



# **Generalized Vandermonde Determinants over the Chebyshev Basis**

**Thorsten Werther\***

**TR-93-024**

**April 1993**

## **Abstract**

It is a well known fact that the generalized Vandermonde determinant can be expressed as the product of the standard Vandermonde determinant and a polynomial with nonnegative integer coefficients. In this paper we generalize this result to Vandermonde determinants over the Chebyshev basis. We apply this result to prove that the number of real roots in  $[1, \infty]$  of a real polynomial is bounded by the number of its nonvanishing coefficients (sparsity) when represented over the Chebyshev basis. This bound on the number of real roots is used to prove finiteness of the Vapnik-Chervonenkis dimension (and thereby uniform learnability) of the class of polynomials of bounded sparsity over the Chebyshev basis.

---

\*Department of Computer Science, University of Bonn, Römerstraße 164, 5300 Bonn, Germany, and International Computer Science Institute, 1947 Center Street, Berkeley, California 94704-1105 (werther@icsi.berkeley.edu).



# 1 Introduction

More than a century ago, Michell [6] proved the striking result that the determinant of the generalized Vandermonde matrix over the standard power basis  $\mathcal{S} = \{1, x, x^2, \dots\}$  can be represented as the product of the standard Vandermonde determinant and a polynomial with nonnegative integer coefficients.

One of the most important consequences of Michell's result is the nonsingularity of the generalized Vandermonde matrix when the indeterminates take on positive and distinct values. This in turn is equivalent to the fact that the number of positive real roots of a polynomial over the real numbers is strictly less than its sparsity (the number of nonvanishing monomials) with respect to  $\mathcal{S}$ . This result also follows immediately from Descartes' Rule of Signs.

These results provide major tools for the recent developments in the area of interpolation [1] and learnability [4] of sparse polynomials over the reals. Numerous generalizations of this setting are proposed in the literature and have attracted a lot of attention. One natural question concerns the common properties of real polynomials which have a sparse representation in bases of  $\mathbb{R}[x]$  other than the standard power basis  $\mathcal{S}$ .

In this paper we study real polynomials that admit sparse representations in the Chebyshev basis  $\mathcal{T} = \{T_0(x), T_1(x), \dots\}$  where  $T_i(x)$  is the  $i$ -th Chebyshev polynomial of the first kind. The main result of the paper states the analogy of Michell's theorem for the Chebyshev case. An immediate consequence of this result is the nonsingularity of Vandermonde matrices over  $\mathcal{T}$  provided that the indeterminates take distinct values from either  $[1, \infty)$  or  $(-\infty, -1]$ .

As an application, we answer the question posed by Lakshman and Saunders [5] about the relationship of the number of real roots of a polynomial and its sparsity with respect to the Chebyshev basis. In fact, the number of real zeros of a polynomial, either to the left or to the right of the interval of orthogonality, does not exceed its sparsity with respect to  $\mathcal{T}$ .

The bound on the number of real roots is used to prove tight lower and upper bounds on the Vapnik-Chervonenkis dimension of the class of polynomials of bounded sparsity over the Chebyshev basis. Surprisingly, these bounds coincide with the bounds given in [4] for the standard power basis.

# 2 Preliminaries and Notation

A polynomial set is a sequence  $\Phi = \{\Phi_n\}_{n \in \mathbb{N}_0}$  in which  $\deg(\Phi_n) = n$  for all  $n \in \mathbb{N}_0$ . Every polynomial set  $\Phi$  represents a basis for the polynomial ring  $\mathbb{R}[x]$ . Hence every polynomial  $f \in \mathbb{R}[x]$  can uniquely be represented as a finite linear combination over  $\Phi$ , i.e.  $f = \sum_{i=0}^n c_i \Phi_i$  with  $c_i \in \mathbb{R}$  and  $n = \deg(f)$ . This representation is called the  $\Phi$ -representation of  $f$ . As usual, we say that  $f$  is  $t$ -sparse with respect to  $\Phi$  if at most  $t$  of the coefficients of the  $\Phi$ -representation of  $f$  are non-zero. The notation generalizes in the usual way to multivariate polynomials.

The Chebyshev polynomials are a special case of orthogonal polynomials, distinguished by their particular simplicity. The  $n$ -th Chebyshev polynomial of the first kind  $T_n(x)$  is defined by

$$T_n(x) = \cos(n \arccos x), \quad |x| \leq 1.$$

The Chebyshev polynomials admit a very simple three term recursion formula:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_n(x) = 2x T_{n-1}(x) - T_{n-2}(x) \text{ for } n > 1.$$

In the following we will frequently use the fact that

$$T_k(x) \cdot T_l(x) = \frac{1}{2} (T_{k+l}(x) + T_{|k-l|}(x)),$$

therefore, we define  $T_k(x) = T_{-k}(x)$ ,  $k < 0$  for simplicity of notation. It is a well known fact that

$$T_n(x) = \sum_{r=0}^{\lfloor n/2 \rfloor} (-1)^r \binom{n}{2r} x^{n-2r} (1-x^2)^r.$$

On the other hand the  $\mathcal{T}$ -representation of powers of  $x$  is given by

$$2^{n-1} x^n = \begin{cases} \sum'_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{\lfloor n/2 \rfloor + i} T_{2i}(x) & \text{if } n \text{ even,} \\ \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{\lfloor n/2 \rfloor + i} T_{2i+1}(x) & \text{if } n \text{ odd,} \end{cases} \quad (1)$$

where  $\sum'$  denotes a sum with the first term halved. For the general theory of orthogonal polynomials the reader is referred to Szegő's classical textbook [7].

Let  $\Phi$  be a polynomial set and  $\mathbf{a} = (a_1, \dots, a_n)$  a vector of distinct nonnegative integers. The generalized Vandermonde determinant  $V_{\Phi, \mathbf{a}}(x_1, \dots, x_n)$  over  $\Phi$  is defined as the determinant of the generalized Vandermonde matrix

$$\begin{pmatrix} \Phi_{a_1}(x_1) & \Phi_{a_2}(x_1) & \dots & \Phi_{a_n}(x_1) \\ \Phi_{a_1}(x_2) & \Phi_{a_2}(x_2) & \dots & \Phi_{a_n}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{a_1}(x_n) & \Phi_{a_2}(x_n) & \dots & \Phi_{a_n}(x_n) \end{pmatrix}$$

where the  $x_i$  are indeterminates. Note that the number of indeterminates of  $V_{\Phi, \mathbf{a}}$  is given by the length of  $\mathbf{a}$ . Let  $\underline{n}$  denote the vector  $(0, 1, \dots, n-1)$ . Then  $V_{\Phi, \underline{n}}$  is the (standard) Vandermonde determinant over  $\Phi$ .

Let  $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}_0^n$ . We say that  $\mathbf{a}$  is increasing if  $a_1 < \dots < a_n$ . If  $\mathbf{a}$  consists of distinct entries there exists a permutation  $\pi \in S_n$  such that  $\pi(\mathbf{a}) = (a_{\pi(1)}, \dots, a_{\pi(n)})$  is increasing and we define  $\text{sgn}(\mathbf{a}) = \text{sgn}(\pi)$ . Let  $\mathbf{b} = (a_{i_1}, \dots, a_{i_r})$  with  $1 \leq i_1 < \dots < i_r \leq n$ . Then  $\mathbf{a} \setminus \mathbf{b} \in \mathbb{N}_0^{n-r}$  denotes the vector  $(a_1, \dots, a_{i_1-1}, a_{i_1+1}, \dots, a_{i_r-1}, a_{i_r+1}, \dots, a_n)$ . Any increasing  $\mathbf{a} = (a_1, \dots, a_n)$  can be written uniquely as  $\mathbf{a} = \underline{n+r} \setminus \mathbf{b}$  for some increasing  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$  with  $b_r < a_n$ . We call the entries of  $\mathbf{b}$  the gaps in  $\mathbf{a}$ .

### 3 Generalized Vandermonde Determinant

Let  $\Phi$  be a polynomial set and  $\mathbf{a} \in \mathbb{N}_0^n$ . Note that the Vandermonde determinant  $V_{\Phi, \mathbf{a}}$  over  $\Phi$  vanishes when  $x_i = x_j$  for some  $i \neq j$ . Since  $V_{\Phi, \underline{n}}$  viewed as a polynomial in  $x_n$  is of degree  $n-1$  and its zeros are  $x_n = x_i$  for  $1 \leq i < n$  there exists a constant  $c_{\Phi, n}$  depending only on  $\Phi$  and  $n$  such that

$$V_{\Phi, \underline{n}} = c_{\Phi, n} \prod_{1 \leq i < j \leq n} (x_j - x_i). \quad (2)$$

Hence,  $V_{\Phi, \mathbf{a}}$  is divisible by  $V_{\Phi, \underline{n}}$  in the polynomial ring  $\mathbb{R}[x_1, \dots, x_n]$  and we write

$$P_{\Phi, \mathbf{a}} = V_{\Phi, \mathbf{a}} / V_{\Phi, \underline{n}}. \quad (3)$$

Note that  $V_{\Phi, \mathbf{a}}(x_1, \dots, x_n) = \text{sgn}(\pi) V_{\Phi, \mathbf{a}}(x_{\pi(1)}, \dots, x_{\pi(n)})$  for any permutation  $\pi \in S_n$  since a transposition in  $\pi$  corresponds to interchanging two rows in the Vandermonde matrix. Hence,  $P_{\Phi, \mathbf{a}}$  is a symmetric polynomial in  $(x_1, \dots, x_n)$ . Likewise,  $V_{\Phi, \mathbf{a}} = \text{sgn}(\pi) V_{\Phi, \pi(\mathbf{a})}$  and therefore  $P_{\Phi, \mathbf{a}} = \text{sgn}(\pi) P_{\Phi, \pi(\mathbf{a})}$  for any permutation  $\pi \in S_n$ .

