to have a direct combinatorial proof.

2.3.3 Proposition (Log-concavity). *Let* $\varphi(f, g, h) := g^2 - fh$, *and let* $S := \{(f, g, h) \in \mathbb{N}^3 \mid g^2 - fh \geq 0\}$. *Then* $\varphi\big(\#\mathsf{P}^{\times 3}_{\in S}\big) \not\subseteq \#\mathsf{P}$ *unless* $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.

Proof. Let $f := 1$, $g := (x + y)$ and $h := 4xy$. Observe that $g^2 - fh = (x - y)^2 \geq 0$ for all $x, y \in \mathbb{R}$. The resulting complete square allows us to use Corollary 2.3.2 and prove the result. We now formalize this approach in the notation above.

Let $\vec{\gamma} : \mathbb{N}^2 \to \mathbb{N}^3$ defined by $(x, y) \mapsto \big(1, (x + y), 4xy\big)$. Then $\vec{\gamma}\big(\#\mathsf{P}^{\times 2}\big) \subseteq \#\mathsf{P}^{\times 3}_{\in S}$. Note that on the left-hand side we have no index anymore, as the image is guaranteed to lie in S. If we have $\varphi\big(\#\mathsf{P}^{\times 3}_{\in S}\big) \subseteq \#\mathsf{P}$, then it follows $\varphi\big(\vec{\gamma}\big(\#\mathsf{P}^{\times 2}\big)\big) \subseteq \#\mathsf{P}$. But we have $\varphi\big(\vec{\gamma}\big(\#\mathsf{P}^{\times 2}\big)\big) = \mathsf{GapP}^2$. We conclude: if $\varphi\big(\#\mathsf{P}^{\times 3}_{\in S}\big) \subseteq \#\mathsf{P}$ then $\mathsf{GapP}^2 \subseteq \#\mathsf{P}$. Hence, by Proposition 2.3.1, we have $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$. □

2.4 Non-monotone closure properties

A map $\varphi : \mathbb{N}^k \to \mathbb{Q}$ is called *monotone* if $\varphi(a_1, \ldots, a_k) \leq \varphi(a'_1, \ldots, a'_k)$ for all integer $a_1 \leq a'_1$, …, $a_k \leq a'_k$. For example, polynomials $x/2$, $x - 1$ and $x + y$ are monotone, but $x^2 - 2x$ and $(x - y)^2$ are not.

2.4.1 Proposition (Non-monotone closure properties). *Fix* $k \geq 1$. *If* $\varphi : \mathbb{N}^k \to \mathbb{N}$ *is a non-monotone closure property of* $\#\mathsf{P}$, *then* $\mathsf{UP} = \mathsf{coUP}$.

Proof. Let φ be a k-variate non-monotone closure property of $\#\mathsf{P}$. Then there exists $\vec{c} \in \mathbb{N}^k$ and $i \in [k]$ with $\varphi(\vec{c}) > \varphi(\vec{c} + \vec{e_i})$, where $\vec{e_i}$ is the i-th standard basis vector. Let $D := \varphi(\vec{c})$, and let $d := \varphi\big(\vec{c} + \vec{e_i}\big)$. Note that

$$\psi \,:\, f \mapsto \begin{pmatrix} \varphi(f \cdot \vec{e_i} + \vec{c}) \\ D \end{pmatrix}$$

is a univariate closure property of $\#\mathsf{P}$.

Now let $f \in \mathsf{UP} = \#\mathsf{P}_{\in\{0,1\}}$ be arbitrary. Let $\beta = f(w)$ for an arbitrary $w \in \{0, 1\}^*$. We have $\beta = 0$ if and only if $\beta \cdot \vec{e_i} + \vec{c} = \vec{c}$, and if and only if $\varphi(\beta \cdot \vec{e_i} + \vec{c}) = D$. Similarly, we have $\beta = 1$ if and only if $\beta \cdot \vec{e_i} + \vec{c} = \vec{c} + \vec{e_i}$, and if and only if $\varphi(\beta \cdot \vec{e_i} + \vec{c}) = d$. Therefore, $\psi(\beta) = 1 - \beta$. Hence, we have seen that $\psi(f) = 1 - f$ and that $\psi(f) \in \mathsf{UP}$. It follows that $f \in 1 - \mathsf{UP} = \mathsf{coUP}$. □

A similar use of binomial coefficients can also be found in [BG92]. Curiously, $x(x - 1)^2 \geqslant_{\#} 0$ since $x(x - 1)^2 = 6\binom{x}{3} + 2\binom{x}{2}$, yet by Proposition 2.4.1 we have:

2.4.2 Corollary. $(x - 1)^2 \not\geqslant_{\#} 0$ *unless* $\mathsf{UP} = \mathsf{coUP}$.

Note that $a^2 + b^2 \geqslant ab$ holds over $\mathbb{N}$, and is halfway between $a^2 + b^2 \geqslant 2ab$ and $a^2 + b^2 \geqslant 0$. So one can ask if $a^2 + b^2 \geqslant_{\#} ab$. Observe that $\varphi(a, b) := a^2 - ab + b^2$ is non-monotone: $\varphi(0, 2) = 4$ and $\varphi(1, 2) = 3$. Proposition 2.4.1 then gives:

2.4.3 Corollary. *$a^2 + b^2 \not\geqslant_{\#} ab$ unless $\mathsf{UP} = \mathsf{coUP}$.*

Recall the *Motzkin polynomial* $M(x, y) := x^2y^4 + x^4y^2 - 3x^2y^2 + 1$. It follows from the AM-GM inequality applied to positive terms, that $M(x, y) \geq 0$ for all $x, y \in \mathbb{R}$. On the other hand, this polynomial is famously not a *sum of squares*, and is a fundamental example in Semidefinite Optimization, see e.g. [Ble13, Mar08]. Now, observe that $M(x, y)$ is not monotone: $M(0, 1) = 1$ and $M(1, 1) = 0$. Proposition 2.4.1 then gives:

2.4.4 Corollary. *$M(x, y) \not\geqslant_{\#} 0$ unless $\mathsf{UP} = \mathsf{coUP}$.*

2.5 The binomial basis theorem

In this section we recall a classical result, describing all relativizing polynomial closure properties of $\#\mathsf{P}$ and GapP. Note that we considered only non-monotone examples in §2.3 and §2.4, while many natural polynomials are monotone. Clearly, every polynomial with integer coefficients is a closure property of GapP, but might not be a closure property of $\#\mathsf{P}$. If all coefficients of φ are nonnegative integers, then φ is clearly a closure property of $\#\mathsf{P}$, but we have seen that there are more, e.g. $\frac{1}{2}x^2 - \frac{1}{2}x = \binom{x}{2} \geqslant_{\#} 0$.

The main tool to shed light onto these issues is the binomial basis for the polynomial ring $\mathbb{Q}[x_1, \ldots, x_k]$, which is given by the polynomials $\beta_{\boldsymbol{a}} \in \mathbb{Q}[x_1, \ldots, x_k]$, $\boldsymbol{a} = (a_1, \ldots, a_k) \in \mathbb{N}^k$, via

$$\beta_{\boldsymbol{a}}(x_1, \ldots, x_k) := \binom{x_1}{a_1} \cdots \binom{x_k}{a_k}.$$

Every polynomial has a unique expression of finite support in this basis. The univariate version is well-known under the name of classical numerical polynomials. The study of the multivariate version goes back to Nagell [Nag19]. This basis explains the behavior we observe, as stated in the following fundamental theorem, for which the proof is split up into the $\#\mathsf{P}$ part, see [HVW95, Thm. 3.13], and the GapP part, see [Bei97, Thm. 6] (see also the bibliographic notes in [HO02, §5.6]). The GapP part can be obtained as a direct consequence of the algebraic properties of the binomial basis, see §4.2. We will reprove the $\#\mathsf{P}$ part as a direct corollary of our much more general Diagonalization Theorem 6.2.1.

Theorem (Binomial basis theorem, Thm. 4.3.2)**.** *The following four properties for a multivariate polynomial φ over $\mathbb{Q}$ are equivalent:*

- *φ is a closure property of GapP*
- *φ is integer-valued*
- *φ is a relativizing closure property of GapP*
- *the expression of φ over the binomial basis has only integer coefficients.*

Moreover, the following are equivalent:

- *φ is a closure property of $\mathsf{GapP}_{\geq 0}$*
- *φ is integer-valued and attains only nonnegative integers*
- *φ is a relativizing closure property of $\mathsf{GapP}_{\geq 0}$*
- *the expression of φ over the binomial basis has integer coefficients and φ attains only nonnegative integers if evaluated at integer points in the nonnegative cone.*

Moreover, the following are equivalent:

- *φ is a relativizing closure property of $\#\mathsf{P}$,*
- *the expression of φ over the binomial basis has only nonnegative integer coefficients.*

Note that even though $\#\mathsf{P}$ and $\mathsf{GapP}_{\geq 0}$ have different relativizing closure properties, this does not unconditionally separate these two classes. Note also that the theorem implies that *all polynomial closure properties of* GapP *and* $\mathsf{GapP}_{\geq 0}$ *relativize*. We conjecture that this is true for $\#\mathsf{P}$ as well, which is Conjecture 4.4.1. Proving this, however, would imply $\#\mathsf{P} \neq \#\mathsf{P}^{\mathsf{NP}}$ (and hence $\mathsf{P} \neq \mathsf{NP}$), even just for the univariate $\varphi = \binom{x-1}{2}$, see Theorem 4.4.3. We get the following sequence of implications:

$$\mathsf{P} = \mathsf{NP} \implies \#\mathsf{P} = \#\mathsf{P}^{\mathsf{NP}} \overset{\text{Thm. 4.4.3}}{\Longrightarrow} \binom{\#\mathsf{P}-1}{2} \subseteq \#\mathsf{P} \overset{\text{Prop. 2.4.1}}{\Longrightarrow} \mathsf{UP} = \mathsf{coUP}.$$

We call polynomials φ whose expression over the binomial basis has only nonnegative integer coefficients *binomial-good*, all others are called *binomial-bad*.

The fact that H_d in §1.3(2) is binomial-bad gives us the described separation, see Proposition 7.2.1. One famous instance of a binomial-good polynomial is the Frobenius map from §1.2(5). There, binomial-goodness can be interpreted as a combinatorial proof of Fermat's little theorem, see Proposition 7.3.1 and its proof by Peterson.

Theorem 4.3.2 is proved in [HVW95, Thm. 3.13] together with [Bei97, Thm. 6], which is in fact an extension of an argument of [CGH+89, Thm 3.1.1] and [OH93, p. 310] about the weakness of $\#\mathsf{P}$ machines in the presence of oracles. We prove it as a corollary of our Diagonalization Theorem 6.2.1 (see §3.1), which greatly extends Theorem 4.3.2.

2.5 (a) The Ahlswede–Daykin inequality

More advanced problems, where the set S is given as a semialgebraic set, are also possible, for example for the Ahlswede–Daykin inequality, see §1.2(3) and §7.4.

2.5.1 Proposition (Ahlswede–Daykin inequality)**.** *Let*

$$S := \left\{ (\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1, \delta_0, \delta_1, h_1, h_2, h_3, h_4) \in \mathbb{N}^{12} \;\middle|\; \begin{array}{l} \alpha_0\beta_0 + h_1 = \gamma_0\delta_0\,,\ \alpha_0\beta_1 + h_2 = \gamma_0\delta_1 \\ \alpha_1\beta_0 + h_3 = \gamma_0\delta_1\,,\ \alpha_1\beta_1 + h_4 = \gamma_1\delta_1 \end{array} \right\}$$

and let

$$\varphi := (\gamma_0 + \gamma_1)(\delta_0 + \delta_1) - (\alpha_0 + \alpha_1)(\beta_0 + \beta_1).$$

Then, under the Univariate Binomial Basis Conjecture 4.4.2, we have $\varphi(\overrightarrow{\#\mathsf{P}}_{\in S}) \not\subseteq \#\mathsf{P}$.

Proof. Define $\vec{\gamma} : \mathbb{N} \to \mathbb{N}^{12}$ via $\gamma(x) = \big(1, 1, x, x, x, 1, 1, x, 0, 2\binom{x}{2}, 2\binom{x}{2}, 0\big)$. Then $\vec{\gamma}(\#\mathsf{P}) \subseteq \#\mathsf{P}^{\times 12}_{\in S}$. Note that on the left-hand side we have no index anymore, as the image is guaranteed to lie in S. Assume for the sake of contradiction that we have an inclusion $\varphi(\#\mathsf{P}^{\times 12}_{\in S}) \subseteq \#\mathsf{P}$. Then it follows that we have an inclusion $\varphi\big(\vec{\gamma}(\#\mathsf{P})\big) \subseteq \#\mathsf{P}$. Conjecture 4.4.2 implies that this inclusion relativizes. Therefore, by Theorem 4.3.2 the univariate polynomial $\varphi \circ \vec{\gamma}$ is binomial-good. But we have $\varphi\big(\vec{\gamma}(f)\big) = f^2 - 2f + 1 = 2\binom{f}{2} - f + 1$, which is binomial-bad, a contradiction. □

3 Main results

In this section we state our main results. It §3.1 we state the Diagonalization Theorem and we give Karamata's inequality as an involved example for its application. In §3.2 we lift these techniques to handle TFNP and its subclasses. We obtain several oracle separations from $\#\mathsf{P}$ in this way, see Figure 1.

3.1 The diagonalization theorem

In Proposition 2.5.1 the set S lies on an affine algebraic variety, and the proof goes by embedding a curve given by binomial-good polynomials. This is a way of finding separations, but it remains unclear if such curves always exist or how we can find them. In general, if S lies on an affine variety Z with vanishing ideal I, then we know that if there exists a polynomial $\xi \in I$ such that $\varphi + \xi$ is binomial-good, then φ is a polynomial closure property of $\#\mathsf{P}$ on S. This is exactly the insight that gives $\textsc{Sperner} - 1 \in \#\mathsf{P}$, where all instances lie on the variety $\{(t_+, t_-) \in \mathbb{N}^2 : t_+ - t_- - 1 = 0\}$.

The reverse is true in the important case of graph varieties (all our examples fall in this category), as we show in the following Diagonalization Theorem. Formally, assume that there exist $\ell \in \{0, \ldots, k\}$, and polynomial maps $\zeta_b : \mathbb{Q}^\ell \to \mathbb{Q}$, where $b \in \{\ell + 1, \ldots, k\}$, such that Z is the image $\big(f_1, \ldots, f_\ell, \zeta_{\ell+1}(f_1, \ldots, f_\ell), \ldots, \zeta_k(f_1, \ldots, f_\ell)\big)$. In this case the vanishing ideal I is generated by the $\zeta_b - f_b$, see §5.3. We call a coset $\varphi + I$ binomial-good, if it contains a binomial-good polynomial, otherwise $\varphi + I$ is binomial-bad.

Theorem (Diagonalization Theorem, *informal version, see* Thm 6.2.1)**.** *Fix k and $0 \le \ell \le k$. Let $\varphi \in \mathbb{Q}[f_1, \ldots, f_k]$. Fix functions $\zeta_b \in \mathbb{Q}[f_1, \ldots, f_\ell]$. Set I to be the ideal generated by the $\zeta_b(f_1, \ldots, f_\ell) - f_b$, where $\ell + 1 \le b \le k$. Fix a function $\textsc{Multiplicities} : 2^{\{0,1\}^*} \to \mathbb{N}^k$ and let $\vec{t} \in S$. For $A = \bigcup_{j \ge 0} A_j$, $A_j \subseteq \{0,1\}^j$, $\widetilde{A}_j$ being the set of length $j-1$ suffixes of the strings in A_j that start with 1, define*

$$p_A(w) := \varphi(\textsc{Multiplicities}(\widetilde{A}_{|w|})) \quad \text{if} \quad A_{|w|}(0^{|w|}) = 1, \quad \text{and} \quad p_A(w) := \varphi(\vec{t}) \quad \text{otherwise.}$$

Assume further technical conditions, e.g., the existence of set-instantiators for $\textsc{Multiplicities}$*. If $\varphi + I$ is binomial-bad, then there exists $A \subseteq \{0,1\}^*$, such that for every nondeterministic polytime Turing machine M there exists j, such that $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$; and whenever $A(0^j) = 1$, we have* $\textsc{Multiplicities}(\widetilde{A}_j) \in S$.

The Diagonalization Theorem 6.2.1 is the technical heart of this paper. It is stated in high generality, and we apply it to a large set of examples of very different flavor, such as for example the Hadamard inequality or $\#\mathsf{PPA} - 1$. Its proof relies on the Witness Theorem 5.4.1, whose proof uses methods from several areas of mathematics including algebraic geometry and Ramsey theory.

As an illustration we now apply the Diagonalization Theorem to Karamata's inequality, see §7.5 for the full details. In the Karamata setting we are given $f_i, g_i \in \#\mathsf{P}$, $1 \le i \le n$, such that the following functions h_i, $1 \le i < n$, are also all in $\#\mathsf{P}$:

$$h_i := f_1 + \ldots + f_i - g_1 - \ldots - g_i,$$

and we are also guaranteed that

$$f_1 + \ldots + f_n - g_1 - \ldots - g_n = 0. \tag{3.1.1}$$

This assumption is called *majorization*, see §7.5. Moreover, the functions

$$d_i := f_i - f_{i+1} \quad \text{and} \quad e_i := g_i - g_{i+1} \tag{3.1.2}$$

are also in $\#\mathsf{P}$ for all $1 \le i < n$. Let $Z \subseteq \mathbb{Q}^{5n-3}$ denote the variety of points that satisfy the constraints (3.1.1) and (3.1.2), and let $S = Z \cap \mathbb{N}^{5n-3}$. Let $\gamma \in \mathsf{GapP}_{\ge 0}$ be any monotone integer-valued convex function. Define the Karamata function as

$$K_{n,\gamma}(\vec{f}, \vec{g}) := \sum_{i=1}^{n} \gamma(f_i) - \sum_{i=1}^{n} \gamma(g_i).$$

Clearly $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in S}) \subseteq \mathsf{GapP}$. In fact, even $K_{n,\gamma}(\overrightarrow{\mathsf{GapP}}) \subseteq \mathsf{GapP}$. Karamata's inequality implies that the answer is always nonnegative for inputs from S. Hence, $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in S}) \subseteq \mathsf{GapP}_{\geq 0}$. For which γ do we have $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in S}) \subseteq \#\mathsf{P}$?

- For affine linear γ we clearly have $K_{n,\gamma} = 0 \in \#\mathsf{P}$.
- For $\gamma(t) = t^2$, we have $K_{2,\gamma}(f_1, f_2, g_1, g_2) = (d_1 + e_1)h_1$ on S. This can be seen by plugging in $d_1 = f_1 - f_2$, $e_1 = g_1 - g_2$, and $g_2 = f_1 + f_2 - g_1$. Clearly $(d_1 + e_1)h_1 \in \#\mathsf{P}$. This has several proofs, for example instead of $(d_1 + e_1)h_1$ we could have taken $2h_1 + 2e_1h_1 + 4\binom{h_1}{2}$ with the same argument.
- For $\gamma(t) = t^2$, we have $K_{3,\gamma}(f_1, f_2, f_3, g_1, g_2, g_3) = (d_1 + e_1)h_1 + (d_2 + e_2)h_2 \in \#\mathsf{P}$ on S.
- For $\gamma(t) = \binom{t}{2}$, we have $K_{2,\gamma}(f_1, f_2, g_1, g_2) = (e_1 + 1)h_1 + 2\binom{h_1}{2} \in \#\mathsf{P}$ on S.
- For $\gamma(t) = \binom{t}{2}$, we observe that for the *double* we have $2K_{3,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in S}) \subseteq \#\mathsf{P}$ via the observation that $2K_{3,\binom{t}{2}} = K_{3,t^2}$ on S (the affine linear parts cancel out).

All inclusions $K_{n,\gamma} \subseteq \#\mathsf{P}$ in this section so far relativize. The next proposition shows that the doubling we just used was in fact necessary, because otherwise we obtain an oracle separation.

Proposition (see Proposition 7.5.5)**.** *There exists $A \subseteq \{0,1\}^*$ such that $K_{3,\binom{t}{2}}(\overrightarrow{\#\mathsf{P}}^A_{\in S}) \not\subseteq \#\mathsf{P}^A$.*

Proof sketch. We use the Diagonalization Theorem 6.2.1. We have $5n - 3 = 12$, so $S = Z \cap \mathbb{N}^{12}$. We fix an arbitrary order of the 12 variables: $(f_1, f_2, f_3, g_1, g_2, g_3, d_1, d_2, e_1, e_2, h_1, h_2)$. The variety Z is then given as the kernel of the linear map given by the following left matrix:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & -2 \end{pmatrix}$$

To obtain the necessary parametrization ζ of Z, we convert it to row echelon form (the fact that the entries are integer is convenient, but not necessary for our techniques to apply), which is the right matrix. We set $\ell = 5$, $k = 12$. Permuting the order of the columns to $(6, 9, 10, 11, 12, 1, 2, 3, 4, 5, 7, 8)$, we obtain affine linear functions $\zeta_8, \ldots, \zeta_{12}$, each ζ_b depends linearly on the first 5 variables. We set $\textsc{Multiplicities} = \textsc{OccurrenceMulti}_{12}$, which is the function defined as follows: on input $w \in \{0,1\}^*$ we split w into 12 parts of roughly the same size, and the output is the vector that specifies how many 1s are in the first part, how many 1s are in the second part, and so on. We set

$$\varphi(f_1, f_2, f_3, g_1, g_2, g_3, d_1, d_2, e_1, e_2, h_1, h_2) \;:=\; f_1^2 + f_2^2 + f_3^2 - g_1^2 - g_2^2 - g_3^2\,.$$

After verifying the technical assumptions, to apply the theorem it remains to show that $\varphi + I$ is binomial-bad. Since all ζ_b are affine linear, this claim boils down to checking that polyhedron $\mathcal{P}_{\varphi,\zeta}$ does not have integer points. We formalize and generalize this implication in the Polyhedron Theorem 5.5.2. Here we have a polyhedron in $\mathbb{Q}^{91}$ given by 21 linear equations intersected with the nonnegative orthant. We use a computer to set up the polyhedron. Indeed, it contains the half-integer point that shows that $2K_{3,\binom{t}{2}}(\overrightarrow{\#\mathsf{P}}_{\in S}) \subseteq \#\mathsf{P}$, but it *does not* contain an integer point, which gives A such that $\varphi(p_A) \notin \#\mathsf{P}^A$, but $\varphi(p_A) \in K_{3,\binom{t}{2}}(\overrightarrow{\#\mathsf{P}}^A_{\in S})$. □

The Polyhedron Theorem 5.5.2 is also what we use for studying counting classes from TFNP (see §3.2), but the corresponding polyhedra there are very simple, see e.g. Theorem 8.3.1 and Theorem 8.9.1. The polyhedron for SPERNER for example has the integer point $(2,0)$ to represent that $\varphi = 2t_- + 0t_+$ on the variety. The technical difficulty in those cases is not the polyhedron, but the existence of set-instantiators. If we just study closure properties of #P, then trivial set-instantiators can be used.

3.2 Counting classes and TFNP

In this section define the counting classes for which we claimed in (4) in §1.3 that many of them coincide with #P, while others are strictly stronger w.r.t. an oracle. In order to do so, we attach oracles to the syntactic subclasses of TFNP.[10]

Consider for example the relation rLONELY. This is for PPA, as the other classes are handled analogously. Let $(C,x) \in$ rLONELY if and only if

$$x \neq 0 \ \wedge \ \big(C'(x) = x \ \vee \ C'(C'(x)) \neq x\big),$$

where C is the description of a polynomially-sized multi-output Boolean circuit that describes the partner function on an exponentially large graph, and C' is the syntactic modification to C that ensures that $C'(0) = 0$.

Now, let rPPA be the set of polynomially balanced relations R for which a pair (α,β) of polytime computable maps exists with $(C,\beta(x)) \in R$ if and only if $(\alpha(C),x) \in$ rLONELY. These are the relations that correspond to search problems in PPA. Let rP denote the set of polynomially balanced relations that can be evaluated in polynomial time. By definition, rPPA $\subseteq$ rP.

For a language $A \subseteq \{0,1\}^*$ we define analogously $(C,x) \in$ rLONELY^A if and only if $\big(C(x) = x \ \vee \ C(C(x)) \neq x\big)$, but now we allow the circuit C to have arbitrary arity oracle gates that query the oracle A. Let rPPA^A be the set of polynomially balanced relations R for which a pair (α,β) of polynomial-time computable maps exists with $(C,\beta(x)) \in R$ if and only if $(\alpha(C),x) \in$ rLONELY^A. Note here that the only difference is that $\alpha(C)$ can have oracle gates. Let rP^A denote the set of polynomially balanced relations that can be evaluated in polynomial time with access to A. By definition, we have rPPA$^A \subseteq$ rP^A.[11]

We define the corresponding counting class #PPA^A via

$$f \in \#\mathsf{PPA}^A \iff \exists R \in \mathsf{rPPA}^A \ : \ f(w) = \textstyle\sum_{y\in\{0,1\}^*} R(w,y).$$

Recall that

$$f \in \#\mathsf{P}^A \iff \exists R \in \mathsf{rP}^A \ : \ f(w) = \textstyle\sum_{y\in\{0,1\}^*} R(w,y).$$

Hence, clearly #PPA$^A \subseteq$ #P^A, and in fact #PPA$^A \subseteq$ #P$^A_{\geq 1}$ for all languages A.[12] [13]

For the study of whether or not a problem is in #P we need the finer viewpoint that is obtained when insisting on (α,β) being a *parsimonious reduction*, i.e., $\big((C,\beta(x)) \in R$ and $(C,\beta(y)) \in R\big)$ implies $x = y$. Since not all PPA-complete problems are equivalent to each other via parsimonious reductions, this gives rise to different counting complexity classes, depending on the PPA-complete

[10] We consider CLS, PLS, PPAD, PPADS, PPA, and PPP here, see e.g. [GP17]. The instances are exponentially large (di)graphs given succinctly by circuits or lists of circuits. For the sake of simplicity, we will assume in this discussion that finite lists of circuits are merged into a single circuit with additional input bits.

[11] We use the r-prefix to avoid notational issues similar to the ones discussed in [HV95, BS01]. We do not claim to have found a particularly good notation, but a suggestive one.

[12] In fact, #P$_{\geq 1}$ can be thought of as the counting analogue of TFNP, i.e., it is reasonable to define #TFNP := #P$_{\geq 1}$.

[13] We choose this approach, which is different from the type-2 complexity approach in [BCE+98, BM04], because we want to compare our counting classes to #P^A.

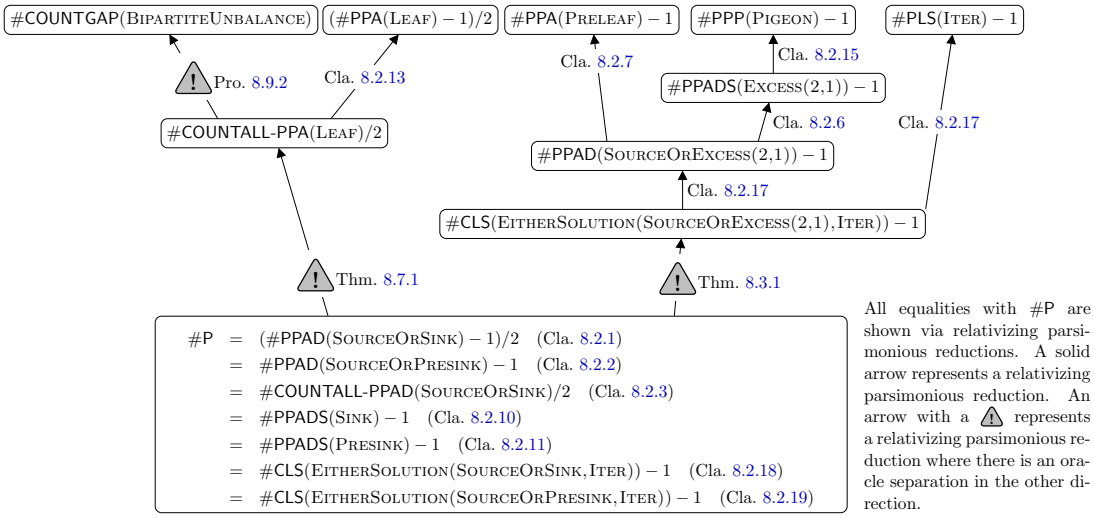


Figure 1: The relativizing equalities and inclusions; and the oracle separations.

problem. We write #PPA(P) to indicate that we mean the counting class defined by problem P under parsimonious reductions.[14] For example, observe that all functions in #PPA(Lonely) output odd integers on every input, while the class #PPA contains more functions than that (as we will see when discussing Preleaf). In fact, for that reason it makes sense to study the class (#PPA(Lonely) + 1)/2 and (#PPA(Lonely) − 1)/2 and ask if they are subsets of #P.

Aside from the classical problems we study slightly adjusted problems that are not equivalent via parsimonious reductions (see the detailed list in §8.1). Each class is defined via parsimonious reductions to a complete problem. The naming prefixes #COUNTALL-PPA and #COUNTGAP are essentially flavor. By definition we have #PPA(P) ⊆ #PPA for all search problems P ∈ PPA, and analogously for all other search problems. The relativizing inclusions and oracle separations that we find are depicted in Figure 1. All equalities with #P are shown via relativizing parsimonious reductions and they are proved in §8.2. A solid arrow represents a relativizing parsimonious reduction. An arrow with a ⚠ represents a relativizing parsimonious reduction where there is an oracle separation in the other direction. This means that all classes that are above #P in the figure strictly contain #P (with respect to an oracle).

We find a surprisingly large number of counting problem classes that, if adjusted properly with "−1" and "/2" are *equal to* #P. This includes the canonical counting versions of PPAD, PPADS and CLS. Only after making slight changes to the problems via non-parsimonious polytime equivalences (similar to the *chessplayer algorithm*, see e.g. [Pap90, BCE+98]), we obtain that the non-parsimonious counting classes strictly contain #P, which puts the new problems outside of #P.

We identify two main "reasons" (i.e., oracle separations), one for "−1" (the decrementation separation, see §8.3), and one for "/2" (the halving separation, see §8.7). There are two versions of #PPA at the top of the diagram, one for each of the two reasons, and they are not easily comparable.[15] It is also noteworthy that we know of no counting version of PPA, PLS or PPP that equals #P.[16]

[14] Note that the prefix PPA is actually superfluous in this case, but we keep it for clarity.

[15] It is not even clear if #PPA(Leaf) − 1 is contained in (#PPA(Leaf) − 1)/2. One would want to just double a Leaf instance, but that will end up creating 2 leaves too many. In other words, the class #PPA(Leaf) − 1 seems to not be closed under the operation of scaling a function by 2.

[16] Note that since PTFNP (see [GP18]) contains CLS, the decrementation separation also shows $\#\mathsf{PTFNP}^{A} - 1 \not\subseteq \#\mathsf{P}^{A}$.

Since the polynomials $x-1$, $\frac{x}{2}$, and $\frac{x-1}{2}$ are monotone, the tool to prove the separations is the Diagonalization Theorem 6.2.1. The main difference from all separations so far is that now the instances are much more involved. In the halving separation we have to "hide" the partner vertex from the #P machine, and in the decrementation separation we have to "hide" which of the solutions is connected to the zero vertex. This is especially difficult for $\#\mathsf{PLS}(\textsc{Iter})-1$ (and hence for $\#\mathsf{CLS}-1$).[17] We formalize our approach in the definition of a set-instantiator in Definition 6.1.3 and the necessary set-instantiators are created in Section 8.

Even though we are mainly interested in membership and non-membership in #P, with only a little extra work our techniques directly give us another oracle separation

$$\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2 \subsetneq \#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A.$$

This is because after doubling we have $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A \subseteq \#\mathsf{P}^A$, while still

$$\#\mathsf{P}^A \subsetneq 2\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A$$

(see Proposition 8.9.2).

3.3 Structure of the paper

Section 4 gives a high-level introduction of the proof ideas that lead to oracle separations and the binomial basis theorem, which is a first step towards our Diagonalization Theorem 6.2.1. Section 5 proves the Witness Theorem 5.4.1, which is a generalization of the naive oracle separation proof approach to all nontrivial graph varieties. It is a key ingredient of the Diagonalization Theorem, which is introduced in Section 6, which uses the notation of set-instantiators, a formal way to treat not only the closure properties of #P, but syntactic subclasses of TFNP.

Section 7 uses the theory that we developed in the earlier chapters to handle the details of the Cauchy inequality, the Alexandrov–Fenchel inequality, the Hadamard inequality, Fermat's little theorem (in #P), the Ahlswede–Daykin inequality, and the Karamata inequality. Section 8 establishes all necessary set-instantiators and proves all oracle separations from Figure 1. We conclude with final remarks and open problems in Section 9.

4 The binomial basis

4.1 Oracle separations and the binomial basis: An informal high-level view

(1) **Fooling polytime Turing machines with oracles.** The initial idea is classical. Consider the following problem ConnectedComponents. A problem instance is similar to an instance to the PPA-complete problem Sink, i.e., we have an exponentially large undirected graph whose nodes have degree at most two, but in the ConnectedComponents problem we are not guaranteed a sink at node zero. The graph is given as a pair of two circuits C_1 and C_2; an edge from x to y is present in G if and only if $\big((C_1(x)=y \text{ or } C_2(x)=y) \text{ and } (C_1(y)=x \text{ or } C_2(y)=x)\big)$. The function ConnectedComponents counts the number of connected components in G, ignoring isolated nodes.

We can quickly see that if we replace the circuits with black boxes (i.e., oracles), then this problem is not in #P. Assume there exists a nondeterministic Turing machine M for which the number of accepting paths $\#\mathsf{acc}_M(G) = \textsc{ConnectedComponents}(G)$ for all G. Let G be an

[17] Notably, PLS is also missing from the oracle separations in [BCE+98].

instance with 2 connected components of superpolynomial size each. Then M has 2 accepting paths, but every path can only access a polynomial number of oracle positions. Therefore we can take two edges $\{u, v\}$ and $\{x, y\}$ of G that lie in different connected components and that are not queried by M, we remove them, and replace them with the two edges $\{u, x\}$, $\{v, y\}$ to obtain a graph G' with only 1 connected component. But since the two accepting paths of M do not query the oracle at these positions, M will have at least two accepting paths on the instance G', which is a contradiction.

(2) **Combinatorial Diagonalization.** This simple idea described above can be used in more sophisticated ways, see [CGH+89, Thm 3.1.1]. We take the counting problem CircuitSat in which we are given a Boolean circuit and the number of accepting paths is supposed to be the number of inputs for which the circuit outputs True. Clearly CircuitSat $\in \#\mathsf{P}$. We modify the problem by defining a function $\varphi : \mathbb{N} \to \mathbb{N}$ with $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(2) = 1$, and the other values are not relevant. Replace the circuit with an oracle, so instead of counting input strings for which the circuit outputs True, we now search for strings for which the oracle query returns 1. We now argue that $\varphi(\text{CircuitSat}) \notin \#\mathsf{P}$ with respect to that oracle (if the oracle is chosen correctly).

To see this, assume the existence of a nondeterministic Turing machine M with $\#\mathsf{acc}_M(C) = \varphi(\text{CircuitSat}(C))$. If a computation path of M does not query any 1 at any oracle position, then it cannot accept. Indeed, if it does accept, then we can change the oracle to all zeros to obtain a contradiction. If there is exactly one oracle position with a 1 (say, at position a), then M must have exactly one accepting path and that path queries the oracle at a. Now, if we change the oracle by setting it to zero at a and to 1 at b, where b was not queried by the accepting path of M before, then we get another accepting path. Let us call this accepting path τ. What happens if we take the oracle that has a 1 at position a and a 1 at position b at the same time? If τ does not query the oracle at position a, then *both* accepting paths will accept on this instance, which is a contradiction, because $\varphi(2) = 1 < 2$. The key point is that this happens *with high probability* if the positions are chosen uniformly at random.

(3) **The binomial basis.** The binomial basis theorem classifies completely the set of polynomials φ for which this argument works. If we pick a polynomial φ in the discussion above, it could have been $\varphi(f) = \frac{1}{6}f^3 - f^2 + \frac{11}{6}f$ for example. This is a monotone function whose values are nonnegative integers, so we have no immediate way rule out membership in $\#\mathsf{P}$. The key insight that relates this polynomial to what we observed is when we write it in its *binomial basis*, the basis of the vector space of polynomials that is spanned by the binomial coefficients: $\varphi(f) = f - \binom{f}{2} + \binom{f}{3}$. The problematic issue is the negative coefficient -1 for $\binom{f}{2}$. This -1 is exactly the amount by which we overshoot with our number of accepting paths when the oracle has two 1s. We call a polynomial binomial-good if all its coefficients in the binomial basis are nonnegative integers, otherwise it is binomial-bad.

