Polynomials that vanish to high order on most of the hypercube

Lisa Sauermann[*]   Yuval Wigderson[†]

September 30, 2020

Abstract

Motivated by higher vanishing multiplicity generalizations of Alon's Combinatorial Nullstellensatz and its applications, we study the following problem: for fixed $k \geq 1$ and n large with respect to k, what is the minimum possible degree of a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ with $P(0, \ldots, 0) \neq 0$ such that P has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$? For $k = 1$, a classical theorem of Alon and Füredi states that the minimum possible degree of such a polynomial equals n. In this paper, we solve the problem for all $k \geq 2$, proving that the answer is $n + 2k - 3$. As an application, we improve a result of Clifton and Huang on configurations of hyperplanes in $\mathbb{R}^n$ such that each point in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$ is covered by at least k hyperplanes, but the point $(0, \ldots, 0)$ is uncovered. Surprisingly, the proof of our result involves Catalan numbers and arguments from enumerative combinatorics.

1 Introduction

Alon's Combinatorial Nullstellensatz [1], which gives a non-vanishing criterion for a polynomial on some grid of points under certain conditions, has had applications to many problems in combinatorics. The following statement due to Alon and Füredi [2] is a now classical example of an application of the Combinatorial Nullstellensatz (even though it historically predated it).

Theorem 1.1 (see [2]). *Let $n \geq 1$. Then any polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ with $P(0, \ldots, 0) \neq 0$ and such that P has zeroes at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$ has degree $\deg P \geq n$.*

The example $P = (x_1 - 1) \cdots (x_n - 1)$ shows that the bound $\deg P \geq n$ is tight. Alon and Füredi [2] proved this theorem in order to solve a problem of Komjáth [6] asking about the minimum possible number m such that there are m hyperplanes in $\mathbb{R}^n$ covering all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$, but not covering $(0, \ldots, 0)$. By considering the product of the linear polynomials defining these hyperplanes, Theorem 1.1 easily implies that at least n hyperplanes are needed. As n hyperplanes are also sufficient, the answer to Komjáth's hyperplane problem is $m = n$.

There has been a lot of work on finding generalizations of both the Combinatorial Nullstellensatz [1] and of Alon and Füredi's result [2] on Komjáth's hyperplane problem to higher vanishing or covering multiplicities [3, 4, 5, 7, 8]. In this spirit, it is natural to also ask for generalizations of Theorem 1.1 to higher vanishing orders for the polynomial P. This leads to the following problem.

Problem 1.2. Let $k \geq 2$ and let n be sufficiently large with respect to k. What is the minimum degree that a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ can have, if $P(0, \ldots, 0) \neq 0$ and if P has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$?

As usual, we say that a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ has a zero of multiplicity at least k at a point $a \in \mathbb{R}^n$ if all derivatives of P up to order $k - 1$ vanish at a. Note that $P(a) = 0$ is equivalent to P having a zero of multiplicity at least 1 at a.

Ball and Serra [3, Theorem 4.1] proved a lower bound of $\deg P \geq n + k - 1$ for Problem 1.2 (in fact, they proved a similar bound for more general grids instead of just $\{0, 1\}^n \subseteq \mathbb{R}^n$). Clifton and Huang [5] proved that the answer to Problem 1.2 is $n + 1$ if $k = 2$ and $n + 3$ if $k = 3$, and they improved the lower bound to $\deg P \geq n + k + 1$ for $k \geq 4$. Clifton and Huang were actually studying the generalization of Komjáth's hyperplane problem [6] mentioned above to higher covering multiplicities (where every point in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ needs to be covered by at least k hyperplanes, while $(0, \ldots, 0)$ must remain uncovered). Their approach was to consider the product of the polynomials for all of the hyperplanes and to prove a lower bound for the degree of this product. This naturally leads to Problem 1.2, even though their lower bound results are not explicitly stated in the setting of Problem 1.2.

In this paper, we resolve Problem 1.2, showing that the answer is $n + 2k - 3$. This is the content of the following theorem.

Theorem 1.3. *Let $k \geq 2$ and $n \geq 2k - 3$. Then any polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ with $P(0, \ldots, 0) \neq 0$ having zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ has degree $\deg P \geq n + 2k - 3$. Furthermore, there exists such a polynomial P with degree $\deg P = n + 2k - 3$.*

Instead of demanding $P(0, \ldots, 0) \neq 0$, one can also ask a more general version of Problem 1.2 where P is required to have a zero of multiplicity exactly ℓ at $(0, \ldots, 0)$ for some $\ell \in \{0, \ldots, k - 1\}$ (note that for $\ell \geq k$, we cannot expect any interesting lower bounds on the degree of P for large n, as the example $P = x_1^\ell (x_1 - 1)^k$ with $\deg P = \ell + k$ shows). Also note that the case $\ell = 0$ corresponds to Problem 1.2 above. Our next theorem resolves this more general problem for $0 \leq \ell \leq k - 2$.

Theorem 1.4. *Let $k \geq 2$ and $n \geq 2k - 3$. Let $P \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial having zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$, and such that P does not have a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$. Then P must have degree $\deg P \geq n + 2k - 3$. Furthermore, for every $\ell = 0, \ldots, k - 2$, there exists a polynomial P with degree $\deg P = n + 2k - 3$ having zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$, and such that P has a zero of multiplicity exactly ℓ at $(0, \ldots, 0)$.*

Note that Theorem 1.4 implies Theorem 1.3. Indeed, it is clear that the first part of Theorem 1.4 (giving the degree bound $\deg P \geq n + 2k - 3$) implies the first part of Theorem 1.3. Furthermore, the second part of Theorem 1.3 is equivalent to the second part of Theorem 1.4 for $\ell = 0$.

Theorem 1.4 does not address the case where P has a zero of multiplicity exactly $k - 1$ at $(0, \ldots, 0)$. This case turns out to have a slightly different answer, as shown in the following theorem.

Theorem 1.5. *Let $k \geq 2$ and $n \geq 1$. Let $P \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial having zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$, and a zero of multiplicity exactly $k - 1$ at $(0, \ldots, 0)$. Then P must have degree $\deg P \geq n + 2k - 2$. Furthermore, there exists such a polynomial P with degree $\deg P = n + 2k - 2$.*

Note that Theorem 1.5 is also true for $k = 1$, and is identical to Alon and Füredi's result in Theorem 1.1. Theorem 1.5 is actually significantly easier to prove then Theorems 1.3 and 1.4.

As an application of Theorem 1.3, we can improve a result of Clifton and Huang [5] concerning collections of hyperplanes in $\mathbb{R}^n$ such that every point in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ is covered by at least k of these hyperplanes, but no hyperplane contains $(0, \ldots, 0)$. Generalizing Komjáth's hyperplane problem [6], Clifton and Huang studied the minimum possible size of such a collection of hyperplanes, where k is fixed and n is assumed to be sufficiently large with respect to k. They proved that the minimum size of such a collection of hyperplanes is $n + 1$ if $k = 2$ and $n + 3$ if $k = 3$. Furthermore, for $k \geq 4$ they proved a lower bound of $n + k + 1$ for the size of any such collection of hyperplanes. As an immediate corollary of Theorem 1.3, we can recover their lower bound for $k \in \{2, 3, 4\}$ and improve it for $k \geq 5$.

Corollary 1.6. *Fix $k \geq 2$ and let $n \geq 2k - 3$. Consider a collection of hyperplanes in $\mathbb{R}^n$ such that every point in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ is covered by at least k of these hyperplanes, but no hyperplane contains $(0, \ldots, 0)$. Then this collection must consist of at least $n + 2k - 3$ hyperplanes.*

Indeed, Corollary 1.6 follows from Theorem 1.3 by taking the linear polynomials corresponding to all of the hyperplanes, and observing that their product P satisfies the assumptions in Theorem 1.3. This was also the approach taken by Clifton and Huang [5] to prove their lower bounds for the size of any such collection of hyperplanes. Clifton and Huang obtained lower bounds on the degree of the resulting product polynomial P by using a punctured higher-multiplicity version of the Combinatorial Nullstellensatz due to Ball and Serra [3] and a relatively involved analysis of the conditions on the coefficients of the polynomial P imposed by its vanishing properties. As mentioned above, their arguments also apply in the setting of Problem 1.2, but their bounds on the degree of P are weaker than the lower bound we obtain in Theorem 1.3.

It is still an interesting open problem to determine the minimum possible size of a collection of hyperplanes as in Corollary 1.6 for $k \geq 4$ (and n sufficiently large with respect to k). While Corollary 1.6 gives the best currently known lower bound, the best known upper bound is $n + \binom{k}{2}$, obtained by a construction due to Clifton and Huang [5]. They conjectured that this upper bound is tight; see also the discussion in Subsection 4.1 in the concluding remarks.

While Theorem 1.5 is relatively easy to prove, the proof of Theorem 1.4 takes up most of this paper. Surprisingly, the proof involves Catalan numbers and arguments from enumerative combinatorics (which is somewhat unexpected, given that Problem 1.2 is a problem in extremal combinatorics).

Notation. Throughout this paper, we work with the usual convention that the binomial coefficients $\binom{n}{m}$ are defined for all integers m and $n \geq 0$, but we have $\binom{n}{m} = 0$ unless $0 \leq m \leq n$. The variables $i, j, k, \ell, m, n, r, s, t$ always refer to integers.

2 Proof of Theorems 1.4 and 1.5

In this section we prove Theorem 1.5, and we also prove Theorem 1.4 apart from the proof of Proposition 2.9 below, which we postpone to Section 3. Recall that Theorem 1.4 implies Theorem 1.3.

For $k \geq 2$, let us say that a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ is *k-reduced*, if $\deg P \leq n + 2k - 3$ and if P does not contain any monomials divisible by $x_{i_1}^2 \cdots x_{i_k}^2$ for some (not necessarily distinct) indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$ (in other words, no monomial of P is divisible by a product of k squares of variables). Let $U_k \subseteq \mathbb{R}[x_1, \ldots, x_n]$ be the vector space of all k-reduced polynomials.

