Combinatorics, Probability and Computing (2019) 28, 253–279.
doi:10.1017/S0963548318000342

Bounding the Number of Common Zeros of Multivariate Polynomials and Their Consecutive Derivatives

O. GEIL[1] and U. MARTÍNEZ-PEÑAS[1,2]

[1]Department of Mathematical Sciences, Aalborg University, Skjernvej 4A, 9220 Aalborg, Denmark
(e-mail: olav@math.aau.dk)
[2]Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario M5S 3G4, Canada
(e-mail: umberto@comm.utoronto.ca)

Received 5 July 2017; revised 15 May 2018; first published online 3 August 2018

We upper-bound the number of common zeros over a finite grid of multivariate polynomials and an arbitrary finite collection of their consecutive Hasse derivatives (in a coordinate-wise sense). To that end, we make use of the tool from Gröbner basis theory known as footprint. Then we establish and prove extensions in this context of a family of well-known results in algebra and combinatorics. These include Alon's combinatorial Nullstellensatz [1], existence and uniqueness of Hermite interpolating polynomials over a grid, estimations of the parameters of evaluation codes with consecutive derivatives [20], and bounds on the number of zeros of a polynomial by DeMillo and Lipton [8], Schwartz [25], Zippel [26, 27] and Alon and Füredi [2]. As an alternative, we also extend the Schwartz–Zippel bound to weighted multiplicities and discuss its connection to our extension of the footprint bound.

2010 *Mathematics subject classification*: Primary 11T06
Secondary 12D10, 13P10

1. Introduction

Estimating the number of zeros of a polynomial over a field $\mathbb{F}$ has been a central problem in algebra, where one of the main inconveniences is counting *repeated zeros*, that is, *multiplicities*. In the univariate case, this is easily solved by defining the multiplicity of a zero as the minimum positive integer r such that the first r *consecutive derivatives* of the given polynomial vanish at that zero. In addition, Hasse derivatives [14] are used instead of classical derivatives in order to give meaningful information over fields of positive characteristic. In this way, the number of zeros of a polynomial, counted with multiplicities, is upper-bounded by its degree. Formally:

$$\sum_{a\in\mathbb{F}} m(F(x),a) \leqslant \deg(F(x)). \tag{1.1}$$

If $\mathcal{V}_{\geqslant r}(F(x))$ denotes the set of zeros of $F(x)$ of multiplicity at least r, then a weaker, but still sharp, bound is the following:

$$\#\mathcal{V}_{\geqslant r}(F(x)) \cdot r \leqslant \deg(F(x)). \tag{1.2}$$

In the multivariate case, the standard approach is to consider the first r consecutive Hasse derivatives as those whose multi-indices have order less than r, where the order of a multi-index $(i_1, i_2, \ldots, i_m)$ is defined as $\sum_{j=1}^{m} i_j$. We will use the terms *standard multiplicities* to refer to this type of multiplicity. In this work, we consider arbitrary finite families $\mathcal{J}$ of multi-indices that are consecutive in a coordinate-wise sense: if $(i_1, i_2, \ldots, i_m)$ belongs to $\mathcal{J}$ and $k_j \leqslant i_j$, for $j = 1, 2, \ldots, m$, then $(k_1, k_2, \ldots, k_m)$ also belongs to $\mathcal{J}$. Obviously, the (finite) family $\mathcal{J}$ of multi-indices of order less than a given positive integer r satisfies this property, hence is a particular case.

Our main contribution is an upper bound on the number of common zeros over a grid of a family of polynomials and their (Hasse) derivatives corresponding to a finite set $\mathcal{J}$ of consecutive multi-indices. This upper bound makes use of the technique from Gröbner basis theory known as *footprint* [11, 16], and can be seen as an extension of the classical *footprint bound* [7, Section 5.3] in the sense of (1.2). A first extension for standard multiplicities has been given as Lemma 2.4 in the expanded version of [24].

We will then show that this bound is sharp for ideals of polynomials, characterize those which satisfy equality, and give as applications extensions of known results in algebra and combinatorics: Alon's combinatorial Nullstellensatz [1, 3, 6, 21, 23], existence and uniqueness of Hermite interpolating polynomials [10, 19, 22], estimations of the parameters of evaluation codes with consecutive derivatives [12, 19, 20], and the bounds by DeMillo and Lipton [8], Zippel [26, 27] and Alon and Füredi [2], and a particular case of the bound given by Schwartz in [25, Lemma 1].