This idea of univariate polynomial closure properties of $\#\mathsf{P}$ generalizes to the *multivariate case*, i.e., for φ a multivariate polynomial. Here we again have a (multivariate) binomial basis, and say that a multivariate polynomial is binomial-good if all its coefficients in this basis are nonnegative integers.

(4) **The multivariate binomial basis.** The multivariate case is of high interest when studying inequalities in combinatorics, but it is also already important when studying concrete instantiations such as $\#\mathsf{PPAD} - 1$. For a SourceOrSink instance (which is parsimoniously equivalent to Sperner) we want to count the number number of nonzero sources f plus the number of sinks g minus 1, i.e. $\varphi(f, g) = f + g - 1$. We know that in the SourceOrSink problem, for all instances we have $f - g + 1 = 0$. This makes it possible to calculate $f + g - 1 = 2f$, which implies $\#\mathsf{PPAD}(\text{SourceOrSink}) - 1 \subseteq \#\mathsf{P}$. It is useful to think of the two functions $f + g - 1$ and $2f$

as being the same function in the quotient ring $\mathbb{Q}[f,g]/\langle f-g+1\rangle$. This ring is known as the *coordinate ring* $\mathbb{Q}[Z]$ of the affine variety $Z = \{(f,g) \mid f-g+1=0\}$.

We will employ this viewpoint in Section 5 to prove the Witness Theorem 5.4.1, which says that in many situations our notion of "binomial-good on an affine variety" is consistent with our intuition. The main challenge in proving this theorem is that some nonnegative integer points on the variety Z might actually *not* correspond to a problem instance. For SourceOrSink they all do, but for example for Iter we have $Z = \mathbb{Q}$ and there is no instance with 0 solutions; or for AllLeaves we also have $Z = \mathbb{Q}$, but there is no instance with an odd number of solutions. The saving grace is that in all our cases the variety Z is a graph variety and hence we can study asymptotic behavior via using a variant of *Ramsey's theorem*.

The case SourceOrSink is a hyperplane, i.e., given by a single equation, but in general we have more equations, for example for #CLS(EitherSolution(SourceOrExcess(2,1),Iter)), or when treating Karamata's inequality in §7.5. In the case where all constraints are affine linear (and φ is arbitrary) we can rephrase the definition of binomial-goodnes of φ on a graph variety given by functions ζ as follows: φ is binomial-good on the graph variety if and only if the polyhedron $\mathcal{P}_{\varphi,\zeta}$ contains an integer point (see the Polyhedron Theorem 5.5.2). These integer points correspond directly to #P algorithms, for example for #PPAD(SourceOrSink) -1 there is an integer point $(2,0)$ to represent that $\varphi = 2f + 0g$ on the variety.

(5) **Set-instantiators.** To make these ideas work for actual problem instances such as SourceOrSink or Iter instead of just for oracles for which we count the 1s, we introduce the notion of a set-instantiator and we construct the set-instantiators for all relevant cases. Intuitively, a set-instantiator mimics the behavior of the oracle idea presented above. For the existence of a set-instantiator one defines how to set up random instances of specified cardinalities that sit nicely in each other so that the Turing machine M is fooled and behaves in the same way as for a CircuitSat instance. This becomes very challenging when the problem does not allow permuting the instance around, which is the case for Iter, where the problem is always directed in a fixed direction. The #PLS(Iter) case is the most challenging (which carries over to #CLS), and we use a refined treatment.

4.2 The binomial basis and integer-valued functions

This section covers some basic properties about the binomial basis. Fix k. Let $S \subseteq \mathbb{N}^k$ and let $D := \{\vec{b} \in \mathbb{N}^k \mid \exists \vec{a} \leqslant \vec{b} : \vec{a} \in S\}$ denote the downwards closure of S. It is instructive to think of $D = \mathbb{N}^k$, although we will need it in greater generality in Section 5, where we work with functions that are not defined on all of $\mathbb{N}^k$.

We are interested in functions $D \to \mathbb{N}$ and functions $\mathbb{Z}^k \to \mathbb{Z}$, and also in function $S_{\leqslant} \to \mathbb{N}$. We focus on $D \to \mathbb{N}$ first. This set of functions lies in the $\mathbb{Q}$-vector space of all functions $D \to \mathbb{Q}$. For $\vec{b} \in D$ we define the binomial function $\beta_{\vec{b}} : D \to \mathbb{Q}$ via

$$\beta_{\vec{b}}(x_1,\ldots,x_k) \;:=\; \binom{x_1}{b_1}\cdots\binom{x_k}{b_k} \;=\; \binom{\vec{x}}{\vec{b}}.$$

The *binomial basis* is defined as the set of all binomial functions $\{\beta_{\vec{b}} \mid \vec{b} \in D\}$. Recall that a function $D \to \mathbb{Q}$ or $\mathbb{Z}^k \to \mathbb{Q}$ is *integer-valued* if it only attains integer values.

4.2.1 Proposition (cf. [Nag19])**.** *Every function $\varphi : D \to \mathbb{Q}$ can be expressed as a (possibly infinite) $\mathbb{Q}$-linear combination of elements from the binomial basis. This expression is unique. The function φ is integer-valued if and only if its coefficients in the binomial basis are all integer.*

If we only consider polynomials φ, then this result is classical and slightly easier to prove. The following is a variation on the original argument by Nagell [Nag19] (see also [Nar95, p. 13]).

Proof. We first show uniqueness. We assume for the sake of contradiction that a function has two distinct expressions in the binomial basis. Then the difference of them is a nontrivial expression of the zero function:

$$0 = \sum_{\vec{b}\in\mathbb{N}^k} c_{\vec{b}}\beta_{\vec{b}}\,, \quad \text{where } c_I \in \mathbb{Q} \text{ and not all } c_{\vec{b}} \text{ are zero.} \tag{$\dagger$}$$

Let $m \in \mathbb{N}$ denote the smallest number for which $c_{\vec{b}} \neq 0$, $|\vec{b}| = m$. Every $\beta_{\vec{b}}$ has a unique monomial of largest total degree, which is $x^{\vec{b}} := x_1^{b_1}\cdots x_k^{b_k}$ of degree $|\vec{b}|$. Moreover,

$$\text{if } \beta_{\vec{a}}(x_1,\ldots,x_k) \neq 0\,, \quad \text{then } \vec{x} \geqslant \vec{a}. \tag{$*$}$$

Therefore ($\dagger$) induces a nontrivial linear combination of 0 in the homogeneous degree m part:

$$0 = \sum_{\vec{b}\in D,\,|\vec{b}|=m} \frac{c_{\vec{b}}}{b_1!\cdots b_k!}\,x^{\vec{b}}\,, \quad \text{where } c_{\vec{b}} \in \mathbb{Q} \text{ and not all } c_{\vec{b}} \text{ are zero.}$$

Note that this is a finite linear combination. But the monomials $x^{\vec{b}}$ are linearly independent functions on D, so all $c_{\vec{b}} = 0$, which is a contradiction. Hence the uniqueness is proved.

It remains to show that every function $\varphi : D \to \mathbb{Q}$ can be expressed over the binomial basis (not necessarily with finite support). The coefficients $c_{\vec{b}}$ for expressing φ over the binomial basis are constructed using the following recursive property. We set

$$c_{0,\ldots,0} := \varphi(0,\ldots,0) \qquad \text{and} \qquad c_{\vec{b}} := \varphi(\vec{b}) - \sum_{\vec{a}\leqslant\vec{b},\,\vec{a}\neq\vec{b}} c_{\vec{a}}\beta_{\vec{a}}(\vec{b}) \tag{4.2.2}$$

This defines all $c_{\vec{b}}$. We have $\varphi = \sum_{\vec{b}\in\mathbb{N}^k} c_{\vec{b}}\beta_{\vec{b}}$, because $\beta_{\vec{b}}(\vec{b}) = 1$, and ($*$) implies that $\beta_{\vec{a}}(\vec{b}) = 0$ for all other $\vec{a}$ over which is not summed.

Since the $\beta_{\vec{b}}$ are integer-valued functions we have that if the coefficients in the binomial basis are integers, then the function is integer-valued. By construction in the proof above (see how $c_{\vec{b}}$ is defined) this works in the other direction as well: if φ is integer-valued, then the coefficients in the binomial basis are all integer. □

For $D = \mathbb{N}^k$, we call polynomials φ whose expression over the binomial basis has only nonnegative integer coefficients *binomial-good*, all others are called *binomial-bad*. For $D \subseteq \mathbb{N}^k$ we call φ *D-good* if the coefficients over the binomial basis are nonnegative integers; otherwise φ is *D-bad*.

We remark that Proposition 4.2.1 generalizes from $D = \mathbb{N}^k$ to functions $\mathbb{Z}^k \to \mathbb{Q}$ as follows, with basically the same proof. For $\vec{b} \in \mathbb{N}^k$, we define the function $\widetilde{\beta}_{\vec{b}} : \mathbb{Z}^k \to \mathbb{Q}$ via

$$\widetilde{\beta}_{\vec{b}}(x_1,\ldots,x_k) := \binom{x_1 - \lfloor b_1\rfloor}{b_1}\cdots\binom{x_k - \lfloor b_k\rfloor}{b_k}.$$

The *shifted binomial basis* is defined as the set of all binomial functions $\{\,\widetilde{\beta}_{\vec{b}} : \vec{b} \in \mathbb{N}^k\,\}$.

4.2.3 Proposition. *Every function $\varphi : \mathbb{Z}^k \to \mathbb{Q}$ can be expressed as a (possibly infinite) $\mathbb{Q}$-linear combination of elements from the shifted binomial basis. This expression is unique. The function φ is integer-valued if and only if its coefficients in the shifted binomial basis are all integer.*

For polynomials it does not matter which of the two bases we use, so we use the simpler (unshifted) one. We remark that if φ is a multivariate polynomial (nonnegative or not) of degree d, then its expression over the binomial basis has finite support. This just follows from the fact that the set of all $\beta_{\vec{b}}$ with $|\vec{b}| \le d$ is linearly independent, and hence (by counting the dimension) is a basis of the space of polynomials of degree $\le d$.

4.3 The binomial basis theorem

We start with the following general definition formalizing and generalizing definitions in §2.2.

4.3.1 Definition. *A function $\varphi : \mathbb{N}^k \to \mathbb{N}$ is a* closure property of $\#\mathsf{P}$ *if the following holds: if $f_1, \dots, f_k \in \#\mathsf{P}$, then the function $\varphi(f_1, \dots, f_k)$ defined via*

$$\big[\varphi(f_1, \dots, f_k)\big](x_1, \dots, x_k) \;:=\; \varphi\big(f_1(x_1), \dots, f_k(x_k)\big)$$

is in $\#\mathsf{P}$. In other words, $\varphi\big(\overrightarrow{\#\mathsf{P}}\big) \subseteq \#\mathsf{P}$. We analogously define closure properties of $\#\mathsf{P}^A$ for any language A as follows.

We say that a closure property φ of $\#\mathsf{P}$ relativizes *if for every language A the function φ is a closure property of $\#\mathsf{P}^A$. A function $\varphi : \mathbb{Z}^k \to \mathbb{Z}$ is a closure property of GapP if $\varphi\big(\overrightarrow{\mathsf{GapP}}\big) \subseteq \mathsf{GapP}$. We say that a closure property φ of GapP* relativizes *if for every language A the function φ is a closure property of GapP^A.*

4.3.2 Theorem (Binomial basis theorem, see [HVW95, Thm. 3.13] and [Bei97, Thm. 6], stated above in §2.5)**.**
The following properties for a multivariate polynomial φ are equivalent:

- *φ is a relativizing closure property of GapP,*
- *φ is a closure property of GapP,*
- *φ is integer-valued,*
- *the expression of φ over the binomial basis has only integer coefficients.*

Moreover, the following are equivalent:

- *φ is a relativizing closure property of $\mathsf{GapP}_{\ge 0}$,*
- *φ is a closure property of $\mathsf{GapP}_{\ge 0}$,*
- *φ is integer-valued and attains only nonnegative integers,*
- *the expression of φ over the binomial basis has integer coefficients and φ attains only nonnegative integers if evaluated at integer points in the nonnegative cone.*

Moreover, the following are equivalent:

- *φ is a relativizing closure property of $\#\mathsf{P}$,*
- *the expression of φ over the binomial basis has only nonnegative integer coefficients.*

Proof. If φ is a closure property of GapP, then clearly φ is integer-valued. If φ is integer-valued, then its expression over the binomial basis has finite support and all coefficients are integers by Proposition 4.2.1. In [FFK94], see closure property 5, they present the proof that GapP^A is closed under taking binomial coefficients, addition, and product. Thus, every integer-valued φ whose its expression over the binomial basis has finite support, and all coefficients are integers is a relativizing closure property of GapP. Clearly, every relativizing closure property of GapP is a closure property of GapP, which shows the equivalence of the first four items. The same argument (with the obvious minimal modifications) shows the equivalence of the second four items.

If the expression of φ over the binomial basis has finite support and all coefficients are non-negative integers, then φ is a relativizing polynomial closure property of $\#\mathsf{P}$, because the $\#\mathsf{P}$ closure proofs for addition, multiplication, and taking binomial coefficients relativize. It remains to show the converse. This is more technical and is a simple application of the much more general Diagonalization Theorem 6.2.1, and hence we postpone the proof to Theorem 6.7.1. □

4.4 The binomial basis conjecture

Note that if the polynomial closure properties of $\#\mathsf{P}$ all relativize, then Theorem 4.3.2 gives a complete classification of all polynomial closure properties of $\#\mathsf{P}$: Those φ whose expression over the binomial basis only nonnegative integers as coefficients. We conjecture that this is indeed the correct classification:

4.4.1 Conjecture (Binomial basis conjecture)**.** *The polynomial closure properties of* $\#\mathsf{P}$ *all relativize.*

Note that Theorem 4.3.2 says that Conjecture 4.4.1 is true if we replace $\#\mathsf{P}$ by GapP or by $\mathsf{GapP}_{\geq 0}$. Note also that Conjecture 4.4.1 is equivalent to saying that the polynomial closure properties of $\#\mathsf{P}$ are exactly the binomial-good polynomials. Sometimes it is sufficient to use the weaker univariate version of the binomial basis conjecture:

4.4.2 Conjecture (Univariate binomial basis conjecture)**.** *The univariate polynomial closure properties of* $\#\mathsf{P}$ *all relativize.*

We do not know if these two conjectures are equivalent. Even the univariate version of the binomial basis conjecture implies $\mathsf{P} \neq \mathsf{NP}$ and will therefore be very difficult to prove.

4.4.3 Theorem. *Conjecture 4.4.2 implies* $\#\mathsf{P} \neq \#\mathsf{P}^{\mathsf{NP}}$ *(and hence, in particular,* $\mathsf{P} \neq \mathsf{NP}$*).*

Proof. This is similar to the idea in [OH93, Thm 3.12]. We claim that if $\#\mathsf{P} = \#\mathsf{P}^{\mathsf{NP}}$, then $\binom{x-1}{2} = \frac{1}{2}x^2 - \frac{3}{2}x + 1 = \binom{x}{0} - \binom{x}{1} + \binom{x}{2}$ is a polynomial closure property of $\#\mathsf{P}$. This is shown as follows.

Given a function $f \in \#\mathsf{P}$ we construct the following $\#\mathsf{P}^{\mathsf{NP}}$ machine: call the NP oracle to see if there is at least one witness. If not, accept. Otherwise, count pairs of distinct witnesses that are both not the smallest witness (an NP oracle call is used to see if there is a smaller witness). The resulting function g is in $\#\mathsf{P}^{\mathsf{NP}}$. By construction, if $f(w) = 0$, then $g(w) = 1$. Moreover, if $f(w) > 0$, then $g(w) = \binom{f(w)-1}{2}$. Since $\binom{0-1}{2} = \frac{(-1)(-2)}{2} = 1$, it follows that $g = \binom{f-1}{2}$. □

5 #P on affine varieties: The witness theorem

In this section we prove the Witness Theorem 5.4.1, which is the crucial theorem for dealing with different sets S, in particular if S is a subset of an affine algebraic variety, such as for example for #PPAD(SourceOrSink), or if S has holes such as for #PPA(Leaf).

5.1 Notation for the witness theorem

Denote $[\![d]\!] := \{0, 1, 2, \ldots, d\}$, which is not to be confused with $[d] = \{1, 2, \ldots, d\}$. Fix k. As before, we write $\vec{f} := (f_1, \ldots, f_k)$ for vectors of length k. Let $\mathbb{O} := \mathbb{N}^k$ be the set of integer points in the *nonnegative orthant*. As before, for $\vec{f}, \vec{g} \in \mathbb{O}$, we write $\vec{g} \leqslant \vec{f}$ if and only if $g_a \leq f_a$ for all $1 \leq a \leq k$. For $\vec{g} \leqslant \vec{f}$, we write

$$\binom{\vec{f}}{\vec{g}} := \prod_{a=1}^{k} \binom{f_a}{g_a}.$$

Let $S \subseteq \mathbb{O}$, and let $\varphi \in \mathbb{Q}[\vec{f}]$. We want to determine whether or not φ is a relativizing multivariate closure property of #P under the guarantee that the cardinalities of the instances only come from the set S. We will see that under certain assumptions on the Zariski closure of S, this only depends on φ and the Zariski closure of S. In fact, in many cases this can be characterized to exactly be the case when $\varphi + I$ contains a binomial-good polynomial, where I is the vanishing ideal of the set S.

Let $I = I(S) := \{\varphi \in \mathbb{Q}[\vec{f}] : \varphi(S) = \{0\}\}$ be the vanishing ideal of S. Let $Z = \overline{S}^{\text{Zar}}$ be the Zariski closure, i.e., $Z = \{\vec{f} \in \mathbb{Q}^k \mid \forall \varphi \in I(S) : \varphi(\vec{f}) = 0\}$.

5.1.1 Remark. The Zariski closure over $\mathbb{Q}$ is the same as the Zariski closure over $\mathbb{C}$ intersected with $\mathbb{Q}^k$.[18] We will work entirely over $\mathbb{Q}$ and all polynomials have coefficients from $\mathbb{Q}$.

5.2 Functions that grow slowly

We will need to make an asymptotic growth behavior analysis, hence we define the vector space of functions that grow slower than φ on S:

$$\mathbb{Q}[Z]_{\leq(\varphi,S)} := \{\psi + I \in \mathbb{Q}[Z] \mid \exists \alpha \in \mathbb{Q}\ \forall \vec{g} \in S : |[\psi + I](\vec{g})| \leq \alpha |[\varphi + I](\vec{g})|\},$$

where $|a|$ denotes the absolute value of $a \in \mathbb{Q}$. Note that this is a linear subspace of $\mathbb{Q}[Z]$. For some applications we are particularly interested in the case where $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$, for a fixed S and for all φ, while for others we also have a fixed φ. Clearly, if S is finite, then $\mathbb{Q}[Z]$ itself is finite dimensional, and hence $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$.

5.2.1 Example. We will mostly study the cases where $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$, but the cases where $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} = \infty$ are also interesting, as the following examples show. Fix $k = 2$. Consider $S = \{(x, y) \in \mathbb{O} \mid y = 2^x\}$, $I = 0$ and $Z = \mathbb{Q}^2$. For $\varphi = y - x$, we have the infinite dimensional linear subspace $\langle 1, x, x^2, x^3, \ldots\rangle \subseteq \mathbb{Q}[Z]_{\leq(\varphi,S)}$. In fact, $y - x = 1 + \binom{x}{2} + \binom{x}{3} + \binom{x}{4} + \ldots$, so $y - x$ has only nonnegative integers in its binomial basis expansion. For $\varphi = y - 2x$, we have the same infinite dimensional linear subspace, but $y - 2x$ is not monotone. Indeed, for $x = 0$ we have $y - 2x = 1$, while for $x = 1$ we have $y - 2x = 0$. Thus if $y - 2x$ is a closure property of #P on S, then UP = coUP by Proposition 2.4.1.

[18] See, e.g., the proof in math.stackexchange.com/questions/279243.

We write suppb for the support in the binomial basis. The next claim shows that in many situations removing a finite amount of points from S does not change whether the dimension of $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)}$ is finite or not.

5.2.2 Claim. *If $S' \subseteq S$ with $S \smallsetminus S' < \infty$ and $\overline{S} = \overline{S'} = Z$, then for all $\varphi \in \mathbb{Q}[Z]$ we have $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$ if and only if $\dim \mathbb{Q}[Z]_{\leq(\varphi,S')} < \infty$.*

Proof. Clearly $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} \leq \dim \mathbb{Q}[Z]_{\leq(\varphi,S')}$, which proves one direction. It suffices to treat the case $|S \smallsetminus S'| = 1$. Let $S \smallsetminus S' = \{\vec{e}\}$. We make a case distinction based on whether or not $\varphi(\vec{e}) = 0$.

If $\varphi(\vec{e}) \neq 0$, then consider any arbitrary $\psi + I \in \mathbb{Q}[Z]_{\leq(\varphi,S')}$. We will see that $\psi + I \in \mathbb{Q}[Z]_{\leq(\varphi,S)}$. Let α be such that $\forall\, \vec{g} \in S' : \big|[\psi + I](\vec{g})\big| \leq \alpha \big|[\varphi + I](\vec{g})\big|$. Similarly, let $\alpha' := \max\big\{\alpha, |\psi(\vec{e})|/|\varphi(\vec{e})|\big\}$. It follows that $\forall\, \vec{g} \in S : \big|[\psi + I](\vec{g})\big| \leq \alpha' \big|[\varphi + I](\vec{g})\big|$, hence $\psi + I \in \mathbb{Q}[Z]_{\leq(\varphi,S)}$.

If $\varphi(\vec{e}) = 0$, then we have $\mathbb{Q}[Z]_{\leq(\varphi,S)} = \big\{\psi \in \mathbb{Q}[Z]_{\leq(\varphi,S')} \mid \psi(\vec{e}) = 0\big\}$. Since $\psi(\vec{e}) = 0$ is a homogeneous linear constraint, the dimension when imposing the constraint either stays the same or goes down by 1. □

Let $\xi_{\vec{e}} := \binom{\vec{f}}{\vec{e}} \in \mathbb{Q}\big[\vec{f}\,\big]$ denote the binomial basis function to the exponent vector $\vec{e} \in \mathbb{O}$. A function $\xi_{\vec{e}}$ is called *small* if

$$\forall\, \vec{f} \in S : \xi_{\vec{e}}(\vec{f}) \leqslant \varphi(\vec{f}).$$

Otherwise, we call $\xi_{\vec{e}}$ *large*. Since functions in I vanish on S, we see that $\xi_{\vec{e}}$ is small if and only if every element of $\xi_{\vec{e}} + I$ is small. In this case we say that $\xi_{\vec{e}} + I$ is a small element of $\mathbb{Q}[Z]$.

5.2.3 Claim. *Fix $\varphi \in \mathbb{Q}[Z]$. Suppose $\psi \in \mathbb{Q}[\vec{f}\,]$ is binomial-good and $\psi + I \notin \mathbb{Q}[Z]_{\leq(\varphi,S)}$. Then there exists $\vec{g} \in S$ and $\vec{e} \in \operatorname{suppb}(\psi)$ with $\xi_{\vec{e}}(\vec{g}) > \varphi(\vec{g})$.*

Proof. Decompose ψ over the binomial basis obtaining nonnegative integer coefficients. Set α to the sum of these coefficients. By definition, there exists $\vec{g} \in S$ with $\psi(\vec{g}) > \alpha\varphi(\vec{g})$. By the pigeonhole principle there exists $\vec{e}$ with $\xi_{\vec{e}}(\vec{g}) > \varphi(\vec{g})$. □

5.3 Graph varieties

In this section we study an important class of varieties for which $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$. We write $\vec{v} \in \mathbb{Q}^\ell$. For $1 \leq b \leq \ell$, we let $f_b := v_b$. Given functions $\zeta_b \in \mathbb{Q}[\vec{v}\,]$ and $\ell + 1 \leq b \leq k$, we define $\mathsf{eq}_b := \zeta_b(\vec{v}) - f_b \in \mathbb{Q}[\vec{f}\,]$. We call each f_b with $1 \leq b \leq \ell$, a *parameter* and each f_b with $\ell + 1 \leq b \leq k$ a *non-parameter*. Note that for parameters we have $f_b = v_b$ (and we use these symbols interchangeably), while for non-parameters there does not exist a v_b. Let

$$Z \;:=\; \big\{\, \vec{f} \in \mathbb{Q}^k \;\big|\; \mathsf{eq}_{\ell+1}(\vec{f}) = \ldots = \mathsf{eq}_k(\vec{f}) = 0 \,\big\}$$

Let $\tau : \mathbb{Q}^\ell \to Z$ be defined via $\tau(v_1, \ldots, v_\ell) := \big(v_1, \ldots, v_\ell, \zeta_{\ell+1}(\vec{v}), \ldots, \zeta_k(\vec{v})\big)$. Clearly, map τ is bijective, where the inverse function is the projection to the first ℓ coordinates. Let τ^* be the pullback algebra homomorphism $\tau^* : \mathbb{Q}[\vec{f}\,] \to \mathbb{Q}[\vec{v}\,]$ defined as $\tau^*(\varphi) := \varphi \circ \tau$, that replaces each occurrence of each non-parameter f_b by the polynomial $\zeta_b(\vec{v})$ in the parameters. Let $I \subseteq \mathbb{Q}[\vec{f}\,]$ be the ideal generated by $\mathsf{eq}_{\ell+1}, \ldots, \mathsf{eq}_k$. We start by establishing some well-known basic facts about graph varieties.

5.3.1 Claim. *We have: $\tau^*(\eta) \in \eta + I$ and $I = \ker \tau^*$.*

Proof. First, observe that $I \subseteq \ker \tau^*$. Indeed, the kernel of every ring homomorphism is always an ideal, so we only have to check this for the generators of I. But clearly $\mathsf{eq}_b \in \ker \tau^*$.

Now let $\eta \in \ker \tau^*$. Consider η as a polynomial in the non-parameters, i.e., as an element of the ring $\big(\mathbb{Q}[\vec{v}]\big)\big[f_{\ell+1}, \ldots, f_k\big]$. Since τ^* is replacing non-parameters by polynomials in parameters only, we can consider the application of τ^* as a process of iteratively substituting (in any order) a single variable in a monomial of η. This decreases this monomial's degree, so this process terminates.

Let τ_j^* be step j in this process of evaluating τ^*. If at step j the variable f_b, $b > \ell$, occurs in the monomial q with a strictly positive power, then we can write

$$\tau_j^*(q) \;=\; \tau_j^*\big(\tfrac{q}{f_b} f_b\big) \;=\; \tfrac{q}{f_b}\, \zeta_b(\vec{v})\,.$$

But we also see that

$$q + \underbrace{\tfrac{q}{f_b}\big(\zeta_b(\vec{v}) - f_b\big)}_{\in I} \;=\; \tfrac{q}{f_b}\, \zeta_b(\vec{v})\,.$$

This implies that $\tau_j^*(q) \in q + I$.

Iterating this process, in every step only elements of I are added. Hence $\tau^*(\eta) \in \eta + I$. Since we assumed that $\eta \in \ker \tau^*$ it follows that $0 \in \eta + I$ and hence $\eta \in I$. □

5.3.2 Claim. *$I = \mathcal{I}(Z)$.*

Proof. Let η be a multivariate polynomial in k variables. The direction $I \subseteq \mathcal{I}(Z)$ is clear: if $\eta \in I$, then η vanishes on Z, because all eq_b vanish on Z.

Now let η vanish on Z. We show that $\eta \in I$. Set $\eta' := \tau^*(\eta)$ and observe that $\eta' \in \eta + I$ by Claim 5.3.1. Since η vanishes on Z and since $I \subseteq \mathcal{I}(Z)$, we have that η' vanishes on Z. But since η' only uses parameter variables, η' vanishes identically on $\mathbb{Q}^\ell$. By multivariate interpolation on $\mathbb{Q}^\ell$ we have $\eta' = 0$, hence $0 = \eta' \in \eta + I$, and therefore $\eta \in I$. □

In other words, we get the short exact sequence

$$0 \;\to\; I \;\to\; \mathbb{Q}[\vec{f}\,] \;\xrightarrow{\tau^*}\; \mathbb{Q}[\vec{v}] \;\to\; 0.$$

Since $\mathbb{Q}[Z] = \mathbb{Q}[\vec{f}\,]/I$, it follows $\mathbb{Q}[Z] \simeq \mathbb{Q}[\vec{v}]$ with the isomorphism induced by τ^*, and its inverse given by $\psi \mapsto \psi + I$.

5.3.3 Claim. *Let S lie Zariski-dense in Z, and let $S' := \tau^{-1}(S)$. Then S' lies Zariski-dense in $\mathbb{Q}^\ell$.*

Proof. Assume that a function $\psi \in \mathbb{Q}[\vec{v}]$ vanishes on S'. Then ψ also vanishes on S. This means that $\psi \in I$. But ψ only uses parameter variables, hence $\psi = 0$. □

We now study asymptotic growth behavior in the parameter space. For $S' = \tau^{-1}(S)$ and $\varphi' = \tau^*(\varphi)$ we define

$$\mathbb{Q}[\vec{v}]_{\leq(\varphi',S')} \;:=\; \big\{\, \psi \in \mathbb{Q}[\vec{v}] \;\big|\; \exists \alpha \in \mathbb{Q}\; \forall\, \vec{g} \in S' \,:\, |\psi(\vec{g})| \,\leq\, \alpha |\varphi'(\vec{g})| \,\big\}.$$

5.3.4 Claim. *Let $\varphi' = \tau^*(\varphi)$ and let $S' = \tau^{-1}(S)$. The isomorphism $\mathbb{Q}[Z] \simeq \mathbb{Q}[\vec{v}]$ induces an isomorphism of linear subspaces $\mathbb{Q}[Z]_{\leq(\varphi,S)} \simeq \mathbb{Q}[\vec{v}]_{\leq(\varphi',S')}$.*

Proof. Let $\psi + I \in \mathbb{Q}[Z]_{\leq(\varphi,S)}$. Let $\alpha \in \mathbb{Q}$ be such that for all $\vec{g} \in S$ we have:

$$\big|[\psi + I](\vec{g})\big| \;\leq\; \alpha\big|[\varphi + I](\vec{g})\big|.$$

Note that

$$[\psi + I](\vec{g}) \;=\; \psi(\vec{g}) \;=\; \psi\big(\tau\big(\tau^{-1}(\vec{g})\big)\big) \;=\; \big[\tau^*(\psi)\big]\big(\tau^{-1}(\vec{g})\big).$$

Therefore, for all $\vec{g} \in S$, we have

$$\big|\big[\tau^*(\psi)\big]\big(\tau^{-1}(\vec{g})\big)\big| \;\leq\; \alpha\,\big|\big[\tau^*(\varphi)\big]\big(\tau^{-1}(\vec{g})\big)\big|.$$

In other words, for all $\vec{g} \in S$, we have

$$\big|[\psi']\big(\tau^{-1}(\vec{g})\big)\big| \;\leq\; \alpha\,\big|[\varphi']\big(\tau^{-1}(\vec{g})\big)\big|,$$

Since $S' = \tau^{-1}(S)$, it follows that $|\psi'(\vec{g})| \leq \alpha|\varphi'(\vec{g})|$ for all $\vec{g} \in S'$. The argument also works in the other direction, because $\tau(S') = S$. □

Let $\mathbb{O}_+ := \mathbb{Q}^\ell_+$. Note that we do not require integrality of the points in $\mathbb{O}_+$. We define the cone $C_{S'}$ of rays in $\mathbb{O}_+$ with infinitely many elements of S', i.e. $C_{S'} := \big\{\vec{v} \in \mathbb{O}_+ \,\big|\, |S' \cap \mathbb{Q}\vec{v}| = \infty\big\}$. Even though $S' \subseteq \mathbb{Q}^\ell$ is Zariski-dense, it can be that $C_{S'} = \varnothing$, for example if $\ell = 2$ and $S' = \{\vec{v} \in \mathbb{O}_+ \mid v_2 \geq v_1^2\}$. Nevertheless, in many important cases, we have $C_{S'}$ is Zariski-dense in $\mathbb{Q}^\ell$.

5.3.5 Claim. *If $C_{S'}$ is Zariski-dense in $\mathbb{Q}^\ell$, then $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$.*

Proof. Let $r_{\varphi'} := \max\big\{|\vec{e}| \,:\, \vec{e} \in \operatorname{suppb}(\varphi')\big\}$. Consider $\psi' \in \mathbb{Q}[\vec{v}]$ with coefficients $c_{\vec{e}}$. Define $E := \big\{\vec{f} \in \operatorname{suppb}(\psi') \,\big|\, r_{\varphi'} < |\vec{f}\,|\big\}$, and suppose $E \neq \varnothing$.

Let $r_E := \max\big\{|\vec{e}| \,:\, \vec{e} \in E\big\}$. By assumption, we have $r_E > r_{\varphi'}$. Denote $E_{\max} := \big\{\vec{e} \in E : r_E = |\vec{e}|\big\}$. Consider a polynomial

$$p(v_1, \ldots, v_\ell) \;:=\; \sum_{\vec{e} \in E_{\max}} c_{\vec{e}}\, v_1^{e_1} \cdots v_\ell^{e_\ell}\,,$$

and define the hypersurface $H = \mathcal{V}(p) \subseteq \mathbb{Q}^\ell$. By definition, all $c_{\vec{e}}$ are nonzero and there is at least one summand, so H is indeed a hypersurface. Since $C_{S'}$ is Zariski-dense in $\mathbb{Q}^\ell$, we conclude that $C_{S'}$ does not lie in any hypersurface. Hence, it follows that there exists $\vec{v}_H \in C_{S'} \smallsetminus H$.

Observe that

$$\begin{aligned} |\psi'(t\vec{v}_\varepsilon)| \;&=\; \left|\sum_{\vec{e} \in E_{\max}} c_{\vec{e}} v_{H,1}^{e_1} \cdots v_{H,\ell}^{e_\ell}\right| t^{|\vec{e}|} \;+\; \text{lower order terms in } t \\ &=\; \underbrace{\left|\sum_{\vec{e} \in E_{\max}} c_{\vec{e}} v_{H,1}^{e_1} \cdots v_{H,\ell}^{e_\ell}\right|}_{\neq 0,\ \text{because } \vec{v}_H \notin H} t^{r_E} \;+\; o(t^{r_E}), \end{aligned}$$

where the lower order terms are bounded by

$$\big|\varphi'(t\vec{v}_H)\big| \;=\; O\big(t^{|\vec{f}_{\varphi'}|}\big) \;=\; o(t^{r_E}).$$

Since there are infinitely many elements of S' along this ray, we have $\psi' \notin \mathbb{Q}[\vec{v}]_{\leq(\varphi',S')}$. Hence, a necessary criterion for $\psi' \in \mathbb{Q}[\vec{v}]_{\leq(\varphi',S')}$ is $\operatorname{suppb}(\psi') \subseteq \big\{\vec{e} \in \mathbb{O}_+ \,:\, r_{\varphi'} \geq |\vec{e}\,|\big\}$, which is a finite set. Hence $\dim \mathbb{Q}[\vec{v}]_{\leq(\varphi',S')} < \infty$. The proof is finished using Claim 5.3.4. □

Recall that $Z := \{\vec{f} \in \mathbb{Q}^k : \mathsf{eq}_{\ell+1}(\vec{f}) = \ldots = \mathsf{eq}_k(\vec{f}) = 0\}$, where $\mathsf{eq}_b := \zeta_b(\vec{v}) - f_b$.

5.3.6 Claim. *We assume that Z contains at least one integer point. Let $S := \mathbb{O} \cap Z$. Furthermore, we assume that for all $\ell+1 \le b \le k$, the polynomial ζ_b is not constant. Let $\zeta_b^{\hom} \in \mathbb{Q}[\vec{v}]$ be the top nonzero homogeneous part of ζ_b, for all $\ell+1 \le b \le k$. Suppose there exists a point $\vec{v} \in \mathbb{O}_+$ that satisfies the strict inequalities*

$$\zeta_b^{\hom}(\vec{v}) > 0 \quad \text{for all} \quad \ell+1 \le b \le k. \tag{$*$}$$

Then $C_{S'}$ lies Zariski-dense in $\mathbb{Q}^\ell$ (and hence $\dim \mathbb{Q}[Z]_{\le(\varphi,S)} < \infty$ by Claim 5.3.5).

We remark that if all ζ_b have an integer as their constant coefficient (for example, if all ζ_b are homogeneous degree ≥ 1 polynomials), then an integer point in Z is obtained by setting all parameters to zero.