Before starting the proofs of Theorems 1.4 and 1.5, we will show a sequence of claims. The first claim shows that in order to prove the first part of Theorems 1.4 and 1.5, we can restrict ourselves to considering k-reduced polynomials P. Furthermore, it will be easy to derive the second part of Theorem 1.4 from this claim.

Claim 2.1. *Let $k \geq 2$ and $n \geq 1$. For every polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$, we can find a polynomial $P \in U_k$ with $\deg P \leq \deg Q$ such that the difference $Q - P$ has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$. Furthermore, if $\deg Q \leq n + 2k - 3$, then we can even choose $P \in U_k$ with $\deg P \leq \deg Q$ such that the difference $Q - P$ has zeroes of multiplicity at least k at all points in $\{0, 1\}^n$.*

Proof. Fix $k \geq 2$ and $n \geq 1$. Note that for any polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$, when defining a polynomial $Q^* \in \mathbb{R}[x_1, \ldots, x_n]$ by either

(a) $Q^* = Q - a \cdot x_{i_1}(x_{i_1} - 1) \cdots x_{i_k}(x_{i_k} - 1) \cdot x_1^{m_1} \cdots x_n^{m_n}$ with $a \in \mathbb{R}$, with non-negative integers $m_1, \ldots, m_n$ and with (not necessarily distinct) indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$, or

(b) $Q^* = Q - a \cdot x_{i_1}(x_{i_1} - 1) \cdots x_{i_{k-1}}(x_{i_{k-1}} - 1) \cdot (x_1 - 1) \cdots (x_n - 1)$ with $a \in \mathbb{R}$ and with (not necessarily distinct) indices $i_1, \ldots, i_{k-1} \in \{1, \ldots, n\}$,

3

the difference $Q - Q^*$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k-1$ at $(0,\ldots,0)$. Furthermore, if we define Q^* as in (a), then $Q - Q^*$ even has zeroes of multiplicity at least k at all points in $\{0,1\}^n$.

Let us say that a monomial in $\mathbb{R}[x_1,\ldots,x_n]$ is *bad* if it is if the form $x_{i_1}^2 \cdots x_{i_k}^2 \cdot x_1^{m_1} \cdots x_n^{m_n}$ with non-negative integers $m_1,\ldots,m_n$ and with (not necessarily distinct) indices $i_1,\ldots,i_k \in \{1,\ldots,n\}$, or of the form $x_{i_1}^2 \cdots x_{i_{k-1}}^2 \cdot x_1 \cdots x_n$ with (not necessarily distinct) indices $i_1,\ldots,i_{k-1} \in \{1,\ldots,n\}$.

Whenever $Q \in \mathbb{R}[x_1,\ldots,x_n]$ contains a bad monomial, we can apply one of the steps (a) or (b) above, replacing the bad monomial with monomials of lower degree. We can repeatedly perform these steps, always replacing a bad monomial of maximum degree by lower-degree monomials, until we arrive at a polynomial $P \in \mathbb{R}[x_1,\ldots,x_n]$ which does not contain any bad monomials. Indeed, note that when repeatedly applying these steps, the process must terminate at some point (since at every step either the maximum degree of the occurring bad monomials decreases or the maximum degree of the occurring bad monomials remains unchanged but the number of different bad monomials of this maximum degree decreases).

By construction of P, we have $\deg P \leq \deg Q$ and the difference $Q - P$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k-1$ at $(0,\ldots,0)$. Let us now check that $P \in U_k$, i.e. that P is k-reduced and has degree $\deg P \leq n + 2k - 3$.

Recall that P does not contain any bad monomials. This means in particular that P does not have any monomials which are divisible by $x_{i_1}^2 \cdots x_{i_k}^2$ for some (not necessarily distinct) indices $i_1,\ldots,i_k \in \{1,\ldots,n\}$. Note that this already implies that $\deg P \leq n + 2k - 2$. Furthermore, if P had a monomial of degree $n + 2k - 2$, then this monomial would need to be of the form $x_{i_1}^2 \cdots x_{i_{k-1}}^2 \cdot x_1 \cdots x_n$ with (not necessarily distinct) indices $i_1,\ldots,i_{k-1} \in \{1,\ldots,n\}$. But such a monomial is also bad, and therefore we must have $\deg P \leq n + 2k - 3$. This proves that $P \in U_k$.

Finally, for the second part of the claim, note that if we have $\deg Q \leq n + 2k - 3$, then throughout the process of obtaining P from Q all polynomials have degree at most $n + 2k - 3$ and we therefore never apply step (b). This means that we will only perform step (a), and hence $Q - P$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n$. $\qquad\square$

For $n \geq 1$ and $t \geq 1$, let us define $M_t(n)$ to be the number of n-tuples $(m_1,\ldots,m_n)$ of non-negative integers with $m_1 + \cdots + m_n < t$. We remark that $M_t(n) = \binom{n+t-1}{n}$, but we will not use this formula. Note that for a polynomial $P \in \mathbb{R}[x_1,\ldots,x_n]$ there are precisely $M_t(n)$ different ways to form a derivative of P of order less than t.

We will be using various dimension-counting arguments. The following claim expresses the dimension of the vector space U_k in terms of the numbers $M_t(n)$ we just defined.

Claim 2.2. *Let $k \geq 2$ and $n \geq 1$. Then the vector space $U_k \subseteq \mathbb{R}[x_1,\ldots,x_n]$ has dimension $\dim U_k = (2^n - 1) \cdot M_k(n) + M_{k-1}(n)$.*

Proof. The dimension of the vector space U_k equals the number of monomials in $\mathbb{R}[x_1,\ldots,x_n]$ of degree at most $n + 2k - 3$ which are not divisible by $x_{i_1}^2 \cdots x_{i_k}^2$ for any (not necessarily distinct) indices $i_1,\ldots,i_k \in \{1,\ldots,n\}$. In other words, $\dim U_k$ is the number of monomials of the form $x_1^{2m_1+r_1} \cdots x_n^{2m_n+r_n}$ with non-negative integers $(m_1,\ldots,m_n)$ with $m_1 + \cdots + m_n < k$ and $r_1,\ldots,r_n \in \{0,1\}$ such that $(2m_1+r_1) + \cdots + (2m_n+r_n) \leq n+2k-3$. If $r_1 + \cdots + r_n \leq n-1$, then there are precisely $M_k(n)$ choices for $(m_1,\ldots,m_n)$. But if $r_1 = \cdots = r_n = 1$, then we must have $m_1 + \cdots + m_n < k-1$, since $m_1 + \cdots + m_n = k - 1$ would imply $(2m_1 + r_1) + \cdots + (2m_n + r_n) = n + 2k - 2 > n + 2k - 3$. Hence, in this case, there are only $M_{k-1}(n)$ choices for $(m_1,\ldots,m_n)$. All in all we obtain $\dim U_k = (2^n - 1) \cdot M_k(n) + M_{k-1}(n)$, as desired. $\qquad\square$

Claim 2.3. *Let $k \geq 2$ and $n \geq 1$, and let $N = (2^n - 1) \cdot M_k(n) + M_{k-1}(n)$. Consider the linear map $\psi_k : U_k \to \mathbb{R}^N$ sending each polynomial $P \in U_k$ to the N-tuple consisting of the derivatives of P of order less than k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and all the derivatives of P of order less than $k-1$ at $(0,\ldots,0)$. Then $\psi_k : U_k \to \mathbb{R}^N$ is an isomorphism.*

4

Proof. By Claim 2.2 we have $\dim U_k = (2^n - 1) \cdot M_k(n) + M_{k-1}(n) = N$, so it suffices to prove that the linear map $\psi_k : U_k \to \mathbb{R}^N$ is surjective. The surjectivity of ψ_k follows from Claim 2.1: indeed, for any N-tuple $\alpha \in \mathbb{R}^N$, we can easily construct a (high-degree) polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ such that the derivatives of Q of order less than k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and the derivatives of Q of order less than $k - 1$ at $(0,\ldots,0)$ form precisely the N-tuple α. Now, by Claim 2.1 there exists a polynomial $P \in U_k$ where this N-tuple of derivatives agrees with the N-tuple for Q. Hence $\psi_k(P) = \alpha$. $\square$

Corollary 2.4. *Let $k \geq 2$ and $n \geq 1$. For every polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ there is at most one polynomial $P \in U_k$ such that $Q - P$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k - 1$ at $(0,\ldots,0)$. In particular, the polynomial P in Claim 2.1 is unique.*

Proof. Suppose that for some polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ there were two different polynomials $P, P^* \in U_k$ with the desired properties. Then the non-zero polynomial $P - P^* \in U_k$ would have zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k - 1$ at $(0,\ldots,0)$, and would therefore be mapped to $(0,0,\ldots,0) \in \mathbb{R}^N$ under the isomorphism $\psi_k : U_k \to \mathbb{R}^N$ in Claim 2.3. This is a contradiction. $\square$

With these preparations, we are now ready to prove Theorem 1.5.

Proof of Theorem 1.5. Suppose for contradiction that there is a polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $\deg Q \leq n + 2k - 3$ having zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$, and a zero of multiplicity exactly $k - 1$ at $(0,\ldots,0)$. By the second part of Claim 2.1, there is a polynomial $P \in U_k$ also having zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$, and a zero of multiplicity exactly $k - 1$ at $(0,\ldots,0)$. In particular, P is not the zero polynomial. However, under the isomorphism $\psi_k : U_k \to \mathbb{R}^N$ in Claim 2.3, the polynomial P is mapped to $(0,0,\ldots,0) \in \mathbb{R}^N$ (recall that the derivatives of order $k - 1$ at $(0,\ldots,0)$ are not recorded by ψ_k). This is a contradiction.

For the second part of the theorem, note that the polynomial

$$P(x_1, \ldots, x_n) = x_1^{k-1}(x_1 - 1)^{k-1} \cdot (x_1 - 1) \cdots (x_n - 1)$$

has degree $\deg P = n + 2k - 2$ and satisfies the desired conditions. $\square$

It remains to prove Theorem 1.4. The second part of Theorem 1.4 (concerning the existence of P with $\deg P = n + 2k - 3$ with the desired conditions) is a direct consequence of Claim 2.1, as we will see below. However, the first part of Theorem 1.4 is much more challenging.