Using our notation we restate Michell's theorem:

**Theorem 1** Let  $\mathcal{S} = \{1, x, x^2, \dots\}$ . For any  $n \in \mathbb{N}$  and any increasing  $\mathbf{a} \in \mathbb{N}_0^n$  the coefficients of the  $\mathcal{S}$ -representation of  $P_{\mathcal{S}, \mathbf{a}}$  are nonnegative integers.

In this section we extend Michell's result to the Chebyshev basis.

**Theorem 2** Let  $\mathcal{T}$  denote the Chebyshev basis. For any  $n \in \mathbb{N}$  and any increasing  $\mathbf{a} \in \mathbb{N}_0^n$  the coefficients of the  $\mathcal{T}$ -representation of  $P_{\mathcal{T},\mathbf{a}}$  are nonnegative integers.

Evans and Isaacs [3] give a very elegant and simple proof of Theorem 1. Their proof is by induction on the number of indeterminates, however, the proof relies on the homogeneity of  $P_{\mathcal{S},\mathbf{a}}$ .

It is easy to see that  $P_{\mathcal{T},\mathbf{a}}$  is not homogeneous in general, so we cannot use their ideas in the Chebyshev setting. Instead, we will use an induction on the number of gaps in the index vector  $\mathbf{a}$ . This approach is used in [11] to study the coefficients of the  $\mathcal{S}$ -representation of  $P_{\mathcal{S},\mathbf{a}}$ . It turns out that these techniques, which only use the symmetry of  $P_{\mathcal{S},\mathbf{a}}$ , can be transferred to the Chebyshev basis. However, the proof of Theorem 2 involves a lot of technical details. Therefore, we sketch the alternative proof of Theorem 1 to provide a guideline for the more complicated proof of Theorem 2.

### 3.1 Vandermonde determinants over the power basis

In this section we give a proof of Theorem 1 which, in contrast to the proof given by Evans and Isaacs, does not depend on the homogeneity of  $P_{\mathcal{S},\mathbf{a}}$ .

Since  $P_{\mathcal{S},\mathbf{a}}$  is a symmetric polynomial it can be written as a polynomial of the elementary symmetric polynomials. Let  $\sigma_k^{(n)} = \sigma_k(x_1, \dots, x_n)$  denote the  $k$ -th elementary symmetric polynomial in the indeterminates  $x_1, \dots, x_n$ , i.e.

$$\sigma_k^{(0)} = \begin{cases} 0 & \text{if } k \neq 0 \\ 1 & \text{if } k = 0, \end{cases} \quad \sigma_k^{(n)} = x_n \sigma_k^{(n-1)} + \sigma_{k-1}^{(n-1)} \quad (4)$$

for  $n \in \mathbb{N}, k \in \mathbb{Z}$ . Note that  $\sigma_k^{(n)} = 0$  for  $k > n$  or  $k < 0$  and that

$$\prod_{i=1}^n (x - x_i) = \sum_{k=0}^n (-1)^{n-k} \sigma_k^{(n)} x^k. \quad (5)$$

For an increasing vector  $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}_0^n$  we will denote the number of gaps in  $\mathbf{a}$  by  $r$ , i.e.  $\mathbf{a} = \underline{n+r} \setminus \mathbf{b}$  for some  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$ . The proof of Theorem 1 is by induction on  $r$ .

Let  $r = 0$ . Then  $\mathbf{a} = \underline{n}$  and the statement of Theorem 1 is trivial since  $P_{\mathcal{S},\underline{n}} = 1$ .

Now, assume that  $r = 1$ . Then  $\mathbf{a}$  has exactly one gap and  $\mathbf{a} = \underline{n+1} \setminus (k)$  for some  $0 \leq k < n$ . Note that by expanding the last row of the standard Vandermonde matrix in  $n+1$  indeterminates, we have

$$V_{\mathcal{S},\underline{n+1}} = \sum_{k=0}^n (-1)^{n-k} V_{\mathcal{S},\underline{n+1} \setminus (k)} x_{n+1}^k$$

and on the other hand by (5),

$$V_{\mathcal{S},\underline{n+1}} = \prod_{i=1}^n (x_{n+1} - x_i) V_{\mathcal{S},\underline{n}} = \sum_{k=0}^n (-1)^{n-k} \sigma_k^{(n)} x_{n+1}^k V_{\mathcal{S},\underline{n}}.$$

Comparing the coefficients of  $x_{n+1}^k$  yields

$$V_{\mathcal{S},\underline{n+1} \setminus (k)} = \sigma_k^{(n)} V_{\mathcal{S},\underline{n}}. \quad (6)$$

Hence,  $P_{\mathcal{S},\underline{n+1} \setminus (k)} = \sigma_k^{(n)}$  and the statement of Theorem 1 follows immediately from the definition (4) of the elementary symmetric polynomials.

Let us now consider arbitrary  $r \in \mathbb{N}$ . Then  $\mathbf{a} = \underline{n+r} \setminus \mathbf{b}$  for some increasing  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$ . It turns out that for any increasing  $\mathbf{b} \in \mathbb{N}_0^r$ ,

$$V_{\mathcal{S}, \underline{n+r} \setminus \mathbf{b}} = \det \begin{pmatrix} \sigma_{b_1}^{(n)} & \cdots & \sigma_{b_r}^{(n)} \\ \sigma_{b_1-1}^{(n)} & \cdots & \sigma_{b_r-1}^{(n)} \\ \vdots & & \vdots \\ \sigma_{b_1-(r-1)}^{(n)} & \cdots & \sigma_{b_r-(r-1)}^{(n)} \end{pmatrix} V_{\mathcal{S}, \underline{n}}. \quad (7)$$

Note that (6) is a special case of (7). In order to prove (7) (and the nonnegativity of the coefficients of the  $\mathcal{S}$ -representation of  $\mathcal{P}_{\mathcal{S}, \underline{n+r} \setminus \mathbf{b}}$ ) we study the determinant of matrices that are slightly more general than the matrix given in (7).

For vectors  $\mathbf{b}, \mathbf{k} \in \mathbb{N}_0^r$  with  $\mathbf{b} = (b_1, \dots, b_r)$  and  $\mathbf{k} = (k_1, \dots, k_r)$  let  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  be defined by

$$C_{\mathbf{k}}^{(n)}(\mathbf{b}) = \det \begin{pmatrix} \sigma_{b_1-k_1}^{(n)} & \cdots & \sigma_{b_r-k_1}^{(n)} \\ \vdots & & \vdots \\ \sigma_{b_1-k_r}^{(n)} & \cdots & \sigma_{b_r-k_r}^{(n)} \end{pmatrix}. \quad (8)$$

Using this notation we may rewrite (7) as  $V_{\mathcal{S}, \underline{n+r} \setminus \mathbf{b}} = C_{\underline{r}}^{(n)}(\mathbf{b}) V_{\mathcal{S}, \underline{n}}$ . Note that  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  vanishes when  $\mathbf{k}$  (or  $\mathbf{b}$ ) does not consist of distinct entries. By expanding  $C_{\mathbf{k}}^{(n+1)}(\mathbf{b})$  according to (4) it is straightforward to prove that for  $n, r \in \mathbb{N}$  and  $\mathbf{b}, \mathbf{k} \in \mathbb{N}_0^r$ ,

$$C_{\mathbf{k}}^{(n+1)}(\mathbf{b}) = \sum_{l=0}^r x_{n+1}^l \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(l)}} C_{\mathbf{q}}^{(n)}(\mathbf{b}), \quad (9)$$

where  $S_{\mathbf{k}}^{(l)}$  denotes the set of vectors  $\mathbf{q} \in \mathbb{N}_0^r$  with distinct entries derived from  $\mathbf{k}$  by fixing  $l$  entries of  $\mathbf{k}$  while increasing the other  $r-l$  entries by 1. Note that for increasing  $\mathbf{k}$  the vectors from  $S_{\mathbf{k}}^{(l)}$  are increasing as well. Hence, by (9) and induction on  $n$ , the coefficients of the  $\mathcal{S}$ -representation of  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  are nonnegative integers for increasing  $\mathbf{b}$  and  $\mathbf{k}$ .