Proof of Claim 5.3.6. Let η be the least common multiple of the denominators of the coefficients of all ζ_b. Consider the stretched polynomial $\eta\zeta_b$, which has integer coefficients. For all $\vec{v} \in \mathbb{Z}^\ell$, we have $\eta\zeta_b(\vec{v}) \equiv 0 \pmod{\eta}$ if and only if $\zeta_b(\vec{v}) \in \mathbb{Z}$. Hence, we can work in the ring $\mathbb{Z}/\eta\mathbb{Z}$. Then, for every integer direction vector $\vec{z} \in \mathbb{Z}^\ell$ we have $\zeta_b(\vec{v}) \in \mathbb{Z}$ if and only if $\zeta_b(\vec{v} + \eta\vec{z}) \in \mathbb{Z}$.

By assumption, Z contains an integer point $\vec{f}$ for all $\vec{y} = \tau^{-1}(\vec{f})$ and for all direction vectors $\vec{z} \in \mathbb{Z}^\ell$. Therefore, $\zeta_b(\vec{y} + \eta\vec{z}) \in \mathbb{Z}$ for all $\ell+1 \le b \le k$. Here we think of the set $\vec{y} + \eta\mathbb{Z}^\ell = \{\vec{y} + \eta\vec{z} : \vec{z} \in \mathbb{Z}^\ell\}$, where $\vec{y} \in \tau^{-1}(Z)$, as an affine shift of the axis parallel orthogonal grid with side length η and offset vector $\vec{y}$. Let us prove that

$$\begin{aligned} &\text{if } \zeta_b(\vec{v}) \in \mathbb{Z} \text{ for all } \ell+1 \le b \le k, \text{ then for infinitely many } s \in \mathbb{N}, \\ &\text{we have } \zeta_b(s\vec{v}) \in \mathbb{Z} \text{ for all } \ell+1 \le b \le k. \end{aligned} \tag{$\dagger$}$$

The proof of ($\dagger$) is similar to those when studying periodic orbits on a ℓ-dimensional rectangular block billiard table torus with side lengths η (see e.g. [Roz19, Tab05]), and goes as follows. We have $\zeta_b(\vec{v}) \in \mathbb{Z}$, if and only if $\eta\zeta_b(\vec{v}) \equiv 0 \pmod{\eta}$. We work over $\mathbb{Z}/\eta\mathbb{Z}$. Since the sequence $s\vec{v} \in (\mathbb{Z}/\eta\mathbb{Z})^\ell$ is periodic with period length at most η^ℓ, it follows that the sequence of vectors $\big(\eta\zeta_b(s\vec{v}) \mod \eta,\ \ell+1 \le b \le k\big) \in (\mathbb{Z}/\eta\mathbb{Z})^{k-\ell}$ is periodic of length at most η^ℓ. This proves ($\dagger$).

Let $\overline{\ast}^{\mathbb{Q}}$ denote the Euclidean closure, and $\overline{\ast}^{\mathrm{Zar}}$ the Zariski closure. Let $B_\varepsilon = \{\vec{u} : |\vec{u}| < \varepsilon\}$ be the open ball with radius $\varepsilon > 0$. Since the strict inequalities ($*$) are open conditions, there exists $\varepsilon \in \mathbb{Q}_{>0}$ such that each element in $\vec{v} + B_\varepsilon$ satisfies the strict inequalities ($*$). We now prove

$$(\vec{v} + B_\varepsilon) \cap \mathbb{O}_+ \subseteq \overline{C_{S'}}^{\mathbb{Q}} \tag{$\ddagger$}$$

This implies $(\vec{v}+B_\varepsilon)\cap\mathbb{O}_+ \subseteq \overline{C_{S'}}^{\mathrm{Zar}}$. Since B_ε is full-dimensional and since each of its affine shifts with center in $\mathbb{O}_+$ has a full-dimensional orthant in $\mathbb{O}_+$, it follows that $\overline{(\vec{v}+B_\varepsilon)\cap\mathbb{O}_+}^{\mathrm{Zar}} = \mathbb{Q}^\ell$. Hence $\mathbb{Q}^\ell = \overline{(\vec{v}+B_\varepsilon)\cap\mathbb{O}_+}^{\mathrm{Zar}} \subseteq \overline{C_{S'}}^{\mathrm{Zar}} \subseteq \mathbb{Q}^\ell$. This implies that the subset relationships are actually all equalities, so in particular $\overline{C_{S'}}^{\mathrm{Zar}} = \mathbb{Q}^\ell$.

It remains to show ($\ddagger$). Let $\vec{w} \in (\vec{v} + B_\varepsilon) \cap \mathbb{O}_+$ be arbitrary, in particular, suppose $\vec{w}$ satisfies ($*$). For all $e \in \mathbb{Q}_{>0}$ we now construct $\vec{w}'' \in C_{S'}$ with distance to $\vec{w}$ at most e. This proves ($\ddagger$).

Let $B_b := \{\vec{u} \in \mathbb{Q}^\ell : \zeta_b^{\hom}(\vec{u}) = 0\}$. Each B_b is a closed set, so if a point is not in B_b, then it has a strictly positive distance to B_b. Let $\Delta(\vec{w})$ be the minimum of the distances from $\vec{w}$ to B_b. By the intercept theorem we have $\Delta(t\vec{w}) = t\Delta(\vec{w})$.

Since $\vec{w}$ satisfies $(*)$, we can find t large enough such that $\Delta(t\vec{w}) > \eta^k$ and $\eta^k/t < e$. Let $\vec{w}'$ be a point in $\vec{y} + \eta \cdot \mathbb{Z}^\ell$ that is closest to $\vec{w}$. By $(\dagger)$, there are infinitely many nonnegative integer scalars $s \in \mathbb{N}$ such that

$$\zeta_b(s\vec{w}') \in \mathbb{Z} \quad \text{for all} \quad \ell + 1 \leq b \leq k.$$

Since $\Delta(\vec{w}') > 0$, we have $\vec{w}'$ also satisfies $(*)$. Moreover, the distance between $\vec{w}'/t$ and $\vec{w}$ is less than e (again, by the intercept theorem). Let $\vec{w}'' := \vec{w}'/t$. Then:

- $\vec{w}''$ satisfies $(*)$,
- the distance between $\vec{w}''$ and $\vec{w}$ is less than e, and
- there are infinitely many nonnegative integer scalars $s \in \mathbb{N}$ such that

$$\zeta_b(s\vec{w}'') \in \mathbb{Z} \quad \text{for all} \quad \ell + 1 \leq b \leq k.$$

It remains to prove that $\vec{w}'' \in C_{S'}$. For this, it suffices to show that for all large enough s we have

$$\zeta_b(s\vec{w}'') > 0 \quad \text{for all} \quad \ell + 1 \leq b \leq k. \tag{$\circledast$}$$

This can be seen as follows. Since

$$\zeta_b^{\hom}(s\vec{w}'') > 0 \quad \text{for all} \quad \ell + 1 \leq b \leq k,$$

it follows that

$$\zeta_b(s\vec{w}'') \;=\; s^{\deg \zeta_d} \underbrace{\zeta_b^{\hom}(\vec{w}'')}_{>0} + o\big(s^{\deg \zeta_d}\big).$$

Hence, for all large enough s we have $(\circledast)$, as desired. □

5.4 The witness theorem

Let $S_{\leqslant} := \{\vec{f} \in \mathbb{O} \mid \exists \vec{g} \leqslant \vec{f} : \vec{g} \in S\}$ denote the downwards closure of S. Recall from Section 4 that for $D \subseteq \mathbb{O}$, a function $\Psi : D \to \mathbb{N}$ is called *D-good* if all its coefficients in the binomial basis are nonnegative integers. Recall that a multivariate polynomial is called *binomial-good* if all coefficients in its binomial basis expansion are nonnegative integers, i.e., it is $\mathbb{O}$-good. Otherwise we call the polynomial *binomial-bad*. A coset of polynomials (for example $\varphi + I$) is called *binomial-good* if at least one of its representatives is binomial-good, otherwise it is called *binomial-bad*.

5.4.1 Theorem (The witness theorem)**.** *Let $S \subseteq \mathbb{O}$, let $I = I(S)$ be the vanishing ideal, and denote $\mathbb{Q}[Z] := \mathbb{Q}[\vec{f}]/I$. Fix $\varphi \in \mathbb{Q}[\vec{f}]$, and suppose that $\dim \mathbb{Q}[Z]_{\leq(\varphi,S)} < \infty$. For an integer $\Delta \in \mathbb{N}$, denote $D = D(\Delta, S) := [\![\Delta]\!]^k \cap S_{\leqslant}$. Then there exists $\Delta \in \mathbb{N}$, such that for every binomial-bad coset $(\varphi + I)$ and a D-good function $\Psi : D \to \mathbb{N}$, there exists $\vec{f} \in S \cap D$ with $\Psi(\vec{f}) \neq \varphi(\vec{f})$.*

Proof. A function in $\mathbb{Q}[Z]$ is completely specified by its evaluations at all points in S, which can be seen as follows. Let $\gamma, \gamma' \in \mathbb{Q}[\vec{f}]$, such that $\gamma|_S = \gamma'|_S$. Then $(\gamma - \gamma')|_S = 0$ which implies $(\gamma - \gamma')|_Z = 0$. Therefore, $\gamma - \gamma' \in I$, and $\gamma + I = \gamma' + I$, i.e., we have equality as functions in $\mathbb{Q}[Z]$. We conclude:

$$\text{the evaluation map } \mathsf{eval} : \mathbb{Q}[Z] \to \mathbb{Q}^S \text{ is injective.} \tag{5.4.2}$$

To simplify the notation, denote $V := \mathbb{Q}[Z]_{\leq(\varphi,S)}$. By (5.4.2), the restriction $\mathsf{eval}|_V : V \to \mathbb{Q}^S$ is injective. Let r denote the rank of the linear map $\mathsf{eval}|_V$, and note that the injectivity implies

that $|S| \geq r$. We now construct a set $S' \subseteq S$. We start with the empty set and iteratively enlarge it point by point from S in such a way that the rank of the evaluation map $\mathsf{eval}|_V : V \to \mathbb{Q}^{S'}$ increases at each step by 1, up to rank r. This gives a set $S' \subseteq S$ of r points such that the map $\mathsf{eval}_{V,S'} : V \to \mathbb{Q}^{S'}$ is injective. This means:

$$\text{an element in } V \text{ is completely specified by its evaluations at } S'. \tag{5.4.3}$$

Let d be the smallest $d \geq \delta$ such that $S' \subseteq [\![d]\!]^k$. For each ξ_e that is large (see above), let $\omega_e \in S$ be a point with

$$\xi_e(\omega_e) \;>\; \varphi(\omega_e). \tag{5.4.4}$$

Let Δ be such that for all large $e \in [\![d]\!]^k$ we have $\omega_e \in [\![\Delta]\!]^k$.

We are now ready to prove the result. Note that we have not considered Ψ so far. Since Ψ is D-good, we can enlarge the domain of definition of Ψ to the whole $[\![\Delta]\!]^k$ in a way such that Ψ is $[\![\Delta]\!]^k$-good. Let Ψ_Δ denote the unique polynomial with multidegree $\leq (\Delta, \dots, \Delta)$, and which is binomial-good.

Let $\widetilde{\Psi}$ be the polynomial that arises from Ψ_Δ by setting all coefficients in the binomial basis whose multidegree is not $\leqslant (d, \dots, d)$ to zero. Clearly $\widetilde{\Psi}$ is also binomial-good. Moreover, for the restrictions to $[\![d]\!]^k$, we have $\Psi|_{[\![d]\!]^k} = \widetilde{\Psi}|_{[\![d]\!]^k}$. On the other hand, on $[\![\Delta]\!]^k$ we have $\Psi|_{[\![\Delta]\!]^k} \geq \widetilde{\Psi}|_{[\![\Delta]\!]^k}$ because Ψ is binomial-good.

Suppose $\varphi|_{S'} \neq \widetilde{\Psi}|_{S'}$. This gives the desired $\vec{f}$, because $S' \subseteq [\![d]\!]^k \subseteq [\![\Delta]\!]^k$ and $\Psi|_{[\![d]\!]^k} = \widetilde{\Psi}|_{[\![d]\!]^k}$. Thus we can consider only the case when

$$\varphi|_{S'} \;=\; \widetilde{\Psi}|_{S'}\,.$$

Assume for the sake of contradiction that $\widetilde{\Psi} + I \in V$. Then, by (5.4.3), we have $\varphi + I = \widetilde{\Psi} + I$. This is a contradiction to $\varphi + I$ being binomial-bad and $\widetilde{\Psi} + I$ being binomial-good. Hence we are left to consider the case that

$$\widetilde{\Psi} + I \;\notin\; V.$$

Since $\widetilde{\Psi}$ is binomial-good, Claim 5.2.3 gives $\vec{e} \in \mathrm{suppb}(\widetilde{\Psi})$ with $\xi_{\vec{e}}$ large. Moreover, since $\widetilde{\Psi}$ is binomial-good, all coefficients in the binomial expansion are nonnegative. Hence $\widetilde{\Psi} \geq \xi_e$ on $\mathbb{O}$. In particular, we have:

$$\big[\widetilde{\Psi} + I\big](\vec{\omega}_e) \;\geq\; \xi_e(\vec{\omega}_e) \overset{(5.4.4)}{>} \varphi(\vec{\omega}_e).$$

Since $\Psi|_{[\![\Delta]\!]^k} \geq \widetilde{\Psi}|_{[\![\Delta]\!]^k}$, we found the desired $\vec{f} := \vec{\omega}_e$. This completes the proof. □

5.5 Integer points in the polyhedron $\mathcal{P}(\varphi, \zeta)$

In the Witness Theorem 5.4.1 we need to have a binomial-bad $\varphi + I$. In this section we discuss the important case of having only affine linear constraints. We use it for many of the cases in this paper.

Formally, given an ideal I generated by $\zeta_b - f_b$ with $b > \ell$, such that ζ_b are affine linear in $f_1, \dots, f_\ell$, and given a polynomial φ, the task in this section is to determine whether or not $\varphi + I$ is binomial-good. We will see that this is the case if and only if the polyhedron $\mathcal{P}(\varphi, \zeta)$ contains an integer point, see below.

Let Z be the vanishing set of I, and let $Z_{\mathbb{Z}}$ denote its integer points. We work under the assumption that $C'_{\tau^{-1}(Z_{\mathbb{Z}})}$ lies Zarsiki-dense in $\mathbb{Q}^\ell$, which is for example guaranteed if all ζ_b have only integer coefficients.

First, note that we can express φ in the parameters by plugging in ζ_b for each f_b, $b > \ell$. We call the resulting polynomial $\varphi' = \tau^*(\varphi)$. Let δ be the degree of φ'. Recall that $\mathbb{O} := \mathbb{N}^k$. Denote $\mathbb{O}_\delta := \{ \vec{e} \in \mathbb{O} \,:\, |\vec{e}| \leq \delta \}$. Take all k variables f_b and consider the binomial basis elements $\binom{\vec{f}}{\vec{e}}$ with $\vec{e} \in \mathbb{O}_\delta$. Each $\vec{e}$ gets a variable $x_{\vec{e}}$, and we define

$$q \;:=\; \sum_{\vec{e} \in \mathbb{O}_\delta} \tau^*\left(\binom{\vec{f}}{\vec{e}} \right) \cdot x_{\vec{e}}\,.$$

Note here that τ^* is only affine linear, so the degree does not increase.

By expressing $\tau^*\left(\binom{\vec{f}}{\vec{e}} \right)$ over the binomial basis in $\vec{v}$, we assign to each $\vec{e} \in \mathbb{O}_\delta$ a coefficient vector $w_{\vec{e}} \in \mathbb{Q}^{\mathbb{N}^\ell_\delta}$, where $\mathbb{N}^\ell_\delta := \{ \vec{v} \in \mathbb{N}^\ell \,:\, |\vec{v}| \leq \delta \}$. Note that if $\vec{e}$ is contained in $\mathbb{N}^\ell$, then $y_{\vec{e}}$ has a single 1 at position $\vec{e}$ and the rest are zeros. We also express φ' over the binomial basis and get coefficients $\varphi'_{\vec{v}}$, for all $\vec{v} \in \mathbb{N}^\ell_\delta$.

We search for a nonnegative integer assignment to the $x_{\vec{e}} \in \mathbb{N}$ such that

$$\varphi' \;=\; \sum_{\vec{e} \in \mathbb{O}_\delta} x_{\vec{e}}\, w_{\vec{e}}\,.$$

Note that the degree restriction $\vec{e} \in \mathbb{O}_\delta$ is not completely obvious and we will talk about it in Claim 5.5.3. We conclude:

$$\varphi'_{\vec{v}} \;=\; \sum_{\vec{e} \in \mathbb{O}_\delta} x_{\vec{e}}\, (w_{\vec{e}})_{\vec{v}} \quad \text{for all} \quad \vec{v} \in \mathbb{N}^\ell_\delta\,, \tag{5.5.1}$$

where the index $\vec{v}$ means taking the coefficient of $\vec{v}$ in the binomial basis.

Now, the question of the existence of such a vector x of nonnegative integers is the question of an integer point in the polyhedron $\mathcal{P}(\varphi, \zeta)$ defined by the equations (5.5.1) and the nonnegativity constraints on all variables. The following theorem states this in the simplified situation when ζ_b have only integer coefficients, while in general we only need that $\tau^{-1}(Z_\mathbb{Z})$ is Zarsiki-dense in $\mathbb{Q}^\ell$, where $Z_\mathbb{Z}$ are the integer points in Z.

5.5.2 Theorem (The polyhedron theorem)**.** *Let I be an ideal generated by the $\zeta_b - f_b$, $\ell < b \leq k$, such that all ζ_b are affine linear in $f_1, \ldots, f_\ell$, with integer coefficients. Fix a polynomial φ. Then $\varphi + I$ is binomial-good if and only if there exists an integer point in the polyhedron $\mathcal{P}(\varphi, \zeta)$.*

Proof. The existence of a nonnegative integer point $x_{\vec{e}}$ that satisfies (5.5.1) implies that $\varphi + I$ is binomial-good by definition. For the reverse, if $\varphi + I$ is binomial-good, then there exist finitely many nonnegative integers $x_{\vec{e}}$ with $\varphi' = \sum_{\vec{e}} x_{\vec{e}}\, w_{\vec{e}}$. The following Claim 5.5.3 implies that here we can indeed assume that $\vec{e} \in \mathbb{O}_\delta$ in this sum, which finishes the proof. □

5.5.3 Claim. *Let $x_{\vec{e}} > 0$ for some $\vec{e} \in \mathbb{N}^k$ with $|\vec{e}| > \delta$. Moreover, assume that all $x_{\vec{e}} \geq 0$ and that $x_{\vec{e}} > 0$ for only finitely many $\vec{e}$. Then $\varphi' \neq \sum_{\vec{e}} x_{\vec{e}}\, w_{\vec{e}}$.*

Proof. We define the vector space of functions that grow slowly:

$$\mathbb{Q}[\vec{v}]_{\leq \varphi'} \;:=\; \left\{ \psi \in \mathbb{Q}[\vec{v}] \;\middle|\; \exists \alpha \in \mathbb{Q}\ \forall \vec{g} \in \mathbb{O}_+ \,:\, |\psi(\vec{g})| \,\leq\, \alpha |\varphi'(\vec{g})| \right\}.$$

Clearly $\varphi' \in \mathbb{Q}[\vec{v}]_{\leq \varphi'}$. Since all functions $\tau^*(\xi_{\vec{f}})$ are eventually nonnegative forever on any ray $\mathbb{Q}_{\geq 0}\vec{v}$ with $\vec{v} \in C_{S'}$, it suffices to prove that $\tau^*(\xi_{\vec{e}}) \notin \mathbb{Q}[\vec{v}]_{\leq \varphi'}$.

Since $|\vec{e}| > \delta$, we have $\deg \xi_{\vec{e}} > \deg \varphi'$. Since τ^* replaces variables by affine linear non-constant polynomials, and since $\xi_{\vec{e}}$ has a single monomial of highest degree, we have $\deg\big(\tau^*(\xi_{\vec{e}})\big) = \deg(\xi_{\vec{e}})$.

Since $C_{S'}$ is dense in $\mathbb{Q}^\ell$, we can choose $\vec{v} \in \mathbb{O}_+$ that is strictly positive in each component. We have

$$\varphi'(t\vec{v}) \;=\; \Theta\big(t^{\deg(\varphi')}\big),$$

whereas

$$\xi'_{\vec{e}}(t\vec{v}) \;=\; \Theta\big(t^{\deg(\xi_{\vec{e}})}\big) \;=\; \omega\big(t^{\deg(\varphi')}\big).$$

This implies the result. □

6 The diagonalization theorem

In this section we use set-instantiators, the Witness Theorem 5.4.1 and a multivariate version of Ramsey's hypergraph theorem to prove the Diagonalization Theorem 6.2.1. The Diagonalization Theorem is our only method of proving oracle separations from #P when the polynomial φ is monotone, such as for example $f-1$, $f/2$ or $(f-1)/2$.

For a subset $B \subseteq \{0,1\}^j$, we write $\widetilde{B} \subseteq \{0,1\}^{j-1}$ do denote the set of suffixes of all strings that start with 1. For a subset $B \subseteq \{0,1\}^{j-1}$, we write $\{1\} \boxplus B \subseteq \{0,1\}^j$ to denote the union of $\{0^j\}$ with the set of strings that start with 1 and continue with a string from B.

6.1 Set-instantiators

We want to consider computation paths of nondeterministic Turing machines, but the actual computational device we are arguing about is a nondeterministic Turing machine with oracle access to an oracle that is defined up to strings of length $< j$, and where the oracle answers with 0 for all oracle queries of length $> j$. We capture this in the following definition.

6.1.1 Definition. *A computation path τ of a nondeterministic Turing machine on some input is defined as the sequence of its nondeterministic choice bits and the answers to its length j oracle queries (both types of bits appear in the same list, ordered chronologically). Formally, it is an element of $\{0,1\}^*$.*

The same Turing machine can yield the same computation path on different inputs (for example, when not the whole input is read) or when having access to different oracles, because the oracles can differ in positions that are not queried. We are especially interested in the case where the input is 0^j and the oracles differ in exactly the set $A_j \subseteq \{0,1\}^j$ of length j strings.

Given a nondeterministic Turing machine M and an oracle $A_{<j} := \bigcup_{j'<j} A_{j'}$ where $A_{j'} \subseteq \{0,1\}^{j'}$, and give a subset $B \subseteq \{0,1\}^{j-1}$, we are interested in the number of accepting paths of M when given oracle access to $A_{<j} \cup (\{1\} \boxplus B)$, where $A_{<j}$ is fixed. We define

$$h_M^B(w) \;:=\; \#\mathsf{acc}_{M^{A_{<j} \cup (\{1\} \boxplus B)}}(w). \tag{6.1.2}$$

It is instructive to think of $A_{<j}$ and M as together forming a computational device that has oracle access to some subset $B \subseteq \{0,1\}^{j-1}$.

For $\vec{b} \in \mathbb{N}^k$, we write $\mathscr{B}(\vec{b}) := \mathscr{B}([b_1]) \times \cdots \times \mathscr{B}([b_k])$, where $\mathscr{B}([a])$ is the set of all subsets of $[a] = \{1, \ldots, a\}$. For $\vec{s}, \vec{t} \in \mathscr{B}(\vec{b})$, we write $\vec{s} \subseteq \vec{t}$ if $s_a \subseteq t_a$ for all $1 \le a \le k$. For an element $\vec{s} \in \mathscr{B}(\vec{b})$, we write $|\vec{s}| := \big(|s_1|, \ldots, |s_k|\big)$.

6.1.3 Definition (Set-instantiator against $(M, j, A_{<j}, S, \vec{b})$)**.** *Let M be a nondeterministic Turing machine, let $j \in \mathbb{N}$, and let $A_{<j} \subseteq \{0,1\}^*$ be a language that contains only strings of length $< j$. Let $S \subseteq \mathbb{N}^k$ be a set and let $\vec{b} \in \mathbb{N}^k$. We set $\mathscr{B}(\vec{b})_S := \{\vec{s} \in \mathscr{B}(\vec{b}) : |\vec{s}| \in S\}$.*

Let $\top$ be a symbolic top element above $\mathscr{B}(\vec{b})$, i.e. $\vec{s} \subsetneq \top$ for all $\vec{s} \in \mathscr{B}(\vec{b})$. A set-instantiator $\mathcal{SI}$ is a pair of

- *an instantiation function $\mathsf{inst}_{\mathcal{SI}} : \mathscr{B}(\vec{b})_S \to \{0,1\}^{[2^j]}$, and*
- *a perception function $\mathsf{perc}_{\mathcal{SI}} : \{0,1\}^* \to \mathscr{B}(\vec{b}) \cup \{\top\}$,*

such that the following property holds for all $\vec{s} \in \mathscr{B}(\vec{b})_S$:

- *$\tau \in \{0,1\}^*$ is an accepting path for the computation $h_M^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ if and only if $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.*

The intuition is that a computation path queries the oracle and sees the existence of several objects (k different types of objects), and then decides to accept or not based solely on the set of objects perceived, independent of whether or not there are actually other unqueried objects in the oracle. The Turing machine might even *know* that there must be other objects for some syntactic reason and can take that information into account.

For example, in Sperner we have $k = 2$, and we consider rainbow triangles of positive/negative orientation. We know that $t_+ - t_- - 1 = 0$, so if we see $t_+ \geq 3$ and $t_- \geq 3$, then we know that there must be at least one rainbow triangle of positive orientation that we have not seen. Note that if an accepting path τ sees an object in the oracle and then we change the oracle, then running the same computation we will at some point get a different oracle answer, and hence τ will not be a computation path of this (input, oracle) combination.

Formally, in the above definition we think of accepting paths as having a perception from $\mathscr{B}(\vec{b})$, while computation paths that never accept on any of the instantiations are given perception $\top$. Note also that from the definition it is immediately clear that from a set-instantiator with a set S, we get a set-instantiator with the same parameters for every subset of S.

We usually do not mention $A_{<j}$ in the context of set-instantiators, as it has no effect on the construction of set-instantiators, and is also understood from the context. When discussing polynomial closure properties of #P, set-instantiators almost trivially exist (see §6.6), but for counting classes coming from TFNP this is not obvious. We create the necessary set-instantiators in §8.4, §8.5, §8.6, §8.8, and §8.10.

6.2 Diagonalization theorem statement

Our main tool for constructing oracles that separate from #P is the following Theorem 6.2.1, which depends on the parameters φ, ζ, S, Multiplicities, and $\vec{t}$. In most situations $\vec{t} \in S$ can be arbitrary, so $\vec{t}$ is often not specified. To use the theorem well, Multiplicities should map into S.

For $A_{<j} \in \{0,1\}^{<j}$, we say that a nondeterministic oracle Turing machine M *answers consistently* for $(j, A_{<j}, \text{Multiplicities})$, if for every $B \subseteq \{0,1\}^{j-1}$ we must have the number $\#\mathsf{acc}_M^{A_{<j} \cup (\{1\} \boxplus B)}(w)$ is the same for all B that have the same Multiplicities(B), for all $w \in \{0,1\}^*$. In other words, we must have:

$$\#\mathsf{acc}_M^{A_{<j} \cup (\{1\} \boxplus B)}(w) \;=\; \#\mathsf{acc}_M^{A_{<j} \cup (\{1\} \boxplus C)}(w)$$

for all $C \in \{0,1\}^{j-1}$ with Multiplicities$(B) =$ Multiplicities(C).

6.2.1 Theorem (Diagonalization Theorem)**.** *Fix $0 \leq \ell \leq k$ and let $\mathbb{O} := \mathbb{N}^k$. We write $\vec{f} = (f_1, \ldots, f_k)$ and $\vec{v} = (v_1, \ldots, v_\ell)$. Fix $\varphi \in \mathbb{Q}[\vec{f}]$. Let $\zeta_b \in \mathbb{Q}[\vec{v}]$ be non-constant functions and set I to be the ideal generated by the $\zeta_b(\vec{v}) - f_b$, where $\ell + 1 \leq b \leq k$.*

Define $Z := \{\vec{f} \in \mathbb{Q}^k \mid \mathsf{eq}_{\ell+1}(\vec{f}) = \ldots = \mathsf{eq}_k(\vec{f}) = 0\}$, *where* $\mathsf{eq}_b := \zeta_b(\vec{v}) - f_b$. *Denote* $T := \mathbb{O} \cap Z$ *and let* $S \subseteq T$. *Consider a map* $\tau : \mathbb{Q}^\ell \to Z$ *defined as*

$$\tau(v_1, \ldots, v_\ell) := \big(v_1, \ldots, v_\ell, \zeta_{\ell+1}(\vec{v}), \ldots, \zeta_k(\vec{v})\big).$$

Set $C'_S := \big\{\vec{v} \in \mathbb{Q}^\ell_{\geq 0} : |\tau^{-1}(S) \cap \mathbb{Q}\vec{v}| = \infty\big\}$. *Assume that*

(1) Z *contains at least one integer point,*
(2) $\overline{C'_S}^{\,\mathrm{Zar}} = \overline{C'_T}^{\,\mathrm{Zar}}$, *and*
(3) *there exists a point* $\vec{v} \in \mathbb{Q}^\ell_{\geq 0}$ *that satisfies strict inequalities*

$$\zeta_b^{\mathrm{hom}}(\vec{v}) > 0 \quad \textit{for all} \quad \ell + 1 \leq b \leq k,$$

where $\zeta_b^{\mathrm{hom}} \in \mathbb{Q}[\vec{v}]$ *is the top nonzero homogeneous part of* ζ_b.

Fix a set of multivariate functions $\textsc{Multiplicities} : \mathscr{B}\big(\{0,1\}^{j-1}\big) \to \mathbb{O}$. *Assume that for every nondeterministic polynomial-time Turing machine* M *and for every* $\vec{f}$, *there exist infinitely many* $j \in \mathbb{N}$, *such that for every* $A_{<j} \in \{0,1\}^{<j}$, *either* M *does not answer consistently for* $\big(j, A_j, \textsc{Multiplicities}\big)$ *or there is a set-instantiator* $\mathcal{SI}$ *against* $\big(M, j, A_{<j}, S, \vec{f}\big)$ *with* $\textsc{Multiplicities}\big(\mathsf{inst}_{\mathcal{SI}}(\vec{s})\big) = |\vec{s}|$ *for all* $\vec{s} \in \mathscr{B}(\vec{f})_S$.

Fix any $\vec{t} \in S$. *For* $A \subseteq \{0,1\}^*$, *we write:*

$$A = \bigcup_{j \geq 0} A_j\,, \quad \textit{where} \quad A_j \subseteq \{0,1\}^j.$$

Define

$$p_A(w) := \begin{cases} \varphi\big(\textsc{Multiplicities}\big(\widetilde{A}_{|w|}\big)\big) & \textit{if } A_{|w|}(0^{|w|}) = 1 \\ \varphi(\vec{t}) & \textit{otherwise,} \end{cases}$$

where $\widetilde{A}_j$ *is the set of length* $j-1$ *suffixes of the strings in* A_j *that start with a 1.*

Finally, suppose $\varphi + I$ *is binomial-bad. Then there exists an oracle* $A \subseteq \{0,1\}^*$ *such that for every nondeterministic polynomial-time Turing machine* M *there exists* j *such that* $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$ *and whenever* $A(0^j) = 1$, *then* $A_j = \{1\} \boxplus \mathsf{inst}_{\mathcal{SI}}(\vec{s})$ *for some* $\vec{s}$ *and one of the* $\mathcal{SI}$ *above.*

Note that the technical conditions (1), (2), and (3) are very easy to check in most situations. They exist to prevent degenerate cases. The rest of this section is devoted to the proof of Theorem 6.2.1.

6.3 Proof setup

In the beginning, we follow the diagonalization framework from Theorem 3.1.1(b) in [CGH+89], to construct an oracle $A \subseteq \{0,1\}^*$ such that the function p_A has the desired properties. The actual implementation of this strategy gets quite technical in some instances: we invoke the Witness Theorem 5.4.1 and use a generalization of Ramsey's theorem.

The computational complexity of the function Multiplicities will not play any role here, but may play a role when invoking the theorem, where it is usually solved by a k-tuple of #P machines. In fact, Multiplicities is usually evaluated on length 2^{j-1} instances. With the standard identification $\mathscr{B}(\{0,1\}^{j-1}) \simeq \{0,1\}^{2^{j-1}}$ a subset $B \subseteq \{0,1\}^{j-1}$ can be interpreted as a length 2^{j-1} instance of the Multiplicities problem.

We construct $A := \bigcup_{j\in\mathbb{N}} A_j$ as the union of sets $A_j \subseteq \{0,1\}^j$, where we will define A_j iteratively for larger and larger j. Once we define A_j, all $A_{j'}$ with $j' < j$ will not be changed again. All those $A_{j'}$ with $j' < j$ and that have not been defined up to that point, are set to $\varnothing$.

We enumerate the set of nondeterministic polynomial-time Turing machines that have oracle access. Note that this enumeration is independent of the specific oracle. Let M_i denote the i-th machine. For a specific oracle A, we denote the corresponding counting function by $H_i^A \in \#\mathsf{P}^A$. We proceed by diagonalizing over i, so assume that even though A is not fully specified, we have enough information to know that $H_{i'}^A \neq p_A$ for all $i' < i$. We show that $H_i^A \neq p_A$ by defining further details of A, i.e., defining A_j for larger j.

Let j be the smallest integer such that A_j has not been already defined, and such that j is large enough for all upcoming claims (the upcoming constructions do not depend on the exact value of j, but just require j to be large). We now define A_j in this case as follows.

Let $A_{<j} := \bigcup_{j'<j} A_{j'}$. If there exist $B \subseteq \{0,1\}^j$ with

$$H_i^{A_{<j}\cup B}(0^j) \;\neq\; p_{A_{<j}\cup B}(0^j),$$

then define $A_j := B$. Also define $A_{j'} := \varnothing$ for all $j < j' \leq j''$, where j'' is the longest length of an oracle query made by the computation $H_i^{A_{<j}\cup B}(0^j)$. In this case we have $H_{i'}^A \neq \varphi(f)$ for all $i' \leq i$, which was the goal.

For treating the other case, it is sufficient to analyze the case where for *all* $B \subseteq \{0,1\}^j$ such that $0^j \in B$. Here we have:

$$H_i^{A_{<j}\cup B}(0^j) \;=\; p_{A_{<j}\cup B}(0^j) \;=\; \varphi\big(\textsc{Multiplicities}(\widetilde{B})\big) \tag{6.3.1}$$

and find a contradiction as follows.

Equation (6.3.1) implies that the number of accepting paths in the computation $H_i^{A_{<j}\cup B}(0^j)$ only depends on the value $\textsc{Multiplicities}(\widetilde{B})$ and not on $\widetilde{B}$ itself. Hence, M_i answers consistently for $(j, A_{<j}, \textsc{Multiplicities})$. We will need this property mainly for the existence of a set-instantiator for $\#\mathsf{PLS}(\textsc{Iter})$, see §8.4.

The case (6.3.1) is then brought to a contradiction in §6.5 via instances generated by set-instantiators, see (6.5.6). We write

$$h_i^B(0^j) \;:=\; H_i^{A_{<j}\cup(\{1\}\boxplus B)}$$

to simplify the notation, which is also a simplification of the notation $h_{M_i}^{\{1\}\boxplus B}(0^j)$ from (6.1.2).

6.4 A Ramsey-type theorem

The following Ramsey-type result is used in the next section to construct set-instantiators. We need the following definitions, some of which we recall from the previous sections.

Fix the *dimension* k and the *number of colors* c. Let $\vec{m} = (m_1,\ldots,m_k) \in \mathbb{N}^k$ and $\vec{n} = (n_1,\ldots,n_k) \in \mathbb{N}^k$. We write $\vec{m} \leqslant \vec{n}$ if $m_i \leq n_i$ for all $1 \leq i \leq k$. Recall that $[n] = \{1,\ldots,n\}$.

Let $X_1,\ldots,X_k$ be an ordered list of finite or countably infinite sets, and write $\vec{X} = (X_1,\ldots,X_k)$. Denote $|\vec{X}| := \big(|X_1|,\ldots,|X_k|\big)$. For $\vec{a} \in \mathbb{N}^k$ we say that $\vec{X}$ is an *$\vec{a}$-list* if $|\vec{X}| = \vec{a}$.

Denote $\vec{Y} \subseteq \vec{X}$ when we have $Y_i \subseteq X_i$ for all $1 \leq i \leq k$. We say that $\vec{Y}$ is a *subset-list* of $\vec{X}$ in this case, and denote by $\mathscr{L}(\vec{X}) := \big\{\vec{Z} : \vec{Z} \subseteq \vec{X}\big\}$ the set of subset-lists of $\vec{X}$. A subset-list that is also an $\vec{a}$-list is called an *$\vec{a}$-subset-list*.