The bound in [25, Lemma 1] can also be derived by those given by DeMillo and Lipton [8], and Zippel [26, Theorem 1], [27, Proposition 3] (see Proposition 5.5 below), and is referred to as the *Schwartz–Zippel bound* in many works in the literature [9, 12, 19, 20]. Interestingly, an extension of such a bound for standard multiplicities in the sense of (1.1) has recently been given in [9, Lemma 8], but as Counterexample 7.4 in [4] shows, no straightforward extension of the footprint bound in the sense of (1.1) seems possible (recall that we will give a footprint bound in the sense of (1.2)). To conclude this work, we give an extension of the Schwartz–Zippel bound in the sense of (1.1) to derivatives with weighted order less than a given positive integer, which we will call *weighted multiplicities*. This bound is inspired by [9, Lemma 8], and we will discuss its connection to our extension of the footprint bound.

The results are organized as follows. We start with some preliminaries in Section 2. We then give the main bound in Section 3, together with some particular cases, an interpretation of the bound, and sharpness and equality conditions. In Section 4, we give a list of applications. Finally, in Section 5 we give an extension of the Schwartz–Zippel bound in the sense of (1.1) to weighted multiplicities, and discuss the connections to the bound in Section 3.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Notation

Throughout this paper, $\mathbb{F}$ denotes an arbitrary field. We denote by $\mathbb{F}[\mathbf{x}] = \mathbb{F}[x_1, x_2, \ldots, x_m]$ the ring of polynomials in the m variables $x_1, x_2, \ldots, x_m$ with coefficients in $\mathbb{F}$. A multi-index is a vector $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$, where $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, and as usual we use the notation $\mathbf{x}^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m}$. We also denote $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$.

In this work, $\preceq$ denotes the coordinate-wise partial ordering in $\mathbb{N}^m$, that is, $(i_1, i_2, \ldots, i_m) \preceq (j_1, j_2, \ldots, j_m)$ if $i_k \leqslant j_k$, for all $k = 1, 2, \ldots, m$. We will use $\preceq_m$ to denote a given monomial ordering in the set of monomials of $\mathbb{F}[\mathbf{x}]$ (see [7, Section 2.2]), and we denote by $\mathrm{LM}_{\preceq_m}(F(\mathbf{x}))$ the leading monomial of $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ with respect to $\preceq_m$, or just $\mathrm{LM}(F(\mathbf{x}))$ if there is no confusion about $\preceq_m$. Finally, the notation $\langle A \rangle$ means ideal generated by A in a ring, and $\langle A \rangle_{\mathbb{F}}$ means vector space over $\mathbb{F}$ generated by A.

2. Consecutive derivatives

In this work, we consider Hasse derivatives, introduced first in [14]. They coincide with usual derivatives except for multiplication with a non-zero constant factor when the corresponding multi-index contains no multiples of the characteristic of the field, and they have the advantage of not being identically zero otherwise.

Definition 2.1 (Hasse derivative [14]). Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial. Given another family of independent variables $\mathbf{z} = (z_1, z_2, \ldots, z_m)$, the polynomial $F(\mathbf{x} + \mathbf{z})$ can be written uniquely as

$$F(\mathbf{x} + \mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^m} F^{(\mathbf{i})}(\mathbf{x}) \mathbf{z}^{\mathbf{i}},$$

for some polynomials $F^{(\mathbf{i})}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, for $\mathbf{i} \in \mathbb{N}^m$. For a given multi-index $\mathbf{i} \in \mathbb{N}^m$, we define the $\mathbf{i}$th Hasse derivative of $F(\mathbf{x})$ as the polynomial $F^{(\mathbf{i})}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$.

We next formalize the concept of zero of a polynomial of at least a given multiplicity as that of common zero of the given polynomial and a given finite family of its derivatives.

Definition 2.2. Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial, let $\mathbf{a} \in \mathbb{F}^m$ be an affine point, and let $\mathcal{J} \subseteq \mathbb{N}^m$ be a finite set. We say that $\mathbf{a}$ is a zero of $F(\mathbf{x})$ of multiplicity at least $\mathcal{J}$ if $F^{(\mathbf{i})}(\mathbf{a}) = 0$, for all $\mathbf{i} \in \mathcal{J}$.

The concept of *consecutive derivatives*, in a coordinate-wise sense, can be formalized by the concept of *decreasing sets* of multi-indices (recall that $\preceq$ denotes the coordinate-wise ordering in $\mathbb{N}^m$).