Let us now turn to the proof of (7). For this purpose, we note that  $S_{\underline{r}}^{(1)}$  consists of the single element  $\underline{r+1} \setminus (1)$ . In consequence of (9) we obtain

$$C_{\underline{r}}^{(n+1)}(\mathbf{b}) = \sum_{l=0}^r x_{n+1}^l C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}). \quad (10)$$

By induction on  $r$  we may assume that (7) holds for any increasing vector  $\mathbf{b} \in \mathbb{N}_0^r$  and  $n \in \mathbb{N}$ . Note that the induction basis is given by (6). Using the Laplace expansion, we have

$$V_{\mathcal{S}, \underline{n+1+r} \setminus \mathbf{b}} = \sum_{\substack{\nu=0 \\ \nu \notin \{b_1, \dots, b_r\}}}^{n+r} (-1)^{n-\nu} \operatorname{sgn}(\nu, b_1, \dots, b_r) V_{\mathcal{S}, (\underline{n+r+1} \setminus \nu) \setminus \nu} x_{n+1}^\nu. \quad (11)$$

Applying the induction hypothesis and using (10) yields

$$\begin{aligned} V_{\mathcal{S}, \underline{n+1+r} \setminus \mathbf{b}} &= C_{\underline{r}}^{(n+1)}(\mathbf{b}) \cdot V_{\mathcal{S}, \underline{n+1}} \\ &= \left( \sum_{l=0}^r x_{n+1}^l C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) \right) \cdot \left( \sum_{\nu=0}^n (-1)^{n-\nu} \sigma_{\nu}^{(n)} x_{n+1}^\nu V_{\mathcal{S}, \underline{n}} \right) \end{aligned}$$

$$\begin{aligned}
&= V_{\mathcal{S}, \underline{n}} \sum_{\nu=0}^{n+r} (-1)^{n-\nu} \left( \sum_{l=0}^r (-1)^l \sigma_{\nu-l}^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) \right) x_{n+1}^\nu \\
&= \sum_{\nu=0}^{n+r} (-1)^{n-\nu} C_{\underline{r+1}}^{(n)}(\nu, b_1, \dots, b_r) V_{\mathcal{S}, \underline{n}} x_{n+1}^\nu.
\end{aligned} \tag{12}$$

Note that we use the fact that  $\sigma_\nu^{(n)} = 0$  for  $\nu > n$  or  $\nu < 0$ . Equating the coefficient of  $x_{n+1}^\nu$  in (11) and (12) yields

$$\text{sgn}(\nu, b_1, \dots, b_r) V_{\mathcal{S}, (\underline{n+r+1} \setminus \nu) \setminus (\nu)} = C_{\underline{r+1}}^{(n)}(\nu, b_1, \dots, b_r) V_{\mathcal{S}, \underline{n}}. \tag{13}$$

Let  $(b_0, b_1, \dots, b_r) \in \mathbb{N}_0^{r+1}$  be an arbitrary increasing vector. Note that  $\text{sgn}(b_0, b_1, \dots, b_r) = 1$ . Then (13) yields

$$V_{\mathcal{S}, \underline{n+r+1} \setminus (b_0, b_1, \dots, b_r)} = C_{\underline{r+1}}^{(n)}(b_0, b_1, \dots, b_r) V_{\mathcal{S}, \underline{n}}$$

which establishes (7). Hence,

$$P_{\mathcal{S}, \underline{n+r} \setminus \mathbf{b}} = C_{\underline{r}}^{(n)}(\mathbf{b})$$

and Theorem 1 follows from the fact that the coefficients of the  $\mathcal{S}$ -representation of  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  are nonnegative integers for increasing  $\mathbf{b}$  and  $\mathbf{k}$ .

### 3.2 Vandermonde determinants over the Chebyshev basis

In this section we give the proof of Theorem 2.

Let us first compute the constant  $c_{\mathcal{T}, n}$  in (2). Since  $T_{n-1}(x_n) = 2^{n-2}x_n^{n-1} + O(x_n^{n-2})$  for  $n > 1$ , we have  $c_{\mathcal{T}, n} = 2^{n-2}c_{\mathcal{T}, n-1}$  and  $c_{\mathcal{T}, 1} = 1$ . Hence,  $c_{\mathcal{T}, n} = 2^{(n-1)(n-2)/2}$  and

$$V_{\mathcal{T}, \underline{n}} = 2^{(n-1)(n-2)/2} \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Note that

$$V_{\mathcal{T}, \underline{n+1}} = 2^{n-1} \prod_{i=1}^n (x_{n+1} - x_i) V_{\mathcal{T}, \underline{n}}. \tag{14}$$

Therefore, the coefficients of the  $\mathcal{T}$ -representation of the polynomial  $2^{n-1} \prod_{i=1}^n (x - x_i) \in \mathbb{R}[x]$  with indeterminates  $x_1, \dots, x_n$  play an important role. Analogous to the case of the standard power basis  $\mathcal{S}$  we call these coefficients the elementary symmetric polynomials over  $\mathcal{T}$  and define  $\tau_k^{(n)} = \tau_k(x_1, \dots, x_n)$  by

$$\tau_k^{(0)} = \begin{cases} 0 & \text{if } k \neq 0 \\ 1 & \text{if } k = 0, \end{cases} \quad \tau_k^{(n)} = \tau_{k+1}^{(n-1)} + 2x_n \tau_k^{(n-1)} + \tau_{k-1}^{(n-1)} \tag{15}$$

for  $n \in \mathbb{N}, k \in \mathbb{Z}$ . Note that  $\tau_k^{(n)} = 0$  for  $|k| > n$  and  $\tau_k^{(n)} = \tau_{-k}^{(n)}$ .

**Lemma 3** Let  $n \in \mathbb{N}_0$ . Then

$$2^{n-1} \prod_{i=1}^n (x - x_i) = \sum_{k=0}^n (-1)^{n-k} \tau_k^{(n)} T_k(x).$$

PROOF. The case  $n = 0$  is trivial. Let  $n > 0$ . Then, by induction,

$$2^{n-1} \prod_{i=1}^n (x - x_i) = 2(x - x_n) 2^{n-2} \prod_{i=1}^{n-1} (x - x_i)$$

$$\begin{aligned}
&= \sum_{k=0}^{n-1} '(-1)^{n-1-k} \tau_k^{(n-1)} 2x \mathbb{T}_k(x) + \sum_{k=0}^{n-1} '(-1)^{n-k} 2x_n \tau_k^{(n-1)} \mathbb{T}_k(x) \\
&= \sum_{k=0}^{n-1} '(-1)^{n-1-k} \tau_k^{(n-1)} \mathbb{T}_{k-1}(x) + \sum_{k=0}^{n-1} '(-1)^{n-1-k} \tau_k^{(n-1)} \mathbb{T}_{k+1}(x) \\
&\quad + \sum_{k=0}^{n-1} '(-1)^{n-k} 2x_n \tau_k^{(n-1)} \mathbb{T}_k(x) \\
&= \sum_{k=-1}^n '(-1)^{n-k} \tau_{k+1}^{(n-1)} \mathbb{T}_k(x) + \sum_{k=1}^n '(-1)^{n-k} \tau_{k-1}^{(n-1)} \mathbb{T}_k(x) \\
&\quad + \sum_{k=0}^n '(-1)^{n-k} 2x_n \tau_k^{(n-1)} \mathbb{T}_k(x) \\
&= \sum_{k=1}^n '(-1)^{n-k} (\tau_{k+1}^{(n-1)} + \tau_{k-1}^{(n-1)} + 2x_n \tau_k^{(n-1)}) \mathbb{T}_k(x) \\
&\quad + (-1)^{n-1} (\frac{1}{2} \tau_0^{(n-1)} + \tau_2^{(n-1)} + \frac{1}{2} \tau_0^{(n-1)} + 2x_n \tau_1^{(n-1)}) \mathbb{T}_1(x) \\
&\quad + (-1)^n (\tau_1^{(n-1)} + x_n \tau_0^{(n-1)}) \mathbb{T}_0(x) \\
&= \sum_{k=0}^n '(-1)^{n-k} \tau_k^{(n)} \mathbb{T}_k(x).
\end{aligned}$$

□

By Lemma 3 and (14) we have

$$V_{\mathcal{T}, \underline{n+1}} = \sum_{k=0}^n '(-1)^{n-k} \tau_k^{(n)} \mathbb{T}_k(x_{n+1}) V_{\mathcal{T}, \underline{n}}, \quad (16)$$

and expansion according to the last row of the standard Vandermonde matrix yields

$$V_{\mathcal{T}, \underline{n+1}} = \sum_{k=0}^n (-1)^{n-k} V_{\mathcal{T}, \underline{n+1} \setminus (k)} \mathbb{T}_k(x_{n+1}). \quad (17)$$

Comparison of the coefficients of  $\mathbb{T}_k(x_{n+1})$  in (16) and (17) establishes

**Lemma 4** Let  $n, k \in \mathbb{N}_0$ . Then

$$V_{\mathcal{T}, \underline{n+1} \setminus (k)} = \begin{cases} \frac{1}{2} \tau_0^{(n)} \cdot V_{\mathcal{T}, \underline{n}} & \text{if } k = 0, \\ \tau_k^{(n)} \cdot V_{\mathcal{T}, \underline{n}} & \text{if } k > 0. \end{cases}$$

From the definition (15) it is obvious that the coefficients of the  $\mathcal{T}$ -representation of  $\tau_k^{(n)}$  are nonnegative integers. This proves Theorem 2 for the case of vectors  $\mathbf{a}$  with one gap.