A *c-coloring* of $\vec{X}$ is a map $\mathscr{L}(\vec{X}) \to [c]$. We say that $\vec{X}$ is *$\vec{m}$-monochromatic* if for all $\vec{a} \leqslant \vec{m}$, all $\vec{a}$-subset-lists $\vec{Y} \subseteq \vec{X}$ have the same color.

6.4.1 Theorem (multipartite hypergraph Ramsey theorem). *Fix $k \geq 1$. For every $\vec{n}$, $\vec{m} \in \mathbb{N}^k$ and every integer $c \geq 1$, there exists a natural number $\mathrm{R} = \mathrm{R}(\vec{n}, \vec{m}, c)$, such that for every c-coloring of a list $\vec{X} = (X_1, \ldots, X_k)$ of finite sets of size $|X_i| \geq \mathrm{R}$, there is a $\vec{m}$–monochromatic subset-list $\vec{Y} \subseteq \vec{X}$ with $|\vec{Y}| \geqslant \vec{n}$.*

Since we do not need the explicit quantitative bounds, the proof below is based on the approach in [GPS12], rather than a more standard argument in Ramsey theory, see e.g. [GRS90, Nes95]. We give it here for the sake of completeness and because it is not covered by the (usual) *hypergraph Ramsey theorem*.

Proof. Given k countably infinite sets $X_1, \ldots, X_k$, and let all subset-lists of cardinality $\leqslant \vec{m}$ be colored with $[c]$. We first prove:

($\Diamond$) there is an infinite $\vec{m}$–monochromatic subset list $\vec{Y} \subseteq \vec{X}$. Here $\vec{Y} = (Y_1, \ldots, Y_k)$, such that $Y_i \subseteq X_i$ are infinite, for all $1 \leq i \leq k$.

Fix $\vec{a} \leqslant \vec{m}$. If we can prove that there are infinite subsets $Y_1 \subseteq X_1, \ldots, Y_k \subseteq X_k$ for which all $\vec{a}$-subset-lists are monochromatic, then we can iterate this argument on $\vec{Y}$ for a another vector $\vec{a}' \leqslant \vec{m}$, and so on, until we treated all $\vec{a} \leqslant \vec{m}$.

We proceed by induction. We assume by induction that for each $\vec{b} \leqslant \vec{a}$, $\vec{b} \neq \vec{a}$, we have a subset-list of infinite subsets such that each $\vec{b}$-subset is monochromatic. We choose $q \in [k]$ with $a_q \neq 0$. Let $\vec{b} := (a_1, \ldots, a_{q-1}, 0, a_{q+1}, \ldots, a_k)$. Take any way of enumerating the set of all $\vec{b}$-subset-lists. Note that every $\vec{b}$-subset-list can be extended to an $\vec{a}$-subset-list in infinitely many ways. Each of these extensions has a color.

For the first $\vec{b}$-subset-list we choose an infinite monochromatic subset in the set of all extensions (which exists by the usual hypergraph Ramsey theorem). We attach the color of this subset to the $\vec{b}$-subset-list and call it its *infcolor*. Now delete all vertices in coordinate q that are not in this monochromatic set and proceed with the next $\vec{b}$-subset-list. We iterate this, and in this way attach an infcolor to each $\vec{b}$-subset-list. By induction we can choose a subset-list of infinite subsets in which all $\vec{b}$-subset-lists are monochromatic w.r.t. infcolor. In the resulting subset-list of infinite subsets, all $\vec{a}$-subset-lists have the same color. This completes the proof of ($\Diamond$).

We now prove the existence of the lower bound $\mathrm{R}(\vec{n}, \vec{m}, c)$ as in the theorem. Assume that there exists no such bound. Then for each $z \in \mathbb{N}$ we find a coloring col_z of the subset-lists of $[z]^k$ such that $[z]^k$ contains no cardinality $\vec{m}$ subset-list in which for all $\vec{a} \leqslant \vec{m}$ all $\vec{a}$-subset-lists are monochromatic.

We enumerate the set of all subset-lists of $\mathbb{N}^k$ that are an $\vec{a}$-subset-list for some $\vec{a} \leqslant \vec{m}$. We assign a color to the first of this list so that the color agrees with the color choice of infinitely many col_z (note that not all col_z might assign a color to this element, but infinitely many do). Now assign a color to the second of the list so that it agrees with the color choice of infinitely many col_z that already agree with the color of the first element. We proceed in this manner and obtain a coloring of all $\vec{a}$-subset-lists, $\vec{a} \leqslant \vec{m}$, of $\mathbb{N}^k$. By construction, no cardinality $\vec{m}$ subset-list has for all $\vec{a} \leqslant \vec{m}$ that all $\vec{a}$-subset-lists are monochromatic. This is a contradiction to ($\Diamond$), as desired. □

6.5 Set-instantiators via Ramsey's theorem

We can now finish the proof of the Diagonalization Theorem 6.2.1 that we started in §6.3. Recall that we invoke the Witness Theorem 5.4.1 with our S and φ to obtain a $\Delta \in \mathbb{N}$ as in the theorem.

Denote $\Delta^{\times k} := (\Delta, \ldots, \Delta) \in \mathbb{N}^k$. Let $\lambda \geq \Delta$ be large enough for Theorem 6.4.1 to hold (note that this also implies that j is large). Let $\mathcal{SI} := (\mathcal{SI})_{\lambda^{\times k}}$ be the set-instantiator with $\vec{b} = \lambda^{\times k}$,

with $M = M_i$, j and $A_{<j}$ being as before, and with S is intersected with $[\![\lambda]\!]^{\times k}$. From now on we have $S \leftarrow S \cap [\![\lambda]\!]^k$. Recall that $\mathscr{B}\big(\lambda^{\times k}\big) = \mathscr{B}([\lambda]) \times \ldots \times \mathscr{B}([\lambda])$, k times.

6.5.1 Definition (Filling the definitional holes with "perception or below"). *For* $\vec{t} \in \mathscr{B}(\lambda^{\times k})$ *we define* $\Phi(\vec{t})$ *as the number of* $\tau \in \{0,1\}^*$ *that satisfy* $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{t}$.

We observe that for $\vec{t} \in \mathscr{B}(\lambda^{\times k})_S$ we have $\Phi(\vec{t}) = h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{t})}(0^j)$. So this definition fills the holes in the domain of definition of the function $\vec{t} \mapsto h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{t})}(0^j)$ that is only defined on $\mathscr{B}(\lambda^{\times k})_S \subseteq \mathscr{B}(\lambda^{\times k})$. Note that

$$\Phi(\vec{t}) \;=\; \big|\big\{\,\tau \in \{0,1\}^* \;:\; \mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{t}\,\big\}\big|, \tag{6.5.2}$$

so in particular the function Φ is monotonic.

By Definition 6.1.3, for every $\vec{t} \in \mathscr{B}(\lambda^{\times k})$ with $|\vec{t}| \in S_{\leqslant}$, we find $\vec{s} \in \mathscr{B}(\lambda^{\times k})_S$ such that $\vec{t} \subseteq \vec{s}$ Take an $\vec{s}$ for which $|\vec{s}|$ is lexicographically smallest. Define $\Omega(\vec{t}) := \varphi(|\vec{s}|)$. Note that by the definition of a set-instantiator, $\Omega(\vec{t})$ is the same number for all $\vec{t}$ that have the same $|\vec{t}|$, so $\overline{\Omega}\big(|\vec{t}|\big) := \Omega(\vec{t})$ is well-defined.

If $\Phi(\vec{t}) > \Omega(\vec{t})$, then also $\Phi(\vec{s}) > \Omega(\vec{t})$ by monotonicity. Hence $h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j) > \varphi(|\vec{s}|)$. By Definition 6.1.3, we have $\textsc{Multiplicities}(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$ It now follows that $h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j) > \varphi(\textsc{Multiplicities}(\mathsf{inst}_{\mathcal{SI}}(\vec{s})))$. This allows us to define $A_j := (\mathsf{inst}_{\mathcal{SI}}(\vec{s}))'$ and get a contradiction to (6.3.1).

It remains to treat the case when this is impossible. Therefore, from now on we assume that for all $\vec{t}$ with $|\vec{t}| \in S_{\leqslant}$ we have $\Phi(\vec{t}) \leq \Omega(\vec{t}) = \varphi\big(|\vec{s}|\big)$. We are not interested in $\vec{t} \notin S_{\leqslant}$. Define $c := 1 + \max\big\{\,\overline{\Omega}(\vec{f}) \;:\; \vec{f} \in [\![\Delta]\!]^k \cap S_{\leqslant}\,\big\}$.

6.5.3 Proposition. *Let* Φ *be a function defined on the set* $\mathscr{B}([\lambda])^k$ *such that* $\Phi(\vec{t}) \in \{0,\ldots,c-1\}$ *for all* $\vec{t}$ *with* $|\vec{t}| \in [\![\Delta]\!]^k$. *Then, there exists* $\Lambda = \Lambda(\Delta, c) \in \mathbb{N}$, *such that for all* $\lambda \geq \Lambda$ *there exist subsets* $Q_1, \ldots, Q_k \subseteq [\lambda]$ *with* $|Q_1| = \ldots = |Q_k| = \Delta$, *with the property:*

$$|\vec{s}| = |\vec{t}| \in S_{\leqslant} \;\; \textit{for some} \;\; \vec{s}, \vec{t} \in \vec{Q} \quad\Longrightarrow\quad \Phi(\vec{s}) \;=\; \Phi(\vec{t}).$$

Proof. We assign to each element $\vec{s} \in \mathscr{B}([\lambda])^k$ with $|\vec{s}| \in [\![\Delta]\!]^k$ the color $\Phi(\vec{s})$ for all $|\vec{s}| \in S_{\leqslant}$. Otherwise, assign the color 0. Let $\Lambda := \mathrm{R}\big(\Delta^{\times k}, \Delta^{\times k}, c\big)$ and use Theorem 6.4.1 to find the desired subsets $Q_1, \ldots, Q_k$. □

From Proposition 6.5.3, we can readily construct a new set-instantiator $\overline{\mathcal{SI}}$ with $\vec{b} = \Delta^{\times k}$ as follows (this is similar to restricting to a subset of S and renaming the elements).

Take a bijection $\beta_a : Q_a \to [\Delta]$ for all $a \in [k]$. This induces a bijection

$$\vec{\beta} \;:\; \mathscr{B}(Q_1) \times \ldots \times \mathscr{B}(Q_k) \;\longrightarrow\; \mathscr{B}\big([\Delta]\big) \times \ldots \times \mathscr{B}\big([\Delta]\big),$$

which we use for identification of the sets. We set $\mathsf{inst}_{\overline{\mathcal{SI}}}(\vec{s}) := \mathsf{inst}_{\mathcal{SI}}(\vec{\beta}^{-1}(\vec{s}))$. Define

$$\mathsf{perc}_{\overline{\mathcal{SI}}}(\tau) := \vec{\beta}\big(\mathsf{perc}_{\mathcal{SI}}(\tau)\big) \quad \text{if} \quad \vec{\beta}\big(\mathsf{perc}_{\mathcal{SI}}(\tau)\big) \in \mathscr{B}\big([\Delta]\big)^k \cup \{\top\},$$

and let $\mathsf{perc}_{\overline{\mathcal{SI}}}(\tau) := \top$ otherwise. With $S \leftarrow S \cap [\![\Delta]\!]^k$, it is easy to check that $\overline{\mathcal{SI}}$ is a set-instantiator.

By analogy to (6.5.2), define

$$\overline{\Phi}(\vec{t}) := \big|\big\{\,\tau \in \{0,1\}^* \;:\; \mathsf{perc}_{\overline{\mathcal{SI}}}(\tau) \subseteq \vec{t}\,\big\}\big| \;=\; \sum_{\vec{s} \subseteq \vec{t}} \big|\big\{\,\tau \in \{0,1\}^* \;:\; \mathsf{perc}_{\overline{\mathcal{SI}}}(\tau) = \vec{s}\,\big\}\big|. \tag{6.5.4}$$

Since $\overline{\Phi}(\vec{t})$ depends only on $|\vec{t}|$ for all $|\vec{t}| \in S_{\leqslant}$, by induction we see that the number

$$\left|\left\{\tau \in \{0,1\}^* \,:\, \mathsf{perc}_{\overline{SI}}(\tau) = \vec{s}\right\}\right|$$

also depends only on $|\vec{s}|$. Denote this number by $c_{\vec{g}}$, where $\vec{g} := |\vec{s}| \in S_{\leqslant}$, and set $\Psi(\vec{f}) := \overline{\Phi}(\vec{t})$ for $\vec{f} = |\vec{t}|$. This means that (6.5.4) simplifies:

$$\Psi(\vec{f}) \;=\; \sum_{\vec{g} \leqslant \vec{f}} \binom{\vec{f}}{\vec{g}} c_{\vec{g}}\,. \tag{6.5.5}$$

The function $\Psi(\vec{f})$ is defined on $S_{\leqslant} \cap [\![\Delta]\!]^k$ and is $S_{\leqslant} \cap [\![\Delta]\!]^k$-good, which is a requirement for the second part of the Witness Theorem 5.4.1 (we already used it to obtain Δ). We invoke the second part of the Witness Theorem 5.4.1 to find a point $\vec{f} \in S \cap [\![\Delta]\!]^k$ with $\Psi(\vec{f}) \neq \varphi(\vec{f})$. We take $\vec{s} \in (\mathscr{B}([\Delta])^k)_S$ with $|\vec{s}| = \vec{f}$, and observe

$$h_i^{\mathsf{inst}_{SI(\beta^{-1}(\vec{s}))}}(0^j) \;=\; h_i^{\mathsf{inst}_{\overline{SI}(\vec{s})}}(0^j) \;=\; \Psi(\vec{f}) \;\neq\; \varphi(\vec{f}). \tag{6.5.6}$$

On the other hand, recall that $\textsc{Multiplicities}(\mathsf{inst}_{\mathcal{SI}}(\vec{t})) = |\vec{t}|$, and that by Definition 6.1.3, we have $|\vec{t}| = |\beta(\vec{t})|$. Hence, for $\vec{s} = \beta(\vec{t})$ we have:

$$h_i^{\mathsf{inst}_{SI(\beta^{-1}(\vec{s}))}}(0^j) \;\neq\; \varphi\big(\textsc{Multiplicities}\big(\mathsf{inst}_{\mathcal{SI}}\big(\beta^{-1}(\vec{s})\big)\big)\big).$$

Finally, define $A_j = \big(\mathsf{inst}_{\mathcal{SI}}\big(\beta^{-1}(\vec{s})\big)\big)'$ to obtain the desired contradiction to (6.3.1). This completes the proof of Theorem 6.2.1. □

6.6 A set-instantiator for OccurrenceMulti

Let $\textsc{OccurrenceMulti}_k : \{0,1\}^* \to \mathbb{N}^k$ be the function defined as follows. On input $w \in \{0,1\}^*$ we split w into k parts of roughly the same size, and the output is the vector that specifies how many 1's are in the first part, how many 1's are in the second part, and so on. We omit the index k when it is clear from the context.

6.6.1 Theorem. *Let M be a given a polynomial time nondeterministic Turing machine. Fix k. For every $\vec{b} \in \mathbb{N}^k$ there exists a threshold $j_0 \in \mathbb{N}$, such that for every $j \geq j_0$ and every $A_{<j} \in \{0,1\}^{<j}$ there exists a set-instantiator $\mathcal{SI}$ against $(M, j, A_{<j}, S \subseteq \mathbb{N}^k, \vec{b})$, such that* $\textsc{OccurrenceMulti}_k(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$.

The rest of this section is devoted to proving this theorem for $S = \mathbb{N}^k$, which immediately proves it for all subsets. We will first define a *creator* whose *creations* will be the instantiations in the end, but the creator is not limited to a single creation for each set. We will then define the set-instantiator from the creator by picking for every subset $\vec{s}$ with $|\vec{s}| \in \mathbb{N}^k$ just any one of its creations for $\vec{s}$.

6.6 (a) Creations

Having 2^{j-1} many bits available, we can encode in a standard way a list of k subsets $\psi = (\psi_1, \ldots, \psi_k)$, where $\psi_a \subseteq [2^{j-2}]$. We ignore any extra bits that are not needed for this encoding. Let $n := j - 2$. For a $(c, x) \in [k] \times [2^n]$ we write $\psi(c, x) := 1$ if $x \in \psi_c$, and $\psi(c, x) := 0$ otherwise. We say that (c, x) is a *hit* if $\psi(c, x) = 1$, otherwise it is a *miss*.

Let $\mathsf{hitvec}(\psi) := \{(c,x) : \psi(c,x) = 1\}$. Let $\mathsf{hit}(\psi_c) \subseteq [2^n]$ denote the set of $x \in [2^n]$ with $\psi(c,x) = 1$. For a vector $\vec{a} \in \mathbb{N}^k$, a list ψ is called an *$\vec{a}$-creation* if $|\mathsf{hit}(\psi_c)| = a_c$ for all $1 \leq c \leq k$. Let $\vec{a}$-$\mathsf{creations}$ denote the set of all $\vec{a}$-creations. Finally, let

$$a^-\text{-}\mathsf{creations} := \bigcup_{1 \leqslant \vec{v} \leqslant \vec{a}} \vec{v}\text{-}\mathsf{creations}$$

These will be used extensively in Section 8.

6.6 (b) Lucky creators

In this section we introduce the concept of a *creator*. A creator is a slightly less restrictive version of a set-instantiator. For each blueprint a creator outputs a creation, where a blueprint is slightly more general than the set $\vec{s}$ for a set-instantiator, but serves the same purpose.

For a set X, let $\binom{X}{c}_{\text{ordered}}$ denote the set of cardinality c ordered subsets of X (i.e., length c lists of pairwise distinct objects of X). For $\vec{b} \in \mathbb{N}^k$, a *$\vec{b}$-creator* ξ is a length k list $\mathsf{hitmatrix}(\xi)$. The c-th entry of this list is a length b_c list of distinct entries from $[2^n]$, i.e., an element from $\binom{[2^n]}{b_c}_{\text{ordered}}$. We can (and will) think of $\mathsf{hitmatrix}(\xi)$ as a set of points in $[k] \times [2^n]$.

Given a $\vec{b}$-creator ξ and an $\vec{L} \in \mathscr{B}(\vec{b})$, we obtain an $|\vec{L}|$-creation $\xi_{\vec{L}}$ by setting $\xi_{\vec{L}}(c, \mathsf{hitmatrix}(\xi)_{c,d}) = 1$ for all $d \in L_c$, and $\xi_{\vec{L}}(c,x) = 0$ otherwise.

6.6.2 Definition. *We call $\xi \in b$-$\mathsf{creators}$ lucky if for all $\vec{L} \in \mathscr{B}(\vec{b})$, all accepting paths of the computation $h_i^{\xi_{\vec{L}}}(0^j)$ do not access the oracle at any point in $\mathsf{hitmatrix}(\xi) \setminus \mathsf{hitvec}(\xi_{\vec{L}})$.*

In a sense, this says that the accepting paths do not access the oracle at positions where lonely nodes are not, but could potentially be.

Suppose for all $j \geq 1$, we have a probability distribution D_j and an event e_j, which satisfy $\lim_{j\to\infty} \Pr_{D_j}[e_j] = 1$. Then we say that e_j *happens with high probability* (w.h.p.) *in* D_j. Given a finite number of events e_j that each happen with high probability in D_j, by the union bound all these events happen simultaneously with high probability in D_j. Let U_X denote the uniform distribution on a finite set X.

6.6.3 Claim. *If ξ is sampled from $U_{b\text{-}\mathsf{creators}}$, then ξ is lucky w.h.p.*

Proof. We observe that for a fixed $L \in \mathscr{B}(\vec{b})$, we have that the $\vec{L}$-creation ξ_L is uniformly distributed from $\vec{L}$-$\mathsf{creations}$. We show that for a fixed $\vec{L} \in \mathscr{B}(\vec{b})$ we have that ξ is lucky w.h.p.

Since there are only constantly many $\vec{L}$, the claim immediately follows from the union bound. Hence, for the rest of the proof fix $\vec{L}$. We have that $\xi_{\vec{L}}$ is uniformly distributed. For $l := \sum_{a=1}^{k} |L_a|$, the probability of an oracle access picking one of these positions is $\leq \frac{kl}{2^n - l}$. By Bernoulli's inequality, we have

$$\left(1 - \frac{kl}{2^n - l}\right)^{t_i(j)} \geq 1 - \frac{klt_i(j)}{2^n - l}.$$

Since k and l are fixed and since n and j are polynomially related, this proves the claim. □

6.6 (c) Defining the set-instantiator

For a $\vec{a}$-creation ψ and a computation path $\tau \in \{0,1\}^*$ of a computation $h_i^{\psi}(0^j)$, let $\mathsf{perception}(\tau) \subseteq \mathsf{hit}(\psi)$ denote the set of accessed oracle positions that are hits.

Let ξ be a lucky $\vec{b}$-creator. We interpret the set $\mathsf{hitmatrix}(\xi)$ as a bijection $\mathsf{bij} : \mathscr{B}(\vec{b}) \to \mathsf{hitmatrix}(\xi)$. Let $S = \mathbb{N}^k$. We set

$$\mathsf{inst}_{\mathcal{SI}}(\vec{s}) := \xi_{\vec{s}},$$

and for $\tau \in \{0,1\}^*$ we set

$$\mathsf{perc}_{\mathcal{SI}}(\tau) = \begin{cases} \mathsf{bij}^{-1}(\mathsf{perception}(\tau)) & \text{if } \tau \text{ is an accepting path of the computation } h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j), \\ \top & \text{otherwise.} \end{cases}$$

The rest of this section is devoted to proving that $\mathcal{SI}$ satisfies the requirements of Definition 6.1.3, which then proves Theorem 6.6.1, because clearly $\textsc{OccurrenceMulti}(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$. Formally:

6.6.4 Proposition. *For all $\vec{s} \in \mathscr{B}(\vec{b})_S$: $\tau \in \{0,1\}^*$ is an accepting path for the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ if and only if $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.*

Proof. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. This implies that $\mathsf{perc}_{\mathcal{SI}}(\tau) = \mathsf{bij}^{-1}(\mathsf{perception}(\tau))$. Clearly $\mathsf{perception}(\tau) \subseteq \mathsf{bij}(\vec{s})$, since otherwise τ would not even be a computational path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ because of its oracle answers when querying lonely nodes in $\mathsf{perception}(\tau) \setminus \mathsf{bij}(\vec{s})$. We conclude: $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.

The argument above is reversible. Indeed, let $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$, so in particular $\mathsf{perc}_{\mathcal{SI}}(\tau) \neq \top$. Then τ is an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$. □

6.7 Binomial-good polynomials and relativizing closure properties of #P

We now draw an important corollary from the Diagonalization Theorem 6.2.1 in a simple subcase that allows us to completely characterize the relativizing multivariate polynomial closure properties of #P.

6.7.1 Theorem. *The relativizing multivariate polynomial closure properties of #P are exactly the binomial-good polynomials.*

Proof. Let φ be binomial-bad. We prove that there exists $A \subseteq \{0,1\}^*$ such that $\varphi(\overrightarrow{\#\mathsf{P}^A}) \not\subseteq \#\mathsf{P}^A$. We use a very simple instantiation of the Diagonalization Theorem 6.2.1 as follows.

Let k be the arity of φ, and let $\ell = k$. We have no functions ζ_b in this case. The ideal $I = \langle 0 \rangle$. Let $S = T = \mathbb{O}$. Then $C'_S = \mathbb{Q}^k_{\geq 0}$. The assumptions of the Diagonalization Theorem are readily verified: (1) $\vec{0} \in Z$, (2) $S = T$, (3) there are no inequalities. We set $\vec{t} = \vec{0}$ and set $\textsc{Multiplicities} = \textsc{OccurrenceMulti}_k$ (for which we have set-instantiators from Theorem 6.6.1).

Since $\varphi + I = \varphi + \langle 0 \rangle$ is binomial-bad, there exists A such that for every M, there is j with $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$. The proof is finished by observing that $p_A \in \varphi(\overrightarrow{\#\mathsf{P}^A})$, which can be seen as follows. There exists a list of oracle Turing machines that on input w first query A at $0^{|w|}$. Then, if $0^{|w|} \notin A$, accept with multiplicities $\vec{t}$; if $0^{|w|} \in A$, then the machines run $\textsc{OccurrenceMulti}$ on $\widetilde{A}_j$. □

7 Applications to classical problems

In this section we first continue the approach in §2.3 to apply Proposition 2.3.1 to complete squares. We then apply the Diagonalization Theorem 6.2.1 and its corollary, Theorem 6.7.1, in several settings.

7.1 Complete squares

The *Cauchy inequality* is the basic inequality which goes into the definition of the *scalar product* in $\mathbb{R}^n$. It is the starting point in [HLP52, Eq. (1.1.1)]:

$$\big(x_1y_1 + \ldots + x_ny_n\big)^2 \le \big(x_1^2 + \ldots + x_n^2\big)\big(y_1^2 + \ldots + y_n^2\big) \tag{7.1.1}$$

Geometrically, the inequality says that cosine of every angle in $\mathbb{R}^n$ is at most 1. It is a special case of the *Hölder inequality* for the L^p-norm when $p = 2$, see e.g. [BB61, §17].

7.1.2 Proposition (Cauchy inequality)**.** *Denote by $C_n = C_n(x_1, \ldots, x_n, y_1, \ldots, y_n)$ the counting function given by (7.1.1), i.e., the difference of the right-hand side and the left-hand side of the inequality. For $n \ge 2$, we have $C_n \not\gtrdot_{\#} 0$ unless $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.*

Proof. Take $n = 2$. We have

$$C_2 \;=\; \big(x_1^2 + x_2^2\big)\big(y_1^2 + y_2^2\big) \;-\; \big(x_1y_1 + x_2y_2\big)^2 \;=\; x_1^2y_2^2 + x_2^2y_1^2 \;-\; 2x_1y_1x_2y_2 \;\ge\; 0.$$

When $y_1 = y_2 = 1$, this is equivalent to $(x_1 - x_2)^2 \geqslant 0$, and the result follows from Corollary 2.3.2. For $n > 2$, let $x_3 = \ldots = x_n = y_3 = \ldots = y_n = 0$. □

The *Minkowski inequality* is another basic inequality, see e.g. [BB61, §21]:

$$\prod_{i=1}^{n}\big(x_i^n \,+\, y_i^n\big) \;\ge\; \left[\prod_{i=1}^{n} x_i \,+\, \prod_{i=1}^{n} y_i\right]^n \tag{7.1.3}$$

This inequality is a special case of the *Brunn–Minkowski inequality* (for bricks in $\mathbb{R}^n$), which is foundational in the theory of geometric inequalities, see e.g. [BZ88, §8]

7.1.4 Corollary (Minkowski inequality)**.** *Denote by $M_n = M_n(x_1, \ldots, x_n, y_1, \ldots, y_n)$ the counting function given by (7.1.3), i.e., the difference of the left-hand side and the right-hand side of the inequality. For $n = 2$, we have $M_2 \not\gtrdot_{\#} 0$ unless $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.*

Proof. Note that $M_2(x_1, x_2, y_1, y_2) = C_2(x_1, y_1, x_2, y_2)$. □

The *Alexandrov–Fenchel inequality* (for *mixed volumes*) is a deep inequality in convex geometry independently proved by Alexandrov (1938) and Fenchel (1936), see e.g. [BZ88, §20] and [Sch14, §7.3]. We refer to [SvH19] for a notable recent proof and its popular exposition in [CP22]. A special case of the inequality for bricks is especially notable as it led to a complete resolution of the long open *van der Waerden conjecture* [vL81, vL82], which in turn led to further combinatorial inequalities, see e.g. [Alon03, Gur08, Sta81].

7.1.5 Theorem (Alexandrov–Fenchel inequality for bricks)**.** *Let $n \geq 2$ and let $x_i, y_i, z_{ij} \geq 0$, for all $1 \leq i, j \leq n$. Then:*

$$\operatorname{per}\begin{pmatrix} x_1 & y_1 & z_{13} & \dots & z_{1n} \\ x_2 & y_2 & z_{23} & \dots & z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & y_n & z_{n3} & \dots & z_{nn} \end{pmatrix}^2 \geq \operatorname{per}\begin{pmatrix} x_1 & x_1 & z_{13} & \dots & z_{1n} \\ x_2 & x_2 & z_{23} & \dots & z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & x_n & z_{n3} & \dots & z_{nn} \end{pmatrix} \operatorname{per}\begin{pmatrix} y_1 & y_1 & z_{13} & \dots & z_{1n} \\ y_2 & y_2 & z_{23} & \dots & z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_n & y_n & z_{n3} & \dots & z_{nn} \end{pmatrix}.$$

7.1.6 Proposition. *Denote by AF_n the polynomial in $k = n^2$ variables given in Theorem 7.1.5 by subtracting the right-hand side from the left-hand side of the inequality. For $n \geq 2$, we have $AF_n \not\geqslant_{\#} 0$ unless $\mathsf{PH} = \Sigma_2^{\mathsf{p}}$.*

Proof. Take $n = 2$, $y_1 = y_2 = 1$. Then the AF_2 inequality becomes:

$$\operatorname{per}\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix}^2 = (x_1 + x_2)^2 \geqslant \operatorname{per}\begin{pmatrix} x_1 & x_1 \\ x_2 & x_2 \end{pmatrix} \operatorname{per}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 4x_1x_2.$$

This is equivalent to $(x_1 - x_2)^2 \geqslant 0$, and the result follows from Corollary 2.3.2. □

7.2 Hadamard inequality

Recall the *Hadamard inequality* (1.3.3) discussed in the introduction. This inequality was proved by Hadamard (1893) and is crucial is the study of *positive definite matrices*, see e.g. [HJ13, §7.8]. Denote by $H_d = H_d(x_{11}, x_{12}, \dots, x_{dd})$ the nonnegative polynomial in $k = d^2$ variables given by (1.3.3). The following result is deduced from our Theorem 6.7.1.

7.2.1 Proposition (Hadamard inequality)**.** *For $d \geq 3$, if $H_d \geqslant_{\#} 0$, then there is an oracle A such that $H_3(\overrightarrow{\#\mathsf{P}}^A) \not\subseteq \#\mathsf{P}^A$.*

Proof. Observe that

$$\begin{aligned} H_3\begin{pmatrix} x & \binom{x}{3} & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} &= \tfrac{1}{12}x^6 - \tfrac{1}{2}x^5 + \tfrac{3}{4}x^4 + \tfrac{8}{3}x^2 \\ &= 3\binom{x}{1} + 6\binom{x}{2} - 3\binom{x}{3} + 28\binom{x}{4} + 90\binom{x}{5} + 60\binom{x}{6}. \end{aligned}$$

Since the coefficient of $\binom{x}{3}$ is negative, by Theorem 6.7.1 this is not a univariate relativizing closure property of $\#\mathsf{P}$. Therefore, since 0, 1, x, and $\binom{x}{3}$ are univariate relativizing closure properties of $\#\mathsf{P}$, it follows that H_3 is not a relativizing closure property of $\#\mathsf{P}$. □

Note that if the (univariate) binomial basis conjecture (Conjecture 4.4.2) is true, then this is not a closure property of $\#\mathsf{P}$, and we could conclude that $H_3(\overrightarrow{\#\mathsf{P}}) \not\subseteq \#\mathsf{P}$.

7.2.2 Remark. One might think that plugging in arbitrary binomial coefficients in the Hadamard-matrix is a valid strategy, but this does not work as the following random choice illustrates:

$$H_3\begin{pmatrix} \binom{x}{4} & \binom{x}{7} & \binom{x}{12} \\ \binom{x}{6} & \binom{x}{3} & \binom{x}{2} \\ \binom{x}{8} & \binom{x}{11} & \binom{x}{5} \end{pmatrix}$$

is indeed a relativizing closure property of $\#\mathsf{P}$.

7.3 Fermat's little theorem

As we mentioned in the introduction, *Fermat's little theorem* states that $p \,|\, a^{p-1}-1$ for all integers a, $p \nmid a$, and prime p. Fermat stated this result in 1640 without proof, and the first published proof was given by Euler in 1736. According to Dickson, "this is one of the fundamental theorems of the theory of numbers" [Dic52, p. V]. Note, see e.g. in [CC16], that Fermat's little theorem can be rephrased to say that the polynomial $\frac{1}{p}(x^p - x)$ is integer valued.

7.3.1 Proposition (Fermat's little theorem is in #P)**.** *If $f \in \#\mathsf{P}$ and p is prime, then*

$$\tfrac{1}{p}(f^p - f) \ \in \#\mathsf{P}.$$

We include a variation on Peterson's original proof [Pet72] for completeness.

Proof. Consider sequences $(a_1, \ldots, a_p)$ of integers $1 \le a_i \le f$ and partition them into orbits under the natural cyclic action of $\mathbb{Z}/p\mathbb{Z}$. Since p is prime, these orbits have either 1 or p elements. There are exactly p orbits with one elements, where $a_1 = \ldots = a_p$. The remaining orbits of size p have a total of $f^p - f$ elements. Since p is fixed, the lex-smallest orbit representative can be found in polytime. □

This proof can be rephrased to show that the polynomial $\frac{1}{p}(x^p - x)$ is binomial-good. For example, for $p = 5$, we have:

$$\frac{1}{5}\,(x^5 - x) \ = \ 24\binom{x}{5} + 48\binom{x}{4} + 30\binom{x}{3} + 6\binom{x}{2}.$$

7.3.2 Remark. Peterson also discovered a similar proof of *Wilson's theorem* [Pet72], see also [Car14, p. 50]. Stanley's elegant proof of the *Lucas congruences* is also combinatorial and in the same spirit [Sta12, Exc. 1.15(c)]. Let us also mention *Kummer's congruences* for the Catalan numbers, which can be proved via group action on binary trees, see e.g. [DS06, KPP94]. We refer to [Ges84] for a survey of combinatorial congruences, and to [RS80, Sag85] for the general group action approach. See also [AZ17] for conjectures on whether there are combinatorial proofs of various binomial congruences.

Finally, note that some congruences have combinatorial proofs for highly nontrivial reasons. Recall the *Ramanujan's congruence* $5 \,|\, p(5n-1)$ for the number of integer partitions (see e.g. [Har40, §6.4]). This congruence was famously interpreted by Dyson [Dys44], by dividing partitions of $(5n-1)$ according to its *rank* (first row minus first column) modulo 5. It remains open to find a direct bijective proof of this equal division, see [Pak06, §2.5.6]. We refer to [AG88, GKS90] for some remarkable generalizations of this approach.

7.4 Ahlswede–Daykin inequality

Let $\mathcal{A} \subseteq 2^{[n]}$ and $\zeta : 2^{[n]} \to \mathbb{R}$. Denote

$$\zeta(\mathcal{A}) \ := \ \sum_{A \in \mathcal{A}} \zeta(A).$$

For $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$, denote

$$\mathcal{A} \cup \mathcal{B} \ := \ \big\{A \cup B \ : \ A \in \mathcal{A},\, B \in \mathcal{B}\big\}, \qquad \mathcal{A} \cap \mathcal{B} \ := \ \big\{A \cap B \ : \ A \in \mathcal{A},\, B \in \mathcal{B}\big\}.$$

7.4.1 Theorem (*Ahlswede–Daykin inequality* [AS16]). *Let $\alpha, \beta, \gamma, \delta : 2^{[n]} \to \mathbb{R}_+$ be such that*

$$\alpha(A)\,\beta(B) \,\le\, \gamma(A \cap B)\,\delta(A \cup B)\,, \qquad \forall\, A, B \in 2^{[n]}. \tag{7.4.2}$$

Then

$$\alpha(\mathcal{A})\,\beta(\mathcal{B}) \,\le\, \gamma\big(\mathcal{A} \cap \mathcal{B}\big)\,\delta\big(\mathcal{A} \cup \mathcal{B}\big)\,, \qquad \forall\, \mathcal{A}, \mathcal{B} \subseteq 2^{[n]}. \tag{7.4.3}$$

This inequality is classical, and is an advanced generalization of the Kleitman inequality, see below. In its most general form it is usually stated for general lattices, not just the Boolean lattice. Among its many applications, let us single out the *FKG inequality* [AS16, §6.2] and the *XYZ inequality* [AS16, §6.4], see also §9.2.