Definition 2.3 (decreasing sets). We say that the set $\mathcal{J} \subseteq \mathbb{N}^m$ is decreasing if, whenever $\mathbf{i} \in \mathcal{J}$ and $\mathbf{j} \in \mathbb{N}^m$ are such that $\mathbf{j} \preceq \mathbf{i}$, it holds that $\mathbf{j} \in \mathcal{J}$.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Decreasing subsets of a partially ordered set are also commonly known as *down sets* in the literature. Observe that the finite set

$$\mathcal{J} = \left\{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : \sum_{j=1}^{m} i_j < r\right\},$$

for a positive integer r, is decreasing. Moreover, if $m = 1$, then these are all possible decreasing finite sets. The concept of weighted orders and weighted multiplicities shows that this is not the case when $m > 1$.

Definition 2.4 (weighted multiplicities). Fix a vector of positive weights

$$\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m.$$

Given a multi-index $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$, we define its weighted order as

$$|\mathbf{i}|_{\mathbf{w}} = i_1 w_1 + i_2 w_2 + \cdots + i_m w_m. \tag{2.1}$$

Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial and let $\mathbf{a} \in \mathbb{F}^m$ be an affine point. We say that $\mathbf{a}$ is a zero of $F(\mathbf{x})$ of weighted multiplicity $r \in \mathbb{N}$, and we write

$$m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) = r,$$

if $F^{(\mathbf{i})}(\mathbf{a}) = 0$, for all $\mathbf{i} \in \mathbb{N}^m$ with $|\mathbf{i}|_{\mathbf{w}} < r$, and $F^{(\mathbf{j})}(\mathbf{a}) \neq 0$, for some $\mathbf{j} \in \mathbb{N}^m$ with $|\mathbf{j}|_{\mathbf{w}} = r$.

We also introduce the definition of weighted degree, which will be convenient for different results in the following sections.

Definition 2.5 (weighted degrees). Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial and let $\mathbf{w} \in \mathbb{N}_+^m$ be a vector of positive weights. We define the weighted degree of $F(\mathbf{x})$ as

$$\deg_{\mathbf{w}}(F(\mathbf{x})) = \max\{|\mathbf{i}|_{\mathbf{w}} : F_{\mathbf{i}} \neq 0\},$$

where $F(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{N}^m} F_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$ and $F_{\mathbf{i}} \in \mathbb{F}$, for all $\mathbf{i} \in \mathbb{N}^m$.

Other interesting sets of consecutive derivatives that we will consider throughout the paper are those given by bounding each index separately, that is, sets of the form

$$\mathcal{J} = \{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : i_j < r_j, j = 1, 2, \ldots, m\},$$

for a given $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$, where $\preceq$ denotes the coordinate-wise partial ordering.

3. The footprint bound for consecutive derivatives

In this section, we will give an extension of the footprint bound [7, Section 5.3] to upper-bound the number of common zeros over a finite grid of a family of polynomials and a given set of their consecutive derivatives, as in Definition 2.2. We give some particular cases and an interpretation of the bound. We conclude by studying its sharpness.

Throughout the section, fix a decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$, an ideal $I \subseteq \mathbb{F}[\mathbf{x}]$ and finite subsets $S_1, S_2, \ldots, S_m \subseteq \mathbb{F}$. Write $S = S_1 \times S_2 \times \cdots \times S_m$, and denote by $G_j(x_j) \in \mathbb{F}[x_j]$ the

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

defining polynomial of S_j, that is, $G_j(x_j) = \prod_{s \in S_j}(x_j - s)$, for $j = 1, 2, \ldots, m$. The three objects involved in our bound are the following.

Definition 3.1. We define the ideal

$$I_{\mathcal{J}} = I + \left\langle \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : (r_1, r_2, \ldots, r_m) \notin \mathcal{J} \right\} \right\rangle$$

and the set of zeros of multiplicity at least $\mathcal{J}$ of the ideal I in the grid $S = S_1 \times S_2 \times \cdots \times S_m$ as

$$\mathcal{V}_{\mathcal{J}}(I) = \{\mathbf{a} \in S : F^{(\mathbf{i})}(\mathbf{a}) = 0, \forall F(\mathbf{x}) \in I, \forall \mathbf{i} \in \mathcal{J}\}.$$

Finally, given a monomial ordering $\preceq_m$, we define the footprint of an ideal $J \subseteq \mathbb{F}[\mathbf{x}]$ as

$$\Delta_{\preceq_m}(J) = \{\mathbf{x}^{\mathbf{i}} : \mathbf{x}^{\mathbf{i}} \notin \langle \mathrm{LM}(J) \rangle\},$$

where $\mathrm{LM}(J) = \{\mathrm{LM}(F(\mathbf{x})) : F(\mathbf{x}) \in J\}$ with respect to the monomial ordering $\preceq_m$. We write $\Delta(J)$ if there is no confusion about the monomial ordering.