For the general case assume that the index vector  $\mathbf{a}$  has  $r$  gaps. Then  $\mathbf{a} = \underline{n+r} \setminus \mathbf{b}$  for some vector  $\mathbf{b} \in \mathbb{N}_0^r$ . Similar to the representation of  $V_{\mathcal{S}, \mathbf{a}}$  given by (7) we will prove in Theorem 5 that

$$V_{\mathcal{T}, \underline{n+r} \setminus \mathbf{b}} = \delta(b_1) \det \begin{pmatrix} \tau_{b_1}^{(n)} & \cdots & \tau_{b_r}^{(n)} \\ \tau_{b_1-1}^{(n)} + \tau_{b_1+1}^{(n)} & \cdots & \tau_{b_r-1}^{(n)} + \tau_{b_r+1}^{(n)} \\ \vdots & & \vdots \\ \tau_{b_1-(r-1)}^{(n)} + \tau_{b_1+(r-1)}^{(n)} & \cdots & \tau_{b_r-(r-1)}^{(n)} + \tau_{b_r+(r-1)}^{(n)} \end{pmatrix} V_{\mathcal{T}, \underline{n}}, \quad (18)$$



where  $\delta(0) = \frac{1}{2}$  and  $\delta(b_1) = 1$  for  $b_1 > 0$ . Note that Lemma 4 is a special case of (18).

The proof of Theorem 5 requires more technical effort than the proof of the corresponding result (7) for the power basis  $\mathcal{S}$ . Again, it is convenient to study the determinant of matrices that are slightly more general than the matrix given in (18). For this purpose we define

$$c_l^{(n)}(\mathbf{b}) = \begin{cases} \tau_{\mathbf{b}}^{(n)} & \text{if } l = 0, \\ \tau_{\mathbf{b}-l}^{(n)} + \tau_{\mathbf{b}+l}^{(n)} & \text{otherwise.} \end{cases} \quad (19)$$

for  $l, \mathbf{b} \in \mathbb{N}_0$ . Let  $\mathbf{k} = (k_1, \dots, k_r) \in \mathbb{N}_0^r$  and  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$ . Then  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  is defined by

$$C_{\mathbf{k}}^{(n)}(\mathbf{b}) = \det \begin{pmatrix} c_{k_1}^{(n)}(\mathbf{b}_1) & \dots & c_{k_r}^{(n)}(\mathbf{b}_1) \\ \vdots & & \vdots \\ c_{k_1}^{(n)}(\mathbf{b}_r) & \dots & c_{k_r}^{(n)}(\mathbf{b}_r) \end{pmatrix}.$$

Note that this definition of  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  should be distinguished from the definition (8). We choose the same notation to point out the similarities in the proofs of Theorem 1 and Theorem 2.

Theorem 5 states that  $\mathcal{P}_{\mathcal{T}, \mathbf{a}}$  is essentially given by  $C_{\underline{\mathbf{1}}}^{(n)}(\mathbf{b})$ . The proof depends on the recursion formula for  $C_{\underline{\mathbf{1}}}^{(n)}(\mathbf{b})$  given by Lemma 8, however, we postpone Lemma 8 to the end of this section.

**Theorem 5** Let  $r \in \mathbb{N}$  and  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$  an increasing vector. Then

$$V_{\mathcal{T}, \underline{\mathbf{n}}+r \setminus \mathbf{b}} = \begin{cases} \frac{1}{2} C_{\underline{\mathbf{1}}}^{(n)}(\mathbf{b}) \cdot V_{\mathcal{T}, \underline{\mathbf{n}}} & \text{if } b_1 = 0, \\ C_{\underline{\mathbf{1}}}^{(n)}(\mathbf{b}) \cdot V_{\mathcal{T}, \underline{\mathbf{n}}} & \text{if } b_1 > 0. \end{cases} \quad (20)$$

**PROOF.** We prove the statement by induction on  $r$ . Note that the induction basis  $r = 1$  is given by Lemma 4. Then, assume that (20) holds for any increasing vector  $\mathbf{b} = (b_1, \dots, b_r) \in \mathbb{N}_0^r$  and  $n \in \mathbb{N}$ .

By expansion according to the last row, we have

$$V_{\mathcal{T}, \underline{\mathbf{n}}+1+r \setminus \mathbf{b}} = \sum_{\substack{\nu=0 \\ \nu \notin \{b_1, \dots, b_r\}}}^{\mathbf{n}+r} (-1)^{n-\nu} \operatorname{sgn}(\nu, b_1, \dots, b_r) V_{\mathcal{T}, \underline{\mathbf{n}}+r+1 \setminus \mathbf{b} \setminus (\nu)} T_{\nu}(x_{\mathbf{n}+1}).$$

On the other hand, by induction, Lemma 3, and Lemma 8,

$$\begin{aligned} V_{\mathcal{T}, \underline{\mathbf{n}}+1+r \setminus \mathbf{b}} &= C_{\underline{\mathbf{1}}}^{(n+1)}(\mathbf{b}) \cdot V_{\mathcal{T}, \underline{\mathbf{n}}+1} \\ &= \left( \sum_{l=0}^r {}' 2 C_{\underline{\mathbf{r}}+1 \setminus (l)}^{(n)}(\mathbf{b}) T_l(x_{\mathbf{n}+1}) \right) \cdot \left( \sum_{\nu=0}^{\mathbf{n}} {}' (-1)^{n-\nu} \tau_{\nu}^{(n)} T_{\nu}(x_{\mathbf{n}+1}) V_{\mathcal{T}, \underline{\mathbf{n}}} \right) \\ &= V_{\mathcal{T}, \underline{\mathbf{n}}} \left( \sum_{\nu=0}^{\mathbf{n}} {}' (-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{\mathbf{r}}+1 \setminus (0)}^{(n)}(\mathbf{b}) T_{\nu}(x_{\mathbf{n}+1}) \right. \\ &\quad \left. + \sum_{l=1}^r \sum_{\nu=0}^{\mathbf{n}} {}' (-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{\mathbf{r}}+1 \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu+l}(x_{\mathbf{n}+1}) \right. \\ &\quad \left. + \sum_{l=1}^r \sum_{\nu=0}^{\mathbf{n}} {}' (-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{\mathbf{r}}+1 \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu-l}(x_{\mathbf{n}+1}) \right) \end{aligned}$$

$$\begin{aligned}
&= V_{\mathcal{T}, \underline{n}} \left( \sum_{\nu=0}^{n+r} '(-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{r+1} \setminus (0)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \right. \\
&\quad + \sum_{l=1}^r \sum_{\nu=l}^{n+r} '(-1)^{n-\nu+l} \tau_{\nu-l}^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \\
&\quad \left. + \sum_{l=1}^r \sum_{\nu=-l}^{n+r} '(-1)^{n-\nu+l} \tau_{\nu+l}^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \right) \\
&= V_{\mathcal{T}, \underline{n}} \left( \sum_{\nu=1}^{n+r} (-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{r+1} \setminus (0)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) + \frac{1}{2} (-1)^n \tau_0^{(n)} C_{\underline{r+1} \setminus (0)}^{(n)}(\mathbf{b}) T_0(x_{n+1}) \right. \\
&\quad + \sum_{l=1}^r \left( \sum_{\nu=l+1}^{n+r} (-1)^{n-\nu+l} (\tau_{\nu-l}^{(n)} + \tau_{\nu+l}^{(n)}) C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \right. \\
&\quad \quad + \frac{1}{2} (-1)^n \tau_0^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_l(x_{n+1}) \\
&\quad \quad + \frac{1}{2} (-1)^n \tau_0^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{-l}(x_{n+1}) \\
&\quad \quad \left. + \sum_{\nu=1}^{l-1} (-1)^{n-\nu+l} (\tau_{-\nu+l}^{(n)} + \tau_{\nu+l}^{(n)}) C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \right) \\
&\quad \left. + \sum_{l=1}^r (-1)^{n+l} \tau_l^{(n)} C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_0(x_{n+1}) \right) \\
&= V_{\mathcal{T}, \underline{n}} \left( \frac{1}{2} (-1)^n \tau_0^{(n)} C_{\underline{r+1} \setminus (0)}^{(n)}(\mathbf{b}) T_0(x_{n+1}) \right. \\
&\quad + \frac{1}{2} \sum_{l=1}^r (-1)^{n+l} (\tau_{-l}^{(n)} + \tau_l^{(n)}) C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_0(x_{n+1}) \\
&\quad + \sum_{\nu=1}^{n+r} (-1)^{n-\nu} \tau_{\nu}^{(n)} C_{\underline{r+1} \setminus (0)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \\
&\quad \left. + \sum_{l=1}^r \sum_{\nu=1}^{n+r} (-1)^{n-\nu+l} (\tau_{\nu-l}^{(n)} + \tau_{\nu+l}^{(n)}) C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) T_{\nu}(x_{n+1}) \right) \\
&= V_{\mathcal{T}, \underline{n}} \sum_{\nu=0}^{n+r} '(-1)^{n-\nu} \left( \sum_{l=0}^r (-1)^l c_l^{(n)}(\nu) C_{\underline{r+1} \setminus (l)}^{(n)}(\mathbf{b}) \right) T_{\nu}(x_{n+1}) \\
&= \sum_{\nu=0}^{n+r} '(-1)^{n-\nu} C_{\underline{r+1}}^{(n)}(\nu, \mathbf{b}_1, \dots, \mathbf{b}_r) V_{\mathcal{T}, \underline{n}} T_{\nu}(x_{n+1}).
\end{aligned}$$

Equating the coefficient of  $T_{\nu}(x_{n+1})$  yields

$$\text{sgn}(\nu, \mathbf{b}_1, \dots, \mathbf{b}_r) V_{\mathcal{T}, \underline{n+r+1} \setminus (\nu)} = C_{\underline{r+1}}^{(n)}(\nu, \mathbf{b}_1, \dots, \mathbf{b}_r) V_{\mathcal{T}, \underline{n}}. \quad (21)$$

Let  $(\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_r) \in \mathbb{N}_0^{r+1}$  be an arbitrary increasing vector. Then  $\text{sgn}(\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_r) = 1$  and by (21) we have

$$V_{\mathcal{T}, \underline{n+r+1} \setminus (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_r)} = C_{\underline{r+1}}^{(n)}(\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_r) V_{\mathcal{T}, \underline{n}}$$

which completes the proof of our statement.  $\square$

Similar to the proof of Theorem 1 in the preceding section we show in Theorem 6 that the coefficients of the  $\mathcal{T}$ -representation of  $C_{\mathbf{k}}^{(n+1)}(\mathbf{b})$  are nonnegative integers for increasing  $\mathbf{b}, \mathbf{k} \in \mathbb{N}_0^*$ . Then, by Theorem 5, the

coefficients of the  $\mathcal{T}$ -representation of  $P_{\mathcal{T},\mathbf{a}}$  for increasing  $\mathbf{a} \in \mathbb{N}_0^r$  are nonnegative integers. This establishes Theorem 2.