Let $n = 1$, and let $\mathcal{A} = \mathcal{B} = \{0, 1\}$. Operations $\cap$ and $\cup$ are replaced with min and max in this case. Let $\alpha_i, \beta_i, \gamma_i, \delta_i$, $i \in \{0, 1\}$, be #P functions satisfying (7.4.2). This is guaranteed by taking functions $h_1, h_2, h_3, h_4 \in \#\mathsf{P}$, such that

$$\alpha_0\beta_0 + h_1 \,=\, \gamma_0\delta_0, \quad \alpha_0\beta_1 + h_2 \,=\, \gamma_0\delta_1, \quad \alpha_1\beta_0 + h_3 \,=\, \gamma_0\delta_1, \quad \alpha_1\beta_1 + h_4 \,=\, \gamma_1\delta_1$$

Let AD_n be the function defined by (7.4.3). Then we have:

$$AD_1 \;=\; (\gamma_0 + \gamma_1)(\delta_0 + \delta_1) \,-\, (\alpha_0 + \alpha_1)(\beta_0 + \beta_1).$$

Now Proposition 2.5.1 gives the oracle separation for this inequality.

7.5 Karamata inequality

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ be nonincreasing sequences of real numbers. We say that $\boldsymbol{x}$ *majorizes* $\boldsymbol{y}$, write $\boldsymbol{x} \trianglerighteq \boldsymbol{y}$, if

$$\begin{aligned} &x_1 + \ldots + x_i \,\ge\, y_1 + \ldots + y_i \quad \text{for all } \ 1 \le i < n, \quad \text{and} \\ &x_1 + \ldots + x_n \,=\, y_1 + \ldots + y_n\,. \end{aligned}$$

In combinatorial context, this is also called the *dominance order*, and appears throughout the area, see e.g. [Bru06, Mac95, Sta12]. See also [Bar07] for a recent connection to the problem of counting *contingency tables*.

7.5.1 Theorem (Karamata inequality). *Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, such that $\boldsymbol{x} \trianglerighteq \boldsymbol{y}$. Then, for every convex function $F : \mathbb{R}^n \to \mathbb{R}$, we have $F(\boldsymbol{x}) \ge F(\boldsymbol{y})$.*

This result is classical, see e.g. [HLP52, §3.17] and [BB61, §28, §30]. Analytically, the inequality can be used to derive *Jensen's inequality*, which in turn implies the AM-GM inequality. See [BP21, PPS20] for some recent applications of the Karamata inequality to combinatorial problems on linear extensions and Young tableaux, respectively. See also [MOA11] for a modern proof, numerous generalizations and further references.

We now convert the Karamata inequality into a counting function problem. Suppose we are given $f_i, g_i \in \#\mathsf{P}$, $1 \le i \le n$, such that the following functions h_i, $1 \le i < n$, are also all in #P:

$$h_i \,:=\, f_1 \,+\, \ldots \,+\, f_i \,-\, g_1 \,-\, \ldots \,-\, g_i \tag{7.5.2}$$

and we are also guaranteed that

$$f_1 \,+\, \ldots \,+\, f_n \,-\, g_1 - \ldots - g_n \,=\, 0. \tag{7.5.3}$$

Moreover, the functions

$$d_i := f_i - f_{i+1} \quad \text{and} \quad e_i := g_i - g_{i+1} \tag{7.5.4}$$

are also in $\#\mathsf{P}$ for all $1 \leq i < n$. Let $Z \subseteq \mathbb{Q}^{5n-3}$ denote the variety of points that satisfy the constraints (7.5.3) and (7.5.4). Let $\gamma \in \mathsf{GapP}_{\geq 0}$ be any convex function.

Define the *Karamata function* as

$$K_{n,\gamma}(\vec{f}, \vec{g}) := \sum_{i=1}^{n} \gamma(f_i) - \sum_{i=1}^{n} \gamma(g_i).$$

Clearly $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}) \subseteq \mathsf{GapP}$. Karamata's inequality implies that the answer is always nonnegative for inputs from Z. This is a semantic guarantee, i.e., we have no information as to why the guarantee holds. This is analogous to $\#\mathsf{P}_{\geq 1}$, but in contrast to problems in Section 8.

We write $\overrightarrow{\#\mathsf{P}}_{\in Z}$ for a tuple of $\#\mathsf{P}$ functions whose function values always lie on Z. Hence, $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in Z}) \subseteq \mathsf{GapP}_{\geq 0}$. Consider the following examples of n, γ for which $K_{n,\gamma}(\overrightarrow{\#\mathsf{P}}_{\in Z}) \subseteq \#\mathsf{P}$:

- For affine linear γ, we clearly have $K_{n,\gamma} = 0 \in \#\mathsf{P}$.
- For $\gamma(t) = t^2$, we have $K_{2,\gamma}(f_1, f_2, g_1, g_2) = (d_1 + e_1)h_1$ as a function on Z. This can be seen by plugging in $d_1 = f_1 - f_2$, $e_1 = g_1 - g_2$, and $g_2 = f_1 + f_2 - g_1$. Clearly $(d_1 + e_1)h_1 \in \#\mathsf{P}$. This has several proofs, for example instead of $(d_1 + e_1)h_1$ we could have taken $2h_1 + 2e_1h_1 + 4\binom{h_1}{2}$ with the same argument.
- For $\gamma(t) = t^2$, we have $K_{3,\gamma}(f_1, f_2, f_3, g_1, g_2, g_3) = (d_1 + e_1)h_1 + (d_2 + e_2)h_2 \in \#\mathsf{P}$ on Z.
- For $\gamma(t) = \binom{t}{2}$, we have $K_{2,\gamma}(f_1, f_2, g_1, g_2) = (e_1 + 1)h_1 + 2\binom{h_1}{2} \in \#\mathsf{P}$ on Z.
- For $\gamma(t) = \binom{t}{2}$, we observe that for the *double* we have $2K_{3,\gamma} \in \#\mathsf{P}$ via the observation that $2K_{3,\binom{t}{2}} = K_{3,t^2}$ on Z (the affine linear parts cancel out).

All inclusions $K_{n,\gamma} \subseteq \#\mathsf{P}$ in this section so far relativize. The next proposition shows that the doubling we just used was in fact necessary, because otherwise obtain an oracle separation.

7.5.5 Proposition. *There exists a language $A \subseteq \{0,1\}^*$ such that $K_{3,\binom{t}{2}}(\overrightarrow{\#\mathsf{P}^A}_{\in Z}) \not\subseteq \#\mathsf{P}^A$.*

Proof. We use the Diagonalization Theorem 6.2.1. We have $5n - 3 = 12$, so in the notation of the theorem, we set $S = Z \cap \mathbb{N}^{12}$. We fix an arbitrary order of the 12 variables: $(f_1, f_2, f_3, g_1, g_2, g_3, d_1, d_2, e_1, e_2, h_1, h_2)$. The variety Z is then given as the kernel of the linear map given by the following matrix:

$$\begin{pmatrix}
1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

We convert it to row echelon form:

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & -1 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 1 & -1 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & -2
\end{pmatrix}$$

We set $\ell = 5$, $k = 12$. Permuting the order of the columns to $(6, 9, 10, 11, 12, 1, 2, 3, 4, 5, 7, 8)$, we obtain affine linear functions $\zeta_8, \ldots, \zeta_{12}$, where each ζ_b depends linearly on the first 5 variables, which we call $(v_1, \ldots, v_\ell) := (g_3, e_1, e_2, h_1, h_2)$. We verify the assumptions of the Diagonalization Theorem 6.2.1:

1. Since all constraints are homogeneous, we have zero is an integer point: $0 \in Z$.
2. All ζ have integer coefficients and all constraints are homogeneous, so the fact that C'_S lies Zariski-dense in $\mathbb{Q}^5$ follows from the next point and having a small open ball that contains $\vec{v}$.
3. The point $\vec{v} = (g_3 = 2, e_1 = 2, e_2 = 1, h_1 = 1, h_2 = 1)$ satisfies all $\zeta_b(\vec{v}) > 0$. Here $\vec{v}$ is the point corresponding to the case $\vec{f} = (6, 3, 1)$ and $\vec{g} = (5, 3, 2)$.

We set $\textsc{Multiplicities} = \textsc{OccurrenceMulti}_{12}$ and use the set-instantiators from Theorem 6.6.1. We fix $\vec{t} \in S$, and set

$$\varphi(f_1, f_2, f_3, g_1, g_2, g_3, d_1, d_2, e_1, e_2, h_1, h_2) \;:=\; f_1^2 + f_2^2 + f_3^2 - g_1^2 - g_2^2 - g_3^2.$$

We now show that $\varphi + I$ is binomial-bad. We use the Polyhedron Theorem 5.5.2, since all our constraints are affine linear. Since $\dim[\mathbb{Q}^{12}]_{\le 2} = \binom{14}{2} = 91$ and $\dim[\mathbb{Q}^5]_{\le 2} = \binom{7}{2} = 21$, this is a polyhedron in $\mathbb{Q}^{91}$ with 21 linear equations intersected with the nonnegative orthant. We use a computer to set up the polyhedron. Indeed, it contains the half-integer point that shows that $2K_{3,\binom{t}{2}} \in \#\mathsf{P}$, but it does *not* contain an integer point, which gives a function p_A with $\varphi(p_A) \notin \#\mathsf{P}^A$.

It remains to show that $\varphi(p_A) \in K_{3,\binom{t}{2}}\big(\overrightarrow{\#\mathsf{P}}^A_{\in Z}\big)$. This can be seen by the existence of the list of oracle Turing machines that on input w first query A at $0^{|A|}$. Then, if $0^{|w|} \notin A$ accept with multiplicities $\vec{t}$; if $0^{|w|} \in A$, then the machines run $\textsc{OccurrenceMulti}$ on $\widetilde{A}_j$. The fact that the output vector is in Z is guaranteed by the fact that if $0^{|w|} \in A$, then $\widetilde{A}_j$ is generated by the set-instantiator, which only generates instances from S. □

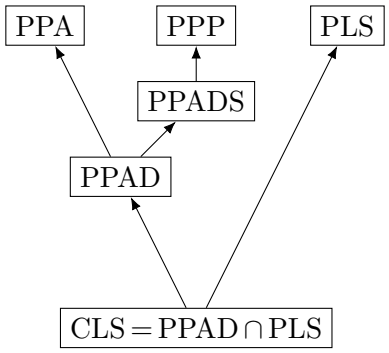
7.5.6 Remark. One can ask if the assumptions underlying Proposition 7.5.5 are reasonable, e.g. whether there is a natural combinatorial problem where we have $h_i \in \#\mathsf{P}$ for all functions defined in (7.5.2). For an example of this, we refer to above mentioned result in [PPS20], where the majorization is proved by a direct injection. In the case of Young diagrams, the "shuffling in the plane" proof in [PPS20, §4] is based on the *wall-equivalence*, see e.g. [Mat10, §23].

We should warn the reader that some applications of the Karamata inequality can in fact be in $\#\mathsf{P}$. For example, the *hook inequality* for the number of increasing trees is proved in [PPS20, §2] via the Karamata inequality, but is a special case of the *Björner–Wachs inequality* (Theorem 9.2.1) known to be in $\#\mathsf{P}$.

8 TFNP and #P

8.1 Definitions and background

Recall the classical inclusion diagram of search complexity classes:



When going to the counting analogs minus 1, as explained in §3.2, one gets the same picture (the right-hand side of the tree in Figure 1), but care must be taken with the complete problems, since for several of the classical problems in CLS, PPAD, and PPADS the counting version minus 1 actually lies in #P.

We study the following classical problems and slightly adjusted problems that are not parsimoniously equivalent. We assume that the reader is familiar with how to encode an exponential graph or digraph via successor or predecessor/successor circuits, see [GW83, MP91].

- The PPAD-complete problem SourceOrSink, which is parsimoniously equivalent to Sperner, see [CD09]. We are given two circuits C_{succ} and C_{pred} that describe a directed graph in which there is an edge from x to y if and only if $C_{\text{succ}}(x) = y$ and $C_{\text{pred}}(y) = x$. We syntactically ensure that the indegree of the 0 vertex is 0 and its outdegree is 1. We search for sources or sinks, i.e., nonzero vertices of (indegree,outdegree) $= (0, 1)$ or (indegree,outdegree) $= (1, 0)$.

- The PPAD-complete problem SourceOrPresink (see Claim 8.2.4). The setup is the same as for SourceOrSink, but we count sources and presinks, where a presink is a vertex that is adjacent to a sink. Note that the two-vertex graph with a single source and sink only counts once, as it has only one node that is a source or a presink (it is actually both a source and a presink).

- The PPAD-complete problem SourceOrExcess(2,1) (see Claim 8.2.5). We are given three circuits C_{succ} and C_{pred_1} and C_{pred_2} that describe a directed graph in which there is an edge from x to y if and only if $C_{\text{succ}}(x) = y$ and $\big(C_{\text{pred}_1}(y) = x \vee C_{\text{pred}_2}(y) = x\big)$. This results in a digraph with indegrees from $\{0, 1, 2\}$ and outdegrees from $\{0, 1\}$. As for SourceOrSink we syntactically ensure that the indegree of the 0 vertex is 0 and its outdegree is 1. We search for sources or excess vertices, i.e., nonzero vertices where the indegree differs from the outdegree. It is crucial here that these digraphs can have double sinks, i.e., vertices with indegree 0 and outdegree 2, that we only count once.

- Instead of being given a source, we can also count AllSourcesOrSinks. This is mainly interesting, because the number is always even, so we can ask if AllSourcesOrSinks/2 $\in$ #P (which it is, but for the undirected analog we have an oracle separation from #P, see Theorem 8.7.1). We call the corresponding counting class #CountAll-PPAD(SourceOrSink).

- The PPADS-complete problem Sink. The setup is the same as for SourceOrSink, but we only count sinks.
- The PPADS-complete problem Presink (see Claim 8.2.8). The setup is the same as for Sink, but we only count presinks, which are vertices that are adjacent to sinks.
- The PPADS-complete problem Excess(2,1) (see Claim 8.2.9). The setup is the same as for SourceOrExcess(2,1), but we only count nodes with indegree greater than outdegree.
- The PPA-complete problem Leaf, which is parsimoniously equivalent to Lonely and Even and Odd, see [BCE+98]. We are given two circuits C_1 and C_2 that describe a graph in which an edge between x and y is present if and only if

$$\big(C_1(x) = y \vee C_2(x) = y\big) \wedge \big(C_1(y) = x \vee C_2(y) = x\big).$$

  We syntactically ensure that the degree of the 0 vertex is 1. We search for leaves, i.e., nonzero vertices of degree 1.
- The hardness of $\big(\textsc{Leaf} - 1\big)/2$ actually comes from the hardness of the easier problem $(\textsc{AllLeaves})/2$. In the problem AllLeaves the setup is the same as for Leaf, but we have no syntactic guarantee about the zero vertex. The number of solutions is always even, and it can be zero. We call the corresponding counting class #COUNTALL-PPA(Leaf).
- The PPA-complete problem Preleaf (see Claim 8.2.12), which has the same setup as Leaf, but we search for vertices that are adjacent to leaves. Note that the line graph with 3 vertices has 2 leaves, but only 1 preleaf.
- A slightly adjusted PPP-complete version of the problem Pigeon (see Claim 8.2.14), where we are given a circuit C and search for vertices x with $C(x) = 0$, or for pairs of vertices (x, y) with $C(x) = C(y)$ and $C(C(x)) \neq 0$. This version gives a cleaner counting problem.
- The classical PLS-complete problem Iter which is parsimoniously equivalent to LocalOpt (see e.g. [FGHS21]), where we are given a circuit C that has the syntactically ensured guarantee that $C(x) \geq x$ and $C(0) > 0$ and $C(C(0)) > C(0)$; and we search for presinks, i.e., for vertices x such that $C(x) \neq x$ and $C(C(x)) = x$.[19] It is important to note that Iter differs significantly from the other problems: its instances cannot be freely permuted. This makes the construction of a set-instantiator in §8.4 quite different from the other set-instantiator constructions.
- A slightly adjusted version of the CLS-complete problem

$$\textsc{EitherSolution}(\textsc{SourceOrSink}, \textsc{Iter})$$

  (see Claim 8.2.16), where in the counting version we are given a pair of a SourceOrSink and an Iter instance and count each solution to either of them, with a slight adjustment: If the SourceOrSink instance has a source or sink at the last possible position, then we say that the instance contains the *last option*. If the Iter instance has a presink at the last possible position (i.e., the second to last vertex), then we say that the instance contains the last option. Note that we can efficiently check if either instance contains the last option. If

[19] Note that if given only a successor circuit, it is difficult to check if a vertex is a sink, so we count presinks instead. Also note that allowing 0 to be a presink would also have been an option.

both instances contain the last option, then we count these two solutions only once. This adjustment is to ensure that it is possible to have a single solution. Otherwise the counting version would be in $\#\mathsf{P}_{\geq 2}$.[20]

- The analogous CLS-complete problem EitherSolution(SourceOrPresink,Iter) (see Claim 8.2.16).
- The analogous CLS-complete problem EitherSolution(SourceOrExcess(2,1),Iter) (ibid.)
- The problem BipartiteUnbalance has no associated search problem. It is the following problem in $\mathsf{GapP}_{\geq 0}$. Let $G = (V, E)$ be a bipartite graph with two parts given by $V = V_- \sqcup V_+$. We say that G is *unbalanced* if for every $(uv) \in E$, $u \in V_-$, $v \in V_+$, we have $\deg(u) \geq \deg(v)$. An instance of the problem is given by a list of polynomially many circuits $C_1, \ldots, C_n$ such that an edge between $u \in V_-$ and $v \in V_+$ exists if and only if $C_i(u) = v$ for some i, and $C_j(v) = u$ for some j. We syntactically ensure that $\deg(u) \geq \deg(v)$ locally by adding vertices from V_- whenever $\deg(u) < \max_{v \in N(u)} \deg(v)$, where $N(u)$ is the neighborhood of u: In this case we add $\max_{v \in N(u)} - \deg(u)$ many vertices in V_- to the graph and connect them to u, but do not connect them to any other vertex. Proposition 8.2.20 below shows that $|V_+| - |V_-| \geq 0$. The problem BipartiteUnbalance is the counting problem with value $|V_+| - |V_-|$. We study this problem in §8.9, cf. Open Problem 9.2(3).

8.2 Simple completeness results, equalities to #P, and simple inclusions

All results in this section are fairly straightforward. They appear here to avoid any oversights, because Figure 1 suggests that #P is strictly contained in classes that differ only slightly in terms of definition.

8.2 (a) #PPAD

8.2.1 Claim. $(\#\mathsf{PPAD}(\text{SourceOrSink}) - 1)/2 = \#\mathsf{P}$ *via relativizing parsimonious reductions.*

Proof. We first note that $(\#\mathsf{PPAD}(\text{SourceOrSink}) - 1)/2 \subseteq \#\mathsf{P}$, because we can just count the nonzero sources, which gives us the correct number. For the other direction, we use that CircuitSat^A is $\#\mathsf{P}^A$-complete. If we are given a Boolean single-output circuit C (with oracle gates) that has c many x for which $C(x)$ is true, then we can simply construct a SourceOrSink instance with the zero vertex immediately going into a sink, the rest of the vertices all being self-loops with one exception: for every position $2x$ for which $C(x)$ is true we add a source-sink edge from $2x$ to $2x + 1$. We end up with an instance of value $2c + 1$, as desired. □

8.2.2 Claim. $\#\mathsf{PPAD}(\text{SourceOrPresink}) - 1 = \#\mathsf{P}$ *via relativizing parsimonious reductions.*

Proof. We have s nonzero sources, t presinks, and a nodes that are both, which we call amalgamations. We count $s + t + a - 1 = 2s + a$, i.e., we count the nonzero sources twice and the amalgamations once. The second direction is the same construction as for Claim 8.2.1, but here the instance ends up with value $c + 1$ instead of $2c + 1$, because the source is also a presink. □

[20] [GHJ+22] give a combinatorial version of CLS, which was pointed out to us by M. Göös after this paper was written.

8.2.3 Claim. $\#\mathsf{COUNTALL\text{-}PPAD}(\textsc{SourceOrSink})/2 = \#\mathsf{P}$ *via relativizing parsimonious reductions.*

Proof. To see the containment in #P, we just count only the sources (or only the sinks). The second direction is the same construction as in Claim 8.2.2, just without the source-sink pair for the zero vertex. □

8.2.4 Claim. SourceOrPresink *is* PPAD*-complete.*

Proof. This follows directly from the fact that if we find a sink, then it is easy to find a presink, and vice versa, and the fact that SourceOrSink is PPAD-complete. □

8.2.5 Claim. SourceOrExcess(2,1) *is* PPAD*-complete.*

Proof. Clearly, every SourceOrSink instance is a SourceOrExcess(2,1) instance. Given a SourceOrExcess(2,1) instance, we replace every double sink with two sinks, and every indegree 2, outdegree 1 vertex with a sink and an indegree 1, outdegree 1 vertex. Note that we replaced only excess vertices and added 1 or 2 sinks each time. Since the construction was local, a solution to this SourceOrSink instance can be converted back into a solution of the SourceOrExcess(2,1) instance. □

8.2.6 Claim.

$$\#\mathsf{PPAD}(\textsc{SourceOrExcess}(2,1)) - 1 \subseteq \#\mathsf{PPADS}(\textsc{Excess}(2,1)) - 1$$

via relativizing parsimonious reductions.

Proof. Given a SourceOrExcess(2,1) instance we create an Excess(2,1) instance of the same value by adding a (source,sink) pair that is not connected to the rest of the instance, for every source vertex in the input. □

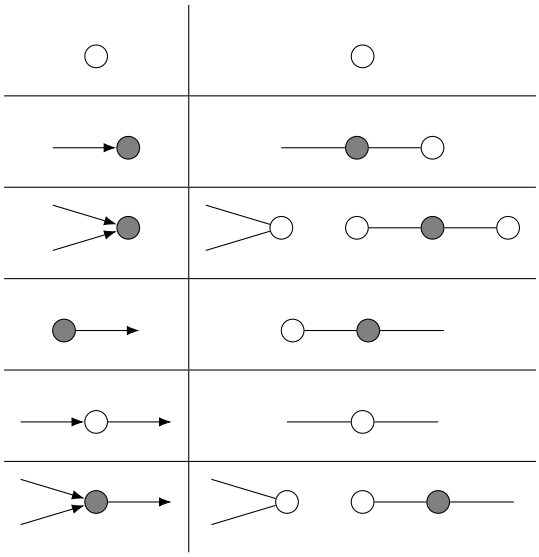
8.2.7 Claim.

$$\#\mathsf{PPAD}(\textsc{SourceOrExcess}(2,1)) - 1 \subseteq \#\mathsf{PPA}(\textsc{Preleaf}) - 1$$

via relativizing parsimonious reductions.

Proof. Given a SourceOrExcess(2,1) instance, we make local graph replacements with gadgets, making sure that no vertex that connects to another gadget is a leaf, so that we know exactly which vertices of a gadget are preleaves. Since the indegree is an element from $\{0,1,2\}$ and the outdegree is an element from $\{0,1\}$, there are $2 \cdot 3 = 6$ possible vertex types in the input SourceOrExcess(2,1) instance.

The transformation table below shows the gadgets that are used to replace the parts in the SourceOrExcess(2,1) instance. Here the sources and excess nodes are marked gray on the left-hand side, preleaves are marked gray on the right-hand side. The proof follows from the fact that in each row the number of gray vertices on the left equals the number of gray vertices on the right. □



8.2 (b) #PPADS

8.2.8 Claim. PRESINK *is* PPADS*-complete.*

Proof. This follows directly from the fact that if we find a sink, then it is easy to find a presink, and vice versa. □

8.2.9 Claim. EXCESS(2,1) *is* PPADS*-complete.*

Proof. Clearly, every SINK instance is a EXCESS(2,1) instance. Given an EXCESS(2,1) instance, we replace every double sink with two sinks, and every indegree 2, outdegree 1 vertex with a sink and an indegree 1, outdegree 1 vertex. Note that we replaced only excess vertices and for each replacement we locally added 1 or 2 sinks. Hence, a solution to this SINK instance can be converted back into a solution for the EXCESS(2,1) instance. □

8.2.10 Claim. #PPADS(SINK) $- 1 =$ #P *via relativizing parsimonious reductions.*

Proof. To see the containment in #P, we just count only the nonzero sources. The second direction is the same construction as in Claim 8.2.2. □

8.2.11 Claim. #PPADS(PRESINK) $- 1 =$ #P *via relativizing parsimonious reductions.*

Proof. This is the same proof as for Claim 8.2.10. □

8.2 (c) #PPA

8.2.12 Claim. PRELEAF *is* PPA*-complete.*

Proof. This follows directly from the fact that if we find a leaf, then we can readily find a preleaf, and vice versa. □

8.2.13 Claim. #COUNTALL-PPA(LEAF)/2 $\subseteq$ (#PPA(LEAF)-1)/2 *via relativizing parsimonious reductions.*

Proof. Given a ALLLEAVES instance G we create a LEAF instance G' as follows. We add a zero leaf vertex and pair it with another newly created leaf, so ALLLEAVES$(G)+1=$ LEAF(G'). Therefore, if ALLLEAVES$(G)/2=k$, then (LEAF$(G')-1)/2=$ (ALLLEAVES$(G)+1-1)/2=k$, which finishes the proof. □

8.2 (d) #PPP

We recall the slightly different definition of the classical PIGEON problem: In our case, if $C(x)=C(y)$ and $C(C(x))=0$, then we do not count this as a solution. However, in this case $z=C(x)$ obviously is a solution, because $C(z)=0$.

8.2.14 Claim. PIGEON *is* PPP*-complete.*

Proof. For the sake of this proof, let CLPIGEON denote the classical pigeon search problem and PIGEON our problem with the slight modification. Given a PIGEON instance. If we find a solution to the PIGEON instance, then this is also a solution for the same instance interpreted as a CLPIGEON instance. If we find a solution to a CLPIGEON problem, then this is also a solution for the same instance interpreted as a PIGEON instance, with one exception: If the solution pair (v,w) satisfies $\varphi(v)=\varphi(w)=x$ and $\varphi(x)=0$. But in this case we immediately find that x is a solution for PIGEON. □

8.2.15 Claim. #PPADS(EXCESS(2,1)) $-1\subseteq$ #PPP(PIGEON) -1 *via relativizing parsimonious reductions.*

Proof. Given an EXCESS(2,1) instance with edges $(v,w)\in\{0,1\}^n\times\{0,1\}^n$, we design a PIGEON instance, which is a map $\varphi:\{0,1\}^n\to\{0,1\}^n$. If the outdegree of w is 1, say (w,v) is an edge, then we set $\varphi(w)=v$. If the outdegree of w is 0, we set $\varphi(w)=w$, with one exception: for every double sink w we set $\varphi(w)=0$. This table explains in which cases the output PIGEON instance has witnesses to count:

(indegree, outdegree) of w	witnesses in φ
(0,0)	no witnesses, $\varphi(w)=w$
(1,0)	yes, $\varphi(v)=\varphi(w)=w$
(2,0)	yes, $\varphi(w)=0$. Note that we do not count the pair (v,v') with $\varphi(v)=\varphi(v')=w$ here, because $\varphi(w)=0$.
(0,1)	no
(1,1)	no
(2,1)	yes, one witness: Either we have $\varphi(v)=\varphi(v')=w$ and $\varphi(w)\neq 0$, or we have $\varphi(w)=0$.

We observe that we have one witness in exactly the excess cases, and no witnesses otherwise. □

8.2 (e) #CLS

The counting problems EitherSolution(∗,Iter) have the technicality that if both component instances contain the last option, then this witness is only counted once. In particular, this is used to show Claim 8.2.17 below. This has no effect on the CLS-completeness of the search problem. For the oracle separation in Theorem 8.3.1 this also plays no role.

8.2.16 Claim. *The search problems to the counting problems*

$$\begin{aligned}&\textsc{EitherSolution}(\textsc{SourceOrSink},\textsc{Iter}),\\&\textsc{EitherSolution}(\textsc{SourceOrPresink},\textsc{Iter}),\ \textit{and}\\&\textsc{EitherSolution}(\textsc{SourceOrExcess}(2,1),\textsc{Iter})\end{aligned}$$

are all CLS*-complete.*

Proof. The search problems are in fact the same as the classical CLS-complete problems. Only the counting versions are different, as in some situations 2 solutions are counted together as 1 solution. □

8.2.17 Claim.

$$\begin{aligned}&\#\mathsf{CLS}(\textsc{EitherSolution}(\textsc{SourceOrExcess}(2,1),\textsc{Iter}))-1\\&\quad\subseteq\ \ \#\mathsf{PLS}(\textsc{Iter})-1\ \cap\ \#\mathsf{PPAD}(\textsc{SourceOrExcess}(2,1))-1\end{aligned}$$

via relativizing parsimonious reductions.

Proof. We first prove that

$$\#\mathsf{CLS}(\textsc{EitherSolution}(\textsc{SourceOrExcess}(2,1),\textsc{Iter}))-1\ \subseteq\ \#\mathsf{PLS}(\textsc{Iter})-1.$$

Given an EitherSolution(SourceOrExcess(2,1),Iter) instance, we construct an Iter instance of the same value. This new instance consists of two parts.

We copy the input Iter instance into the first part of the output Iter instance. Then for every source or excess vertex in the input SourceOrExcess instance we construct a (source,sink) pair that consists of a single edge in the second part of the output Iter instance, with one exception: if both the SourceOrExcess input and the Iter input contain the last option, then we do not create the last (source,sink) pair. The value of the resulting instance is the sum of the two values of the input instances in the case where not both use the last option, and is 1 less otherwise. That is exactly the correct amount, by definition of EitherSolution(SourceOrExcess(2,1),Iter).

The proof for

$$\begin{aligned}&\#\mathsf{CLS}(\textsc{EitherSolution}(\textsc{SourceOrExcess}(2,1),\textsc{Iter}))-1\\&\quad\subseteq\ \ \#\mathsf{PPAD}(\textsc{SourceOrExcess}(2,1))-1\end{aligned}$$

is only slightly more tricky. Given an EitherSolution(SourceOrExcess(2,1),Iter) instance, we construct a SourceOrExcess(2,1) instance of the same value. This new instance consists of two parts, where we have the zero vertex in the second part.

We copy the input SourceOrExcess(2,1) instance into the first part of the output SourceOrExcess(2,1) instance. Then we create the second part as a long line starting from the zero vertex and pointing from x to $x+2$ leaving 1 loop in between, until the last vertex maps to the source of the original zero source of the instance that is embedded into the first part. Then

for every presink in the input ITER instance at position x, we construct a source in the second part at $x+1$ pointing at x. This makes x an excess vertex.

We have one exception to this rule: If both the SOURCEOREXCESS input and the ITER input contain the last option, then then we do not create the last excess vertex. The value of the resulting instance is the sum of the two values of the input instances in the case where not both use the last option, and is 1 less otherwise. That is exactly the correct amount, by definition of EITHERSOLUTION(SOURCEOREXCESS(2,1),ITER). □

8.2.18 Claim. $\#\mathsf{CLS}(\text{EITHERSOLUTION}(\text{SOURCEORSINK},\text{ITER})) - 1 = \#\mathsf{P}$ *via relativizing parsimonious reductions.*

Proof. We first prove the inclusion $\#\mathsf{P} \subseteq \#\mathsf{CLS}(..) - 1$. Given a CIRCUITSAT instance, we create a EITHERSOLUTION(SOURCEORSINK,ITER) instance with one more solution: the SOURCEORSINK part of our output instance points from the zero vertex to the last vertex, hence using the last option. The ITER part of our output instance points from the zero vertex to the vertex before the last vertex and from there to the last vertex, hence also using the last option. Now, we add (source,sink) pairs to the ITER part at all position where the CIRCUITSAT instance yields True. This gives a EITHERSOLUTION(SOURCEORSINK,ITER) instance of value 1 more than the CIRCUITSAT instance.

We now show the opposite inclusion $\#\mathsf{CLS}(..) - 1 \subseteq \#\mathsf{P}$. The proof is similar as for SPERNER. The $\#\mathsf{P}$ machine first checks if both parts use the last option. If yes, then not two, but only one of the nondeterministic computation paths accepts. The rest of the computation counts all SOURCEORSINK nonzero sources twice and all ITER presinks once. □

8.2.19 Claim. $\#\mathsf{CLS}(\text{EITHERSOLUTION}(\text{SOURCEORPRESINK},\text{ITER})) - 1 = \#\mathsf{P}$ *via relativizing parsimonious reductions.*

Proof. The inclusion $\#\mathsf{P} \subseteq \#\mathsf{CLS}(..) - 1$ is the same as in Claim 8.2.18, but in the SOURCEORPRESINK part instead of directly pointing to the last vertex we create the edge from the zero vertex to the second to last vertex and then an edge to the last vertex, with the same effect.

The proof of the opposite inclusion $\#\mathsf{CLS}(..) - 1 \subseteq \#\mathsf{P}$ is only slightly more subtle than for Claim 8.2.18, because a SOURCEORPRESINK instance can have vertices that are both source and presink, which we call amalgamations. The $\#\mathsf{P}$ machine first checks if both parts use the last option. If yes, then not two, but only one of the nondeterministic computation paths accepts. The rest of the computation counts all SOURCEORPRESINK nonzero sources twice, all SOURCEORPRESINK amalgamations, and all ITER presinks once. □

8.2 (f) BIPARTITEUNBALANCE

Recall the setting of this problem. Let $G = (V, E)$ be a bipartite graph with two parts given by $V = V_- \sqcup V_+$. We say that G is *unbalanced* if for every $(uv) \in E$, $u \in V_-$ and $v \in V_+$, we have $\deg(u) \geq \deg(v)$.

8.2.20 Proposition. *Let $G = (V, E)$, $V = V_- \sqcup V_+$, be an unbalanced bipartite graph as above. Then $|V_+| \geq |V_-|$.*

Proof. The proof is one line:

$$|V_+| \;=\; \sum_{v\in V_+} 1 \;=\; \sum_{v\in V_+}\sum_{(u,v)\in E} \frac{1}{\deg(v)} \;=\; \sum_{(u,v)\in E} \frac{1}{\deg(v)} \;\geq$$
$$\geq \sum_{(u,v)\in E} \frac{1}{\deg(u)} \;=\; \sum_{u\in V_-}\sum_{(u,w)\in E} \frac{1}{\deg(u)} \;=\; \sum_{u\in V_-} 1 \;=\; |V_-|\,,$$

where the only inequality is by the unbalanced condition. □

This first insight shows that BipartiteUnbalance $\in \mathsf{GapP}_{\geq 0}$. We study this problem further in §8.9.

8.3 The decrementation separation

In this section we prove the following technical result:

8.3.1 Theorem. *We have* $\#\mathsf{P} \subseteq \#\mathsf{CLS}(\textsc{EitherSol}(\textsc{SourceOrSink}(2{,}1),\textsc{Iter})) - 1$ *via a parsimonious relativizing reduction. Moreover, there exists a language* $A \subseteq \{0,1\}^*$ *with respect to which* $\#\mathsf{CLS}(\textsc{EitherSol}(\textsc{SourceOrSink}(2{,}1),\textsc{Iter}))^A - 1 \not\subseteq \#\mathsf{P}^A$.

Proof. The inclusion $\#\mathsf{P}^A \subseteq \#\mathsf{CLS}(..)^A - 1$ is of the same level of difficulty as similar inclusions in §8.2. Given a CircuitSat instance, we construct an instance of $\#\mathsf{CLS}(..)$ as follows. The SourceOrSink(2,1) instance jumps directly from the zero vertex to the last vertex. The Iter instance also does that. Now wherever the CircuitSat circuit yields true, we add a (source,sink) pair on the Iter side. We obtain an instance of value $f+1$, where the $+1$ comes from the two last options that are counted together as 1. We are done, because $(f+1)-1 = f$.

For the other direction we use the Diagonalization Theorem 6.2.1 and the set-instantiator for Multi(NZSource,Excess,Iter), see Theorem 8.6.1, with

$$k = 3\,, \quad S \;=\; \left\{\, \vec{f} \in \mathbb{N}^k \;:\; 2f_2 = f_1 + 1,\; f_2 \geq 1,\; f_3 \geq 1 \,\right\}.$$

We switch the positions 1 and 3 to get the set-instantiator for Multi(Iter,Excess,NZSource) with

$$k = 3\,, \quad S \;=\; \left\{\, \vec{f} \in \mathbb{N}^k \;:\; 2f_2 - 1 = f_3,\; f_2 \geq 1,\; f_1 \geq 1 \,\right\}.$$

We set $\ell = 2$ and $\zeta_3(v_1,v_2) = 2v_2 - 1$, so that $I = \langle 2v_2 - 1 - f_3 \rangle$ and

$$Z \;=\; \left\{\, \vec{f} \in \mathbb{Q}^3 \;:\; 2f_2 - 1 - f_3 = 0 \,\right\}.$$

We choose any $\vec{t} \in S$ such that $C_\perp$ exists with Multi(Iter,Excess,NZSource)$(C_\perp) = \vec{t}$, and $C_\perp$ does not use both last options. Now we verify the preconditions of the Diagonalization Theorem 6.2.1:

1. Z contains the integer point $(0,0,-1)$.
2. We prove that C'_S lies Zariski-dense in $\mathbb{Q}^2$ by giving infinitely many distinct rays from the origin that lie in C'_S. This proves the claim, because we are in a 2-dimensional situation. In fact, for every $(1,f_2)$, $f_2 \in \mathbb{N}_{\geq 1}$, the corresponding ray $\{(\alpha, \alpha f_2) \;:\; \alpha \in \mathbb{Q}_{\geq 0}\}$ is in C'_S, because $(\alpha, \alpha f_2, 2\alpha f_2 - 1) \in S$ for all $\alpha \in \mathbb{N}_{\geq 1}$.
3. $\vec{v} = (0,1)$ satisfies $2v_2 > 0$.