3.1. The general bound

Theorem 3.2. *For any monomial ordering, it holds that*

$$\#\mathcal{V}_{\mathcal{J}}(I) \cdot \#\mathcal{J} \leqslant \#\Delta(I_{\mathcal{J}}). \tag{3.1}$$

The rest of the subsection is devoted to the proof of this result. The first auxiliary tool is the Leibniz formula, which follows by a straightforward computation (see also [15, pp. 144–155]).

Lemma 3.3 (Leibniz formula). *Let $F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_s(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and let $\mathbf{i} \in \mathbb{N}^m$. It holds that*

$$\left(\prod_{j=1}^{s} F_j(\mathbf{x}) \right)^{(\mathbf{i})} = \sum_{\mathbf{i}_1 + \mathbf{i}_2 + \cdots + \mathbf{i}_s = \mathbf{i}} \left(\prod_{j=1}^{s} F_j^{(\mathbf{i}_j)}(\mathbf{x}) \right).$$

The second auxiliary tool is the existence of Hermite interpolating polynomials with Hasse derivatives. For our purposes, a *separated-variables* extension of univariate Hermite interpolation over grids is enough. This extension is straightforward and seems to be known in the literature (see [22, Section 3.1]), but we give a short proof in the Appendix for the convenience of the reader.

Definition 3.4. We define the evaluation map on a finite set $T \subseteq \mathbb{F}^m$ with derivatives corresponding to multi-indices in $\mathcal{J}$ as

$$\begin{aligned} \mathrm{Ev} : \mathbb{F}[\mathbf{x}] &\longrightarrow \mathbb{F}^{\#T \cdot \#\mathcal{J}} \\ F(\mathbf{x}) &\mapsto ((F^{(\mathbf{i})}(\mathbf{a}))_{\mathbf{i} \in \mathcal{J}})_{\mathbf{a} \in T}. \end{aligned} \tag{3.2}$$

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Lemma 3.5 (Hermite interpolation). *The evaluation map* $\mathrm{Ev} : \mathbb{F}[\mathbf{x}] \longrightarrow \mathbb{F}^{\#T \cdot \#\mathcal{J}}$ *defined in* (3.2) *is surjective, for all finite sets* $T \subseteq \mathbb{F}^m$ *and* $\mathcal{J} \subseteq \mathbb{N}^m$.

Proof. See the Appendix. □

With these tools, we may now prove Theorem 3.2.

Proof of Theorem 3.2. Fix multi-indices $\mathbf{r} = (r_1, r_2, \ldots, r_m) \notin \mathcal{J}$ and $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathcal{J}$, and define $G(\mathbf{x}) = \prod_{j=1}^m G_j(x_j)^{r_j}$. By Lemma 3.3, it holds that

$$G^{(\mathbf{i})}(\mathbf{x}) = \prod_{j=1}^m (G_j(x_j)^{r_j})^{(i_j)}. \tag{3.3}$$

Furthermore, if $r > i$ and $F(x) \in \mathbb{F}[x]$, then there exists $H(x) \in \mathbb{F}[x]$ such that

$$(F(x)^r)^{(i)} = \sum_{i_1+i_2+\cdots+i_r=i} \left(\prod_{j=1}^r F^{(i_j)}(x) \right) = H(x) F(x)^{r-i}, \tag{3.4}$$

again by Lemma 3.3, since at least $r - i > 0$ indices i_j must be equal to 0, for each $(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$ such that $\sum_{j=1}^m i_j = i$. Finally, since $\mathcal{J}$ is decreasing, it holds that $\mathbf{r} - \mathbf{i}$ has at least one positive coordinate. Hence, combining (3.3) and (3.4), we see that $G^{(\mathbf{i})}(\mathbf{a}) = 0$, for all $\mathbf{a} \in \mathcal{V}_{\mathcal{J}}(I) \subseteq S$. This implies that

$$\mathrm{Ev}(F(\mathbf{x})) = \mathbf{0}, \quad \text{for all } F(\mathbf{x}) \in I_{\mathcal{J}},$$

by the definition of the ideal $I_{\mathcal{J}}$ and the set $\mathcal{V}_{\mathcal{J}}(I)$, and where we consider $T = \mathcal{V}_{\mathcal{J}}(I)$ in the definition of Ev (Definition 3.4).