The proof of Theorem 6 uses the following recursion formula for  $c_l^{(n+1)}(\mathbf{b})$  which follows immediately from (15) and (19). Let  $l \in \mathbb{N}_0$ . Then

$$c_l^{(n+1)}(\mathbf{b}) = \begin{cases} c_1^{(n)}(\mathbf{b}) + 2x_{n+1}c_0^{(n)}(\mathbf{b}) & \text{if } l = 0, \\ 2c_0^{(n)}(\mathbf{b}) + 2x_{n+1}c_1^{(n)}(\mathbf{b}) + c_2^{(n)}(\mathbf{b}) & \text{if } l = 1, \\ c_{l-1}^{(n)}(\mathbf{b}) + 2x_{n+1}c_l^{(n)}(\mathbf{b}) + c_{l+1}^{(n)}(\mathbf{b}) & \text{if } l > 1. \end{cases} \quad (22)$$

**Theorem 6** For any  $r \in \mathbb{N}$  and increasing  $\mathbf{b}, \mathbf{k} \in \mathbb{N}_0^r$  the coefficients of the  $\mathcal{T}$ -representation of  $C_{\mathbf{k}}^{(n+1)}(\mathbf{b})$  are nonnegative integers.

PROOF. Note that  $C_{\mathbf{k}}^{(n)}(\mathbf{b})$  vanishes when  $\mathbf{k}$  (or  $\mathbf{b}$ ) does not consist of distinct entries. Expand  $C_{\mathbf{k}}^{(n+1)}(\mathbf{b})$  according (22) and collect the terms corresponding to powers of  $x_{n+1}$ . Then

$$C_{\mathbf{k}}^{(n+1)}(\mathbf{b}) = \sum_{l=0}^r 2^l x_{n+1}^l \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(l)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}). \quad (23)$$

where  $S_{\mathbf{k}}^{(l)}$  is the set of vectors  $\mathbf{q} \in \mathbb{N}_0^r$  with distinct entries derived from  $\mathbf{k}$  by fixing  $l$  entries of  $\mathbf{k}$  while increasing or decreasing the other  $r - l$  entries by 1. The factor  $\delta_{\mathbf{q}}$  is always 1 except in the case  $k_i = 1$  and  $q_i = 0$  where  $\delta_{\mathbf{q}}$  equals 2.

We call an entry  $q_i$  of  $\mathbf{q} \in S_{\mathbf{k}}^{(l)}$  active if  $q_i = k_i$ . A subsequence  $(q_i, q_{i+1}, \dots, q_j)$  of active entries of  $\mathbf{q}$  is called active sequence if  $q_{i+1} = q_i + 1, q_{i+2} = q_{i+1} + 1, \dots, q_j = q_{j-1} + 1$ . An active pair of  $\mathbf{q}$  is an active sequence of length 2.

Applying (1) to (23) and collecting the coefficients of the  $T_l(x_{n+1})$ 's, we have

$$\begin{aligned} C_{\mathbf{k}}^{(n+1)}(\mathbf{b}) &= \sum_{\substack{0 \leq l \leq r \\ l \text{ even}}} \sum_{j=0}^{\lfloor l/2 \rfloor} 2^{\binom{l}{\lfloor l/2 \rfloor + j}} T_{2j}(x) \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(l)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \\ &+ \sum_{\substack{0 \leq l \leq r \\ l \text{ odd}}} \sum_{j=0}^{\lfloor l/2 \rfloor} 2^{\binom{l}{\lfloor l/2 \rfloor + j}} T_{2j+1}(x) \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(l)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \\ &= \sum_{m=0}^{\lfloor r/2 \rfloor} \sum_{j=0}^m 2^{\binom{2m}{m+j}} T_{2j}(x) \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(2m)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \\ &+ \sum_{m=0}^{\lfloor \frac{r-1}{2} \rfloor} \sum_{j=0}^m 2^{\binom{2m+1}{m+1+j}} T_{2j+1}(x) \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(2m+1)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \\ &= \sum_{j=0}^{\lfloor r/2 \rfloor} 2^{\binom{r}{j}} T_{2j}(x) \underbrace{\left( \sum_{m=j}^{\lfloor r/2 \rfloor} \binom{2m}{m+j} \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(2m)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \right)}_{=: K_{2j}(\mathbf{k})} \end{aligned}$$

$$+ \sum_{j=0}^{\lfloor \frac{r-1}{2} \rfloor} 2T_{2j+1}(x) \underbrace{\left( \sum_{m=j}^{\lfloor \frac{r-1}{2} \rfloor} \binom{2m+1}{m+1+j} \sum_{\mathbf{q} \in S_{\mathbf{k}}^{(2m+1)}} \delta_{\mathbf{q}} C_{\mathbf{q}}^{(n)}(\mathbf{b}) \right)}_{=: K_{2j+1}(\mathbf{k})} \quad (24)$$

Suppose that for every  $m = j, \dots, \lfloor r/2 \rfloor$  each  $\mathbf{q} \in S_{\mathbf{k}}^{(2m)}$  is increasing. Then, by induction, the coefficients of the  $\mathcal{T}$ -representation of  $C_{\mathbf{q}}^{(n)}(\mathbf{b})$  are nonnegative integers and so are the coefficients of  $K_{2j}(\mathbf{k})$ . However, if  $\mathbf{q}$  has an active pair  $(q_i, q_{i+1})$ , there exists  $\mathbf{p} \in S_{\mathbf{k}}^{(2(m-1))}$  with  $p_i = q_{i+1}$  and  $p_{i+1} = q_i$  ( $p_j = q_j$  for  $j \neq i, i+1$ ). We say  $\mathbf{p}$  is derived from  $\mathbf{q}$  by flipping the pair  $(q_i, q_{i+1})$ .

In general, if  $\mathbf{p} \in S_{\mathbf{k}}^{(2(m-s))}$  derived from  $\mathbf{q}$  by flipping the  $s$  distinct active pairs, then  $\mathbf{p}$  is non-increasing, however

$$C_{\mathbf{p}}^{(n)}(\mathbf{b}) = (-1)^s C_{\mathbf{q}}^{(n)}(\mathbf{b}). \quad (25)$$

On the other hand, every not increasing  $\mathbf{p}$  corresponds by means of (25) to some unique increasing  $\mathbf{q}$ .

Let  $Q_{\mathbf{k}}^{(1)}$  denote the set of increasing elements from  $S_{\mathbf{k}}^{(1)}$ . By induction, the coefficients of the  $\mathcal{T}$ -representation of  $C_{\mathbf{q}}^{(n)}(\mathbf{b})$  are nonnegative integers for  $\mathbf{q} \in Q_{\mathbf{k}}^{(1)}$ . Therefore, using (25), we replace in (24) the sum over  $S_{\mathbf{k}}^{(1)}$  by a sum over  $Q_{\mathbf{k}}^{(1)}$  by introducing appropriate weights. It then remains to show that these weights are nonnegative integers.