We fix $\textsc{Multiplicities} = \textsc{Multi}(\textsc{Iter},\textsc{Excess},\textsc{NZSource})$ for which we know for every M the existence of a set-instantiator against M. We set $\varphi(f_1,f_2,f_3) = f_1+f_2+f_3-1$. We verify that $\varphi + I = f_1+f_2+f_3-1+\langle 2f_2-1-f_3\rangle$ is binomial-bad. This is proved using the Polyhedron Theorem 5.5.2.

In the notation of the Polyhedron Theorem 5.5.2, we have $\varphi' = f_1+3f_2-2$ and $\delta = 1$, so $\mathbb{O}_\delta$ has only 3 elements. Hence the polyhedron $\mathcal{P}(\varphi,\zeta)$ will be defined using 3 equations. Let $(1,v_1,v_2)$ be the order of basis vectors of basis vectors of $\mathbb{Q}[\vec{v}]_{\leq 1}$. We express the 4 polynomials $y_{\vec{e}}$ over this basis, and obtain coefficient vectors $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(-1,0,2)$.

Now, the polyhedron given by

$$\begin{array}{rrrcr} x_0 & & -x_3 & = & -2 \\ & x_1 & & = & 1 \\ & & x_2 \quad +2x_3 & = & 3 \\ & & x_0,x_1,x_2,x_3 & \geq & 0 \end{array}$$

has no integer points. Hence $\varphi+I$ is binomial-bad by the Polyhedron Theorem 5.5.2.[21]

From the Diagonalization Theorem 6.2.1 it follows that for every Turing machine M, we have $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$, and hence $p_A \notin \#\mathsf{P}^A$. It remains to show that

$$p_A \in \#\mathsf{CLS}(\textsc{EitherSolution}(\textsc{SourceOrExcess}(2{,}1),\textsc{Iter}))^A - 1.$$

Let $C_\perp$ be an input to $\textsc{Multi}(\textsc{Iter},\textsc{Excess},\textsc{NZSource})$ (i.e., a circuit, see footnote 10) that yields output $\vec{t}$. Let ℓ be the number of inputs of the circuit $C_\perp$. Let $\nu_j(w) =$ True if and only if the string w has only zeros at positions $j+1, j+2, \ldots$ Finally, let C_j be an input to $\textsc{Multi}(\textsc{Iter},\textsc{Excess},\textsc{NZSource})^A$ that describes the $(j-1)$-input circuit that consists of a single arity j oracle gate that takes a constant 1 into its first input, takes all inputs into its remaining $j-1$ inputs, and forwards its output directly to the circuit output.

We construct a circuit $\alpha(x) := D_{|x|}$ with $\max\{\ell,|x|\}$ many inputs and with oracle gates as follows ($w \in \{0,1\}^{\max\{\ell,|x|\}}$):

$$\begin{array}{rcl} D_j(w) & := & \Big((A(0^{|w|+1})=0) \text{ and } \nu_\ell(w) \text{ and } C_\perp(w_1,\ldots,w_\ell)\Big) \\ & \text{and} & \Big((A(0^{|w|+1})=1) \text{ and } \nu_{|x|}(w) \text{ and } C_{|x|+1}(w_1,\ldots,w_{|x|})\Big) \end{array}$$

We define the polynomially balanced relation R via $R(w,y)$ if and only if

$$\textsc{rEitherSolution}(\textsc{SourceOrExcess}(2{,}1),\textsc{Iter})(\alpha(w),y).$$

We set β to be the identity function. By definition, $R \in \mathsf{rCLS}$. The following function is thus in $\#\mathsf{CLS}$:

$$\sum_{y\in\{0,1\}^*} R(x,y) = \sum_{y\in\{0,1\}^*} \textsc{rEitherSolution}(\textsc{SourceOrExcess}(2{,}1),\textsc{Iter})(\alpha(x),y)$$

which equals (because the set-instantiator does not create instances that use the last option)

$$\begin{cases} \textsc{Iter}(D_{|x|}) + \textsc{Excess}(D_{|x|}) + \textsc{NZSource}(D_{|x|}) & \text{if } A(0^{|x|+1})=1 \\ t_1+t_2+t_3 & \text{otherwise} \end{cases}$$

which equals $p_A(w)+1$. This finishes the proof. □

[21] In fact, even the LP relaxation is empty in this case, so every positive integer multiple of the problem $\textsc{EitherSolution}(\textsc{SourceOrExcess}(2{,}1),\textsc{Iter})$ can also be separated from $\#\mathsf{P}$ via an oracle. This is not always the case, and we use this property to separate two classes that are outside of $\#\mathsf{P}$ (w.r.t. an oracle), see Proposition 8.9.2.

8.4 A set-instantiator for #PLS(ITER)

The set-instantiator in this section is the most complicated we create. The problem is that instances cannot be permuted arbitrarily, so one has to view $[2^j]$ as a $[2^{j/2}] \times [2^{j/2}]$ square, and permute within the rows. But that does not immediately give a way to mask from the #P machine which presink belongs to the zero source, a property which crucially must be achieved.

8.4.1 Theorem. *For every $b \in \mathbb{N}$, there exists a threshold $j_0 \in \mathbb{N}$, such that for every $j \geq j_0$, every $A_{<j} \in \{0,1\}^{<j}$ and every polynomial time Turing machine M that answers consistently for $(j, A_{<j}, \text{ITER})$, there exists a set-instantiator $\mathcal{SI}$ against $(M, j, A_{<j}, S = \mathbb{N}_{\geq 1}, b)$, such that $\text{ITER}(\mathsf{inst}_{\mathcal{SI}}(s)) = |s|$ for all s with $|s| \in S$.*

Intuitively, this says that for all our purposes #PLS behaves exactly as $\#\mathsf{P}_{\geq 1}$. The rest of this section is devoted to proving this theorem. We will first define a *creator* whose *creations* will be the instantiations in the end, but the creator is not limited to a single creation for each set. We will then define the set-instantiator from the creator by picking for every subset $\vec{s}$ just any one of its creations for $\vec{s}$.

8.4 (a) Creations

Having 2^{j-1} many bits available, we can encode in a standard way a function $[2^{j'}] \to [2^{j'}]$, where j' and j are polynomially related. We ignore any extra bits that are not needed for this encoding. We think of $[2^{j'}]$ arranged as a square, so set $n := \lfloor \log_2(\sqrt{2^{j'}}) \rfloor$ and interpret 2^n as the number of rows, i.e., assume we have a polynomial time computable injective map $[2^n] \times [2^n] \to [2^{j'}]$. We ignore numbers in $[2^{j'}]$ that are outside this square (i.e., outside the image of this injective map). We sometimes call elements of $[2^n] \times [2^n]$ nodes, as is usual for ITER instances.

The function ψ is given by the j-th layer A_j of the oracle j, so can say that a computation "queries ψ" when we mean it queries A_j. We will use these interchangeably. For every $x = (x_1, x_2) \in [2^n] \times [2^n]$, let $\mathsf{row}(x) = x_1$. Let $\mathsf{toprows} := [1, 2^{n-1}]$ (the upper half of the rows), $\mathsf{presinkrow} := 2^n - 1$ (one from the very bottom), and let $\mathsf{sinkrow} := 2^n$ (at the very bottom). All our instances will have all sinks in row $\mathsf{sinkrow}$ and all presinks in row $\mathsf{presinkrow}$.

A map $\psi : [2^n] \times [2^n] \to [2^n] \times [2^n]$ is called *row-layered* if $\psi(x) > x$ implies that $\mathsf{row}(\psi(x)) = \mathsf{row}(x) + 1$. A point $x \in [2^n] \times [2^n]$ with $\psi(x) = x$ is called a *fixed point*. A point $x \in [2^n] \times [2^n]$ for which $\exists y \in [2^n] \times [2^n] \setminus \{x\} : \psi(y) = x$ is called a *sink*. A fixed point that is not a sink is called a *singleton*. A point $y \in [2^n] \times [2^n]$ for which $\psi(y)$ is a sink is called a *presink*. Note that when given oracle access to ψ it is easy to check if a specific y is a presink, but it is not easy to check if a specific x is a sink (because it is hard to distinguish it from a singleton).

A *path part* in ψ is a set of vertices $v_1, \ldots, v_q$ such that $\psi(v_i) = v_{i+1}$, $\psi(v_q) \neq v_q$ and $\psi(\psi(v_q)) = \psi(v_q)$. The node v_1 is called the *source*. By definition, the node v_q is a presink and each path part has exactly one presink. $\psi(v_q)$ is called the *sink*. We consider the presink to be part of the path part, while we consider the sink to *not* be part of the path part. It is clear that a single vertex with a loop is not a path part. A *path* is a path part that is maximal with respect to inclusion.

A map $\psi : [2^n] \times [2^n] \to [2^n] \times [2^n]$ is called a *set-of-paths* if every node x has no predecessor or exactly one predecessor in ψ, i.e., we partition $[2^n] \times [2^n]$ into a disjoint union of paths and sinks and singletons. A set-of-paths is called *rooted* if one of its sources is at position $(1,1)$. This path is called the *main path*. All other paths are called non-main paths.

For $a \geq 1$, a rooted set-of-paths is called an *a-creation* if it consists of exactly a many paths, and the rows of the sources of all non-main paths are in row $\mathsf{presinkrow}$, and all paths have their

sink in row sinkrow in the same column as their presink (in particular, those paths consist only of a single edge). Let a-creations denote the set of all a-creations. Finally, let a^--creations $:= \bigcup_{1\le c\le a} c$-creations.

8.4 (b) The unaccessed row η

Fix $b \in \mathbb{N}$. From this point on, η will depend on b, but b will remain constant. Let $t_i(j)$ be a polynomial in j, defined as the upper bound on the number of computation steps that M makes on inputs of length j (this is independent of to which oracle M has access). Since M answers consistently, for every $\psi \in b^-$-creations with $\textsc{Iter}(\psi) = a$, there exist exactly $\varphi(a)$ many accepting computation paths for the computation $h_i^\psi(0^j)$. Each of these paths queries the oracle ψ at most $\varphi(a)t_i(j)$ many positions. Thus, together they access at most an $\frac{\varphi(a)t_i(j)}{2^{j/2}}$ fraction of toprows.

From above, for a uniformly random $\eta \in$ toprows and any $\psi \in b^-$-creations, we have w.h.p. that no accepting computation path accesses row η. This holds for every $\psi \in b^-$-creations, so it also holds when sampling ψ from any distribution E on b^--creations: When sampling (η, ψ) from $U_{\text{toprows}} \times E$, then w.h.p. none of the accepting paths of $h_i^\psi(0^j)$ accesses row η. This can be seen, for example, by first sampling $\psi \in E$, and then sampling $\eta \in U_{\text{toprows}}$ independently. Therefore, using the union bound, for every finite set of distributions $E_1, \dots, E_b$ on b^--creations, there exists $\eta_E \in$ toprows, such that for all $1 \le a \le b$ we have:

$$\begin{aligned} &\text{when sampling } \psi \text{ from } E_a, \text{ we have w.h.p. that} \\ &\text{no accepting path of } h_i^\psi(0^j) \text{ accesses row } \eta_E. \end{aligned} \tag{8.4.2}$$

For the rest of this proof, we fix

$$\eta = \eta_{(U_{1\text{-creations}}, \dots, U_{b\text{-creations}})}.$$

We remark that it is an important technical difficulty when constructing set-instantiators that the sampling process of the b^--creations must be independent of η.

8.4 (c) Lucky creators

In this section we introduce the concept of a *creator*. A creator is a slightly less restrictive version of a set-instantiator. For each blueprint a creator outputs a creation, where a blueprint is slightly more general than the set $\vec{s}$ for a set-instantiator, but serves the same purpose.

A layered path part with source $(1,1)$ to a presink in row $\eta - 1$ is called a *head*. A layered path part with source row = presink row = η is called a *fork* (just one edge). A layered path part with source row $\eta+1$ to a presink in row presinkrow -1 is called a *tail*. A layered path part x with source row = presink row = presinkrow, where column(x) = column$(\psi(x))$ is called a *sink part* (just one edge). Recall that by definition an a-creation consists of 1 head, 1 fork, 1 tail, a many sink parts, and is uniquely defined by these. If the sink of a path part equals the source of another path part, then their union is another path part. For every tail there is a unique way to take a union with a sink part, which is by adding an edge that stays in the same column.

For $b \ge 1$ a *b-creator* ξ is a 3-tuple $(\text{head}(\xi), \text{tails}(\xi), \text{presinks}(\xi))$, where $\text{head}(\xi)$ is a head, $\text{tails}(\xi)$ is a cardinality b ordered set of tails whose nodes are pairwise disjoint, and $\text{presinks}(\xi)$ is a cardinality b ordered set of sink parts, so that the tails line up properly with the corresponding sink parts: adding an edge that stays in the same column to a tail is the same as taking the union with the sink part corresponding to the tail. Let b-creators denote the set of all b-creators. Instead of writing "presink of ξ" we sometimes write ξ-presink. Analogously for tails.

Given a b-creator ξ and $o' \in s' \subseteq [b]$, we define the $|s'|$-creation $\xi_{s',o'}$ via removing all tails but keeping the tail that connects to the presink $\mathsf{presinks}(\xi)_{o'}$, and removing all presinks outside of $\{\mathsf{presinks}(\xi)_a \mid a \in s'\}$, defining the function ϕ that has exactly those path parts, and then setting $\phi(x) = y$ for the fork x and the source of the tail y. Clearly $\xi_{s',o'}$ is an $|s'|$-creation.

Given a b-creator ξ, define $\xi_{o'} := \xi_{[b],o'}$. Clearly $\xi_{o'}$ is a b-creation.

8.4.3 Definition. *We call $\xi \in b$-$\mathsf{creators}$ lucky if for all subsets $L \subseteq [b]$, $|L| \geq 1$, and all $o \in L$, all accepting paths of the computation $h_i^{\xi_{L,o}}(0^j)$*

(i) do not access the oracle in row η,

(ii) do not access the oracle in any ξ-presink besides the presinks that correspond to indices in L (they do not have to access all *L-presinks), and*

(iii) do not access the oracle in any tails besides the tail tails_o.

Informally, this says that the accepting paths do not access the oracle at positions where presinks and tails are not, but could potentially be.

8.4.4 Claim. *Let ξ be sampled from $U_{b\text{-}\mathsf{creators}}$. Then ξ is lucky w.h.p.*

Proof. We describe a way of sampling from $U_{b\text{-}\mathsf{creators}}$. Let heads be the set of all heads, let tails be the set of all tails, and let

$$\mathsf{tail}\text{-}b\text{-}\mathsf{tuples} = \left\{ \vec{q} \in \mathsf{tails}^b \;:\; \text{all } q_1, \ldots, q_b \text{ have pairwise disjoint nodes and sinks} \right\}.$$

We first sample $(h, \vec{q})$ from $U_{\mathsf{heads}} \times U_{\mathsf{tail}\text{-}b\text{-}\mathsf{tuples}}$. Choosing the matching presinks, this process samples a random variable $\mathsf{crea}(h, \vec{q}) \in b$-$\mathsf{creators}$ according to $U_{b\text{-}\mathsf{creators}}$. We observe that for fixed $L \subseteq [b]$ and $o \in L$, and $\xi := \mathsf{crea}(h, \vec{q})$, the $|L|$-creation $\xi_{L,o}$ is uniformly distributed from $|L|$-$\mathsf{creations}$.

We show that for a fixed $L \subseteq [b]$, $|L| \geq 1$, and $o \in L$ we have that $(h, \vec{q})_{L,o}$ satisfies all three properties of Definition 8.4.3 with high probability. Since there are only $2^b - 1$ many possible sets L, since there are only b many possible values for o, and since b is fixed, the claim then immediately follows from the union bound. Hence, for the rest of this proof we fix $L \subseteq [b]$ and $o \in L$. In this fixed case we will also use the union bound.

Property (i). Since $(h, \vec{q})_{L,o}$ is distributed as $U_{|L|\text{-}\mathsf{creations}}$, (8.4.2) proves (i) in Definition 8.4.3 with high probability.

Properties (ii) and (iii). For a fixed $\psi \in a$-$\mathsf{creations}$, if we draw $(h, \vec{q})$ from $U_{\mathsf{heads}} \times U_{\mathsf{tail}\text{-}b\text{-}\mathsf{tuples}}$ under the assumption that $\mathsf{crea}(h, \vec{q})_{L,o} = \psi$, then we see that the tails q_l, $l \notin L$, are *uniformly* distributed (and hence this also holds for the corresponding presinks). Moreover, by definition, each (tail,presink) pair only contains at most 1 node in each row. Hence each oracle query accesses such a tail or presink with probability $< \frac{b}{2^n - t_i(j)} \leq \frac{b}{2^{n-1}}$. Here $t_i(j)$ is subtracted in the denominator, because that many positions could have been queried already and there is no reasone to query the oracle twice at the same position). By Bernoulli's inequality we have

$$\left(1 - \frac{b}{2^{n-1}}\right)^{t_i(j)} \;\geq\; 1 - \frac{b \cdot t_i(j)}{2^{n-1}}.$$

This proves the property (ii) and (iii).

Since each property (i)–(iii) is satisfied with high probability, the union bound gives that they are simultaneously satisfied with high probability. This finishes the argument for fixed L and o. According to the previous discussion about the union bound the overall claim is proved. □

The following lemma handles the technical difficulty of the tails.

8.4.5 Lemma. *Let ξ be lucky, let $b \geq 1$, and suppose τ is an accepting path of $h_i^{\xi_{s,o}}(0^j)$ for every $o \in s \subseteq [b]$. Then τ does not query any of the b many tails of ξ.*

Proof. Let there be c many accepting paths of the computation $h_i^{\xi_{s,o}}(0^j)$, let τ be one of those accepting paths, and w.l.o.g. assume that τ queries the tail q_1. Let ψ be obtained from the instance $\xi_{s,o}$ by removing tail q_1 and adding tail q_2. Clearly ψ is also a c-creation. Hence, since ξ is lucky, and since τ queries q_1, we conclude that τ does not query q_2 (Definition 8.4.3).

Since M answers consistently, there are c many accepting paths for the computation $h_i^{\psi}(0^j)$. Since ψ is a c-creation and ξ is lucky, these accepting paths do not query q_1. Therefore, if we add q_1 into the instance (so that now both q_1 and q_2 are present), all these a many paths are still accepting paths, and τ is also accepting (and is obviously different from all these a many paths). Thus there are now at least $a+1$ many accepting paths for this instance, which must have (because M answers consistently) exactly a many accepting paths. This is a contradiction. □

8.4 (d) Defining the set-instantiator

For an a-creation ψ and a computation path $\tau \in \{0,1\}^*$ of a computation $h_i^{\psi}(0^j)$, let $\mathsf{perception}(\tau) \subseteq \mathsf{presinks}(\psi)$ denote the set of oracle positions with presinks that get accessed by τ.

Let ξ be a lucky b-creator. We interpret the list $\mathsf{presinks}(\xi)$ as a bijection $\mathsf{bij} : [b] \to \mathsf{presinks}(\xi)$. Let $S = \mathbb{N}_{\geq 1}$. For every $\vec{s} \in \mathscr{B}(\vec{b})_S$, we choose a $o_{\vec{s}} \in \mathsf{presinks}(\xi_{\mathsf{bij}(\vec{s})})$. Set

$$\mathsf{inst}_{\mathcal{SI}}(\vec{s}) \;:=\; \xi_{\mathsf{bij}(\vec{s}),o_{\vec{s}}}$$

and for $\tau \in \{0,1\}^*$ set

$$\mathsf{perc}_{\mathcal{SI}}(\tau) \;:=\; \begin{cases} \mathsf{bij}^{-1}(\mathsf{perception}(\tau)) & \text{if } \tau \text{ is an accepting path of the computation } h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j) \\ \top & \text{otherwise.} \end{cases}$$

The rest of this section is devoted to proving that $\mathcal{SI}$ satisfies the requirements of Definition 6.1.3, which then proves Theorem 8.4.1, because clearly $\textsc{Iter}(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$. Formally, we prove the following result.

8.4.6 Proposition. *For all $\vec{s} \in \mathscr{B}(\vec{b})_S$ we have: $\tau \in \{0,1\}^*$ is an accepting path for the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ if and only if $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.*

Proof. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$, since ξ is lucky, and since τ does not query any tail (Lemma 8.4.5), we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. This implies that $\mathsf{perc}_{\mathcal{SI}}(\tau) = \mathsf{bij}^{-1}(\mathsf{perception}(\tau))$. Clearly $\mathsf{perception}(\tau) \subseteq \mathsf{bij}(\vec{s})$, because otherwise τ would not even be a computational path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ because of its oracle answers when querying presinks in $\mathsf{perception}(\tau) \setminus \mathsf{bij}(\vec{s})$. We conclude: $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.

The argument is reversible: let $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$, so in particular $\mathsf{perc}_{\mathcal{SI}}(\tau) \neq \top$. Then τ is an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$, since ξ is lucky, and since τ does not query any tail (Lemma 8.4.5), we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$. □

8.5 A set-instantiator for #PPAD(SourceOrExcess(2,1))

Let $S_{\text{line}} := \{(f_1, f_2) : 2f_2 - f_1 - 1 = 0\}$. Think of f_1 as the number of sources, and f_2 as the number of double sinks. Let Multi(NZSource, Excess) be the bivariate counting problem of counting sources and counting excess nodes in a SourceOrExcess(2,1) instance. This set $S = S_{\text{line}}$ will be sufficient for our set-instantiator.[22]

8.5.1 Theorem. *Fix a polynomial time nondeterministic Turing machine M. Then, for every $\vec{b} \in \mathbb{N}^2$, there exists a threshold $j_0 \in \mathbb{N}$, such that for every $j \geq j_0$ and every $A_{<j} \in \{0,1\}^{<j}$, there exists a set-instantiator $\mathcal{SI}$ against $(M, j, A_{<j}, S = S_{\text{line}}, \vec{b})$, which satisfies:*

$$\text{Multi}(\text{NZSource}, \text{Excess})(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}| \quad \textit{for all} \quad \vec{s} \ \textit{with} \ |\vec{s}| \in S.$$

The rest of this section is devoted to proving this theorem. We will first define a *creator* whose *creations* will be the instantiations in the end, but the creator is not limited to a single creation for each set. We will then define the set-instantiator from the creator, by picking for every subset $\vec{s}$ with $|\vec{s}| \in S$, just any one of its creations for $\vec{s}$.

8.5 (a) Creations

Having 2^{j-1} many bits available, we can encode in a standard way a tuple of two predecessor functions $\psi_{\text{pred}_1} : [2^n] \to [2^n]$, $\psi_{\text{pred}_2} : [2^n] \to [2^n]$, and a successor function $\psi_{\text{succ}} : [2^n] \to [2^n]$, such that n and j are polynomially related. We ignore any extra bits that are not needed for this encoding.

We interpret $\psi = (\psi_{\text{pred}_1}, \psi_{\text{pred}_2}, \psi_{\text{succ}})$ as a directed graph (with some additional information, because the map from the set of function triples to the set of digraphs is not injective). Here an edge from $x \in [2^n]$ to $y \in [2^n]$ is present if and only if $\psi_{\text{succ}}(x) = y$, and either $\psi_{\text{pred}_1}(y) = x$ or $\psi_{\text{pred}_2}(y) = x$. Each node in the resulting graph has indegree ≤ 2 and outdegree ≤ 1.

A node whose indegree exceeds its outdegree is called an *excessive node*. A node with indegree 0 and outdegree 1 is called a *source*. A node with indegree 2 and outdegree 0 is called a *double sink*. We assume that it the encoding is made in a way that the 0 node is always at source. A source that is not the 0 node is called a *nonzero source*.

Let $\mathsf{nzsource}(\psi)$ denote the set of nonzero sources. Let $\mathsf{dsink}(\psi)$ denote the set of double sinks. For $\vec{a} \in S$, ψ is called an *$\vec{a}$-creation* if $|\mathsf{nzsource}(\psi)| = a_1$ and $|\mathsf{dsink}(\psi)| = a_1$. Let a-creations denote the set of all a-creations. Let a^--creations $:= \bigcup_{1 \leq c \leq a} c$-creations.

8.5 (b) Lucky creators

In this section we introduce the concept of a *creator*. A creator is a slightly less restrictive version of a set-instantiator. For each blueprint a creator outputs a creation, where a blueprint is slightly more general than the set $\vec{s}$ for a set-instantiator, but serves the same purpose.

For two disjoint sets X and Y, let $\binom{X}{2}_{\text{ordered}} \times Y$ denote the set of triples, where the first and second element are distinct elements from X, and the third element is from Y. We say that two triples are *disjoint* if they share no common element. Let $T(X, Y, c)$ denote the set c-tuples of pairwise disjoint triples.

For $\vec{b} \in S$, a *$\vec{b}$-creator* ξ is a triple consisting of:

1. a length b_1 ordered list $\mathsf{nzsources}(\xi) \subseteq [2^n]$,

[22] Larger S are possible, but not necessary for our purposes, because already we have that $f_1 + f_2 - 1 + \langle 2f_2 - f_1 - 1 \rangle$ is binomial-bad, as can be seen via the Polyhedron Theorem 5.5.2.

2. a length b_2 ordered list $\mathsf{dsinks}(\xi) \subseteq [2^n]$ of even numbers, and

3. a map $\mathsf{straight}(\xi) : [2^n] \to [2^n]$ such that all nodes have indegree at most 2, the 0 node is a source, the set of all other sources is exactly $\mathsf{nzsources}(\xi)$, the set of all double sinks is exactly $2\mathsf{dsinks}(\xi)$, the set of all loops is exactly $2\mathsf{dsinks}(\xi)+1$, all other nodes hat indegree = outdegree = 1, $\mathsf{straight}(\xi)(x) \neq x+1$ for x even, $\mathsf{straight}(\xi)(x+1) \neq x$ for x even, and $\mathrm{pred}_2(x) = x$ for all nodes that are not double sinks.

This means that each double sink comes with a loop at the next position, and no node maps to its paired neighbor.

Given a $\vec{b}$-creator ξ and an $\vec{L} \in T\big([b_1]\times\{1\},[b_2]\times\{2\},c\big)$, we obtain a (b_2-c)-creation $\psi : [2^n] \to [2^n]$ by performing the following step for each triple $(x_1,x_2,y) \in \vec{L}$: set

$$\psi_{\mathrm{succ}}\big(\mathsf{dsinks}(\xi)_y\big) := \mathsf{nzsources}(\xi)_{x_1}\,, \quad \psi_{\mathrm{succ}}\big(\mathrm{pred}_1(\mathsf{dsinks}(\xi)_y)\big) := \mathsf{dsinks}(\xi)_y + 1,$$
$$\text{and} \quad \psi_{\mathrm{succ}}\big(\mathsf{dsinks}(\xi)_y + 1\big) := \mathsf{nzsources}(\xi)_{x_2}\,.$$

In other words, remove the double sink and the loop and redirect the two paths to the two sources.

8.5.2 Definition. *We call $\xi \in b$-creators lucky if for all c and for all $\vec{L} \in T\big([b_1]\times\{1\},[b_2]\times\{2\},c\big)$, all accepting paths of the computation $h_i^{\xi_{\vec{L}}}(0^j)$ do not access the oracle at any point in $\mathsf{nzsources}(\xi)\setminus\mathsf{nzsource}(\xi_{\vec{L}})$, nor any point in $\mathsf{dsinks}(\xi)\setminus\mathsf{dsink}(\xi_{\vec{L}})$, nor any point in $(\mathsf{dsinks}(\xi)+1)\setminus(\mathsf{dsink}(\xi_{\vec{L}})+1)$ or any directly adjacent nodes.*

This says that the accepting paths do not access the oracle at positions where sources or double sinks (or their loops) are not, but could potentially be.

8.5.3 Claim. *Let ξ be sampled from $U_{b\text{-creators}}$. Then ξ is lucky w.h.p.*

Proof. It is crucial to observe that for a fixed $\vec{L} \in T\big([b_1]\times\{1\},[b_2]\times\{2\},c\big)$, we have that the $\vec{L}$-creation $\xi_{\vec{L}}$ is uniformly distributed from (L_2-c)-creations. We show that for a fixed $\vec{L} \in T\big([b_1]\times\{1\},[b_2]\times\{2\},c\big)$, we have ξ is lucky w.h.p. Since there are only constantly many $\vec{L}$, the claim immediately follows from the union bound. Thus, for the rest of the proof, we fix $\vec{L}$.

We have that $\xi_{\vec{L}}$ is uniformly distributed. The, the probability of an oracle access picking one of the forbidden positions is $\leq \frac{8(L_1+L_2)}{2^{n-1}} \leq \frac{1}{2^{n-2}}$. By Bernoulli's inequality we have $(1-\frac{1}{2^{n-2}})^{t_i(j)} \geq 1-\frac{t_i(j)}{2^{n-2}}$. Since n and j are polynomially related, this proves the claim. □

8.5 (c) Defining the set-instantiator

For a $\vec{a}$-creation ψ and a computation path $\tau \in \{0,1\}^*$ of a computation $h_i^{\psi}(0^j)$, let $\mathsf{perception}_1(\tau) \subseteq \mathsf{nzsource}(\psi)$ denote the set of accessed oracle positions that are nonzero sources or adjacent. Let $\mathsf{perception}_2(\tau) \subseteq \mathsf{dsink}(\psi)$ denote the set of accessed oracle positions that are double sinks or adjacent (if the loop is accessed, this counts as an access of the double sink).

Let ξ be a lucky $\vec{b}$-creator. We interpret the list $\mathsf{nzsources}(\xi)$ as a bijection $\mathsf{bij}_1 : \mathscr{B}(b_1) \to \mathsf{nzsources}(\xi)$. Similarly, we interpret the list $\mathsf{dsinks}(\xi)$ as a bijection $\mathsf{bij}_2 : \mathscr{B}(b_2) \to \mathsf{dsinks}(\xi)$. Let $S = S_{\mathrm{line}}$. We set

$$\mathsf{inst}_{\mathcal{SI}}(\vec{s}) := \xi_{\vec{s}}\,.$$

Finally, for $\tau \in \{0,1\}^*$ we set

$$\mathsf{perc}_{\mathcal{SI}}(\tau) := \big(\mathsf{bij}_1^{-1}(\mathsf{perception}_1(\tau)), \mathsf{bij}_2^{-1}(\mathsf{perception}_2(\tau))\big)$$

if τ is an accepting path for the computation $h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$, and $\mathsf{perc}_{\mathcal{SI}}(\tau) := \top$, otherwise.

We can now prove that $\mathcal{SI}$ satisfies the requirements of Definition 6.1.3, which then proves Theorem 8.5.1, because clearly $\textsc{Multi}(\textsc{Source},\textsc{Excess})(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$. Formally, we have:

8.5.4 Proposition. *For all $\vec{s} \in \mathscr{B}(\vec{b})_S$, we have: $\tau \in \{0,1\}^*$ is an accepting path for the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ if and only if $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.*

Proof. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. This implies that $\mathsf{perc}_{\mathcal{SI}}(\tau) = \mathsf{bij}^{-1}(\mathsf{perception}(\tau))$. Clearly $\mathsf{perception}(\tau) \subseteq \mathsf{bij}(\vec{s})$, since otherwise τ would not even be a computational path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ because of its oracle answers when querying lonely nodes in $\mathsf{perception}(\tau) \setminus \mathsf{bij}(\vec{s})$. We conclude that $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.

The argument is reversible: let $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$, in particular $\mathsf{perc}_{\mathcal{SI}}(\tau) \neq \top$. Then τ is an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$. □

8.6 Combining set-instantiators for handling #CLS − 1

Let $S_{\mathrm{CLS}} := \left\{ \vec{f} \in \mathbb{N}^3 \,:\, 2f_2 = f_1 + 1,\ f_2 \geq 1,\ f_3 \geq 1 \right\}$.

8.6.1 Theorem. *Given a polynomial time nondeterministic Turing machine M. For every $\vec{b} \in \mathbb{N}^3$ there exists a threshold $j_0 \in \mathbb{N}$ such that for every $j \geq j_0$ and every $A_{<j} \in \{0,1\}^{<j}$ there exists a set-instantiator $\mathcal{SI}$ against $(M, j, A_{<j}, S = S_{\mathrm{CLS}}, \vec{b})$ such that $\forall \vec{s}$ with $|\vec{s}| \in S$: $\textsc{Multi}(\textsc{NZSource},\textsc{Excess},\textsc{Iter})(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$.*

Proof. The construction is basically a combination of the set-instantiator for $\textsc{Multi}(\textsc{NZSource},\textsc{Excess})$, and the set-instantiator for $\textsc{Iter}$. We can ignore the subtlety of the interaction between both problems that was introduced to be able to have instances of value 1, and for this set-instantiator we just care about instances that have value at least 2, which is reflected in the set S by having $f_2 \geq 1$ and $f_3 \geq 1$. This is similar to choosing S_{line} in §8.5: it already gives instances that are difficult enough.

We create a set-instantiator that creates pairs of instances that are glued together in a non-sophisticated way by having the $\textsc{Multi}(\textsc{NZSource},\textsc{Excess})$ instance in the first half and the $\textsc{Iter}$ instance in the second half. Both set-instantiators were obtained by proving that a uniformly random creator is lucky with high probability, because it just has to satisfy a finite number of constraints. We put both sets of constraints together and readily obtain a creator with a finite number of constraints that also with high probability is lucky. As usual, the set-instantiator is then obtained by picking any single instance from the creator. □

8.7 The halving separation

In this section we prove the following result:

8.7.1 Theorem. *We have $\#\mathsf{P} \subseteq \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})/2$ via a parsimonious relativizing reduction. Moreover, there exists an oracle A with respect to which $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2 \not\subseteq \#\mathsf{P}^A$.*

Proof. The inclusion $\#\mathsf{P}^A \subseteq \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2$ is of the same level of difficulty as the considerations in §8.2. Given a CircuitSat instance we construct a AllLeaves instance of as follows: Wherever the circuit yields True, we add two vertices that are connected by an edge. We add no other vertices. We obtain an instance of value $2f$. We are done, because $(2f)/2 = f$.

For the other direction, we use the Diagonalization Theorem 6.2.1 and the set-instantiator for AllLonely (see Theorem 8.8.1), with

$$k = 1\,, \quad S = 2\mathbb{N}.$$

Set $\ell = k = 1$ and have no functions ζ.

We can now verify the preconditions of the Diagonalization Theorem 6.2.1:

1. Z contains the integer point 0,
2. $C'_S = \mathbb{Q}_{\geq 0}$ lies Zariski-dense in $\mathbb{Q}$, and
3. the last point is a vacuous truth.

We fix Multiplicities = AllLonely. The necessary set-instantiators are given by Theorem 8.8.1. We set $\varphi(f_1) = f_1/2$. Clearly $\varphi + I = f_1/2 + \langle 0 \rangle$ is binomial-bad. From the Diagonalization Theorem 6.2.1, it follows that for every M we have $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$. Hence $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Lonely})^A \neq \#\mathsf{P}^A$. The statement follows from the relativizing parsimonious equivalence between Lonely and Leaf, once we observe that $p_A \in \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Lonely})^A/2$.

Let $C_\perp$ be an input to AllLonely (i.e., a circuit[23]) that yields output $\vec{t}$. Let ℓ be the number of inputs of the circuit $C_\perp$. Let $\nu_j(w) =$ True if and only if the string w has only zeros at positions $j+1, j+2, \ldots$ Finally, let C_j be an input to $\textsc{AllLonely}^A$ that describes the $(j-1)$-input circuit that consists of a single arity j oracle gate that takes a constant 1 into its first input, takes all inputs into its remaining $(j-1)$ inputs, and forwards its output directly to the circuit output.