Therefore, the evaluation map Ev can be extended to the quotient ring

$$\mathrm{Ev} : \mathbb{F}[\mathbf{x}]/I_{\mathcal{J}} \longrightarrow \mathbb{F}^{\#\mathcal{V}_{\mathcal{J}}(I) \cdot \#\mathcal{J}},$$

which is again surjective, since the original evaluation map is surjective by Lemma 3.5. Since the domain and codomain of this map are $\mathbb{F}$-linear vector spaces and the map itself is also $\mathbb{F}$-linear, we conclude that

$$\#\mathcal{V}_{\mathcal{J}}(I) \cdot \#\mathcal{J} = \dim_{\mathbb{F}}(\mathbb{F}^{\#\mathcal{V}_{\mathcal{J}}(I) \cdot \#\mathcal{J}}) \leqslant \dim_{\mathbb{F}}(\mathbb{F}[\mathbf{x}]/I_{\mathcal{J}}).$$

Finally, Proposition 4 in [7, Section 5.3] says that the monomials in $\Delta(J)$ constitute a basis of $\mathbb{F}[\mathbf{x}]/J$, for an ideal $J \subseteq \mathbb{F}[\mathbf{x}]$. This fact implies that

$$\dim_{\mathbb{F}}(\mathbb{F}[\mathbf{x}]/I_{\mathcal{J}} r) = \#\Delta(I_{\mathcal{J}}),$$

and the result follows. □

3.2. Some particular cases

In this subsection, we derive some particular cases of Theorem 3.2. We start with the classical form of the footprint bound (see Proposition 8 in [7, Section 5.3], and [11, 16]).

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Corollary 3.6 ([7, 11, 16]). *Setting* $\mathcal{J} = \{\mathbf{0}\}$, *we obtain that*

$$\#\mathcal{V}(I) \leqslant \#\Delta(I + \langle G_1(x_1), G_2(x_2), \ldots, G_m(x_m)\rangle),$$

where $\mathcal{V}(I)$ *denotes the set of zeros of the ideal* I *in* S.

The case of zeros of standard multiplicity at least a given positive integer was first obtained as Lemma 2.4 in the extended version of [24], and reads as follows.

Corollary 3.7 ([24]). *Given an integer* $r \in \mathbb{N}_+$, *and setting*

$$\mathcal{J} = \left\{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : \sum_{j=1}^{m} i_j < r\right\},$$

we obtain that

$$\#\mathcal{V}_{\geqslant r}(I) \cdot \binom{m+r-1}{m} \leqslant \#\Delta\left(I + \left\langle \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : \sum_{j=1}^{m} r_j = r \right\} \right\rangle\right),$$

where $\mathcal{V}_{\geqslant r}(I)$ *denotes the set of zeros of multiplicity at least* r *of the ideal* I *in* S.

Another particular case is obtained when upper-bounding each coordinate of the multi-indices separately.

Corollary 3.8. *Given a multi-index* $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$, *and setting*

$$\mathcal{J} = \{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : i_j < r_j, j = 1, 2, \ldots, m\},$$

we obtain that

$$\#\mathcal{V}_{\mathcal{J}}(I) \cdot \prod_{j=1}^{m} r_j \leqslant \#\Delta(I + \langle G_1(x_1)^{r_1}, G_2(x_2)^{r_2}, \ldots, G_m(x_m)^{r_m}\rangle).$$

Finally, we obtain a footprint bound for weighted multiplicities.

Corollary 3.9. *Given an integer* $r \in \mathbb{N}_+$, *a vector of positive weights*

$$\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+,$$

and setting

$$\mathcal{J} = \{\mathbf{i} \in \mathbb{N}^m : |\mathbf{i}|_{\mathbf{w}} < r\},$$

we obtain that

$$\#\mathcal{V}_{\geqslant r,\mathbf{w}}(I) \cdot \mathrm{B}(\mathbf{w}; r) \leqslant \#\Delta\left(I + \left\langle \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : \sum_{j=1}^{m} r_j w_j \geqslant r \right\} \right\rangle\right),$$

where $\mathcal{V}_{\geqslant r,\mathbf{w}}(I)$ *denotes the set of zeros of weighted multiplicity at least* r *of the ideal* I *in* S, *and where* $\mathrm{B}(\mathbf{w}; r) = \#\{\mathbf{i} \in \mathbb{N}^m : |\mathbf{i}|_{\mathbf{w}} < r\}$.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

To conclude, we give a more explicit form of the bound in the previous corollary by estimating the number $B(\mathbf{w};r)$.