For $k \geq 2$, let $V_k \subseteq U_k \subseteq \mathbb{R}[x_1, \ldots, x_n]$ be the vector space of all k-reduced polynomials which have zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$. In order to prove the first part of Theorem 1.4, our goal will be to show that all non-zero polynomials $P \in V_k$ have degree $n + 2k - 3$. We continue with our sequence of claims.

Claim 2.5. *Let $k \geq 2$ and $n \geq 1$. Then the vector space $V_k \subseteq \mathbb{R}[x_1, \ldots, x_n]$ has dimension $M_{k-1}(n)$.*

Proof. Note that V_k consists precisely of those polynomials in U_k which are mapped to N-tuples under the isomorphism $\psi_k : U_k \to \mathbb{R}^N$ in Claim 2.3 where the first $(2^n - 1) \cdot M_k(n)$ entries are zero. Since the subspace of such N-tuples has dimension $N - (2^n - 1) \cdot M_k(n) = M_{k-1}(n)$, the subspace $V_k \subseteq U_k$ also has dimension $M_{k-1}(n)$. $\square$

Claim 2.6. *Let $k \geq 3$ and $n \geq 1$. Then for each $j \in \{1, \ldots, n\}$ and each polynomial $P \in V_{k-1}$, we have $x_j(x_j - 1) \cdot P \in V_k$.*

Proof. Recall that $P \in V_{k-1} \subseteq U_{k-1}$ is $(k-1)$-reduced, meaning that $\deg P \leq n + 2k - 5$ and no monomial of P is divisible by $x_{i_1}^2 \cdots x_{i_{k-1}}^2$ for any (not necessarily distinct) indices $i_1, \ldots, i_{k-1} \in \{1, \ldots, n\}$. Then clearly the polynomial $x_j(x_j - 1) \cdot P$ has degree at most $n + 2k - 3$, and it is also not hard to see that no monomial of $x_j(x_j - 1) \cdot P$ is divisible by $x_{i_1}^2 \cdots x_{i_k}^2$ for any (not necessarily distinct) indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Hence $x_j(x_j - 1) \cdot P \in U_k$.

Since $P \in V_{k-1}$, the polynomial P has zeroes of multiplicity at least $k - 1$ at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$. Hence $x_j(x_j - 1) \cdot P$ has zeroes of multiplicity at least k at all these points, and we can conclude that $x_j(x_j - 1) \cdot P \in V_k$. $\square$

Now, in order to show that all non-zero polynomials $P \in V_k$ have degree $n + 2k - 3$, let us consider the linear map $\varphi_k : V_k \to \mathbb{R}[x_1, \ldots, x_n]$ sending each polynomial in V_k to its homogeneous degree $n + 2k - 3$ part. Our goal is to show that this map is injective. To show this, it is sufficient to prove that the image $\varphi_k(V_k)$ has dimension $\dim V_k = M_{k-1}(n)$.

Let $W_k \subseteq \mathbb{R}[x_1, \ldots, x_n]$ be the subspace spanned by all polynomials of the form

$$x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_n^{2d_n} \tag{2.1}$$

for non-negative integers $(m, d_1, \ldots, d_n)$ with $m + 2(d_1 + \cdots + d_n) = 2k - 3$. Note that all polynomials in W_k are homogeneous of degree $n + 2k - 3$. We will later show that $\varphi_k(V_k) = W_k$.

Claim 2.7. *Let $k \geq 2$ and $n \geq k - 1$. Then the polynomials in (2.1) form a basis of W_k.*

Proof. By definition of W_k, the polynomials in (2.1) span W_k. It remains to show that these polynomials are linearly independent. Since we can divide all the polynomials by $x_1 \cdots x_n$, it suffices to show that the polynomials of the form

$$Q_{m,d_1,\ldots,d_n} = (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_n^{2d_n}$$

for non-negative integers $(m, d_1, \ldots, d_n)$ with $m + 2(d_1 + \cdots + d_n) = 2k - 3$ are linearly independent.

Note that for each such polynomial $Q_{m,d_1,\ldots,d_n}$ we have that m is odd and that $d_1 + \cdots + d_n \leq k - 2 < n$. Hence all monomials of $Q_{m,d_1,\ldots,d_n}$ have exactly one variable with an odd exponent, this odd exponent is always at least m, and there exists some monomial where the odd exponent is equal to m (namely each of the monomials $x_i^m \cdot x_1^{2d_1} \cdots x_n^{2d_n}$ for those $i \in \{1, \ldots, n\}$ with $d_i = 0$, which must exist since $d_1 + \cdots + d_n < n$).

Now suppose we had a linear dependence relationship

$$\sum_{(m,d_1,\ldots,d_n)} \lambda_{m,d_1,\ldots,d_n} Q_{m,d_1,\ldots,d_n} = 0$$

for some coefficients $\lambda_{m,d_1,\ldots,d_n} \in \mathbb{R}$ not all of which are zero. Then let m^* be the minimum value of m for which some coefficient $\lambda_{m,d_1,\ldots,d_n}$ is non-zero (and note that m^* is odd, since m is always odd). Furthermore, fix $i \in \{1, \ldots, n\}$ such that we have $\lambda_{m,d_1,\ldots,d_n} \neq 0$ for some $(m, d_1, \ldots, d_n)$ with $m = m^*$ and $d_i = 0$.

Now, in the polynomials $Q_{m,d_1,\ldots,d_n}$ with $\lambda_{m,d_1,\ldots,d_n} \neq 0$ all monomials have exactly one variable with an odd exponent, and this odd exponent is always at least m^*. Let us consider all the monomials where x_i has exponent m^*. These monomials can only appear in polynomials $Q_{m,d_1,\ldots,d_n}$ with $\lambda_{m,d_1,\ldots,d_n} \neq 0$ where $m = m^*$ and $d_i = 0$. But note that each such polynomial contains precisely one monomial where x_i has exponent $m = m^*$, namely the monomial $x_i^{m^*} \cdot x_1^{2d_1} \cdots x_n^{2d_n}$. Hence these monomials cannot cancel out between different polynomials $Q_{m,d_1,\ldots,d_n}$ with $\lambda_{m,d_1,\ldots,d_n} \neq 0$. Thus, a linear dependence relationship between the polynomials $Q_{m,d_1,\ldots,d_n}$ is not possible. $\square$

Corollary 2.8. *Let $k \geq 2$ and $n \geq k - 1$. Then the vector space $W_k \subseteq \mathbb{R}[x_1, \ldots, x_n]$ has dimension $M_{k-1}(n)$.*

6

Proof. By Claim 2.7, the dimension of W_k equals the number of polynomials of the form as in (2.1), i.e. the number of choices of non-negative integers $(m, d_1, \ldots, d_n)$ with $m + 2(d_1 + \cdots + d_n) = 2k - 3$. We must have $d_1 + \cdots + d_n < k - 1$. So there are $M_{k-1}(n)$ possibilities to choose $(d_1, \ldots, d_n)$, and for each of them $m = 2k - 3 - 2(d_1 + \cdots + d_n)$ is uniquely determined. $\qquad \square$

A key step in the proof of Theorem 1.4 is to prove the following proposition, from which we will deduce that the map $\varphi_k : V_k \to \mathbb{R}[x_1, \ldots, x_n]$ is injective.

Proposition 2.9. *Let $k \geq 2$ and $n \geq 2k - 3$. Then there exists a polynomial $P \in V_k$ such that the homogeneous degree $n + 2k - 3$ part $\varphi_k(P)$ of P satisfies $\varphi_k(P) \in W_k$ as well as the following condition: if we write $\varphi_k(P) \in W_k$ in terms of the basis in (2.1), then the coefficient of the basis element $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$ is non-zero.*

We postpone the proof of Proposition 2.9 to Section 3. The proposition implies the following corollaries.

Corollary 2.10. *Let $k \geq 2$ and $n \geq 2k - 3$. Then we have $W_k \subseteq \varphi_k(V_k)$.*

Proof. We prove the corollary by induction on k. For the base case $k = 2$ note that the space W_2 is the one-dimensional space spanned by the polynomial $x_1 \cdots x_n \cdot (x_1 + \cdots + x_n)$. By Proposition 2.9 for $k = 2$, there exists a polynomial $P \in U_2$ such that $\varphi_2(P) \in W_2$ is a non-zero scalar multiple of $x_1 \cdots x_n \cdot (x_1 + \cdots + x_n)$. This shows that $W_2 \subseteq \varphi_2(V_2)$.

Now let $k \geq 3$ and assume that we have already shown that $W_{k-1} \subseteq \varphi_{k-1}(V_{k-1})$. Let $W_k' \subseteq W_k$ be the subspace spanned by all polynomials of the form (2.1) with $d_1 + \cdots + d_n \geq 1$. In other words, $W_k' \subseteq W_k$ is spanned by all of the form (2.1) except the polynomial $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$. Note that $\dim W_k' = \dim W_k - 1$.

Let us first show that $W_k' \subseteq \varphi_k(V_k)$. So fix a polynomial $x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_n^{2d_n}$ of the form (2.1) where $d_1 + \cdots + d_n \geq 1$. Now, let $j \in \{1, \ldots, n\}$ be chosen such that $d_j \geq 1$. Note that then (using the induction hypothesis $W_{k-1} \subseteq \varphi_{k-1}(V_{k-1})$) we have

$$x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_{j-1}^{2d_{j-1}} x_j^{2d_j - 2} x_{j+1}^{2d_{j+1}} \cdots x_n^{2d_n} \in W_{k-1} \subseteq \varphi_{k-1}(V_{k-1}).$$

Hence there exists a polynomial $P^* \in V_{k-1}$ of degree $n + 2k - 5$ whose homogeneous degree $n + 2k - 5$ part is $\varphi_{k-1}(P^*) = x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_{j-1}^{2d_{j-1}} x_j^{2d_j - 2} x_{j+1}^{2d_{j+1}} \cdots x_n^{2d_n}$. By Claim 2.6 we have $x_j(x_j - 1) \cdot P^* \in V_k$, and note that the homogeneous degree $n + 2k - 3$ part of $x_j(x_j - 1) \cdot P^*$ is precisely $x_j^2 \cdot \varphi_{k-1}(P^*) = x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_n^{2d_n}$. Hence φ_k maps the polynomial $x_j(x_j - 1) \cdot P^* \in V_k$ to $x_1 \cdots x_n \cdot (x_1^m + \cdots + x_n^m) \cdot x_1^{2d_1} \cdots x_n^{2d_n}$. This establishes that $W_k' \subseteq \varphi_k(V_k)$.