Let  $\alpha_k(\mathbf{q})$  denote the number of distinct vectors that can be derived by flipping  $k$  active pairs of  $\mathbf{q}$ . Note that this number depends only on the number and length of active sequences of  $\mathbf{q}$ . Assume  $\mathbf{q}$  consists of  $s$  active sequences of lengths  $N_1, \dots, N_s$ . Then  $\alpha_k(\mathbf{q}) = \alpha_k(N_1, \dots, N_s)$  where

$$\alpha_k(N_1) = \begin{cases} \binom{N_1-k}{k} & \text{if } N_1 \geq k \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \alpha_k(N_1, \dots, N_s) = \sum_{i=0}^k \alpha_i(N_1) \alpha_{k-i}(N_2, \dots, N_s).$$

Applying the addition law for binomial coefficients,

$$\alpha_k(N_1, \dots, N_i, \dots, N_s) = \alpha_k(N_1, \dots, N_i - 1, \dots, N_s) + \alpha_{k-1}(N_1, \dots, N_i - 2, \dots, N_s). \quad (26)$$

In order to deal with the sum in (24) we have to take the multiplicity  $\delta_{\mathbf{p}}$  ( $\mathbf{p}$  derived from  $\mathbf{q}$ ) into account. Suppose  $(0, 1)$  is an active pair of  $\mathbf{q}$ , so  $(0, 1)$  is a subsequence of the first active sequence of  $\mathbf{q}$ . Then the number of vectors derived by flipping  $k$  active pairs of  $\mathbf{q}$  counted with multiplicity is denoted by  $\beta_k(\mathbf{q}) = \beta_k(N_1, \dots, N_s)$  where

$$\beta_k(N_1, N_2, \dots, N_s) = 2\alpha_{k-1}(N_1 - 2, N_2, \dots, N_s) + \alpha_k(N_1 - 1, N_2, \dots, N_s).$$

Note that  $\beta_k$  is not symmetric, however, using (26),

$$\beta_k(N_1, \dots, N_i, \dots, N_s) = \beta_k(N_1, \dots, N_i - 1, \dots, N_s) + \beta_{k-1}(N_1, \dots, N_i - 2, \dots, N_s). \quad (27)$$

Finally, let  $\gamma_k(\mathbf{q})$  equal  $\beta_k(\mathbf{q})$  or  $\alpha_k(\mathbf{q})$  depending on whether  $(0, 1)$  is an active pair of  $\mathbf{q}$  or not. Then for  $j = 0, \dots, \lfloor r/2 \rfloor$

$$K_{2j}(\mathbf{k}) = \sum_{m=j}^{\lfloor r/2 \rfloor} \sum_{\mathbf{q} \in Q_{\mathbf{k}}^{(2m)}} \delta_{\mathbf{q}} \omega_{j,m}(\mathbf{q}) C_{\mathbf{q}}^{(n)}(\mathbf{b}) \quad (28)$$

where

$$\omega_{j,m}(\mathbf{q}) = \sum_{i=j}^m (-1)^{m-i} \binom{2i}{i+j} \gamma_{m-i}(\mathbf{q}) = \sum_{k \in \mathbb{Z}} (-1)^k \binom{2m-2k}{m-j-k} \gamma_k(\mathbf{q}).$$

In a similar manner, we have for  $j = 0, \dots, \lfloor \frac{r-1}{2} \rfloor$

$$K_{2j+1}(\mathbf{k}) = \sum_{m=j}^{\lfloor \frac{r-1}{2} \rfloor} \sum_{\mathbf{q} \in Q_{\mathbf{k}}^{(2m+1)}} \delta_{\mathbf{q}} \omega'_{j,m}(\mathbf{q}) C_{\mathbf{q}}^{(n)}(\mathbf{b}) \quad (29)$$

where

$$\omega'_{j,m}(\mathbf{q}) = \sum_{k \in \mathbb{Z}} (-1)^k \binom{2m-2k+1}{m-k-j+1} \gamma_k(\mathbf{q}).$$

Fortunately, we don't have to compute the weights  $\omega_{j,m}$ , however, we have to show that  $\omega_{j,m}(\mathbf{q}) \geq 0$  for  $\mathbf{q} \in Q_{\mathbf{k}}^{(2m)}$  and  $\omega'_{j,m}(\mathbf{q}) \geq 0$  for  $\mathbf{q} \in Q_{\mathbf{k}}^{(2m+1)}$ . (From the definition of  $\gamma_k(\cdot)$  it is obvious that the weights are integers). Note that  $\mathbf{q} \in Q_{\mathbf{k}}^{(l)}$  implies that the sum of the lengths of the active sequences of  $\mathbf{q}$  is at most  $l$ . Let

$$\omega_{j,m}^{(\alpha)}(N_1, \dots, N_s) = \sum_{k \in \mathbb{Z}} (-1)^k \binom{2m-2k}{m-j-k} \alpha_k(N_1, \dots, N_s) \quad (30)$$

and define  $\omega_{j,m}^{(\alpha)}(\cdot)$ ,  $\omega_{j,m}^{(\beta)}(\cdot)$ ,  $\omega_{j,m}^{(\beta)}(\cdot)$  likewise.

From (26) and (30) we derive that

$$\omega_{j,m}^{(\alpha)}(N_1, \dots, N_i - 1, \dots, N_s) = \omega_{j,m}^{(\alpha)}(N_1, \dots, N_i, \dots, N_s) + \omega_{j,m-1}^{(\alpha)}(N_1, \dots, N_i - 2, \dots, N_s). \quad (31)$$

Using (31) twice, we have

$$\begin{aligned} & \omega_{j,m}^{(\alpha)}(N_1 - 1, N_2 + 1, \dots) \\ &= \omega_{j,m}^{(\alpha)}(N_1, N_2, \dots) + \omega_{j,m-1}^{(\alpha)}(N_1 - 2, N_2, \dots) - \omega_{j,m-1}^{(\alpha)}(N_1 - 1, N_2 - 1, \dots). \end{aligned} \quad (32)$$

Assume that  $\omega_{j,m}^{(\alpha)}(2m)$  is a nonnegative integer for  $m \geq j$ . Then, by fixing  $j$  and induction on  $m$  and  $N$  we derive from (31) that for  $0 < N < 2m$

$$\omega_{j,m}^{(\alpha)}(N - 1) \geq \omega_{j,m}^{(\alpha)}(N) \geq 0. \quad (33)$$

Likewise from (32) we conclude that

$$\omega_{j,m}^{(\alpha)}(N_1 - 1, N_2 + 1, \dots, N_s) \leq \omega_{j,m}^{(\alpha)}(N_1, N_2, \dots, N_s)$$

provided that  $N_1 < N_2$  and  $\sum N_i \leq 2m$ . Hence, under the same assumption,

$$\omega_{j,m}^{(\alpha)}(N_1, N_2, \dots, N_s) \geq \omega_{j,m}^{(\alpha)}(\sum N_i)$$

which is a nonnegative integer by (33).

The same conclusions are valid for  $\omega_{j,m}^{(\beta)}$  provided that  $\omega_{j,m}^{(\beta)}(2m+1)$  is a nonnegative integer. Furthermore, because of (27), the above reasoning can be applied for  $\omega_{j,m}^{(\beta)}$  and  $\omega_{j,m}^{(\beta)}$ . From the following Lemma 7 we conclude

$$\begin{aligned} \omega_{j,m}(\mathbf{q}) &\in \mathbb{N}_0 \quad \text{for } \mathbf{q} \in Q_{\mathbf{k}}^{(2m)}, m = j, \dots, \lfloor r/2 \rfloor, \text{ and} \\ \omega'_{j,m}(\mathbf{q}) &\in \mathbb{N}_0 \quad \text{for } \mathbf{q} \in Q_{\mathbf{k}}^{(2m+1)}, m = j, \dots, \lfloor \frac{r-1}{2} \rfloor, \end{aligned}$$

which completes the proof.  $\square$

**Lemma 7** Let  $j, m \in \mathbb{N}$  with  $m \geq j$ . We use the above notation. Then

$$\omega_{j,m}^{(\alpha)}(2m) = \omega_{j,m}^{(\alpha)}(2m+1) = 1$$

and

$$\omega_{j,m}^{(\beta)}(2m) = \omega_{j,m}^{(\beta)}(2m+1) = \begin{cases} 1 & \text{if } j = m, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. Using basic binomial coefficient identities, we have

$$\begin{aligned} \omega_{j,m}^{(\alpha)}(2m) &= \sum_k (-1)^k \binom{2m-2k}{m-j-k} \binom{2m-k}{k} = \sum_{k \leq 2m} (-1)^k \binom{2m-2k}{m-j-k} \binom{2m-k}{2m-2k} \\ &= \sum_{k \leq 2m} (-1)^k \binom{2m-k}{m-j} \binom{m-j}{k} = (-1)^{m-j} \binom{-1}{m+j} = 1. \end{aligned}$$

Note that  $\beta_k(2m) = 2 \binom{2m-2k-1}{k-1} + \binom{2m-2k-1}{k} = \binom{2m-2k}{k-1} + \binom{2m-2k}{k}$ . Let  $m > j$ . Then

$$\begin{aligned} \omega_{j,m}^{(\beta)}(2m) &= \sum_k (-1)^k \binom{2m-2k}{m-j-k} \left( \binom{2m-k-1}{k-1} + \binom{2m-k}{k} \right) = 1 + \sum_k (-1)^k \binom{2m-2k}{m-j-k} \binom{2m-k-1}{2m-2k} \\ &= 1 + \sum_{k \leq 2m-1} (-1)^k \binom{2m-1-k}{m-1-j} \binom{m-1-j}{k-1} = 1 + (-1)^{m-j} \binom{-1}{m+j-1} = 0, \end{aligned}$$

and

$$\omega_{m,m}^{(\beta)}(2m) = 1 + \sum_k (-1)^k \binom{2m-1-k}{-1} \binom{-1}{k-1} = 1 + 0 = 1.$$

The identities for  $\omega_{j,m}^{(\alpha)}(2m+1)$  and  $\omega_{j,m}^{(\beta)}(2m+1)$  are established in the same way.  $\square$

**Lemma 8** Let  $n, r \in \mathbb{N}$  and  $\mathbf{b} \in \mathbb{N}_0^r$ . Then

$$C_{\underline{r}}^{(n+1)}(\mathbf{b}) = \sum_{l=0}^r {}' 2 T_l(x_{n+1}) C_{\underline{r+1} \setminus \{l\}}^{(n)}(\mathbf{b}).$$