Corollary 3.10. *Given an integer $r \in \mathbb{N}_+$ and a vector of positive weights*

$$\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+,$$

it holds that

$$\binom{m+r-1}{m} \leqslant w_1 w_2 \cdots w_m B(\mathbf{w};r). \tag{3.5}$$

In particular, we deduce from the previous corollary that

$$\#\mathcal{V}_{\geqslant r,\mathbf{w}}(I) \cdot \binom{m+r-1}{m} \leqslant w_1 w_2 \cdots w_m \cdot \#\Delta\left(I + \left\langle \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : \sum_{j=1}^{m} r_j w_j \geqslant r \right\} \right\rangle\right).$$

Proof. Define the map $T_{\mathbf{j}} : \mathbb{N}^m \longrightarrow \mathbb{N}^m$ by

$$T_{\mathbf{j}}(\mathbf{i}) = (i_1 w_1 + j_1, i_2 w_2 + j_2, \ldots, i_m w_m + j_m),$$

for all $\mathbf{i} = (i_1, i_2, \ldots, i_m), \mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$. Now define

$$\mathcal{J}(\mathbf{w};r) = \{\mathbf{i} \in \mathbb{N}^m : |\mathbf{i}|_{\mathbf{w}} < r\}.$$

By the Euclidean division, we see that

$$\mathcal{J}((1,1,\ldots,1);r) \subseteq \bigcup_{\mathbf{j} \in \prod_{k=1}^{m} [0, w_k)} T_{\mathbf{j}}(\mathcal{J}(\mathbf{w};r)).$$

By counting elements on both sides of the inclusion, the result follows. □

3.3. Interpretation of the bound and illustration of the set $\Delta(I_{\mathcal{J}})$

In this subsection, we give a graphical description of the footprint $\Delta(I_{\mathcal{J}})$ which will allow us to provide an interpretation of the bound (3.1).

Observe that the ideal $I_{\mathcal{J}}$ always contains the polynomials $\prod_{i=1}^{m} G_i(x_i)^{r_i}$, for $(r_1, r_2, \ldots, r_m) \notin \mathcal{J}$. Hence the set $\Delta(I_{\mathcal{J}})$ is contained in a larger set $\mathcal{J}_S \subseteq \mathbb{N}^m$, which is independent of I and consists in removing the monomials that are multiples of the leading monomials of $\prod_{i=1}^{m} G_i(x_i)^{r_i}$, for $(r_1, r_2, \ldots, r_m) \notin \mathcal{J}$.

Definition 3.11. We define the set

$$\mathcal{J}_S = \{\mathbf{i} \in \mathbb{N}^m : \mathbf{i} \not\geq (r_1 \# S_1, r_2 \# S_2, \ldots, r_m \# S_m), \forall (r_1, r_2, \ldots, r_m) \notin \mathcal{J}\}.$$

For clarity, we now give a description of this set by a positive defining condition.

Lemma 3.12. *It holds that*

$$\mathcal{J}_S = \{(p_1 \# S_1 + t_1, p_2 \# S_2 + t_2, \ldots, p_m \# S_m + t_m) \in \mathbb{N}^m : (p_1, p_2, \ldots, p_m) \in \mathcal{J}, 0 \leqslant t_j < \#S_j, \forall j = 1, 2, \ldots, m\}.$$

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Proof. Take $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathcal{J}_S$. Performing the Euclidean division of i_j by $\#S_j$, there exist $p_j, t_j \in \mathbb{N}$ such that $0 \leqslant t_j < \#S_j$ and $i_j = p_j \#S_j + t_j$, for $j = 1, 2, \ldots, m$. Since $\mathbf{i} \in \mathcal{J}_S$ and $i_j \geqslant p_j \#S_j$, for $j = 1, 2, \ldots, m$, then it must hold that $(p_1, p_2, \ldots, p_m) \in \mathcal{J}$ by definition of $\mathcal{J}_S$.

Conversely, take $\mathbf{i} = (p_1 \#S_1 + t_1, p_2 \#S_2 + t_2, \ldots, p_m \#S_m + t_m)$, where $\mathbf{p} = (p_1, p_2, \ldots, p_m) \in \mathcal{J}$ and $0 \leqslant t_j < \#S_j$, for $j = 1, 2, \ldots, m$. Assume that $\mathbf{i} \geq (r_1 \#S_1, r_2 \#S_2, \ldots, r_m \#S_m)$, for some $\mathbf{r} = (r_1, r_2, \ldots, r_m) \notin \mathcal{J}$. Since $t_j < \#S_j$, it must hold that $p_j \geqslant r_j$, for $j = 1, 2, \ldots, m$. Since $\mathbf{p} \in \mathcal{J}$ and $\mathcal{J}$ is decreasing, we deduce that $\mathbf{r} \in \mathcal{J}$, a contradiction. Thus $\mathbf{i} \in \mathcal{J}_S$. □

We may then state the fact that the footprint is bounded by this set as follows.