Thus, $\varphi_k(V_k) \cap W_k$ is a subspace of W_k with $W_k' \subseteq \varphi_k(V_k) \cap W_k \subseteq W_k$. Now, Proposition 2.9 states that there exists a polynomial $P \in V_k$ with $\varphi_k(P) \in W_k$ and $\varphi_k(P) \notin W_k'$. Hence $\varphi_k(V_k) \cap W_k \neq W_k'$. Since $\dim W_k' = \dim W_k - 1$, this implies that $\varphi_k(V_k) \cap W_k = W_k$ and consequently $W_k \subseteq \varphi_k(V_k)$. $\qquad \square$

Corollary 2.11. *Let $k \geq 2$ and $n \geq 2k - 3$. Then $\varphi_k(V_k) = W_k$, and the linear map $\varphi_k : V_k \to W_k$ is an isomorphism.*

Proof. Recall that $\dim V_k = M_{k-1}(n) = \dim W_k$ by Claim 2.5 and Corollary 2.8. Furthermore recall that by Corollary 2.10 we have $W_k \subseteq \varphi_k(V_k)$. Since $\dim \varphi_k(V_k) \leq \dim V_k = \dim W_k$, we must have $\varphi_k(V_k) = W_k$ and the linear map $\varphi_k : V_k \to W_k$ must be an isomorphism. $\qquad \square$

Using Corollary 2.11, we are now ready to prove Theorem 1.4.

Proof of Theorem 1.4. Fix $k \geq 2$ and $n \geq 2k - 3$, and let $Q \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial having zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$, and such that Q does not have a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$. We need to show that $\deg Q \geq n + 2k - 3$. Let P be a polynomial as in Claim 2.1. Then we have $P \in U_k$ and $\deg P \leq \deg Q$. Furthermore, P also has zeroes

7

of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$, which implies that $P \in V_k$. Finally, P does not have a zero of multiplicity at least $k-1$ at $(0,\ldots,0)$, which in particular means that P is not the zero polynomial. Recall that by Corollary 2.11 the map $\varphi_k : V_k \to W_k$ is an isomorphism (and in particular injective). Thus, we have $\varphi_k(P) \neq 0$, which means that the homogeneous degree $n + 2k - 3$ part of P is non-zero. So we can conclude that $\deg Q \geq \deg P \geq n + 2k - 3$.

It remains to prove the second part of Theorem 1.4. Fix $0 \leq \ell \leq k-2$. It is not hard to construct a (high-degree) polynomial Q having zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$, and such that Q has a zero of multiplicity exactly ℓ at $(0,\ldots,0)$, for example the polynomial $Q = x_1^\ell \cdot (x_1 - 1)^k \cdots (x_n - 1)^k$. By Claim 2.1, there exists a polynomial $P \in U_k$ which also has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$, and a zero of multiplicity exactly ℓ at $(0,\ldots,0)$ (recall that $\ell \leq k-2$). It remains to check that $\deg P = n + 2k - 3$. Since $P \in U_k$, we must have $\deg P \leq n + 2k - 3$. Furthermore, the first part of Theorem 1.4 proved above implies that $\deg P \geq n + 2k - 3$. Hence $\deg P = n + 2k - 3$, and the polynomial P has all of the desired properties. $\qquad\square$

3 Proof of Proposition 2.9

In this section, we will prove Proposition 2.9 by constructing a polynomial $P \in V_k$ satisfying the desired conditions. Our polynomial P will be a symmetric polynomial.

Recall that we can write any symmetric polynomial $R \in \mathbb{R}[x_1,\ldots,x_n]$ in terms of the power sum symmetric polynomials $x_1^m + \cdots + x_n^m$ for $1 \leq m \leq n$. More precisely, we can write R as a linear combination of products of power sum symmetric polynomials, i.e. as a linear combination of terms of the form $(x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ for positive integers $m_1,\ldots,m_\ell$.

Claim 3.1. *Let $n \geq 1$. Suppose that $d_1,\ldots,d_t$ are positive integers with $d_1 + \cdots + d_t \leq n$, such that exactly one of $d_1,\ldots,d_t$ is odd. Then the symmetric polynomial*

$$\sum_{\substack{i_1,\ldots,i_t \in \{1,\ldots,n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t}$$

is a linear combination of terms of the form $(x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ for $m_1,\ldots,m_\ell \in \{1,\ldots,n\}$ with $m_1 + \cdots + m_\ell = d_1 + \cdots + d_t$ and such that exactly one of $m_1,\ldots,m_\ell$ is odd.

Proof. Let us prove the claim by induction on t. If $t = 1$, then d_1 is odd and the symmetric polynomial $x_1^{d_1} + \cdots + x_n^{d_1}$ is already of the desired form. So let us now assume that $t \geq 2$ and that the claim holds for $t - 1$. As only one of $d_1,\ldots,d_t$ is odd, we may assume without loss of generality that d_t is even. Note that

$$\sum_{\substack{i_1,\ldots,i_t \in \{1,\ldots,n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t} = (x_1^{d_t} + \cdots + x_n^{d_t}) \cdot \sum_{\substack{i_1,\ldots,i_{t-1} \in \{1,\ldots,n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_{t-1}}^{d_{t-1}}$$
$$- \sum_{s=1}^{t-1} \sum_{\substack{i_1,\ldots,i_{t-1} \in \{1,\ldots,n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_{s-1}}^{d_{s-1}} x_{i_s}^{d_s + d_t} x_{i_{s+1}}^{d_{s+1}} \cdots x_{i_{t-1}}^{d_{t-1}}.$$

As exactly one of $d_1,\ldots,d_{t-1}$ is odd, we can apply the induction hypothesis to the sum $\sum x_{i_1}^{d_1} \cdots x_{i_{t-1}}^{d_{t-1}}$. We can then conclude that $(x_1^{d_t} + \cdots + x_n^{d_t}) \cdot \sum x_{i_1}^{d_1} \cdots x_{i_{t-1}}^{d_{t-1}}$ is a linear combination of terms of the form $(x_1^{d_t} + \cdots + x_n^{d_t}) \cdot (x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ for positive integers $m_1,\ldots,m_\ell$ with $d_t + m_1 + \cdots + m_\ell = d_1 + \cdots + d_t$ and such that exactly one of $d_t, m_1,\ldots,m_\ell$ is odd. Furthermore, for each $s = 1,\ldots,t-1$ we can also apply the induction hypothesis to the sum $\sum x_{i_1}^{d_1} \cdots x_{i_{s-1}}^{d_{s-1}} x_{i_s}^{d_s + d_t} x_{i_{s+1}}^{d_{s+1}} \cdots x_{i_{t-1}}^{d_{t-1}}$ and we find that each of these sums is a linear combination of terms of the desired form as well. $\qquad\square$

Claim 3.2. *Let $k \geq 2$ and let $n \geq 2k - 3$. Let $P \in U_k$ be a symmetric polynomial and let $\overline{P}$ be the homogeneous degree $n + 2k - 3$ part of P. Suppose that the polynomial $\overline{P}$ is divisible by $x_1 \cdots x_n$. Then we have $\overline{P} \in W_k$. Furthermore, when we write $\overline{P}$ in terms of the basis of W_k in (2.1), then the coefficient of $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$ is the same as the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when writing the symmetric polynomial $\overline{P}/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials.*

Proof. Recall that $P \in U_k$ does not contain any monomials which are divisible by $x_{i_1}^2 \cdots x_{i_k}^2$ for some (not necessarily distinct) indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Therefore any monomial of P must be of the form $x_1^{2m_1 + r_1} \cdots x_n^{2m_n + r_n}$ with non-negative integers $(m_1, \ldots, m_n)$ with $m_1 + \cdots + m_n \leq k - 1$ and $r_1, \ldots, r_n \in \{0, 1\}$. If $(2m_1 + r_1) + \cdots + (2m_n + r_n) = n + 2k - 3$, then this is only possible if $m_1 + \cdots + m_n = k - 1$ and $r_1 + \cdots + r_n = n - 1$. This means that any monomial of P of degree $n + 2k - 3$ must have an odd exponent for exactly $n - 1$ of the n variables $x_1, \ldots, x_n$.

Now let $R = \overline{P}/(x_1 \cdots x_n)$, and note that $R \in \mathbb{R}[x_1, \ldots, x_n]$ is homogeneous of degree $2k - 3$. Furthermore, since P is symmetric, the polynomials $\overline{P}$ and R are also symmetric. We saw above that every monomial in $\overline{P}$ has odd exponents for exactly $n - 1$ of the variables $x_1, \ldots, x_n$ (and an even exponent for the remaining variable). Hence every monomial in R has an odd exponent for exactly one variable. Thus, we can apply Claim 3.1 to all different types of monomials appearing in R (recalling that $n \geq 2k - 3$). We obtain that R is a linear combination of terms of the form $(x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ for $m_1, \ldots, m_\ell \in \{1, \ldots, n\}$ with $m_1 + \cdots + m_\ell = 2k - 3$ and such that exactly one of $m_1, \ldots, m_\ell$ is odd. By renaming $m_1, \ldots, m_\ell$, we can always assume that m_1 is odd and $m_2, \ldots, m_\ell$ are even.