We construct a circuit $\alpha(x) := D_{|x|}$ with $\max\{\ell, |x|\}$ many inputs and with oracle gates defined as follows $\left(w \in \{0,1\}^{\max\{\ell,|x|\}}\right)$:

$$\begin{aligned} D_j(w) \quad &:= \quad \Big((A(0^{|w|+1}) = 0) \text{ and } \nu_\ell(w) \text{ and } C_\perp(w_1, \ldots, w_\ell)\Big) \\ &\text{and} \quad \Big((A(0^{|w|+1}) = 1) \text{ and } \nu_{|x|}(w) \text{ and } C_{|x|+1}(w_1, \ldots, w_{|x|})\Big) \end{aligned}$$

We define the polynomially balanced relation R via $R(w, y)$ if and only if $\textsc{rAllLonely}(\alpha(w), y)$. We set β to be the identity function. By definition, $R \in \mathsf{rCOUNTALL\text{-}PPA}$. The following function is thus in $\#\mathsf{COUNTALL\text{-}PPA}$:

$$\sum_{y \in \{0,1\}^*} R(x, y) \;=\; \sum_{y \in \{0,1\}^*} \textsc{rAllLonely}(\alpha(x), y)$$

which equals

$$\begin{cases} \textsc{AllLonely}(D_{|x|}) & \text{if } \; A(0^{|x|+1}) = 1, \\ t & \text{otherwise.} \end{cases}$$

This function equals $2p_A(w)$, which finishes the proof. □

[23]Note that in the case of Lonely it is naturally a single circuit, not a list of circuits, cf. footnote 10.

8.8 A set-instantiator for #COUNTALL-PPA(LEAF)

8.8.1 Theorem. *Let M be a polynomial time nondeterministic Turing machine. For every $b \in \mathbb{N}$, there exists a threshold $j_0 \in \mathbb{N}$, such that for every $j \geq j_0$ and every $A_{<j} \in \{0,1\}^{<j}$, there exists a set-instantiator $\mathcal{SI}$ against $\big(M, j, A_{<j}, S = 2\mathbb{N}, b\big)$ such that $\textsc{AllLonely}(\mathsf{inst}_{\mathcal{SI}}(s)) = |s|$ for all s with $|s| \in S$.*

Intuitively, this says that for all our purposes ALLLONELY behaves exactly as $\#\mathsf{P}_{\text{even}}$. The rest of this section is devoted to proving this theorem. We will first define a *creator* whose *creations* will be the instantiations in the end, but the creator is not limited to a single creation for each set. We will then define the set-instantiator from the creator by picking for every subset $\vec{s}$ with $|\vec{s}| \in 2\mathbb{N}$ just any one of its creations for $\vec{s}$.

We crucially use the parsimonious polynomial-time reductions from ALLLONELY to ALLLEAVES and back, so we can focus on ALLLONELY here. The reductions are gadget-based (with extremely simple gadgets), but ALLLONELY fits perfectly for our proof technique, while it is slightly more trouble to work with ALLLEAVES.

8.8 (a) Creations

Having 2^{j-1} many bits available, we can encode in a standard way a function $\psi : [2^n] \to [2^n]$, where n and j are polynomial related. We ignore any extra bits that are not needed for this encoding. A *paired node* in ψ is an $x \in [2^n]$ for which $y \in [2^n] \setminus \{x\}$ exists such that $\psi(x) = y$ and $\psi(y) = x$. A *lonely node* is a node that is not paired. Since 2^n is even, clearly the number of lonely nodes is always even.

For $a \in 2\mathbb{N}$, a function ψ is called an *a-creation* if it consists of exactly a many lonely nodes, and for each lonely node x we have $\psi(x) = x$. Let a-$\mathsf{creations}$ denote the set of all a-creations. Let a^--$\mathsf{creations} := \bigcup_{1 \leq c \leq a} c$-$\mathsf{creations}$. Finally, let $\mathsf{lonely}(\psi) \subseteq [2^n]$ denote the set of all lonely nodes.

8.8 (b) Lucky creators

In this section we introduce the concept of a *creator*. A creator is a slightly less restrictive version of a set-instantiator. For each blueprint a creator outputs a creation, where a blueprint is slightly more general than the set $\vec{s}$ for a set-instantiator, but serves the same purpose.

For a set X, let $\binom{X}{2,\dots,2}$ denote the set of all set partitions of X into pairs. We call these elements *pairings*. For $b \in 2\mathbb{N}$, a *b-creator* ξ is a finite ordered set $\mathsf{lonelies}(\xi)$ of b many distinct elements of $[2^n]$ together with a pairing $\mathsf{pairing}(\xi) \in \binom{[2^n] \setminus \mathsf{lonelies}(\xi)}{2,\dots,2}$. Let b-$\mathsf{creators}$ denote the set of all b-creators.

Given a b-creator ξ, a subset $L \subseteq [b]$ with $|L|$ even, and a pairing $o \in \binom{L}{2,\dots,2}$, we obtain a $(b - |L|)$-creation $\xi_{L,o}$ by taking the pairing $\mathsf{pairing}(\xi)$, enlarging it pairing up the nodes as indicated by o (i.e., if a and b are paired in o, then $\mathsf{lonelies}(\xi)_a$ and $\mathsf{lonelies}(\xi)_b$ get paired), and setting $\xi_{L,o}(x) = x$ for all nodes that are still unpaired.

8.8.2 Definition. *We call $\xi \in b$-$\mathsf{creators}$ lucky if for all $L \subseteq [b]$ with $|L|$ even, and all pairings $o \in \binom{L}{2,\dots,2}$, we have: all accepting paths of the computation $h_i^{\xi_{L,o}}(0^j)$ do not access the oracle at any point in $\mathsf{lonelies}(\xi) \setminus \mathsf{lonely}(\xi_{L,o})$.*

In other words, this says that the accepting paths do not access the oracle at positions where lonely nodes are not, but could potentially be.

8.8.3 Claim. *If ξ is sampled from $U_{b\text{-}\mathsf{creators}}$, then ξ is lucky w.h.p.*

Proof. We describe a way of sampling from $U_{b\text{-creators}}$. Let $\binom{X}{b}_{\text{ordered}}$ denote the set of length b lists of distinct elements from X. We first sample $\mathsf{lonelies}(\xi)$ from $U_{\binom{[2^n]}{b}_{\text{ordered}}}$. We then sample $\mathsf{pairing}(\xi)$ from $U_{\binom{[2^n]\setminus \mathsf{lonelies}(\xi)}{2,\dots,2}}$. This process samples a random variable ξ from $U_{b\text{-creators}}$.

Observe that for a fixed (L, o), where $L \subseteq [b]$ and $o \in \binom{L}{2,\dots,2}$, we have that $\xi_{L,o}$ is uniformly distributed from $(b - |L|)$-creations. It remains to show that for such fixed (L, o), we have that ξ is lucky w.h.p. Note that since there are only constantly many such (L, o), the claim immediately follows from the union bound. Hence, for the rest of the proof fix (L, o).

We have that $\xi_{L,o}$ is uniformly distributed. Therefore, the set of $\mathsf{lonelies}(\xi) \setminus \mathsf{lonely}(\xi_{L,o})$ is uniformly distributed over a set of size $\binom{2^n-b+|L|}{|L|}$. By the union bound, an oracle query accesses such a position with probability $\leq \frac{|L|}{2^n-b+|L|}$. By Bernoulli's inequality, we have

$$\left(1 - \frac{|L|}{2^n - b + |L|}\right)^{t_i(j)} \;\geq\; 1 - \frac{|L| \cdot t_i(j)}{2^n - b + |L|}\,.$$

Since n and j are polynomially related, this proves the claim. □

8.8 (c) Defining the set-instantiator

For an a-creation ψ and a computation path $\tau \in \{0,1\}^*$ of a computation $h_i^{\psi}(0^j)$, let $\mathsf{perception}(\tau) \subseteq \mathsf{lonely}(\psi)$ denote the set of accessed oracle positions that are lonely nodes.

Let ξ be a lucky b-creator. We interpret the list $\mathsf{lonelies}(\xi)$ as a bijection $\mathsf{bij} : [b] \to \mathsf{lonelies}(\xi)$. Let $S = 2\mathbb{N}$. For every $\vec{s} \in \mathscr{B}(\vec{b})_S$ (actually, $\vec{s} = s$ is univariate), we let $L_s := [b] \setminus s$ and we choose a $o_s \in \binom{L}{2,\dots,2}$. We set

$$\mathsf{inst}_{\mathcal{SI}}(s) \;:=\; \xi_{L_s,o_s}\,.$$

Finally, for $\tau \in \{0,1\}^*$, let

$$\mathsf{perc}_{\mathcal{SI}}(\tau) \;:=\; \begin{cases} \mathsf{bij}^{-1}(\mathsf{perception}(\tau)) & \text{if } \tau \text{ is an accepting path of the computation } h_i^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j), \\ \top & \text{otherwise.} \end{cases}$$

The rest of this section is devoted to proving that $\mathcal{SI}$ satisfies the requirements of Definition 6.1.3, which then proves Theorem 8.8.1, because clearly $\textsc{Lonely}(\mathsf{inst}_{\mathcal{SI}}(\vec{s})) = |\vec{s}|$. Formally, we have the following result.

8.8.4 Proposition. *For all $\vec{s} \in \mathscr{B}(\vec{b})_S$, we have: $\tau \in \{0,1\}^*$ is an accepting path for the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ if and only if $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.*

Proof. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. This implies that $\mathsf{perc}_{\mathcal{SI}}(\tau) = \mathsf{bij}^{-1}(\mathsf{perception}(\tau))$. Clearly $\mathsf{perception}(\tau) \subseteq \mathsf{bij}(\vec{s})$, because otherwise τ would not even be a computational path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$ because of its oracle answers when querying lonely nodes in $\mathsf{perception}(\tau) \setminus \mathsf{bij}(\vec{s})$. We conclude $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$.

The argument is reversible. Indeed, let $\mathsf{perc}_{\mathcal{SI}}(\tau) \subseteq \vec{s}$, so in particular $\mathsf{perc}_{\mathcal{SI}}(\tau) \neq \top$. Then τ is an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$. Since τ is an accepting path of $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{b})}(0^j)$, and since ξ is lucky, we conclude that τ is also an accepting path of the computation $h^{\mathsf{inst}_{\mathcal{SI}}(\vec{s})}(0^j)$, as desired. □

8.9 The unbalanced flow separation

In this section we stud the BipartiteUnbalance problem, and prove that

$$\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2\,.$$

For clarity of exposition, in notation of §8.2 (f), the vertices in V_+ are called *white* and the vertices in V_- are called *dark*.

While all other classes in Figure 1 are contained in

$$\mathsf{PolynP}_{\text{univariate}} \;:=\; \big\{\varphi(\#\mathsf{P}) \,:\, \varphi \text{ univariate}\big\},$$

it is not clear if the problem BipartiteUnbalance lies in that class. It does lie in

$$\mathsf{GapP} \;\subseteq\; \mathsf{PolynP}_{\text{bivariate}} \;:=\; \big\{\varphi(\overrightarrow{\#\mathsf{P}}) \,:\, \varphi \text{ bivariate}\big\}.$$

To put it into the complexity class inclusion diagram in Figure 1, the definition of its corresponding complexity class $\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})$ is done in analogy to the definitions in §3.2.

Consider the two relations

$$\textsc{rBipartiteUnbalanceWhite} \quad\text{and}\quad \textsc{rBipartiteUnbalanceDark},$$

defined as follows. Let $(C,w) \in \textsc{rBipartiteUnbalanceWhite}$ if and only if C describes a graph in which w is a white vertex. Let the relation rBipartiteUnbalanceDark be defined analogously.

Let $\mathsf{rCOUNTGAP}(\textsc{BipartiteUnbalance})$ be the set of pairs of polynomially balanced relations $R=(R_1,R_2)$ for which a pair (α,β) of polynomial-time computable maps exists with $(C,\beta(x)) \in R_1$ if and only if $(\alpha(C),x) \in \textsc{rBipartiteUnbalanceWhite}$ and $(C,\beta(x)) \in R_2$ if and only if $(\alpha(C),x) \in \textsc{rBipartiteUnbalanceDark}$. In addition, we require that $(C,\beta(x)) \in R_1$ and $(C,\beta(y)) \in R_1$ implies $x=y$, and we require that $(C,\beta(x)) \in R_2$ and $(C,\beta(y)) \in R_2$ implies $x=y$.

Let $\mathsf{rP}\times\mathsf{rP}$ denote the set of pairs of polynomially balanced relations that can be evaluated in polynomial time. By definition, $\mathsf{rCOUNTGAP}(\textsc{BipartiteUnbalance}) \subseteq \mathsf{rP}\times\mathsf{rP}$. We attach an oracle completely analogous to §3.2, to obtain $\mathsf{rCOUNTGAP}(\textsc{BipartiteUnbalance})^A \subseteq \mathsf{rP}^A\times\mathsf{rP}^A$. We define the corresponding counting class $\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A$ via

$$\begin{aligned} &f \in \#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \qquad \text{if and only if}\\ &\exists R \in \mathsf{rCOUNTGAP}(\textsc{BipartiteUnbalance})^A \;:\; f(w) \;=\; \sum_{y\in\{0,1\}^*} \big(R_1(w,y) - R_2(w,y)\big). \end{aligned}$$

By definition and Proposition 8.2.20 we have $\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \subseteq \mathsf{GapP}^A_{\geq 0}$. The main result of this subsection (Proposition 8.9.2) follows from the following theorem.

8.9.1 Theorem. *There exists $A \subseteq \{0,1\}^*$ with*

$$2\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{P}^A.$$

The following proposition is the inclusion and separation shown in Figure 1.

8.9.2 Proposition. $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})/2 \subseteq \#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})$ *via a relativizing parsimonious reduction. Moreover, there exists an oracle A with respect to which* $\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2$.

Proof. We prove the first inclusion via a relativizing reduction that preserves the function value (in the same way a parsimonious reduction preserves the function value). Given a Leaf instance, replace each vertex that has nonzero degree by a white vertex and replace each edge by a black vertex connecting the corresponding two white vertices. The value of the resulting BipartiteUnbalance instance is exactly the number of connected components in the original instance.

The (non-)inclusion in the second part follows directly from dividing Theorem 8.9.1 by 2. Formally, for all oracles A, we have $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A \subseteq \#\mathsf{P}^A_{\text{even}}$. Therefore, we have $\#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2 \subseteq \#\mathsf{P}^A_{\text{even}}/2$.

We use Theorem 8.9.1 to obtain an oracle A such that

$$2\,\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{P}^A.$$

In particular, we have

$$2\,\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{P}^A_{\text{even}}.$$

We divide by 2:

$$\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \not\subseteq \#\mathsf{P}^A_{\text{even}}/2.$$

But if

$$\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \subseteq \#\mathsf{COUNTALL\text{-}PPA}(\textsc{Leaf})^A/2$$

were true, then

$$\#\mathsf{COUNTGAP}(\textsc{BipartiteUnbalance})^A \subseteq \#\mathsf{P}^A_{\text{even}}/2$$

were also true, which we know is false. □

Proof of Theorem 8.9.1. We use the Diagonalization Theorem 6.2.1 and the set-instantiator for Multi(1Source,3Sink), see Theorem 8.10.1), with

$$k = 2, \quad S = \{\vec{f} \in \mathbb{N}^2 \,:\, f_1 = 3f_2 + 6\}.$$

We switch the positions to get the set-instantiator for Multi(3Sink,1Source), with

$$k = 2, \quad S = \{\vec{f} \in \mathbb{N}^2 \,:\, f_2 = 3f_1 - 6\}.$$

We set $\ell = 1$ and $\zeta_2(v_1) = 3v_1 - 6$, so that $I = \langle 3v_1 - 6 - f_2\rangle$ and $Z = \{\vec{f} \in \mathbb{Q}^2 \,:\, 3f_1 - 6 - f_2 = 0\}$. We verify the preconditions of the Diagonalization Theorem 6.2.1:

1. Z contains the integer point $(0, 6)$,
2. C'_S lies Zariski-dense in $\mathbb{Q}$ because $(f_1, 3f_1 - 6) \in S$ for all $f_1 \in \mathbb{N}_{\geq 2}$, and
3. $\vec{v} = (1)$ satisfies $3v_1 > 0$.

We fix Multiplicities = Multi(3Sink, 1Source). The necessary set-instantiators are given by Theorem 8.10.1. We set $\varphi(f_1, f_2) = 8-4f_1+2f_2$, which is the value of a 2BipartiteUnbalance instance with one 6-source, f_1 many 3-sinks, and f_2 many 1-sources, as created by the set-instantiator in §8.10.

We verify that $\varphi + I = -4f_1 + 2f_2 + 8 + \langle 3f_1 - f_2 - 6\rangle$ is binomial-bad by using the Polyhedron Theorem 5.5.2. In the notation of the Polyhedron Theorem 5.5.2, we have $\varphi' = 2f_1 - 4$, $\delta = 1$, so that $\mathbb{O}_\delta$ has only 2 elements. Hence, the polyhedron $\mathcal{P}(\varphi, \zeta)$ will be defined using 2 equations.

Let $(1, v_1)$ be the order of basis vectors of basis vectors of $\mathbb{Q}[\vec{v}]_{\leq 1}$. We express the 3 polynomials $y_{\vec{e}}$ over this basis and obtain coefficient vectors $(1, 0)$, $(0, 1)$, $(-6, 3)$. The polyhedron given by

$$\begin{array}{rrrcr} x_0 & & -6x_2 & = & -4 \\ & x_1 & +3x_2 & = & 2 \\ & x_0,\, x_1,\, x_2 & & \geq & 0 \end{array}$$

has no integer points. Therefore, $\varphi + I$ is binomial-bad by the Polyhedron Theorem 5.5.2.[24] From the Diagonalization Theorem 6.2.1 it follows that for every M we have $p_A(0^j) \neq \#\mathsf{acc}_{M^A}(0^j)$. Hence $p_A \notin \#\mathsf{P}^A$. It remains to show that $p_A \in 2\#\mathsf{COUNTGAP}(\text{BipartiteUnbalance})^A$.

Let $C_\perp$ be an input to Multi(3Sink, 1Source) (i.e., a circuit merged from multiple circuits that describe the exponentially large graph, cf. footnote 10) that yields output $\vec{t}$. Let ℓ be the number of inputs of the circuit $C_\perp$. Let $\nu_j(w) =$ True if and only if the string w has only zeros at positions $j+1, j+2, \ldots$

Finally, let C_j be an input to Multi(3Sink, 1Source)^A that describes the $(j-1)$-input circuit that consists of a single arity j oracle gate that takes a constant 1 into its first input, takes all inputs into its remaining $j-1$ inputs, and forwards its output directly to the circuit output.

We construct a circuit $\alpha(x) := D_{|x|}$ with $\max\{\ell, |x|\}$ many inputs, and with oracle gates as follows ($w \in \{0,1\}^{\max\{\ell,|x|\}}$):

$$\begin{aligned} D_j(w) \quad &:= \quad \Big((A(0^{|w|+1}) = 0) \quad \mathsf{and} \quad \nu_\ell(w) \quad \mathsf{and} \quad C_\perp(w_1, \ldots, w_\ell)\Big) \\ &\quad\ \mathsf{and} \quad \Big((A(0^{|w|+1}) = 1) \quad \mathsf{and} \quad \nu_{|x|}(w) \quad \mathsf{and} \quad C_{|x|+1}(w_1, \ldots, w_{|x|})\Big) \end{aligned}$$

We define the polynomially balanced relations $R = (R_1, R_2)$ as follows: let $R_1(w, y)$ if and only if rBipartiteUnbalanceWhite$(\alpha(w), y)$. Similarly, let $R_2(w, y)$ if and only if rBipartiteUnbalanceDark$(\alpha(w), y)$. We set β to be the identity function.

By definition, $R \in \mathsf{rCOUNTGAP}(\text{BipartiteUnbalance})$. Therefore, the following function is in $\#\mathsf{COUNTGAP}(\text{BipartiteUnbalance})$:

$$\begin{aligned} &\sum_{y\in\{0,1\}^*} R(x, y) \\ &= \sum_{y\in\{0,1\}^*} (\text{rBipartiteUnbalanceWhite} - \text{rBipartiteUnbalanceDark})(\alpha(x), y) \\ &= \begin{cases} |V_+|(D_{|x|}) - |V_-|(D_{|x|}) & \text{if} \quad A(0^{|x|+1}) = 1, \\ t_1 - t_2 & \text{otherwise.} \end{cases} \end{aligned}$$

Thus, this function equals $p_A(w) + 1$, as desired. This finishes the proof.[25] □

[24] Note that the LP relaxation has solutions, for example $(0, 0, \frac{2}{3})$, so a if a computation path could not just accept or reject, but accept with a fractional contribution, then this would work out.

[25] We remark that the graph gadgets that make this separation work must be carefully chosen, and for several other choices of gadgets we do not get a separation.

8.10 A set-instantiator for 2#COUNTGAP(BIPARTITEUNBALANCE)

8.10.1 Theorem. *Let M be a polynomial time nondeterministic Turing machine. Fix $k = 2$. Then, for every $\vec{b} \in \mathbb{N}^2$, there exists a threshold $j_0 \in \mathbb{N}$, such that for every $j \geq j_0$ and every $A_{<j} \in \{0,1\}^{<j}$, there exists a set-instantiator $\mathcal{SI}$ against $(M, j, A_{<j}, S = S_{\mathrm{BUline}}, \vec{b})$, so that* MULTI(1SOURCE, 3SINK)($\mathsf{inst}_{\mathcal{SI}}(\vec{s})$) $= |\vec{s}|$ *for all $\vec{s}$ with $|\vec{s}| \in S$.*

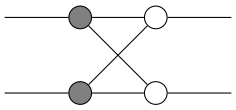
Proof. The set-instantiator is constructed analogously to the construction for MULTI(NZSOURCE, EXCESS) in Theorem 8.5.1, but with the source at zero, while the other sources, the directed edges and the double sinks are replaced by graph gadgets. The rest of the proof then works analogously.

We have slightly different parameters in this case. Let

$$S_{\mathrm{BUline}} := \big\{(f_1, f_2) \in \mathbb{N}^2 : f_1 - 3f_2 + 6 = 0\big\}.$$

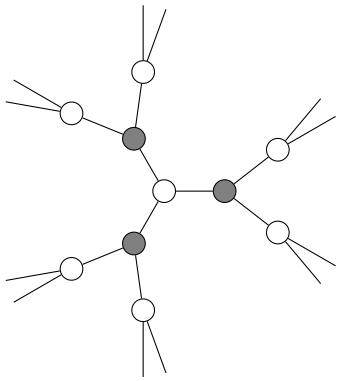
The graph gadgets are as follows, where the dark nodes have degree at least as high as all adjacent white nodes.

First, a *directed edge* is replaced by a *double edge*, where each edge connects a white node of degree 2 or 3, and a dark node of degree 3. An *indegree = outdegree* 1 node is replaced by the following gadget:



Note that $2(\#\text{white nodes} - \#\text{dark nodes}) = 0$ in this case.

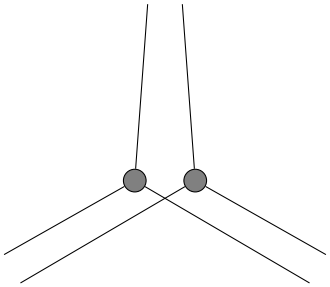
Next, the *zero source* is replaced by the following gadget (in the original directed problem this would be a 6-fold source):



Note that $2(\#\text{white nodes} - \#\text{dark nodes}) = 8$ in this case. The *source* is replaced by the following simple gadget:

Note that $2(\#\text{white nodes} - \#\text{dark nodes}) = 2$.

Finally, there are no double sink nodes, but for a triple sink we use the following gadget:



Note that $2(\#\text{white nodes} - \#\text{dark nodes}) = -4$ in this case. This finishes the proof. □

9 Open problems

9.1 Counting subgraphs

Let $G = (V, E)$ be a simple graph. Denote by $m_k = m_k(G)$ the number of *k-matchings*, i.e. sets of k edges with disjoint vertices. As we mentioned in the introduction, the Heilmann–Lieb theorem states [HL72]:

$$\delta(k, G) := m_k^2 - m_{k+1} m_{k-1} \geq 0 \quad \text{for all } k \geq 1,$$

where we assume $m_0 = 1$. Clearly, $\delta \in \mathsf{GapP}_{\geq 0}$.

9.1.1 Proposition ([Pak19]). *In the notation above, $\delta \in \#\mathsf{P}$.*

The proof in [Pak19] is an easy adaptation of the proof by Krattenthaler [Kra96]. Many other log-concavity problems remain open. Proposition 2.3.3 suggests some of them might not be in $\#\mathsf{P}$.

9.1.2 Theorem ([AHK18]). *Let $G = (V, E)$ be a simple graph, and denote by $\tau_k = \tau_k(G)$ the number of spanning forests in G with k edges. Then:*

$$\tau_k^2 \geq \tau_{k+1} \tau_{k-1} \quad \textit{for all } k \geq 1.$$

This is a celebrated result by Adiprasito, Huh and Katz, which holds for all matroids, not just truncations of the graphical matroids. We refer to [Huh18] for a survey of the algebraic approach, and to [CP21, CP22] for an elementary approach using linear algebra.

(1) Is it true that $\tau_k^2 \geqslant_{\#} \tau_{k+1} \tau_{k-1}$? This remains open and is a major challenge in the area, see e.g. [Pak19] and [CP21, §17.17].

For many other unimodality and log-concavity results, see [Brä15, Huh18, Sta80]. Following [Pak19] and the approach in this paper, most of these results can be rephrased as questions about corresponding counting functions in $\#\mathsf{P}$.

9.2 Counting linear extensions

Let $P = (X, \prec)$ be a poset on $|X| = n$ elements. A *linear extension* is a order-preserving bijection $f : X \to [n]$, i.e. $f(x) < f(y)$ for all $x \prec y$. As in §1.2(1), denote by $\mathcal{E}(P)$ and $e(P) := |\mathcal{E}(P)|$ the set and the number of linear extensions, respectively. The problem of computing $e(P)$ is famously $\#\mathsf{P}$-complete [BW91], even in special cases of posets of height two, or of width two [DP18].

There are some surprising combinatorial inequalities for the numbers of linear extensions. For each element $x \in X$, let $B(x) := \{y \in X : y \succcurlyeq x\}$ be the *upper order ideal* generated by x, and let $b(x) := |B(x)|$.

9.2.1 Theorem (*Björner–Wachs inequality* [BW89]). *In the notation above, we have:*

$$e(P) \prod_{x \in X} b(x) \ \geqslant_{\#} \ n!$$

The Björner–Wachs inequality is an equality for ordered forests, where it is called the *hook-length formula*, see e.g. [SY89, Sta12]. The original proof in [BW89] uses an explicit injection which is easily computable in polytime, see e.g. [CPP22].

For an element $x \in X$ and integer $1 \leq a \leq n$, let $\mathcal{E}(P, x, a)$ be the set of linear extensions $f \in \mathcal{E}(P)$ such that $f(x) = a$. Denote by $e(P, x, a) := \big|\mathcal{E}(P, x, a)\big|$ the number of such linear extensions.

9.2.2 Theorem (*Stanley inequality* [Sta81]). *In the notation above, we have:*

$$e(P, x, a)^2 \ \geq \ e(P, x, a+1)\, e(P, x, a-1).$$

The original proof in [Sta81] uses the Alexandrov–Fenchel inequalities applied to *order polytopes* (see §7.1), suggesting that a direct combinatorial proof might not exist. Recently, Stanley's inequality has been the subject of intense recent investigation, see [CPP21, SvH20]. Notably, an elementary linear algebraic proof of the inequality is given in [CP21].

(2) Is it true that $e(P, x, a)^2 \geqslant_{\#} e(P, x, a+1)\, e(P, x, a-1)$?

9.2.3 Theorem (*XYZ inequality* [She82]). *Let $P = (X, \prec)$, $x, y, z \in X$ incomparable elements. Let $P_{xy} := P \cup \{x \prec y\}$, $P_{xz} := P \cup \{x \prec z\}$ and $P_{xyz} := P \cup \{x \prec y, x \prec z\}$. Then:*

$$e\big(P\big)\, e\big(P_{xyz}\big) \ \geq \ e\big(P_{xy}\big)\, e\big(P_{xz}\big).$$

This correlation inequality is derived by Shepp from the *FKG inequality* mentioned in §7.4, again suggesting that a direct combinatorial proof might not exist.

(3) Is it true that $e\big(P\big)\, e\big(P_{xyz}\big) \geqslant_{\#} e\big(P_{xy}\big)\, e\big(P_{xz}\big)$?

The closest to resolving the problem in the positive is the combinatorial (but not injective!) proof in [BT02], which involves an application of BipartiteUnbalance, see §8.1 and §8.9.

9.3 Ranks of matrices

In [GKPT16], the authors show the #P-hardness of the dimension $\dim \partial^* f$ of the space of partial derivatives of a polynomial f, when the input is given in the sparse presentation (i.e., as a list of monomials that have nonzero coefficients). They explicitly state as an open question whether or not $\dim \partial^* f$ or $\dim \partial^{=k} f$ are in #P. When f is homogeneous, then $\partial^{=k} f$ is a special case of a so-called *Young flattening*.

For a partition λ, let $S^\lambda \mathbb{C}^n$ denote the irreducible GL_n-representation of type λ, i.e., for $\lambda = (n)$, we have $S^{(d)}\mathbb{C}^n = \mathbb{C}[x_1, \ldots, x_n]_d$. Similarly, for $\lambda = (1^d)$, we have $S^{(1^d)}\mathbb{C}^n = \bigwedge^d \mathbb{C}^n$. Pieri's rule states that if the Young diagrams of two partitions $\lambda \subseteq \mu$ differ by d boxes, at most by 1 in each column, and if they both have least n many rows, then there is a unique nonzero GL_n-equivariant map $S^{(d)}\mathbb{C}^n \otimes S^\lambda \mathbb{C}^n \to S^\mu \mathbb{C}^n$. For every $f \in S^{(d)}\mathbb{C}^n$ this induces a linear map

$\mathcal{F}_{\lambda,\mu} : S^{(d)} \to \mathrm{End}(S^\lambda \mathbb{C}^n, S^\mu \mathbb{C}^n)$, see e.g. [Sam09, HI21]. The rank of this map has been used to derive lower bounds in algebraic complexity theory, see e.g. [LO15, Lan15, Far16].

(4) Is the map $(\lambda, \mu, d, n, f) \mapsto \mathrm{rank}(\mathcal{F}_{\lambda,\mu}(f))$ in $\#\mathsf{P}$, where f is given in the sparse presentation?

In [BIM+20], a very similar lower bounds technique, based on [HL16], is used to show that a polynomial can be computed by an algebraic branching program of a certain format *only when approximations are allowed*.

9.4 Orbit closures and representation theoretic multiplicities

The Kronecker and the plethysm coefficients have interpretations as *dimensions of highest weight vector spaces* as follows. For a GL_m-representation V a *weight vector* is a vector v that is rescaled by the action of the algebraic torus: $\mathrm{diag}(\alpha_1, \ldots, \alpha_m)v = \alpha_1^{\lambda_1} \cdots \alpha_n^{\lambda_m} v$. The list $\lambda \in \mathbb{Z}^m$ is called the *weight* of v.

A *highest weight vector* is a weight vector that is fixed under the action of a maximal unipotent subgroup U (for example the upper triangular matrices with 1s on the main diagonal). In this case, λ is always a partition. Highest weight vectors of weight λ form a vector space that we denote by $\mathsf{HWV}_\lambda(V)$. In general, we have $\dim \mathsf{HWV}_\lambda(V) = \mathrm{mult}_\lambda(V)$, where $\mathrm{mult}_\lambda(V)$ is the number of irreducible representations of type λ in any decomposition of V into irreducibles (this number does not depend on the actual decomposition). We have:

$$p_\lambda(d,n) \;=\; \dim \mathsf{HWV}_\lambda\big(S^{(d)}(S^{(n)}\mathbb{C}^m)\big) \;=\; \mathrm{mult}_\lambda\big(S^{(d)}(S^{(n)}\mathbb{C}^m)\big)$$

for m large enough. In fact, the dimension is zero if the number of rows of λ is larger than m, and it is a fixed number otherwise. An analog argument can be done for the triple product $\mathrm{GL}_m \times \mathrm{GL}_m \times \mathrm{GL}_m$ on the space of order 3 tensors $\bigotimes^3 \mathbb{C}^m$:

$$g(\lambda,\mu,\nu) \;=\; \dim \mathsf{HWV}_{(\lambda,\mu,\nu)}\big(S^{(d)}(\textstyle\bigotimes^3\mathbb{C}^m)\big) \;=\; \mathrm{mult}_{(\lambda,\mu,\nu)}\big(S^{(d)}(\textstyle\bigotimes^3\mathbb{C}^m)\big).$$

For a GL_m-orbit closure[26] $Z \subseteq S^{(n)}\mathbb{C}^m$, the plethysm coefficient decomposes:

$$p_\lambda(d,n) \;=\; \mathrm{mult}_\lambda\big(I(Z)_d\big) \;+\; \mathrm{mult}_\lambda\big(\mathbb{C}[Z]_d\big),$$

We write $p_\lambda(d,n)_{Z^+} := \mathrm{mult}_\lambda(\mathbb{C}[Z]_d)$ and $p_\lambda(d,n)_{Z^-} := \mathrm{mult}_\lambda(I(Z)_d)$. Analogously for the tensor setting, for a GL_m^3-orbit closure Z, with d being the number of boxes in λ and also in μ and ν, we have:

$$g(\lambda,\mu,\nu) \;=\; \mathrm{mult}_{(\lambda,\mu,\nu)}\big(I(Z)_d\big) \;+\; \mathrm{mult}_{(\lambda,\mu,\nu)}\big(\mathbb{C}[Z]_d\big)$$

We write $g(\lambda,\mu,\nu)_{Z^+} := \mathrm{mult}_{(\lambda,\mu,\nu)}(\mathbb{C}[Z]_d)$ and $g(\lambda,\mu,\nu)_{Z^-} := \mathrm{mult}_{(\lambda,\mu,\nu)}(I(Z)_d)$. These numbers are zero when m is too small, but are constant otherwise.

If Z is an orbit closure of a point q, then an upper bound for $p_\lambda(d,n)_{Z^+}$ (and for $g(\lambda,\mu,\nu)_{Z^+}$ in the tensor setting), is given by the dimension of the invariant space of the irreducible GL_m-representation V_λ of type λ under the action of the stabilizer H of q, i.e. $p_\lambda(d,n)_{Z^+} \le \dim V_\lambda^H$ or $g(\lambda,\mu,\nu)_{Z^+} \le \dim(V_\lambda \otimes V_\mu \otimes V_\nu)^H$, see [BLMW11]. We denote this quantity by $u_\lambda(q)$.

The *geometric complexity* paper [MS08] pioneered studying a subset of the following multiplicities, where input partitions and input numbers are always given in unary. Let

$$\det_n \;:=\; \sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi) \prod_{i=1}^n x_{i\,\pi(i)} \quad \text{and let} \quad \mathrm{per}_n \;:=\; \sum_{\pi\in\mathfrak{S}_n} \prod_{i=1}^n x_{i\,\pi(i)}\,.$$

[26] indeed, for any GL_m-variety that is closed under rescaling, i.e., a projective variety

(5) For $Z = \overline{\mathrm{GL}_{n^2}\det_n}$, is $(d,n,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P? Is $(d,n,\lambda) \mapsto u_\lambda(\det_n)$ in #P? These problems are studied in [Kum15, IP17, BIP19].

(6) For $Z = \overline{\mathrm{GL}_{n^2}\mathrm{per}_n}$, is $(d,n,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P? Is $(d,n,\lambda) \mapsto u_\lambda(\mathrm{per}_n)$ in #P? See [BLMW11, eq (5.5.2)].

Let

$$\mathrm{IMM}_m^{(n)} := \mathrm{tr}\left[\begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,m}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{m,1}^{(1)} & \cdots & x_{m,m}^{(1)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(n)} & \cdots & x_{1,m}^{(n)} \\ \vdots & \ddots & \vdots \\ x_{m,1}^{(n)} & \cdots & x_{m,m}^{(n)} \end{pmatrix}\right] \quad \text{and let} \quad \mathrm{Pow}_m^{(n)} := \mathrm{tr}\left[\begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,m}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{m,1}^{(1)} & \cdots & x_{m,m}^{(1)} \end{pmatrix}^n\right]$$

(7) For $Z = \overline{\mathrm{GL}_{m^2 n}\mathrm{IMM}_m^{(n)}}$, is $(d,n,m,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P? Is $(d,n,m,\lambda) \mapsto u_\lambda(\mathrm{IMM}_m^{(n)})$ in #P?

(8) For $Z = \overline{\mathrm{GL}_{m^2}\mathrm{Pow}_m^{(n)}}$, is $(d,n,m,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P? Is $(d,n,m,\lambda) \mapsto u_\lambda(\mathrm{Pow}_m^{(n)})$ in #P? This last quantity has been studied in [GIP17].