Lemma 3.13. *It holds that*

$$\Delta(I_{\mathcal{J}}) \subseteq \{\mathbf{x}^{\mathbf{i}} : \mathbf{i} \in \mathcal{J}_S\}.$$

Moreover, the set $\mathcal{J}_S$ can be easily seen as the union of $\#\mathcal{J} m$-dimensional grids in $\mathbb{N}^m$ whose sides have lengths $\#_1, \#S_2, \ldots, \#S_m$, respectively. In particular, we obtain the following.

Lemma 3.14. *It holds that*

$$\#\mathcal{J}_S = \#S \cdot \#\mathcal{J}. \tag{3.6}$$

The footprint bound (3.1) can then be interpreted as follows. Consider the set $\mathcal{J}_S \subseteq \mathbb{N}^m$. For each $\mathbf{x}^{\mathbf{i}} \in \mathrm{LM}(I_{\mathcal{J}})$, remove from $\mathcal{J}_S$ all points $\mathbf{j}$ such that $\mathbf{i} \preceq \mathbf{j}$. The remaining points correspond to the multi-indices in $\Delta(I_{\mathcal{J}})$, and thus there are $\#\Delta(I_{\mathcal{J}})$ of them.

In particular, if $F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_t(\mathbf{x}) \in I$, then we may only remove the points corresponding to $\mathrm{LM}(F_i(\mathbf{x}))$, for $i = 1, 2, \ldots, t$, and we obtain an upper bound on $\#\Delta(I_{\mathcal{J}})$.

Example 3.15. Let us assume now that $m = 2$, $\#S_1 = \#S_2 = 2$, and

$$\mathcal{J} = \{(0,1), (1,1), (2,1), (0,0), (1,0), (2,0), (3,0), (4,0), (5,0)\}.$$

In Figure 1(a) we represent by black dots the monomials whose multi-indices belong to $\mathcal{J}_S$, among which medium-sized dots correspond to multi-indices that belong to $\mathcal{J}$ when each coordinate is multiplied by 2. Blank dots correspond to multi-indices that do not belong to $\mathcal{J}_S$, and the largest ones correspond to minimal multi-indices that do not belong to $\mathcal{J}_S$.

In Figure 1(b) we represent in the same way the set $\Delta(I_{\mathcal{J}})$, whenever $\langle \mathrm{LM}(I_{\mathcal{J}}) \rangle$ is generated by $x_1^2 x_2^3$, $x_1^8 x_2$, and the leading monomials of $G_1(x_1)^{r_1} G_2(x_2)^{r_2}$, for minimal $(r_1, r_2) \notin \mathcal{J}$, which in this case are x_2^4, $x_1^6 x_2^2$ and x_1^{12}.

In conclusion, the bound (3.1) says that the number of zeros in S of I of multiplicity at least $\mathcal{J}$ is at most 3.

As a consequence of this interpretation, we may deduce the following useful fact.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