Note that this representation of R as a linear combination of such terms is the representation of R in terms of power sum symmetric polynomials. Hence the coefficient λ of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when expressing $\overline{P}/(x_1 \cdots x_n) = R$ in terms of power sum symmetric polynomials is precisely the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when expressing R as a linear combination as above.

By multiplying with $x_1 \cdots x_n$, we can now express $\overline{P} = x_1 \cdots x_n \cdot R$ as a linear combination of terms of the form $x_1 \cdots x_n \cdot (x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ for $m_1, \ldots, m_\ell \in \{1, \ldots, n\}$ with $m_1 + \cdots + m_\ell = 2k - 3$ such that m_1 is odd and $m_2, \ldots, m_\ell$ are even. Note that the only such term with $m_1 = 2k - 3$ is the term $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$, and the coefficient of this term equals λ.

Note that each of these terms $x_1 \cdots x_n \cdot (x_1^{m_1} + \cdots + x_n^{m_1}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ (with $m_1 + \cdots + m_\ell = 2k - 3$ and such that m_1 is odd and $m_2, \ldots, m_\ell$ are even) can be written as a linear combination of terms of the form (2.1) with $m = m_1$ (indeed, when multiplying out $(x_1^{m_2} + \cdots + x_n^{m_2}) \cdots (x_1^{m_\ell} + \cdots + x_n^{m_\ell})$ all variables always appear with even exponents). All in all, this shows that $\overline{P}$ can be written as a linear combination of terms of the form (2.1), so $\overline{P} \in W_k$. Furthermore, when writing $\overline{P}$ in terms of the basis of W_k in (2.1), the coefficient of $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$ equals λ, which we defined as the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when expressing $\overline{P}/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials. $\square$

Let us now prove Proposition 2.9. The proof of the proposition depends on three lemmas, whose proofs we will postpone to the next two subsections.

Proof of Proposition 2.9. Fix $k \geq 2$ and $n \geq 2k - 3$. In order to construct the desired polynomial $P \in V_k$, let us first define the symmetric polynomial

$$Q(x_1, \ldots, x_n) = (-1)^{(k-1)n} \cdot (x_1 - 1)^k \cdots (x_n - 1)^k. \tag{3.1}$$

Note that Q has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$.

Let us now define $P \in U_k$ to be the unique polynomial in U_k such that $Q - P$ has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$ (such a polynomial P exists by Claim 2.1 and it is unique by Corollary 2.4). Then P has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$, and hence $P \in V_k$. Furthermore, since Q is

symmetric and P is unique, P must also be a symmetric polynomial (otherwise we could permute the variables in P and obtain another polynomial with the desired properties, contradicting uniqueness).

Recall that $\varphi_k(P)$ is the homogeneous degree $n + 2k - 3$ part of P. We need to show that $\varphi_k(P) \in W_k$ and that when writing $\varphi_k(P) \in W_k$ in terms of the basis in (2.1), then the coefficient of the basis element $x_1 \cdots x_n \cdot (x_1^{2k-3} + \cdots + x_n^{2k-3})$ is non-zero.

We wish to apply Claim 3.2 to P, but this requires the homogeneous degree $n + 2k - 3$ part $\varphi_k(P)$ of P to be divisible by $x_1 \cdots x_n$.

Lemma 3.3. *The polynomial $\varphi_k(P)$ is divisible by $x_1 \cdots x_n$.*

We postpone the proof of Lemma 3.3 to Subsection 3.1. Assuming Lemma 3.3, by Claim 3.2 we obtain $\varphi_k(P) \in W_k$, which establishes the first of our two required properties for P. Furthermore, we also obtain from Claim 3.2 that the desired coefficient of $\varphi_k(P)$ (when writing it in terms of the basis elements of W_k) is equal to the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when writing $\varphi_k(P)/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials. So it remains to prove that this coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ is non-zero. The following lemma determines this coefficient.

Lemma 3.4. *When writing $\varphi_k(P)/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials, the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ is equal to*

$$\sum_{(m_1,\ldots,m_t)} (-1)^t \cdot \binom{k-1-m_1}{m_1-1} \binom{k-1-m_2}{m_2} \cdots \binom{k-1-m_t}{m_t},$$

where the sum is over all sequences $(m_1, \ldots, m_t)$ of positive integers with $m_1 + \cdots + m_t = k - 1$.

We also postpone the proof of Lemma 3.4 to Subsection 3.1. Surprisingly, the value of the sum in Lemma 3.4 is (up to its sign) equal to the Catalan number C_{k-2}, see the following lemma. Recall that the Catalan numbers $C_0, C_1, \ldots$ are given by the explicit formula $C_n = \binom{2n}{n}/(n+1)$ for all $n \geq 0$.

Lemma 3.5. *For any $\ell \geq 1$, we have*

$$\sum_{(m_1,\ldots,m_t)} (-1)^t \cdot \binom{\ell-m_1}{m_1-1} \binom{\ell-m_2}{m_2} \cdots \binom{\ell-m_t}{m_t} = (-1)^\ell C_{\ell-1},$$

where the sum is over all sequences $(m_1, \ldots, m_t)$ of positive integers with $m_1 + \cdots + m_t = \ell$.

We postpone the proof of Lemma 3.5 to Subsection 3.2. Applying the lemma with $\ell = k - 1$, we see that the sum in Lemma 3.4 equals $(-1)^{k-1} C_{k-2} = (-1)^{k-1} \binom{2k-4}{k-2}/(k-1) \neq 0$. Thus, the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when writing $\varphi_k(P)/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials is indeed non-zero, finishing the proof of the Proposition 2.9. $\qquad\square$

3.1 Proof of Lemmas 3.3 and 3.4

In this subsection, we prove Lemmas 3.3 and 3.4 by analyzing the polynomial P and its homogeneous degree $n + 2k - 3$ part $\varphi_k(P)$. As before, let $k \geq 2$ and $n \geq 2k - 3$ be fixed.

In order to calculate $\varphi_k(P)$, let us first rewrite the definition of the polynomial Q in (3.1) as

$$Q(x_1, \ldots, x_n) = \prod_{i=1}^{n} \left((-1)^{k-1} \cdot (x_i - 1)^k \right).$$

For $i = 1, \ldots, n$, let us now write $(-1)^{k-1} \cdot (x_i - 1)^k$ as a linear combination of $(x_i - 1), x_i(x_i - 1),$ $x_i(x_i - 1)^2, x_i^2(x_i - 1)^2, x_i^2(x_i - 1)^3, \ldots$, as in the following claim.

Claim 3.6. *For any $\ell \geq 1$ and $i \in \{1, \ldots, n\}$, we have*

$$(-1)^{\ell-1} \cdot (x_i - 1)^\ell = \sum_{m=0}^{\lfloor (\ell-1)/2 \rfloor} \binom{\ell-1-m}{m} x_i^m (x_i - 1)^{m+1} - \sum_{m=1}^{\lfloor \ell/2 \rfloor} \binom{\ell-1-m}{m-1} x_i^m (x_i - 1)^m.$$

Proof. First, note that we may equivalently write the desired equation as

$$(-1)^{\ell-1} \cdot (x_i - 1)^\ell = \sum_{m=0}^{\infty} \binom{\ell-1-m}{m} x_i^m (x_i - 1)^{m+1} - \sum_{m=1}^{\infty} \binom{\ell-1-m}{m-1} x_i^m (x_i - 1)^m,$$

since for all terms appearing in the infinite sums that did not appear in the original equation, the respective binomial coefficients are zero.

Let us now prove this equation by induction on ℓ. For $\ell = 1$, the equation is easy to check. Let us now assume it is true for some $\ell \geq 1$, and let us check it for $\ell + 1$. We have

$$(-1)^{\ell}(x_i - 1)^{\ell+1} = -(x_i - 1) \sum_{m=0}^{\infty} \binom{\ell-1-m}{m} x_i^m (x_i - 1)^{m+1} + (x_i - 1) \sum_{m=1}^{\infty} \binom{\ell-1-m}{m-1} x_i^m (x_i - 1)^m$$

$$= -\sum_{m=0}^{\infty} \binom{\ell-1-m}{m} x_i^{m+1} (x_i - 1)^{m+1} + \sum_{m=0}^{\infty} \binom{\ell-1-m}{m} x_i^m (x_i - 1)^{m+1}$$

$$+ \sum_{m=1}^{\infty} \binom{\ell-1-m}{m-1} x_i^m (x_i - 1)^{m+1}$$

$$= -\sum_{m=1}^{\infty} \binom{\ell-m}{m-1} x_i^m (x_i - 1)^m + \sum_{m=0}^{\infty} \binom{\ell-m}{m} x_i^m (x_i - 1)^{m+1},$$

as desired. This proves the claim. $\qquad\square$

From Claim 3.6, we now obtain that $Q(x_1, \ldots, x_n)$ equals

$$\prod_{i=1}^{n} \left(\sum_{m=0}^{\lfloor (k-1)/2 \rfloor} \binom{k-1-m}{m} x_i^m (x_i - 1)^{m+1} - \sum_{m=1}^{\lfloor k/2 \rfloor} \binom{k-1-m}{m-1} x_i^m (x_i - 1)^m \right). \tag{3.2}$$

Using this equation, we can calculate the polynomial P. Recall that P is the unique polynomial in U_k such that $Q - P$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$.

When expanding the product over all i in (3.2), each of the terms that we obtain is a product of factors of the form $x_i^{m_i}(x_i - 1)^{m_i+1}$ or $x_i^{m_i}(x_i - 1)^{m_i}$ for each $i = 1, \ldots, n$ (with some coefficient given as the product of certain binomial coefficients). Note that each such term is divisible by the product $x_1^{m_1}(x_1 - 1)^{m_1} \cdots x_n^{m_n}(x_n - 1)^{m_n}$, and if $m_1 + \cdots + m_n \geq k$ this means that the term has zeroes of multiplicity at least k at all points in $\{0,1\}^n$. We can omit all such terms from Q without violating the defining property of $P \in U_k$ (i.e. the property that $Q - P$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$).