PROOF. We use the notation from the proof of Theorem 6. For each  $0 \leq l \leq r$  the set  $Q_r^{(l)}$  consists of the single element  $\underline{r+1} \setminus \{l\}$ . Note that  $\delta_{\underline{r+1} \setminus \{l\}} = 1$  and  $(0, 1)$  is an active pair of  $\underline{r+1} \setminus \{l\}$ . Furthermore,  $\underline{r+1} \setminus \{l\}$  consists of one active sequence of length  $l$ . Hence, by (28) and Lemma 7,

$$K_{2j}(\underline{r}) = \sum_{m=j}^{\lfloor r/2 \rfloor} \sum_{\mathbf{q} \in Q_{\underline{r}}^{(2m)}} \delta_{\mathbf{q}} \omega_{j,m}^{(\beta)}(\mathbf{q}) C_{\mathbf{q}}^{(n)}(\mathbf{b}) = \sum_{m=j}^{\lfloor r/2 \rfloor} \omega_{j,m}^{\beta}(2m) C_{\underline{r+1} \setminus \{2m\}}^{(n)}(\mathbf{b}) = C_{\underline{r+1} \setminus \{2j\}}^{(n)}(\mathbf{b})$$

and likewise by (29)

$$K_{2j}(\underline{r}) = \sum_{m=j}^{\lfloor \frac{r-1}{2} \rfloor} \omega_{j,m}^{\beta}(2m+1) C_{\underline{r+1} \setminus \{2m+1\}}^{(n)}(\mathbf{b}) = C_{\underline{r+1} \setminus \{2j+1\}}^{(n)}(\mathbf{b}).$$

Hence

$$\begin{aligned} C_{\underline{r}}^{(n+1)}(\mathbf{b}) &= \sum_{j=0}^{\lfloor r/2 \rfloor} {}' 2 T_{2j}(x) C_{\underline{r+1} \setminus \{2j\}}^{(n)}(\mathbf{b}) + \sum_{j=0}^{\lfloor \frac{r-1}{2} \rfloor} {}' 2 T_{2j+1}(x) C_{\underline{r+1} \setminus \{2j+1\}}^{(n)}(\mathbf{b}) \\ &= \sum_{l=0}^r {}' 2 T_l(x_{n+1}) C_{\underline{r+1} \setminus \{l\}}^{(n)}(\mathbf{b}). \end{aligned}$$

$\square$

## 4 Number of Real Zeros of Sparse Polynomials

One of the basic results in the theory of orthogonal polynomials is that the zeros of these polynomials are all real, simple and lie in the interval of orthogonality. By Theorem 2 the sign of the coefficients of the  $\mathcal{T}$ -representation of  $P_{\mathcal{T},\mathbf{a}}$  equals  $\text{sgn}(\mathbf{a})$ , hence,  $P_{\mathcal{T},\mathbf{a}}$  does not vanish when the indeterminates take values in either  $(-\infty, -1]$  or  $[1, \infty)$ . Furthermore, the standard Vandermonde matrix is nonsingular for distinct indeterminates. We conclude

**Theorem 9** The generalized Vandermonde matrix over  $\mathcal{T}$  is nonsingular provided that the indeterminates take distinct values from either  $(-\infty, -1]$  or  $[1, \infty)$ .

As an application, we answer the question posed by Lakshman and Saunders [5] about the relationship of the number of real roots of a polynomial and its sparsity with respect to the Chebyshev basis.

**Theorem 10** The number of real zeros of a nonvanishing polynomial in the interval  $[1, \infty)$  (or  $(-\infty, -1]$ ) does not exceed its sparsity with respect to  $\mathcal{T}$ .

PROOF. Let  $f \not\equiv 0$  be an  $m$ -sparse polynomial with  $\mathcal{T}$ -representation

$$f(x) = \sum_{i=1}^m c_i T_{a_i}(x)$$

with  $\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{N}_0^m$ . Assume that there exists distinct  $x_1, \dots, x_m \in [1, \infty)$  such that  $f(x_i) = 0$  for  $i = 1, \dots, m$ . Then

$$\begin{pmatrix} T_{a_1}(x_1) & \dots & T_{a_m}(x_1) \\ \vdots & & \vdots \\ T_{a_1}(x_m) & \dots & T_{a_m}(x_m) \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Note that the left-hand side matrix is nonsingular by Theorem 9. Therefore,  $(c_1, \dots, c_m) = \mathbf{0}$  and  $f \equiv 0$ , contradicting the assumption.  $\square$

## 5 VC Dimension of Sparse Polynomials

Given a set  $X$  and a finite subset  $S$ , a collection  $C$  of subsets of  $X$  *shatters*  $S$  if for every subset  $T$  of  $S$  there is a  $c \in C$  with  $T = c \cap S$ . The *Vapnik-Chervonenkis dimension* of the class  $C$  is the largest (possible infinite) integer  $d$  such that some set  $S \subset X$  of cardinality  $d$  is shattered by  $C$ . If the VC dimension of  $C$  is finite,  $C$  is called a VC class.

This notation was introduced by Vapnik and Chervonenkis [9] to give sufficient conditions on a class  $C$  of events so that the relative frequency of an event in the class converges to its probability. Recently, the VC dimension proved to be useful in the field of uniform and distribution-free learnability. In fact, it turns out that the VC classes are exactly the concept classes which are learnable in Valiant's so-called PAC model of learning [8] (cmp. [2] for this relationship as well as for bounds on the sample complexity in terms of the VC dimension).

Let  $\mathcal{F}$  be a collection of real-valued functions on a set  $X$ . Let  $\text{pos}(\mathbf{y} - \mathcal{F})$  denote the collection of all sets  $\text{pos}(\mathbf{y} - f) = \{x \in X \mid y - f(x) > 0\}$  for  $f \in \mathcal{F}$ . We identify the VC dimension of  $\mathcal{F}$  with the VC dimension of  $\text{pos}(\mathbf{y} - \mathcal{F})$ .

Wenocur and Dudley [10] proved that for vector spaces  $\mathcal{F}$  of real-valued functions the VC dimension of  $\text{pos}(\mathbf{y} - \mathcal{F})$  coincides with the vector space dimension of  $\mathcal{F}$ . The VC dimension of general collections of

real-valued functions might therefore be thought of as a measure of the degree of freedom in the absence of an underlying vector space structure.

Let  $\mathcal{P}_{\Phi, t} \subset \mathbb{R}[x]$  denote the set of univariate polynomials with  $t$ -sparse  $\Phi$ -representation. A labeling of a finite set  $S \subset \mathbb{R}^2$  is a mapping  $\sigma : S \mapsto \{0, 1\}$ . A polynomial  $f \in \mathcal{P}_{\Phi, t}$  is said to satisfy the labeling  $\sigma$  on  $S$  if  $\sigma((x, y)) = 1 \iff y > f(x)$  for every  $(x, y) \in S$ . Hence, the set  $S$  is shattered by  $\mathcal{P}_{\Phi, t}$  if there is a satisfying  $f_\sigma \in \mathcal{P}_{\Phi, t}$  for every labeling  $\sigma$  of  $S$ . For a fixed set  $S$  of cardinality  $d$  we may identify the set of all labelings of  $S$  with  $\{0, 1\}^d$ .

In [4] Karpinski and Werther prove linear bounds on the VC dimension of sparse polynomials over the standard power basis. In this section we extend their result to sparse polynomials over the Chebyshev basis.

**Theorem 11** The VC dimension of  $\mathcal{P}_{\mathcal{T}, t}$  on  $[1, \infty) \times \mathbb{R}$  equals  $2t$  and is infinite on  $(-1, 1) \times \mathbb{R}$ .

Theorem 11 follows from the three lemmas below. The first lemma gives a lower bound on the VC dimension of sparse polynomials over arbitrary polynomial sets.

**Lemma 12** Let  $\Phi = \{\Phi_n\}$  be an arbitrary polynomial set. For any  $t \in \mathbb{N}$  and  $b \in \mathbb{R}$  there exists a set  $S_t \subset [b, \infty) \times \mathbb{R}$  of size  $2t$  that is shattered by polynomials of degree less than  $2t$  and  $t$ -sparse  $\Phi$ -representation.

PROOF. We prove the statement by induction on  $t$ .

Let  $t = 1$ . Then  $\Phi_0(x) = c$  and  $\Phi_1 = ax + d$ . We may assume  $a > 0$ . It is easily verified that the set  $S_1 = \{(b, \Phi_1(b) + 1), (b + \frac{4}{a}, \Phi_1(b) + 3)\}$  is shattered by the 1-sparse polynomials  $f_{00}(x) = \Phi_1(b) + 4$ ,  $f_{01}(x) = \Phi_1(b) + 2$ ,  $f_{10}(x) = \Phi_1(x)$ , and  $f_{11}(x) = \Phi_1(b) - 1$ , each of degree at most 1.

For purpose of induction let  $k = 2t$  and assume that  $S_t = \{(x_i, y_i)\}_{i=1, \dots, k}$  is shattered by the set  $F_t = \{f_\sigma\}_{\sigma \in \{0, 1\}^k}$  of  $t$ -sparse polynomials of degree at most  $k - 1$ . We assume that the leading coefficients of  $\Phi_k$  and  $\Phi_{k+1}$  are positive.