(9) For $Z = \overline{\mathrm{GL}_m(x_1^n + x_2^n + \cdots + x_m^n)}$, is $(d,n,m,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P? Is $(d,n,m,\lambda) \mapsto u_\lambda(x_1^n + x_2^n + \cdots + x_m^n) \in$ #P?

(10) For $Z = \overline{\mathrm{GL}_n(x_1 \cdots x_n)}$, is $(d,n,\lambda) \mapsto p_\lambda(d,n)_{Z^\pm}$ in #P?

The multiplicities in (9) and (10) are used in [DIP20, IK20], see also [BI18, Thm. 22.4.1]. The multiplicities in (10) play a central role in the Foulkes–Howe approach to the *Foulkes conjecture*, see e.g. [McK08, CIM17]. For a partition λ with $|\lambda| = nd$, we actually have $u_\lambda(x_1 \cdots x_n) = p_\lambda(n,d)$.

Let $e_i \in \mathbb{C}^m$ denote the i-th standard basis vector, and let $\{e_{i,j} : 1 \le i,j \le m\}$ be a basis of $\mathbb{C}^{m^2}$. Consider the unit tensor and the *matrix multiplication tensor*:

$$E_m := e_1^{\otimes 3} + \cdots + e_m^{\otimes 3} \quad \text{and} \quad M_m := \sum_{i,j,k=1}^{m} e_{i,j} \otimes e_{j,k} \otimes e_{k,i}$$

(11) For $Z = \overline{\mathrm{GL}_m^3(E_m)}$, is $(d,m,\lambda,\mu,\nu) \mapsto g(\lambda,\mu,\nu)_{Z^\pm}$ in #P? Is $(d,m,\lambda,\mu,\nu) \mapsto q_{(\lambda,\mu,\nu)}(E_m) \in$ #P?

(12) For $Z = \overline{\mathrm{GL}_{m^2}^3(M_m)}$, is $(d,m,\lambda,\mu,\nu) \mapsto g(\lambda,\mu,\nu)_{Z^\pm}$ in #P? Is $(d,m,\lambda,\mu,\nu) \mapsto q_{(\lambda,\mu,\nu)}(M_m) \in$ #P?

The coefficients in open problems (11) and (12) were used in [BI11, BI13a] to study the *complexity of matrix multiplication*. For (11), see also [BI18, Thm. 22.4.3]. See [BIL+21] for more tensor setting examples.

In [BHIM22], the authors give software to compute $p_\lambda(d,n)_{Z^\pm}$ in small cases. The papers [BIJL18] and [GIM+20] point out the hardness of computing these quantities in general. Notably, Grochow [Gro15] reinterprets numerous lower bounds results in *algebraic complexity theory* as group orbit (closure) separations, and thus they fall into the category of this paragraph, but the focus in [Gro15] is not directly on the multiplicities, but on the more general approach of finding separating modules.

9.5 Algorithms for the binomial basis on an affine variety

Expressing a polynomial in its binomial basis can be done using a greedy algorithm starting from the top total degrees and proceeding to the lower degrees. Hence, given a polynomial φ in its monomial basis, one can check in polynomial time whether or not φ is binomial-good. We have

seen that it is important to understand when a coset $\varphi + I$ contains a binomial-good element (we say that $\varphi + I$ is binomial-good), where I is an ideal.

(13) If we assume that a set of generating elements of I is presented in the binomial basis, what is the complexity of determining whether or not $\varphi + I$ is binomial-good? In which cases do we get efficient algorithms?

Acknowledgements

We are grateful to Swee Hong Chan, Nikita Gladkov and Greta Panova for numerous helpful discussions and remarks on the subject. We thank Joshua Grochow and Greg Kuperberg for many inspiring conversations on computational complexity over the years. We learned about [HVW95] and [Bei97] only after the paper was written; we thank Lane Hemaspaandra who kindly for pointed out these references. We thank Markus Bläser and Paul Goldberg for helpful comments on a draft version of this paper.

We are grateful to the Oberwolfach Research Institute for Mathematics where we started this project in the innocent times of February of 2020. The first author was partially supported by the DFG grant IK 116/2-1 and the EPSRC grant EP/W014882/1. The second author was partially supported by the NSF grant CCF-2007891.

References

[Aar16] S. Aaronson, $\mathsf{P} \stackrel{?}{=} \mathsf{NP}$, in *Open problems in mathematics*, Springer, Cham, 2016, 1–122.

[AHK18] K. Adiprasito, J. Huh and E. Katz, Hodge theory for combinatorial geometries, *Annals of Math.* **188** (2018), 381–452.

[AD78] R. Ahlswede and D. E. Daykin, An inequality for the weights of two families of sets, their unions and intersections, *Z. Wahrsch. Verw. Gebiete* **43** (1978), 183–185.

[Alon03] N. Alon, Problems and results in extremal combinatorics I, *Discrete Math.* **273** (2003), 31–53.

[AS16] N. Alon and J. H. Spencer, *The probabilistic method* (Fourth ed.), John Wiley, Hoboken, NJ, 2016, 375 pp.

[AG88] G. E. Andrews and F. G. Garvan, Dyson's crank of a partition, *Bull. AMS* **18** (1988), 167–171.

[AZ17] M. Apagodu and D. Zeilberger, Using the "freshman's dream" to prove combinatorial congruences, *Amer. Math. Monthly* **124** (2017), no. 7, 597–608.

[AB09] S. Arora and B. Barak, *Computational complexity. A modern approach*, Cambridge Univ. Press, Cambridge, 2009, 579 pp.

[Bar07] A. Barvinok, Brunn–Minkowski inequalities for contingency tables and integer flows, *Adv. Math.* **211** (2007), 105–122.

[BCE+98] P. Beame, S. Cook, J. Edmonds, R. Impagliazzo and T. Pitassi, The relative complexity of NP search problems, in *Proc. 27th STOC* (1995), 303–314; *J. Comput. System Sci.* **57** (1998), 3–19.

[BB61] E. Beckenbach and R. Bellman, *Inequalities*, Springer, Berlin, 1961, 198 pp.

[Bei97] R. Beigel, Closure properties of GapP and $\#\mathsf{P}$, in *Proc. 5th Israeli Symposium on Theory of Computing and Systems* (1997), 144–146.

[BG92] R. Beigel and J. Gill, Counting classes: Thresholds, parity, mods, and fewness, in *Proc. 7th STACS* (1990), 49–57; *Theoret. Comput. Sci.* **103** (1992), 3–23.

[BW89] A. Björner and M. L. Wachs, q-hook length formulas for forests, *J. Combin. Theory, Ser. A* **52** (1989), 165–187.

[BG83] A. Blass and Yu. Gurevich, Equivalence relations, invariants, and normal forms. II, in *Symposium on Recursive Combinatorics*, Springer, Berlin, 1983, 24–42.

[BG84] A. Blass and Yu. Gurevich, Equivalence relations, invariants, and normal forms, *SIAM J. Comput.* **13** (1984), 682–689.

[BI18] M. Bläser and C. Ikenmeyer, Introduction to geometric complexity theory, preprint (2018), to appear in *Theory of Computing Graduate Surveys*.

[BIJL18] M. Bläser, C. Ikenmeyer, G. Jindal and V. Lysikov, Generalized matrix completion and algebraic natural proofs, in *Proc. 50th STOC* (2018), 1193–1206.

[BIL+21] M. Bläser, C. Ikenmeyer, V. Lysikov, A. Pandey and F.-O. Schreyer, On the orbit closure containment problem and slice rank of tensors, in *Proc. 32nd SODA* (2021), 2565–2584.

[BIM+20] M. Bläser, C. Ikenmeyer, M. Mahajan, A. Pandey and N. Saurabh, Algebraic branching programs, border complexity, and tangent spaces, in *Proc. 35th CCC* (2020), Art. 21, 24 pp.; `arXiv:2003.04834`.

[Ble13] G. Blekherman, Nonnegative polynomials and sums of squares, in *Semidefinite optimization and convex algebraic geometry* (Edited by G. Blekherman, P. A. Parrilo and R. R. Thomas), SIAM, Philadelphia, PA, 2013, 159–202.

[BP21] I. A. Bochkov and F. V. Petrov, The bounds for the number of linear extensions via chain and antichain coverings, *Order* **38** (2021), no. 2, 323–328.

[BS01] B. Borchert and R. Silvestri, Dot operators, in *Proc. 12th CCC* (1997), 36–44; *Theoret. Comput. Sci.* **262** (2001), 501–523.

[BDO15] C. Bowman, M. De Visscher and R. Orellana, The partition algebra and the Kronecker coefficients, *Trans. AMS* **367** (2015), 3647–3667.

[Brä15] P. Brändén, Unimodality, log-concavity, real-rootedness and beyond, in *Handbook of enumerative combinatorics*, CRC Press, Boca Raton, FL, 2015, 437–483.

[BHIM22] P. Breiding, R. Hodges, C. Ikenmeyer and M. Michałek, Equations for GL invariant families of polynomials, *Vietnam J. Math.* (2022), 1–12.

[BT02] G. Brightwell and W. T. Trotter, A combinatorial approach to correlation inequalities, *Discrete Math.* **257** (2002), 311–327.

[BW91] G. Brightwell and P. Winkler, Counting linear extensions, in *Proc. 39th STOC* (1991), 175–181; *Order* **8** (1991), 225–247.

[Bri93] M. Brion, Stable properties of plethysm: on two conjectures of Foulkes, *Manuscripta Math.* **80** (1993), 347–371.

[Bru06] R. A. Brualdi, *Combinatorial matrix classes*, Cambridge Univ. Press, Cambridge, 2006, 544 pp.

[BZ88] Yu. D. Burago and V. A. Zalgaller, *Geometric inequalities*, Springer, Berlin, 1988, 331 pp.

[BM04] J. Buresh-Oppenheim and T. Morioka, Relativized NP search problems and propositional proof systems, in *Proc. 19th CCC* (2004), 54–67.

[BI11] P. Bürgisser and C. Ikenmeyer, Geometric complexity theory and tensor rank, in *Proc. 43rd STOC* (2011), 509–518.

[BI13a] P. Bürgisser and C. Ikenmeyer, Explicit lower bounds via geometric complexity theory, in *Proc. 45th STOC* (2013), 141–150.

[BI13b] P. Bürgisser and C. Ikenmeyer, Deciding positivity of Littlewood–Richardson coefficients, in *Proc. 21st FPSAC* (2009), 265–276; *SIAM J. Discrete Math.* **27** (2013), 1639–1681.

[BIP19] P. Bürgisser, C. Ikenmeyer and G. Panova, No occurrence obstructions in geometric complexity theory, in *Proc. 57th FOCS* (2016), 386–395; *Jour. AMS* **32** (2019), 163–193.

[BLMW11] P. Bürgisser, J. M. Landsberg, L. Manivel and J. Weyman, An overview of mathematical issues arising in the geometric complexity theory approach to VP≠VNP, *SIAM J. Comput.* **40** (2011), 1179–1209.

[CC16] P.-J. Cahen and J.-L. Chabert, What you should know about integer-valued polynomials, *Amer. Math. Monthly* **123** (2016), 311–337.

[CGH+89] J.-Yi Cai, T. Gundermann, J. Hartmanis, L. A. Hemachandra, V. Sewelson, K. Wagner and G. Wechsung, The Boolean hierarchy. II. Applications, *SIAM J. Comput.* **18** (1989), 95–111.

[Cam01] K. Cameron, Thomason's algorithm for finding a second Hamiltonian circuit through a given edge in a cubic graph is exponential on Krawczyk's graphs, *Discrete Math.* **235** (2001), 69–77.

[CW95] E. R. Canfield and S. G. Williamson, A loop-free algorithm for generating the linear extensions of a poset, *Order* **12** (1995), 57–75.

[Car14] R. D. Carmichael, *The theory of numbers*, John Wiley, New York, 1914, 94 pp.

[CP21] S. H. Chan and I. Pak, Log-concave poset inequalities, preprint (2021), 71 pp; `arXiv:2110.10740`.

[CP22] S. H. Chan and I. Pak, Introduction to the combinatorial atlas, preprint (2022), 28 pp.; `arXiv:2203.01533`.

[CPP21] S. H. Chan, I. Pak and G. Panova, Extensions of the Kahn–Saks inequality for posets of width two, preprint (2021), 25 pp.; `arXiv:2106.07133`.

[CPP22] S. H. Chan, I. Pak and G. Panova, Effective poset inequalities, preprint (2022), 36 pp.; `arXiv:2205.02798`.

[CD09] Xi Chen and X. Deng, On the complexity of 2D discrete fixed point problem, in *Proc. 33rd ICALP* (2006), 489–500; *Theoret. Comput. Sci.* **410** (2009), 4448–4456.

[CIM17] M.-W. Cheung, C. Ikenmeyer and S. Mkrtchyan, Symmetrizing tableaux and the 5th case of the Foulkes conjecture, *J. Symbolic Comput.* **80** (2017), 833–843.

[CS08] E. F. Cornelius, Jr. and P. Schultz, Multinomial points, *Houston J. Math.* **34** (2008), 661–676.

[Cur16] R. Curticapean, Parity separation: A scientifically proven method for permanent weight loss, in *Proc. 43rd ICALP* (2016), Art. 47, 14 pp.

[DGP09] C. Daskalakis, P. W. Goldberg and C. H. Papadimitriou, The complexity of computing a Nash equilibrium, in *Proc. 38th STOC* (2006), 71–78; *SIAM J. Comput.* **39** (2009), 195–259.

[DFHM22] A. Deligkas, J. Fearnley, A. Hollender and T. Melissourgos, Constant inapproximability for PPA, in *Proc. 54th STOC* (2022), to appear; `arXiv:2201.10011`.

[DEF+21] X. Deng, J. R. Edmonds, Z. Feng, Z. Liu, Qi Qi and Z. Xu, Understanding PPA-completeness, in *Proc. 31st CCC* (2016), Art. 23, 25 pp.; *J. Comput. System Sci.* **115** (2021), 146–168.

[DS06] E. Deutsch and B. E. Sagan, Congruences for Catalan and Motzkin numbers and related sequences, *J. Number Theory* **117** (2006), 191–215.

[Dic52] L. E. Dickson, *History of the theory of numbers. Vol. I: Divisibility and primality*, Chelsea, New York, 1952, 486 pp.

[DP18] S. Dittmer and I. Pak, Counting linear extensions of restricted posets, preprint (2018), 33 pp.; `arXiv:1802.06312`.

[DIP20] J. Dörfler, C. Ikenmeyer and G. Panova, On geometric complexity theory: Multiplicity obstructions are stronger than occurrence obstructions, in *Proc. 46th ICALP* (2019), Art. 51, 14 pp.; *SIAM J. Appl. Algebra Geom.* **4** (2020), 354–376.

[Dys44] F. J. Dyson, Some guesses in the theory of partitions, *Eureka* **8** (1944), 10–15; available at tinyurl.com/3apae925

[Far16] C. Farnsworth, Koszul–Young flattenings and symmetric border rank of the determinant, *J. Algebra* **447** (2016), 664–676.

[FGHS21] J. Fearnley, P. W. Goldberg, A. Hollender and R. Savani, The complexity of gradient descent: CLS = PPAD ∩ PLS, in *Proc. 53rd STOC* (2021), 46–59.

[FFK94] S. A. Fenner, L. Fortnow and S. A. Kurtz, Gap-definable counting classes, in *Proc. 6th CCC* (1991), 30–42; *J. Comput. System Sci.* **48** (1994), 116–148.

[FRH+20] A. Filos-Ratsikas, A. Hollender, K. Sotiraki and M. Zampetakis, Consensus-halving: Does it ever get easier?, in *Proc. 21st ACM-EC* (2020), 381–399.

[FRG18] A. Filos-Ratsikas and P. W. Goldberg, Consensus halving is PPA-complete, in *Proc. 50th STOC* (2018), 51–64.

[FI20] N. Fischer and C. Ikenmeyer, The computational complexity of plethysm coefficients, *Comput. Complexity* **29** (2020), no. 2, Paper 8, 43 pp.

[For97] L. Fortnow, Counting complexity, in *Complexity theory retrospective* II, Springer, New York, 1997, 81–107.

[For09] L. Fortnow, A simple proof of Toda's theorem, *Theory Comput.* **5** (2009), 135–140.

[FG11] L. Fortnow and J. A. Grochow, Complexity classes of equivalence problems revisited, *Inform. and Comput.* **209** (2011), 748–763.

[Ful98] W. Fulton, Eigenvalues of sums of Hermitian matrices (after A. Klyachko), *Astérisque* **252** (1998), No. 845, 255–269.

[GW83] H. Galperin and A. Wigderson, Succinct representations of graphs, *Inform. and Control* **56** (1983), 183–198.

[GKPT16] I. Garcia-Marco, P. Koiran, T. Pecatte and S. Thomassé, On the complexity of partial derivatives, in *Proc. 34th STACS* (2016), Art. 37, 13 pp.; `arXiv:1607.05494`.

[GIM+20] A. Garg, C. Ikenmeyer, V. Makam, R. Oliveira, M. Walter and A. Wigderson, Search problems in algebraic complexity, gct, and hardness of generators for invariant rings, in *Proc. 35th CCC* (2020), 1–17.

[GKS90] F. Garvan, D. Kim and D. Stanton, Cranks and t-cores, *Invent. Math.* **101** (1990), 1–17.

[GPS12] W. Gasarch, A. Parrish and S. Sinai, Three proofs of the hypergraph Ramsey theorem (an exposition), preprint (2012), 39 pp.; `arXiv:1206.4257`.

[Ges84] I. M. Gessel, Combinatorial proofs of congruences, in *Enumeration and Design* (Edited by D. M. Jackson and S. A. Vanstone), Academic Press, Orlando, 1984, 157–197; available at tinyurl.com/bxvszced

[GIP17] F. Gesmundo, C. Ikenmeyer and G. Panova, Geometric complexity theory and matrix powering, *Differential Geom. Appl.* **55** (2017), 106–127.

[God93] C. D. Godsil, *Algebraic combinatorics*, Chapman & Hall, New York, 1993, 362 pp.

[GH21] P. W. Goldberg and A. Hollender, The hairy ball problem is PPAD-complete, in *Proc. 46th ICALP* (2019), Art. 65, 14 pp.; *J. Comput. System Sci.* **122** (2021), 34–62.

[GP17] P. W. Goldberg and C. H. Papadimitriou, TFNP: an update, in *Algorithms and complexity*, Springer, Cham, 2017, 3–9.

[GP18] P. W. Goldberg and C. H. Papadimitriou, Towards a unified complexity theory of total functions, *J. Comput. System Sci.* **94** (2018), 167–192.

[Gol56] S. W. Golomb, Combinatorial Proof of Fermat's "Little" Theorem, *Amer. Math. Monthly* **63** (1956), no. 10, 718.

[GHJ+22] M. Göös, A. Hollender, S. Jain, G. Maystre, W. Pires, R. Robere and R. Tao, Further collapses in TFNP, 2022, 14 pp.; `arXiv: 2202.07761`.

[GRS90] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey theory* (Second ed.), Wiley, New York, 1990, 196 pp.

[Gre93] F. Green, On the power of deterministic reductions to $C_=P$, *Math. Systems Theory* **26** (1993), 215–233.

[Gri01] M. Grigni, A Sperner lemma complete for PPA, *Inform. Process. Lett.* **77** (2001), 255–259.

[Gri76] G. R. Grimmett, An upper bound for the number of spanning trees of a graph, *Discrete Math.* **16** (1976), 323–324.

[Gro15] J. A. Grochow, Unifying known lower bounds via geometric complexity theory, in *Proc. 29th CCC* (2014), 274–285; *Computational Complexity* **24** (2015), 393–475.

[GS88] J. Grollmann and A. L. Selman, Complexity measures for public-key cryptosystems, in *Proc. 25th FOCS* (1984), 495–503; *SIAM J. Comput.* **17** (1988), 309–335.

[Gup95] S. Gupta, Closure properties and witness reduction, *J. Comput. System Sci.* **50** (1995), 412–432.

[GW87] T. Gundermann and G. Wechsung, Counting classes with finite acceptance types, *Comput. Artificial Intelligence* **6** (1987), 395–409.

[Gur08] L. Gurvits, Van der Waerden/Schrijver–Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all, in *Proc. 38th STOC* (2006), 417–426; *Electron. J. Combin.* **15** (2008), no. 1, RR 66, 26 pp.

[HI21] L. J. Haas and C. Ikenmeyer, Young flattenings in the Schur module basis, preprint (2021), 16 pp.; `arXiv: 2104.02363`.

[Har40] G. H. Hardy, *Ramanujan. Twelve lectures on subjects suggested by his life and work*, Cambridge Univ. Press, Cambridge, UK, 1940, 236 pp.

[HLP52] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities* (Second ed.), Cambridge Univ. Press, 1952, 324 pp.

[HL72] O. J. Heilmann and E. H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Phys.* **25** (1972), 190–243.

[HO02] L. A. Hemaspaandra and M. Ogiwara, *The complexity theory companion*, Springer, Berlin, 2002, 369 pp.

[HR00] L. A. Hemaspaandra and J. Rothe, A second step towards complexity-theoretic analogs of Rice's theorem, in *Proc. 23rd MFCS* (1998), 418–426; *Theoret. Comput. Sci.* **244** (2000), 205–217.

[HV95] L. A. Hemaspaandra and H. Vollmer, The satanic notations: counting classes beyond #P and other definitional adventures, *SIGACT News* **26** (1995), no. 1, 2–13.

[HVW95] U. Hertrampf, H. Vollmer and K. Wagner, On the power of number-theoretic operations with respect to counting, in *Proc. 10th CCC* (1995), 299–314.

[HT03] C. M. Homan and M. Thakur, One-way permutations and self-witnessing languages, in *Proc. 2nd IFIP TCS* (2002), 243–254; *J. Comput. System Sci.* **67** (2003), 608–622.

[HJ13] R. A. Horn and C. R. Johnson, *Matrix analysis* (Second ed.), Cambridge Univ. Press, Cambridge, 2013, 643 pp.

[Huh18] J. Huh, Combinatorial applications of the Hodge–Riemann relations, in *Proc. ICM Rio de Janeiro*, vol. IV, World Sci., Hackensack, NJ, 2018, 3093–3111.

[HL16] J. Hüttenhain and P. Lairez, The boundary of the orbit of the 3-by-3 determinant polynomial, *C. R. Math. Acad. Sci. Paris* **354** (2016), no. 9, 931–935.

[IK20] C. Ikenmeyer and U. Kandasamy, Implementing geometric complexity theory: On the separation of orbit closures via symmetries, in *Proc. 52nd STOC* (2020), 713–726.

[IMW17] C. Ikenmeyer, K. D. Mulmuley and M. Walter, On vanishing of Kronecker coefficients, *Comput. Complexity* **26** (2017), 949–992.

[Ike16] C. Ikenmeyer, Small Littlewood–Richardson coefficients, *J. Algebraic Cominatorics* **44** (2016), 1–29.

[IP17] C. Ikenmeyer and G. Panova, Rectangular Kronecker coefficients and plethysms in geometric complexity theory, in *Proc. 57th FOCS* (2016), 396–405; *Adv. Math.* **319** (2017), 40–66.

[IR82] K. F. Ireland and M. I. Rosen, *A classical introduction to modern number theory*, Springer, New York, 1982, 341 pp.

[Jeř16] E. Jeřábek, Integer factoring and modular square roots, *J. Comput. System Sci.* **82** (2016), 380–394.

[KM18] T. Kahle and M. Michałek, Obstructions to combinatorial formulas for plethysm, *Electron. J. Combin.* **25** (2018), no. 1, Paper 1.41, 9 pp.

[KVVY93] R. Kannan, H. Venkateswaran, V. Vinay and A. C. Yao, A circuit-based proof of Toda's theorem, *Inform. and Comput.* **104** (1993), 271–276.

[Kle66] D. J. Kleitman, Families of non-disjoint subsets, *J. Combin. Theory* **1** (1966), 153–155.

[Kly98] A. A. Klyachko, Stable bundles, representation theory and Hermitian operators, *Selecta Math.* **4** (1998), 419–445.

[Knu16] A. Knutson, Schubert calculus and puzzles, in *Schubert calculus*, Math. Soc. Japan, Tokyo, 2016, 185–209.

[KT99] A. Knutson and T. Tao, The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products I: Proof of the saturation conjecture, *Jour. AMS* **12** (1999), 1055–1090.

[KZ20] A. Knutson and P. Zinn-Justin, Schubert puzzles and integrability I: Invariant trilinear forms, preprint (2020), 51 pp.; `arXiv:1706.10019v6`.

[Ko85] K.-I. Ko, On some natural complete operators, *Theoret. Comp. Sci.* **37** (1985), 1–30.

[Kra96] C. Krattenthaler, Combinatorial proof of the log-concavity of the sequence of matching numbers, *J. Combin. Theory Ser. A* **74** (1996), 351–354.

[Kra99] A. Krawczyk, The complexity of finding a second Hamiltonian cycle in cubic graphs, *J. Comput. System Sci.* **58** (1999), 641–647.

[Kum15] S. Kumar, A study of the representations supported by the orbit closure of the determinant, *Compos. Math.* **151** (2015), 292–312.

[KPP94] A. G. Kuznetsov, I. Pak and A. E. Postnikov, Increasing trees and alternating permutations, *Russian Math. Surveys* **49** (1994), no. 6, 79–114.

[Lan15] J. M. Landsberg, Geometric complexity theory: an introduction for geometers, *Ann. Univ. Ferrara Sez. VII Sci. Mat.* **61** (2015), 65–117.

[LO15] J. M. Landsberg and G. Ottaviani, New lower bounds for the border rank of matrix multiplication, *Theory Comput.* **11** (2015), 285–298.

[LS82] A. Lascoux and M.-P. Schützenberger, Polynômes de Schubert (in French), *C. R. Acad. Sci. Paris Sér I, Math.* **294** (1982), 447–450.

[LR34] D. E. Littlewood and A. R. Richardson, Group characters and algebra, *Phil. Trans. Royal Soc. London, Ser. A* **233** (1934), 99–141.

[Loe11] N. A. Loehr, *Bijective combinatorics*, CRC Press, Boca Raton, FL, 2011, 590 pp.

[Mac91] I. G. Macdonald, *Notes on Schubert polynomials*, Publ. LaCIM, UQAM, Montreal, 1991, 116 pp.

[Mac95] I. G. Macdonald, *Symmetric functions and Hall polynomials* (Second ed.), Oxford U. Press, New York, 1995, 475 pp.

[Man01] L. Manivel, *Symmetric functions, Schubert polynomials and degeneracy loci*, SMF/AMS, Providence, RI, 2001, 167 pp.

[Mat10] J. Matoušek, *Thirty-three miniatures*, AMS, Providence, RI, 2010, 182 pp.

[MSS13] A. Marcus, D. A. Spielman and N. Srivastava, Interlacing families I: Bipartite Ramanujan graphs of all degrees, in *Proc. 54th FOCS* (2013), 529–537; *Ann. of Math.* **182** (2015), 307–325.

[MOA11] A. Marshall, I. Olkin and B. C. Arnold, *Inequalities: theory of majorization and its applications* (Second ed.), Springer, New York, 2011.

[Mar08] M. Marshall, *Positive polynomials and sums of squares*, AMS, Providence, RI, 2008, 187 pp.

[McK08] T. McKay, On plethysm conjectures of Stanley and Foulkes, *J. Algebra* **319** (2008), 2050–2071.

[MP91] N. Megiddo and C. H. Papadimitriou, On total functions, existence theorems and computational complexity, *Theoret. Comput. Sci.* **81** (1991), 317–324.

[MPP14] K. Mészáros, G. Panova and A. Postnikov, Schur times Schubert via the Fomin–Kirillov algebra, *Electron. J. Combin.* **21** (2014), no. 1, Paper 1.39, 22 pp.

[MM11] C. Moore and S. Mertens, *The nature of computation*, Oxford Univ. Press, Oxford, 2011, 985 pp.

[Mul07] K. D. Mulmuley, Geometric Complexity Theory VII: Nonstandard quantum group for the plethysm problem, preprint (2007), 59 pp.; arXiv:0709.0749.

[Mul09] K. D. Mulmuley, Geometric Complexity Theory VI: The flip via saturated and positive integer programming in representation theory and algebraic geometry, preprint (2009), 139 pp.; arXiv:0704.0229v4.

[MNS12] K. D. Mulmuley, H. Narayanan and M. Sohoni, Geometric Complexity Theory III: On Deciding Positivity of Littlewood–Richardson Coefficients, *J. Algebraic Cominatorics* **36** (2012), 103–110.

[MS08] K. D. Mulmuley and M. Sohoni, Geometric Complexity Theory II: Towards explicit obstructions for embeddings among class varieties, *SIAM J. Comput.* **38** (2008), 1175–1206.

[Mur38] F. D. Murnaghan, The analysis of the direct product of irreducible representations of the symmetric groups, *Amer. J. Math.* **60** (1938), 44–65.

[Nag19] T. Nagell, Über zahlentheoretische Polynome (in German), *Norsk. Mat. Tidsskr* **1** (1919), 14–23; available at https://tinyurl.com/3cv69sbm

[Nar95] W. Narkiewicz, *Polynomial mappings*, Springer, Berlin, 1995, 130 pp.

[Nes95] J. Nešetřil, Ramsey theory, in *Handbook of combinatorics*, Vol. 2 (Edited by R. L. Graham, M. Grötschel and L. Lovász), Elsevier, Amsterdam, 1995, 1331–1403.

[Noy14] M. Noy, Random planar graphs and beyond, in *Proc. ICM Seoul*, Vol. IV, 2014, 407–430.

[OH93] M. Ogiwara and L. A. Hemachandra, A complexity theory for feasible closure properties, in *Proc. 6th CCC* (1991), 16–29; *J. Comput. System Sci.* **46** (1993), 295–325.

[Pak03] I. Pak, Tile invariants: new horizons, *Theoret. Comput. Sci.* **303** (2003), 303–331.

[Pak06] I. Pak, Partition bijections, a survey, *Ramanujan J.* **12** (2006), 5–75.

[Pak18] I. Pak, Complexity problems in enumerative combinatorics, in *Proc. ICM Rio de Janeiro*, Vol. IV, World Sci., Hackensack, NJ, 2018, 3153–3180; an expanded version of the paper is available at arXiv:1803.06636 and tinyurl.com/2p8uxbzs

[Pak19] I. Pak, Combinatorial inequalities, *Notices AMS* **66** (2019), 1109–1112; an expanded version of the paper is available at tinyurl.com/py8sv5v6

[PP17] I. Pak and G. Panova, On the complexity of computing Kronecker coefficients, *Computational Complexity* **26** (2017), 1–36.

[PPY19] I. Pak, G. Panova and D. Yeliussizov, On the largest Kronecker and Littlewood–Richardson coefficients, *J. Combin. Theory, Ser. A* **165** (2019), 44–77.

[PPS20] I. Pak, F. Petrov and V. Sokolov, Hook inequalities, *Math. Intelligencer* **42** (2020), no. 2, 1–8.

[Pap90] C. H. Papadimitriou, On graph-theoretic lemmata and complexity classes, in *Proc. 31st FOCS* (1990), 794–801.

[Pap94a] C. H. Papadimitriou, On the complexity of the parity argument and other inefficient proofs of existence, *J. Comput. System Sci.* **48** (1994), 498–532.

[Pap94b] C. H. Papadimitriou, *Computational Complexity*, Addison-Wesley, Reading, MA, 1994, 523 pp.

[Pet72] J. Peterson, Beviser for Wilsons og Fermats Theoremer (in Danish, Proofs of the theorems of Wilson and Fermat), *Tidsskr. Math.* **2** (1872), 64–65.

[Pop62] K. R. Popper, *Conjectures and refutations: The growth of scientific knowledge*, Basic Books, New York, 1962, 412 pp.

[Pri77] W. L. Price, A topological transformation algorithm which relates the Hamiltonian circuits of a cubic planar map, *J. London Math. Soc.* **15** (1977), 193–196.

[RS80] G.-C. Rota and B. Sagan, Congruences derived from group action, *European J. Combin.* **1** (1980), 67–76.

[Roz19] U. A. Rozikov, *An introduction to mathematical billiards*, World Sci., Hackensack, NJ, 2019, 204 pp.

[Sam09] S. V. Sam, Computing inclusions of Schur modules, *J. Softw. Algebra Geom.* **1** (2009), 5–10.

[Sag85] B. E. Sagan, Congruences via abelian groups, *J. Number Theory* **20** (1985), 210–237.

[SY89] B. E. Sagan and Y. N. Yeh, Probabilistic algorithms for trees, *Fibonacci Quart.* **27** (1989), 201–208.

[Sch15] G. Schaeffer, Planar maps, in *Handbook of enumerative combinatorics*, CRC Press, Boca Raton, FL, 2015, 335–395.

[Sch14] R. Schneider, *Convex bodies: the Brunn–Minkowski theory* (Second ed.), Cambridge Univ. Press, Cambridge, UK, 2014, 736 pp.

[SvH19] Y. Shenfeld and R. van Handel, Mixed volumes and the Bochner method, *Proc. AMS* **147** (2019), 5385–5402.

[SvH20] Y. Shenfeld and R. van Handel, The extremals of the Alexandrov–Fenchel inequality for convex polytopes, *Acta Math.*, to appear, 82 pp.; arXiv:2011.04059.

[She82] L. A. Shepp, The XYZ conjecture and the FKG inequality, *Ann. Probab.* **10** (1982), 824–827.

[Spe11] D. E Speyer, How can I show a GapP problem is outside #P?, `CSTheoryStackExchange` (2011); csthe-ory.stackexchange.com/q/8178/

[Sta80] R. P. Stanley, Unimodal sequences arising from Lie algebras, in *Lecture Notes in Pure and Applied Math.* **57**, Dekker, New York, 1980, 127–136.

[Sta81] R. P. Stanley, Two combinatorial applications of the Aleksandrov–Fenchel inequalities, *J. Combin. Theory, Ser. A* **31** (1981), 56–65.

[Sta00] R. P. Stanley, Positivity problems and conjectures in algebraic combinatorics, in *Mathematics: frontiers and perspectives*, AMS, Providence, RI, 2000, 295–319.

[Sta12] R. P. Stanley, *Enumerative Combinatorics*, vol. 1 (Second ed.) and vol. 2, Cambridge Univ. Press, 2012 and 1999, 626 pp. and 581 pp.

[SW86] D. Stanton and D. White, *Constructive combinatorics*, Springer, New York, 1986, 183 pp.

[Syl82] J. J. Sylvester, A constructive theory of partitions, arranged in three acts, an interact and an exodion, *Amer. J. Math.* **5** (1882), 251–330.

[Tab05] S. Tabachnikov, *Geometry and billiards*, AMS, Providence, RI, 2005, 176 pp.

[Tar91] J. Tarui, Randomized polynomials, threshold circuits, and the polynomial hierarchy, in *Proc. 8th STACS* (1991), 238–250.

[Tho78] A. G. Thomason, Hamiltonian cycles and uniquely edge colourable graphs, *Ann. Discrete Math.* **3** (1978), 259–268.

[Toda91] S. Toda, PP is as hard as the polynomial-time hierarchy, *SIAM J. Comput.* **20** (1991), 865–877.

[Tut46] W. T. Tutte, On Hamiltonian circuits, *J. London Math. Soc.* **21** (1946), 98–101.

[Tut63] W. T. Tutte, A census of planar maps, *Canadian J. Math.* **15** (1963), 249–271.

[Val76] L. G. Valiant, Relative complexity of checking and evaluating, *Inf. Process. Lett.* **5** (1976), 20–23.

[Val79] L. G. Valiant, The complexity of computing the permanent, *Theor. Comput. Sci.* **8** (1979), 189–201.

[VV85] L. G. Valiant and V. V. Vazirani, NP is as easy as detecting unique solutions, in *Proc. 7th STOC* (1985), 458–463; *Theoret. Comput. Sci.* **47** (1986), 85–93.

[vL81] J. H. van Lint, Notes on Egoritsjev's proof of the van der Waerden conjecture, *Linear Algebra Appl.* **39** (1981), 1–8.

[vL82] J. H. van Lint, The van der Waerden conjecture: two proofs in one year, *Math. Intelligencer* **4** (1982), no. 2, 72–77.

[Vie16] X. G. Viennot, *The Art of Bijective Combinatorics*, a video-book (2016); viennot.org/abjc.html

[Wig19] A. Wigderson, *Mathematics and computation*, Princeton Univ. Press, Princeton, NJ, 2019, 418 pp.

[Wor18] N. Wormald, Asymptotic enumeration of graphs with given degree sequence, in *Proc. ICM Rio de Janeiro*, Vol. 3, 2018, 3229–3248.