So let us now imagine that we expand the product over i in (3.2), but we only keep the terms which are products of factors of the form $x_i^{m_i}(x_i - 1)^{m_i+1}$ or $x_i^{m_i}(x_i - 1)^{m_i}$ for each $i = 1, \ldots, n$ such that $m_1 + \cdots + m_n \leq k - 1$. Note that all of these terms have degree at most $n + 2k - 2$. Furthermore, any such term of degree $n + 2k - 2$ must be of the form $x_1^{m_1}(x_1 - 1)^{m_1+1} \cdots x_n^{m_n}(x_n - 1)^{m_n+1} = (x_1 - 1) \cdots (x_n - 1) \cdot x_1^{m_1}(x_1 - 1)^{m_1} \cdots x_n^{m_n}(x_n - 1)^{m_n}$ with $m_1 + \cdots + m_n = k - 1$. Hence any such term of degree $n + 2k - 2$ has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0,\ldots,0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$. We can also omit all of these terms from Q without violating the defining property of $P \in U_k$.

11

Now, let Q^* be the polynomial consisting only of the remaining terms in the expansion of the product over i in (3.2), i.e. of the terms which are products of factors of the form $x_i^{m_i}(x_i-1)^{m_i+1}$ or $x_i^{m_i}(x_i-1)^{m_i}$ for each $i = 1, \ldots, n$ such that $m_1 + \cdots + m_n \le k - 1$ and which have degree at most $n + 2k - 3$. Then $Q^* - P$ has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$ and a zero of multiplicity at least $k - 1$ at $(0, \ldots, 0)$. However, the polynomial Q^* is also k-reduced, so we have $Q^* \in U_k$. By applying Corollary 2.4 to the polynomial Q^*, we can therefore conclude that $Q^* = P$ (since $P, Q^* \in U_k$ both satisfy the conditions in Corollary 2.4).

In other words, we just showed that P can be obtained by expanding the product over i in (3.2), but only keeping the terms which are products of factors of the form $x_i^{m_i}(x_i - 1)^{m_i+1}$ or $x_i^{m_i}(x_i - 1)^{m_i}$ for each $i = 1, \ldots, n$ such that $m_1 + \cdots + m_n \le k - 1$ and which have degree at most $n + 2k - 3$. Now, the homogeneous degree $n + 2k - 3$ part $\varphi_k(P)$ of P can be calculated by only considering the products of degree exactly $n + 2k - 3$ (and by only taking the homogeneous degree $n + 2k - 3$ part of these products). In other words, $\varphi_k(P)$ is obtained by expanding

$$\prod_{i=1}^{n} \left(\sum_{m=0}^{\lfloor (k-1)/2 \rfloor} \binom{k-1-m}{m} x_i^{2m+1} - \sum_{m=1}^{\lfloor k/2 \rfloor} \binom{k-1-m}{m-1} x_i^{2m} \right), \tag{3.3}$$

but only keeping the terms which are products of factors of the form $x_i^{2m_i+1}$ or $x_i^{2m_i}$ for each $i = 1, \ldots, n$ such that $m_1 + \cdots + m_n \le k - 1$ and which have degree exactly $n + 2k - 3$. Note that such a term can only have degree $n + 2k - 3$ if it consists of $n - 1$ factors of the form $x_i^{2m_i+1}$ and one factor of the form $x_i^{2m_i}$ (with $m_1 + \cdots + m_n = k - 1$). Hence $\varphi_k(P)$ can be obtained by expanding (3.3), but only keeping the monomials of degree exactly $n + 2k - 3$ in which exactly one variable has an even exponent.

To simplify notation, let us define $a_0, \ldots, a_{k-1}$ to be the coefficients such that (3.3) can be written as $\prod_{i=1}^{n}(a_{k-1}x_i^k + a_{k-2}x_i^{k-1} + \cdots + a_0 x_i)$. Furthermore, for $d \ge k$, let us define $a_d = 0$. Note that then for all $m \ge 0$ we have $a_{2m} = \binom{k-1-m}{m}$, and for all $m \ge 1$ we have $a_{2m-1} = -\binom{k-1-m}{m-1}$. Also note that $a_0 = 1$. Using this notation, we now obtain that

$$\varphi_k(P) = \sum_{(d_1, \ldots, d_n)} a_{d_1} \cdots a_{d_n} \cdot x_1^{d_1+1} \cdots x_n^{d_n+1}, \tag{3.4}$$

where the sum is over all sequences $(d_1, \ldots, d_n)$ of non-negative integers with $(d_1+1)+\cdots+(d_n+1) = n + 2k - 3$ such that exactly one of $d_1 + 1, \ldots, d_n + 1$ is even. These conditions are equivalent to demanding that $d_1 + \cdots + d_n = 2k - 3$ and exactly one of $d_1, \ldots, d_n$ is odd. Now, Lemma 3.3 follows easily.

Proof of Lemma 3.3. All monomials appearing in (3.4) are divisible by $x_1 \cdots x_n$, and hence $\varphi_k(P)$ is divisible by $x_1 \cdots x_n$, as desired. $\qquad\square$

From (3.4), we now obtain that

$$\varphi_k(P)/(x_1 \cdots x_n) = \sum_{(d_1, \ldots, d_n)} a_{d_1} \cdots a_{d_n} \cdot x_1^{d_1} \cdots x_n^{d_n},$$

where the sum is over all sequences $(d_1, \ldots, d_n)$ of non-negative integers with $d_1 + \cdots + d_n = 2k - 3$ such that exactly one of $d_1, \ldots, d_n$ is odd. Recalling that $a_0 = 1$, we can rewrite this equation as

$$\varphi_k(P)/(x_1 \cdots x_n) = \sum_{(d_1, \ldots, d_t)} \left(a_{d_1} \cdots a_{d_t} \sum_{1 \le i_1 < \cdots < i_t \le n} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t} \right), \tag{3.5}$$

where the sum is over all sequences $(d_1, \ldots, d_t)$ of positive integers with $d_1 + \cdots + d_t = 2k - 3$ such that exactly one of $d_1, \ldots, d_t$ is odd. Note that when we consider all $t!$ permutations of one fixed such

sequence $(d_1, \ldots, d_t)$ (some of which may be equal to the original sequence), then the resulting part of the sum in (3.5) is the symmetric polynomial $a_{d_1} \cdots a_{d_t} \sum_{i_1, \ldots, i_t \text{ distinct}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t}$. By averaging over all permutations of $(d_1, \ldots, d_t)$, we can therefore rewrite (3.5) as

$$\varphi_k(P)/(x_1 \cdots x_n) = \sum_{(d_1, \ldots, d_t)} \left(\frac{a_{d_1} \cdots a_{d_t}}{t!} \sum_{\substack{i_1, \ldots, i_t \in \{1, \ldots, n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t} \right), \qquad (3.6)$$

where the sum is again over all sequences $(d_1, \ldots, d_t)$ of positive integers with $d_1 + \cdots + d_t = 2k - 3$ such that exactly one of $d_1, \ldots, d_t$ is odd.

In order to prove Lemma 3.4, we need to find the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when expressing $\varphi_k(P)/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials. We will use the following claim, which can be derived from more general statements in the theory of symmetric polynomials. For the reader's convenience we provide a simple self-contained proof.

Claim 3.7. *For any sequence $(d_1, \ldots, d_t)$ of positive integers with $d_1 + \cdots + d_t \leq n$, when expressing*

$$\sum_{\substack{i_1, \ldots, i_t \in \{1, \ldots, n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t}$$

in terms of power sum symmetric polynomials, the coefficient of $x_1^{d_1 + \cdots + d_t} + \cdots + x_n^{d_1 + \cdots + d_t}$ equals $t! \cdot (-1)^{t-1}/t$.

Proof. For $t = 1$, the claim is trivially true. Let us now assume that $t \geq 2$, and that we already proved the claim for $t - 1$. Note that

$$\sum_{\substack{i_1, \ldots, i_t \in \{1, \ldots, n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_t}^{d_t} = (x_1^{d_t} + \cdots + x_n^{d_t}) \cdot \sum_{\substack{i_1, \ldots, i_{t-1} \in \{1, \ldots, n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_{t-1}}^{d_{t-1}}$$

$$- \sum_{s=1}^{t-1} \sum_{\substack{i_1, \ldots, i_{t-1} \in \{1, \ldots, n\} \\ \text{distinct}}} x_{i_1}^{d_1} \cdots x_{i_{s-1}}^{d_{s-1}} x_{i_s}^{d_s + d_t} x_{i_{s+1}}^{d_{s+1}} \cdots x_{i_{t-1}}^{d_{t-1}}.$$

Let us now imagine that we express the sums on the left-hand side in terms of power sum symmetric polynomials. The terms contributed from the first part (before the minus sign), all contain a factor $x_1^{d_t} + \cdots + x_n^{d_t}$, and in particular this first part does not contribute any $x_1^{d_1 + \cdots + d_t} + \cdots + x_n^{d_1 + \cdots + d_t}$ terms. For the second part (after the minus sign), the coefficient of $x_1^{d_1 + \cdots + d_t} + \cdots + x_n^{d_1 + \cdots + d_t}$ is by the induction hypothesis equal to $- \sum_{s=1}^{t-1} (t-1)! \cdot (-1)^{t-2}/(t-1) = (t-1)! \cdot (-1)^{t-1} = t! \cdot (-1)^{t-1}/t$. This finishes the proof of the claim. $\qquad \square$

Proof of Lemma 3.4. Let Y denote the coefficient of $x_1^{2k-3} + \cdots + x_n^{2k-3}$ when writing $\varphi_k(P)/(x_1 \cdots x_n)$ in terms of power sum symmetric polynomials. We need to prove that Y equals the sum in the statement of Lemma 3.4.