We construct the set  $S_{t+1}$  by adding a set  $S = \{(x_{k+1}, y_{k+1}), (x_{k+2}, y_{k+2})\}$ ,  $x_k < x_{k+1} < x_{k+2}$  to  $S_t$ . We will specify  $S$  such that  $S$  is shattered by a set  $G = \{g_\gamma\}_{\gamma \in \{0, 1\}^2}$  of 1-sparse polynomials of degree at most  $k + 1$  with the following property: For each labeling  $\sigma \gamma \in \{0, 1\}^k \times \{0, 1\}^2$  the polynomial  $f_\sigma + g_\gamma$  satisfies  $\sigma$  on  $S_t$  and  $\gamma$  on  $S$ . Hence the set  $S_{t+1} = S_t \cup S$  of size  $2t + 2$  is shattered by the set  $F_{t+1} = \{f + g \mid f \in F_t, g \in G\}$ . Note that each  $f \in F_{t+1}$  is  $t + 1$ -sparse and of degree at most  $k + 1$ .

To be more precise we define

$$\epsilon_i = \min_{f \in F_t} |f(x_i) - y_i|$$

for  $i = 1, \dots, k$  and

$$\epsilon_i = \min_{g \in G} |g(x_i) - y_i|$$

for  $i = k + 1, k + 2$ . Since  $\deg(f) < k$ ,  $f \in F_t$  there exists a constant  $c \in \mathbb{R}^+$  such that

$$\forall_{f \in F_t} |f(x)| \leq c |\Phi_{k-1}(x)|$$

for  $x > x_k$ . Clearly, the above requirement is fulfilled if  $|g_\gamma(x_i)| < \epsilon_i$  for  $\gamma \in \{0, 1\}^2$  and  $i = 1, \dots, k$  and  $c |\Phi_{k-1}(x_i)| < \epsilon_i$  for  $i = k + 1, k + 2$ . Let

$$c_{00} = \frac{1}{2} \min_{\substack{i=1, \dots, k \\ \Phi_k(x_i) \neq 0}} \frac{\epsilon_i}{|\Phi_k(x_i)|} \quad \text{and} \quad c_{01} = \frac{1}{2} c_{00}.$$

Now we set  $g_{00} = c_{00} \Phi_k$ ,  $g_{01} = c_{01} \Phi_k$ , and  $g_{11} = 0$ . By definition  $|g_\gamma(x_i)| < \epsilon_i$  for  $\gamma = (0, 0), (0, 1), (1, 1)$  and  $i = 1, \dots, k$ . Since  $\deg(\Phi_k) > \deg(\Phi_{k-1})$  there exists a  $x_{k+1} \in \mathbb{R}$  such that



$g_{01}(x_{k+1}) \geq 3c |\Phi_{k-1}(x_{k+1})|$  and  $\Phi_k(x) > |\Phi_{k-1}(x)|$  for  $x > x_{k+1}$ . Let  $y_{k+1} = 2c |\Phi_{k-1}(x_{k+1})|$ . We set

$$c_{10} = \min \left\{ \frac{1}{2} \min_{\substack{i=1, \dots, k \\ \Phi_{k+1}(x_i) \neq 0}} \frac{\epsilon_i}{|\Phi_{k+1}(x_i)|}, c \frac{|\Phi_{k-1}(x_{k+1})|}{|\Phi_{k+1}(x_{k+1})|} \right\}$$

and  $g_{10} = c_{10} \Phi_{k+1}$ . Note that  $|g_{10}(x_i)| < \epsilon_i$  for  $i = 1, \dots, k$  and  $|g_{10}(x_{k+1})| + c |\Phi_{k-1}(x_{k+1})| \leq y_{k+1}$ . Hence  $\epsilon_{k+1} > c |\Phi_{k-1}(x_{k+1})|$ .

Since  $\deg(\Phi_{k+1}) > \deg(\Phi_k)$  there exists a  $x_{k+2} \in \mathbb{R}$  such that  $g_{01}(x_{k+2}) \geq g_{00}(x_{k+2})$ . Let  $y_{k+1} = \frac{1}{2}(g_{00}(x_{k+2}) + g_{10}(x_{k+2}))$ . Then  $\epsilon_{k+2} > c |\Phi_{k-1}(x_{k+2})|$  and the claimed properties are established.  $\square$

The upper bound depends on the number of real roots of sparse polynomials.

**Lemma 13** Let  $S \subset [1, \infty) \times \mathbb{R}$  be a set shattered by  $\mathcal{P}_{\mathcal{T}, t}$ . Then  $|S| \leq 2t$ .

PROOF. Let  $S = \{(x_i, y_i)\}_{i=1, \dots, d}$ ,  $b \leq x_1 < x_2 < \dots < x_d$ , be a set of points shattered by  $\mathcal{P}_{\mathcal{T}, t}$ . Let  $f_1, f_2 \in \mathcal{P}_{\mathcal{T}, t}$  satisfy the two alternating labelings  $\sigma_1 = (1, 0, 1, 0, \dots)$  and  $\sigma_2 = (0, 1, 0, 1, \dots)$ . Then  $F = (f_1 - f_2)$  is  $2t$ -sparse. Furthermore,  $F(x_i) \cdot F(x_{i+1}) < 0$  for  $i = 1, \dots, d-1$ , forcing  $F$  to have at least  $d-1$  real roots in the interval  $(x_1, x_d) \subset [1, \infty)$ . By Theorem 10,  $d-1 < 2t$ .  $\square$

**Lemma 14** The class  $\mathcal{P}_{\mathcal{T}, t}$  is of infinite VC dimension on  $(-1, 1) \times \mathbb{R}$ .

PROOF. We construct for each  $m \in \mathbb{N}$  a set  $S_m \subset (-1, 1) \times \mathbb{R}$  of size  $m$  shattered by  $\mathcal{T} = \mathcal{P}_{\mathcal{T}, 1}$ .

Let  $x \in (-1, 1)$ . Then  $T_n(x) = \cos(n \arccos(x))$ . Note that  $T_n(x) = 1$  if  $x = \cos(\frac{2k}{n}\pi)$  for some  $k \in \mathbb{N}$  and  $T_n(x) < 1$  otherwise.

Let  $p_i > 2$ ,  $i = 1, \dots, m$ , denote distinct prime numbers and let  $x_i = \cos(\frac{2}{p_i}\pi)$ ,  $i = 1, \dots, m$ . For each labeling  $\sigma \in \{0, 1\}^m$  we define  $n_\sigma = \prod_{i=1}^m p_i^{\sigma(i)}$ . Then for all  $\sigma \in \{0, 1\}^m$ ,

$$T_{n_\sigma}(x_i) = 1 \iff \frac{2}{p_i}\pi = \frac{2k}{n_\sigma}\pi \text{ for some } k \in \mathbb{N} \iff p_i \mid n_\sigma \iff \sigma(i) = 1$$

and  $T_{n_\sigma}(x_i) < 1$  if  $\sigma(i) = 0$ . Define  $\epsilon > 0$  such that  $T_{n_\sigma}(x_i) < 1 - \epsilon$  for all labelings  $\sigma$  and all  $i$  with  $\sigma(i) = 0$ . Let  $S_m = \{(x_i, 1 - \epsilon)\}_{i=1, \dots, m}$ . Then for each  $\sigma \in \{0, 1\}^m$ ,  $T_{n_\sigma}$  satisfies  $\sigma$  on  $S_m$ , i.e. the set  $S_m$  is shattered by  $\mathcal{T}$ .  $\square$

## References

- [1] M. BEN-OR AND P. TIWARI, *A deterministic algorithm for sparse multivariate polynomial interpolation*, in Proc. 20th Annual ACM Symposium on Theory of Computing, Association for Computing Machinery, New York, 1988, pp. 301–309.
- [2] A. BLUMER, A. EHRENFEUCHT, D. HAUSSLER, AND M. WARMUTH, *Learnability and the vapnik-chervonenkis dimension*, J. ACM, 36 (1989), pp. 929–965.
- [3] R.J. EVANS AND I.M. ISAACS, *Generalized Vandermonde Determinants and Roots of Unity of Prime Order*, Proc. Amer. Math. Soc., 58 (1976), pp. 51–54.
- [4] M. KARPINSKI AND T. WERTHER, *VC Dimension and Sampling Complexity of Learning Sparse Polynomials and Rational Functions*, ICSI Technical Report TR-91-044, August 1991, to appear in SIAM J. Comput. 22, No. 6 (1993).

- [5] Y. LAKSHMAN AND D. SAUNDERS, *Sparse Polynomial Interpolation in Non-standard Bases*, Technical Report 93-11, Dept. of Computer & Information Sciences, University of Delaware, September 1992.
- [6] O.H. MICHELL, *Note on Determinants of Powers*, Amer. J. Math. 4 (1881), pp. 341–344.
- [7] G. SZEGÖ *Orthogonal Polynomials*, Amer. Math. Soc., Colloquium Public., vol. 23, 4th ed., 1975.
- [8] L. VALIANT, *A theory of the learnable*, Comm. ACM, 27 (1984), pp. 1134–1142.
- [9] V. VAPNIK AND A. CHERVONENKIS, *On the uniform convergence of relative frequencies of events and their probabilities*, Th. Prob. and its Appl., 16 (1971), pp. 264–280.
- [10] R. WENOCUR AND R. DUDLEY, *Some special vapnik-chervonenkis classes*, Discrete Math., 33 (1981), pp. 313–318.
- [11] T. WERTHER, *A Note on Generalized Vandermonde Determinants*, manuscript, 1990.