Recalling that we assumed $n \geq 2k - 3$, we can apply Claim 3.7 to the terms on the right-hand side of (3.6), and obtain

$$Y = \sum_{(d_1, \ldots, d_t)} \frac{a_{d_1} \cdots a_{d_t}}{t!} \cdot \frac{t! \cdot (-1)^{t-1}}{t} = \sum_{(d_1, \ldots, d_t)} a_{d_1} \cdots a_{d_t} \cdot \frac{(-1)^{t-1}}{t},$$

where the sums are over all sequences $(d_1, \ldots, d_t)$ of positive integers with $d_1 + \cdots + d_t = 2k - 3$ such that exactly one of $d_1, \ldots, d_t$ is odd. Note that all t cyclic permutations of one fixed such sequence

13

$(d_1, \ldots, d_t)$ contribute the same amount to the sum above, and for exactly one of these permutations d_1 is odd. Hence we can conclude that

$$Y = \sum_{(d_1,\ldots,d_t)} t \cdot a_{d_1} \cdots a_{d_t} \cdot \frac{(-1)^{t-1}}{t} = \sum_{(d_1,\ldots,d_t)} (-1)^{t-1} a_{d_1} \cdots a_{d_t},$$

where this time the sums are over all sequences $(d_1, \ldots, d_t)$ of positive integers with $d_1 + \cdots + d_t = 2k-3$ such that d_1 is odd and $d_2, \ldots, d_t$ are even.

Let us now change variables, writing $d_1 = 2m_1 - 1$ and $d_j = 2m_j$ for $j = 2, \ldots, t$. Then we obtain

$$Y = \sum_{(m_1,\ldots,m_t)} (-1)^{t-1} a_{2m_1-1} a_{2m_2} \cdots a_{2m_t} = \sum_{(m_1,\ldots,m_t)} (-1)^t \cdot (-a_{2m_1-1}) \cdot a_{2m_2} \cdots a_{2m_t}$$

$$= \sum_{(m_1,\ldots,m_t)} (-1)^t \cdot \binom{k-1-m_1}{m_1-1} \binom{k-1-m_2}{m_2} \cdots \binom{k-1-m_t}{m_t},$$

where the sums are over all sequences $(m_1, \ldots, m_t)$ of positive integers with $m_1 + \cdots + m_t = k - 1$. This proves Lemma 3.4. $\qquad\square$

3.2   Proof of Lemma 3.5

In this subsection, we prove Lemma 3.5. We will use the formula for Catalan numbers stated in the following claim. This formula follows from work of Riordan [10] and is also stated as Theorem 12.1 in a book on Catalan numbers by Koshy [9]. For the reader's convenience we give a self-contained proof here.

Claim 3.8. *For all $s \geq 1$, we have*

$$\sum_{i=0}^{s} (-1)^i C_i \cdot \binom{i+1}{s-i} = 0. \tag{3.7}$$

Proof. We prove the desired formula by induction on s. For $s = 1$ and $s = 2$, the formula is easy to check. So let us now assume that $s \geq 3$ and that the formula is true for $s - 1$ and $s - 2$.

Recall that for $i \geq 0$ we have $C_i = \binom{2i}{i}/(i+1)$, which implies that $C_i = C_{i-1} \cdot 2(2i-1)/(i+1)$ for $i \geq 1$. Consequently, for all $i \geq 1$ we have

$$C_i \cdot \binom{i+1}{s-i} = 2C_{i-1} \cdot (2i-1) \cdot \frac{1}{i+1} \binom{i+1}{s-i}$$

$$= 2C_{i-1} \cdot \left(\frac{2s-4}{s+1}(s-i) + \frac{2s-1}{s+1}(2i+1-s) \right) \cdot \frac{1}{i+1} \binom{i+1}{s-i}$$

$$= \frac{4s-8}{s+1} \cdot C_{i-1} \cdot \frac{s-i}{i+1} \binom{i+1}{s-i} + \frac{4s-2}{s+1} \cdot C_{i-1} \cdot \frac{2i+1-s}{i+1} \binom{i+1}{s-i}$$

$$= \frac{4s-8}{s+1} \cdot C_{i-1} \cdot \binom{i}{s-i-1} + \frac{4s-2}{s+1} \cdot C_{i-1} \cdot \binom{i}{s-i}.$$

Now, recalling that $s \geq 3$, we obtain that

$$\sum_{i=0}^{s} (-1)^i C_i \cdot \binom{i+1}{s-i} = \sum_{i=1}^{\infty} (-1)^i C_i \cdot \binom{i+1}{s-i}$$

$$= -\frac{4s-8}{s+1} \sum_{i=1}^{\infty} (-1)^{i-1} C_{i-1} \cdot \binom{i}{s-1-i} - \frac{4s-2}{s+1} \sum_{i=1}^{\infty} (-1)^{i-1} C_{i-1} \cdot \binom{i}{s-i}$$

14

$$= -\frac{4s-8}{s+1}\sum_{i=0}^{\infty}(-1)^i C_i \cdot \binom{i+1}{s-2-i} - \frac{4s-2}{s+1}\sum_{i=0}^{\infty}(-1)^i C_i \cdot \binom{i+1}{s-1-i}$$

$$= -\frac{4s-8}{s+1}\sum_{i=0}^{s-2}(-1)^i C_i \cdot \binom{i+1}{s-2-i} - \frac{4s-2}{s+1}\sum_{i=0}^{s-1}(-1)^i C_i \cdot \binom{i+1}{s-1-i} = 0,$$

using the induction hypothesis for $s-2$ and $s-1$ in the last step. $\qquad\square$

We will also use the following well-known formula for binomial coefficients (which is easy to prove, for example by double-counting): for all non-negative integers m, n and s, we have

$$\binom{n+s+1}{m-s} = \sum_{j=s}^{\infty}\binom{s+1}{j-s}\binom{n}{m-j}. \tag{3.8}$$

From (3.7) and (3.8) we can derive the following statement, which will be used in the proof of Lemma 3.5.

Claim 3.9. *For any non-negative integers m, n and s, we have*

$$\binom{n}{m} = \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s+1}^{2s+1}\left(\sum_{i=0}^{s}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}.$$

Proof. First, note that for any $j > 2s+1$, we have $\binom{i+1}{j-i} = 0$ for all $i = 0,\ldots,s$. We can therefore equivalently write the desired equation as

$$\binom{n}{m} = \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s+1}^{\infty}\left(\sum_{i=0}^{s}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}.$$

We prove this equation by induction on s. Note that for $s = 0$, it simply states that (recall that $C_0 = 1$)

$$\binom{n}{m} = \binom{n+1}{m} - \binom{n}{m-1},$$

which is true. Now let us assume that $s \geq 1$ and that the desired equation is true for $s-1$. Then

$$\binom{n}{m} = \sum_{j=0}^{s-1}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s}^{\infty}\left(\sum_{i=0}^{s-1}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}$$

$$= \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - (-1)^s C_s \cdot \binom{n+s+1}{m-s} - \sum_{j=s}^{\infty}\left(\sum_{i=0}^{s-1}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}$$

$$= \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s}^{\infty}\left((-1)^s C_s \cdot \binom{s+1}{j-s} + \sum_{i=0}^{s-1}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}$$

$$= \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s}^{\infty}\left(\sum_{i=0}^{s}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j}$$

$$= \sum_{j=0}^{s}(-1)^j C_j \cdot \binom{n+j+1}{m-j} - \sum_{j=s+1}^{\infty}\left(\sum_{i=0}^{s}(-1)^i C_i \cdot \binom{i+1}{j-i}\right)\binom{n}{m-j},$$

where for the third equation we used (3.8), and in the last equation we used (3.7). This finishes the proof of the claim. $\qquad\square$

Using Claim 3.9, we now prove Lemma 3.5.

Proof of Lemma 3.5. Fix $\ell \geq 1$. For convenience, let us denote the left-hand side of the equation in Lemma 3.5 by Z. We then need to show that $Z = (-1)^{\ell} C_{\ell-1}$.

We have

$$Z = \sum_{\substack{(m_1,\ldots,m_t) \\ m_1,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell}} (-1)^t \cdot \binom{\ell - m_1}{m_1 - 1} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}.$$

Applying Claim 3.9 to the first binomial coefficient in the product (with $n = \ell - m_1$ and $m = m_1 - 1$ and $s = \ell - 1$), we obtain

$$Z = \sum_{j=0}^{\ell-1} \sum_{\substack{(m_1,\ldots,m_t) \\ m_1,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell}} (-1)^t \cdot (-1)^j C_j \binom{\ell - m_1 + j + 1}{m_1 - 1 - j} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}$$

$$- \sum_{j=\ell}^{2\ell-1} \sum_{\substack{(m_1,\ldots,m_t) \\ m_1,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell}} (-1)^t \cdot \left(\sum_{i=0}^{\ell-1} (-1)^i C_i \binom{i+1}{j-i} \right) \binom{\ell - m_1}{m_1 - 1 - j} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}.$$

Note that the sum after the minus sign in the above equation is zero. Indeed, for all terms appearing in this sum we have $m_1 \leq \ell \leq j$, so the binomial coefficient $\binom{\ell - m_1}{m_1 - 1 - j}$ is zero. Hence

$$Z = \sum_{j=0}^{\ell-1} \sum_{\substack{(m_1,\ldots,m_t) \\ m_1,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell \\ t \geq 1}} (-1)^{j+t} C_j \binom{\ell - m_1 + j + 1}{m_1 - j - 1} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}.$$

With a change of variables, replacing $m_1 - j - 1$ by m_1, we can rewrite this as

$$Z = \sum_{j=0}^{\ell-1} \sum_{\substack{(m_1,\ldots,m_t) \\ m_1 \geq 0,\, m_2,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell - j - 1 \\ t \geq 1}} (-1)^{j+t} C_j \binom{\ell - m_1}{m_1} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}$$

(here, a priori the condition on m_1 would be $m_1 > -j - 1$, but since the binomial coefficient $\binom{\ell - m_1}{m_1}$ vanishes for negative m_1, we can instead take the condition $m_1 \geq 0$). Note that in the above equation, the contribution for $j = \ell - 1$ is just $(-1)^{\ell-1+1} C_{\ell-1} \binom{\ell-0}{0} = (-1)^{\ell} C_{\ell-1}$. Thus,

$$Z = (-1)^{\ell} C_{\ell-1} + \sum_{j=0}^{\ell-2} \sum_{\substack{(m_1,\ldots,m_t) \\ m_1 \geq 0,\, m_2,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell - j - 1}} (-1)^{j+t} C_j \binom{\ell - m_1}{m_1} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t}.$$

Now, in order to prove the lemma, it suffices to show that for any fixed $j \in \{0, \ldots, \ell - 2\}$ we have

$$\sum_{\substack{(m_1,\ldots,m_t) \\ m_1 \geq 0,\, m_2,\ldots,m_t > 0 \\ m_1 + \cdots + m_t = \ell - j - 1}} (-1)^t \cdot \binom{\ell - m_1}{m_1} \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t} = 0. \tag{3.9}$$

16

Indeed, by distinguishing whether m_1 is positive or zero (and noting that for $m_1 = 0$ we have $\binom{\ell - m_1}{m_1} = \binom{\ell - 0}{0} = 1$), we can rewrite the left-hand size of (3.9) as

$$\sum_{\substack{(m_1, \ldots, m_t) \\ m_1, m_2, \ldots, m_t > 0 \\ m_1 + \cdots + m_t = \ell - j - 1}} (-1)^t \cdot \binom{\ell - m_1}{m_1} \cdots \binom{\ell - m_t}{m_t} + \sum_{\substack{(m_2, \ldots, m_t) \\ m_2, \ldots, m_t > 0 \\ m_2 + \cdots + m_t = \ell - j - 1}} (-1)^t \cdot \binom{\ell - m_2}{m_2} \cdots \binom{\ell - m_t}{m_t} = 0.$$

This proves (3.9), finishing the proof of Lemma 3.5. $\qquad\square$

4 Concluding remarks

4.1 Clifton and Huang's hyperplane problem

As mentioned in the introduction, Clifton and Huang [5] studied the minimum size of a collection of hyperplanes in $\mathbb{R}^n$ such that every point in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$ is covered by at least k of these hyperplanes, but no hyperplane contains $(0, \ldots, 0)$, where $k \geq 2$ is fixed and n is large with respect to k. While Theorem 1.3 improves their lower bound for this problem to $n + 2k - 3$, the best known upper bound is still $n + \binom{k}{2}$. It would be very interesting to close this gap.

Clifton and Huang [5] conjectured that their upper bound $n + \binom{k}{2}$ for this hyperplane problem is tight if n is sufficiently large with respect to k. Theorem 1.3 shows that this conjecture cannot be proved by following the approach of Clifton and Huang and only being more careful with the analysis of the coefficients of the polynomial they call u (which is obtained by applying the punctured higher-multiplicity version of the Combinatorial Nullstellensatz due to Ball and Serra [3] to the product of the linear polynomials corresponding to the hyperplanes). Indeed, the second part of Theorem 1.3 implies that there exist polynomials of degree at most $2k - 3$ which have all the properties of the polynomial u in the approach of Clifton and Huang, and hence one cannot prove a lower bound of $\binom{k}{2}$ on the degree of u just by a more careful coefficient analysis. Instead, a different approach would be needed to prove their conjecture.

Note that Clifton and Huang's hyperplane problem is equivalent to the following problem: for fixed $k \geq 1$ and n large with respect to k, what is the minimum possible degree of a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ with $P(0, \ldots, 0) \neq 0$ such that P has zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$ and such that P can be written as a product of linear polynomials? Without this last condition, this is precisely Problem 1.2. By the results of Alon and Füredi [2], the answers for both problems agree if $k = 1$. The answers also agree for $k = 2$ and $k = 3$, since then $n + 2k - 3 = n + \binom{k}{2}$. However, if Clifton and Huang's conjecture is true, then by Theorem 1.3 the answers must be different for $k \geq 4$. It would be interesting to prove (or disprove) that the two problems have different answers for sufficiently large k.

One can also study a variant of Clifton and Huang's hyperplane problem, where one replaces the condition that no hyperplane contains $(0, \ldots, 0)$ by the condition that $(0, \ldots, 0)$ is covered by exactly ℓ hyperplanes for some given $0 \leq \ell \leq k - 1$. The case of $\ell = 0$ corresponds to Clifton and Huang's original hyperplane problem. As before, we can equivalently rephrase the problem in terms of polynomials, asking for the minimum possible degree of a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ with a zero of multiplicity exactly ℓ at $(0, \ldots, 0)$ and zeroes of multiplicity at least k at all points in $\{0,1\}^n \setminus \{(0, \ldots, 0)\}$ such that P can be written as a product of linear polynomials. Again, one may ask whether the answer to this problem changes by omitting the last condition that P is a product of linear polynomials. Theorems 1.4 and 1.5 determine the answer of the problem where the last condition is omitted (the answer is $n + 2k - 3$ for $0 \leq \ell \leq k - 2$ and $n + 2k - 2$ for $\ell = k - 1$). By finding examples for the polynomial P with the desired properties and of the appropriate degree such that P is a product of linear polynomials, one can show that the answers for both problems agree for $k - 3 \leq \ell \leq k - 1$. However, it is not clear what happens for smaller ℓ.

17

Clifton and Huang also studied their hyperplane problem in the opposite parameter range, where the dimension n is fixed and k is large. They proved that for any fixed dimension n, the answer is of the form $(1 + \frac{1}{2} + \cdots + \frac{1}{n} + o(1)) \cdot k$ as k goes to infinity. It might also be interesting to study Problem 1.2 for fixed dimension n and large k.

4.2   Problem 1.2 over other fields

One can also consider Problem 1.2 over other fields than $\mathbb{R}$. For an arbitrary field $\mathbb{F}$, let us say that a polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ has a zero of multiplicity at least k at a point $(a_1, \ldots, a_n) \in \mathbb{F}^n$ if the following holds: when expanding the polynomial $P(x_1 + a_1, \ldots, x_n + a_n) \in \mathbb{F}[x_1, \ldots, x_n]$, all monomials occurring in $P(x_1 + a_1, \ldots, x_n + a_n)$ have degree at least k. Note that for $\mathbb{F} = \mathbb{R}$ this agrees with our definition in terms of the derivatives of P at $(a_1, \ldots, a_n)$.

Our proof of Theorems 1.3 to 1.5 works for every field of characteristic 0, but it is also interesting to consider Problem 1.2 over fields of positive characteristic. Since Alon and Füredi's result [2] stated in Theorem 1.1 is valid over any field, one might also expect the answer for Problem 1.2 to be independent of the characteristic of the field.

Interestingly, this is not the case, and the answer for Problem 1.2 does depend on the characteristic of the field. Specifically, for a field $\mathbb{F}$ of characteristic $p > 3$ and $k = (p+5)/2$, there exists a polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $\deg P \leq n + 2k - 4$ with $P(0, \ldots, 0) \neq 0$ and such that P has zeroes of multiplicity at least k at all points in $\{0, 1\}^n \setminus \{(0, \ldots, 0)\}$. In particular, the statement in Theorem 1.3 fails to hold in this case. Theorem 1.3 similarly fails for a field of characteristic 2 and $k = 4$ and for a field of characteristic 3 and $k = 7$.

For example, for $k = 4$ and the field $\mathbb{F} = \mathbb{F}_2$, one can check that the following polynomial of degree $n + 2k - 4 = n + 4$ has zeroes of multiplicity at least 4 at all points in $\mathbb{F}_2^n \setminus \{(0, \ldots, 0)\}$, but does not vanish at $(0, \ldots, 0)$:

$$\left(\prod_{\ell=1}^n (x_\ell + 1) \right) \cdot \left(1 + \sum_{i=1}^n (x_i^3 + x_i^2 + x_i) + \sum_{1 \leq i \neq j \leq n} (x_i^3 + x_i^2) x_j + \sum_{1 \leq i < j \leq n} x_i x_j + \sum_{1 \leq i < j < k \leq n} x_i x_j x_k \right).$$

The reason for having counterexamples to the statement Theorem 1.3 for $k = (p+5)/2$ in characteristic $p > 3$, for $k = 4$ in characteristic $p = 2$, and for $k = 7$ in characteristic $p = 3$ is as follows: these numbers k are chosen to be the smallest values such that the Catalan number C_{k-2} is divisible by p. By using the arguments from our proof (together with some additional analysis of our map φ_k), one can show that Theorem 1.3 fails for this value of k (but holds for all smaller values). It is worth pointing out that the only point in our proof which is dependent on the characteristic of the field $\mathbb{F}$ is the assertion that the Catalan numbers are non-zero in $\mathbb{F}$.

Since the first part of Theorem 1.3 fails to hold over fields of positive characteristic, the first part of Theorem 1.4 (which is a more general statement) also fails for over fields of positive characteristic. Theorem 1.5, however, holds over any field.

References

[1] N. Alon, *Combinatorial Nullstellensatz*, Combin. Probab. Comput. **8** (1999), 7–29.

[2] N. Alon and Z. Füredi, *Covering the cube by affine hyperplanes*, European J. Combin. **14** (1993), 79–83.

[3] S. Ball and O. Serra, *Punctured combinatorial Nullstellensätze*, Combinatorica **29** (2009), 511–522.

[4] G. Batzaya and G. Bayarmagnai, *A generalized combinatorial Nullstellensatz for multisets*, European J. Combin. **83** (2020), 5 pp.

[5] A. Clifton and H. Huang, *On almost k-covers of hypercubes*, Combinatorica **40** (2020), 511–526.

[6] P. Komjáth, *Partitions of vector spaces*, Period. Math. Hungar. **28** (1994), 187–193.

[7] G. Kós and L. Rónyai, *Alon's Nullstellensatz for multisets*, Combinatorica **32** (2012), 589–605.

[8] G. Kós, T. Mészáros, and L. Rónyai, *Some extensions of Alon's Nullstellensatz*, Publ. Math. Debrecen **79** (2011), 507–519.

[9] T. Koshy, *Catalan numbers with applications*, Oxford University Press, Oxford, 2009.

[10] J. Riordan, *A note on Catalan parentheses*, Amer. Math. Monthly **80** (1973), 904–906.