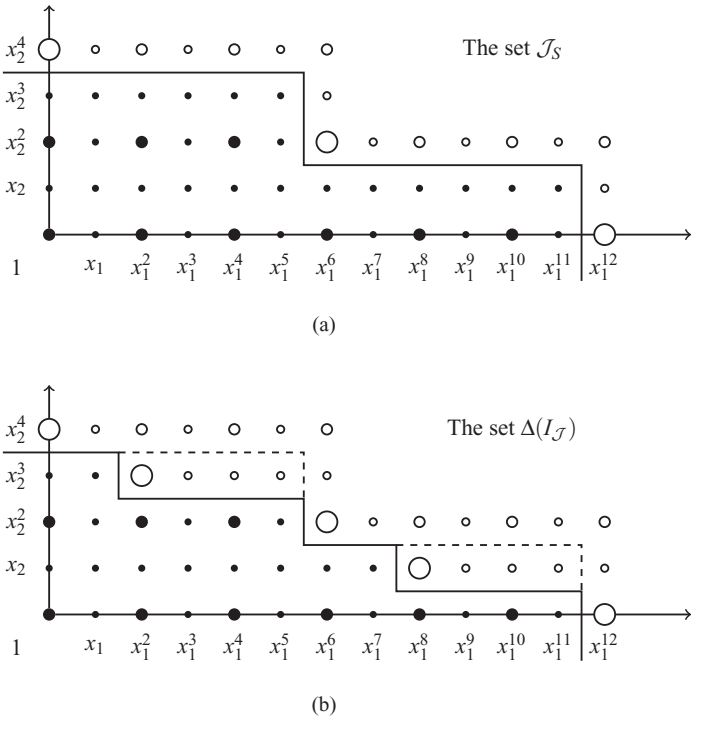

Figure 1. Illustration of the sets $\mathcal{J}_S$ (a) and $\Delta(I_{\mathcal{J}})$ (b) in $\mathbb{N}^m$.

Lemma 3.16. *Assume that the finite set $\mathcal{J} \subseteq \mathbb{N}^m$ is decreasing and $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$ with respect to some monomial ordering, for some polynomial $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$. If $\mathbf{i} \in \mathcal{J}_S$, then it holds that*

$$\#\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}) < \#S \cdot \#\mathcal{J}. \tag{3.7}$$

We conclude with a simple description of $\mathcal{J}_S$ in the cases of multi-indices bounded by weighted orders and multi-indices bounded on each coordinate separately, which follow by straightforward calculations.

Remark 3.17. Given a vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, a positive integer $r \in \mathbb{N}_+$, and

$$\mathcal{J} = \{\mathbf{r} \in \mathbb{N}^m : |\mathbf{r}|_{\mathbf{w}} < r\},$$

it holds that

$$\mathcal{J}_S = \left\{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : \sum_{j=1}^{m} \left\lfloor \frac{i_j}{\#S_j} \right\rfloor w_j < r \right\}.$$

On the other hand, given $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$ and

$$\mathcal{J} = \{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : i_j < r_j, j = 1, 2, \ldots, m\},$$

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

it holds that

$$\mathcal{J}_S = \{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : i_j < r_j \# S_j, j = 1, 2, \ldots, m\}.$$

3.4. Sharpness and equality conditions

To conclude the section, we study the sharpness of the bound (3.1). We will give sufficient and necessary conditions on the ideal I for (3.1) to be an equality, and we will see that (3.1) is the sharpest bound that can be obtained as a strictly increasing function of the size of the footprint $\Delta(I_{\mathcal{J}})$.

We start by defining the ideal associated with a set of points and a set of multi-indices.

Definition 3.18. Given $\mathcal{V} \subseteq \mathbb{F}^m$, we define

$$I(\mathcal{V}; \mathcal{J}) = \{F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}] : F^{(\mathbf{i})}(\mathbf{a}) = 0, \forall \mathbf{a} \in \mathcal{V}, \forall \mathbf{i} \in \mathcal{J}\}.$$

In the next proposition we show that this set is indeed an ideal and gather other properties similar to those of ideals and algebraic sets in algebraic geometry.

Proposition 3.19. *Given a set of points $\mathcal{V} \subseteq \mathbb{F}^m$, the set $I(\mathcal{V}; \mathcal{J})$ in the previous definition is an ideal in $\mathbb{F}[\mathbf{x}]$. Moreover, the following properties hold.*

(1) $I \subseteq I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J})$.
(2) $\mathcal{V} \subseteq \mathcal{V}_{\mathcal{J}}(I(\mathcal{V}; \mathcal{J}))$.
(3) $I = I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J})$ *if and only if* $I = I(\mathcal{W}; \mathcal{J})$, *for some set* $\mathcal{W} \subseteq \mathbb{F}^m$.
(4) $\mathcal{V} = \mathcal{V}_{\mathcal{J}}(I(\mathcal{V}; \mathcal{J}))$ *if and only if* $\mathcal{V} = \mathcal{V}_{\mathcal{J}}(K)$, *for some ideal* $K \subseteq \mathbb{F}[\mathbf{x}]$.

Proof. The fact that $I(\mathcal{V}; \mathcal{J})$ is an ideal follows from the Leibniz formula (Lemma 3.3) and the fact that $\mathcal{J}$ is decreasing. The properties in items (1), (2), (3) and (4) follow as in classical algebraic geometry and are left to the reader. □

The following is the main result of the subsection.

Theorem 3.20. *Fixing a monomial ordering, the bound* (3.1) *is an equality if and only if*

$$I_{\mathcal{J}} = I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}). \tag{3.8}$$

In particular, for any choice of decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$ and a finite set of points $\mathcal{V} \subseteq \mathbb{F}^m$, there exists an ideal, $I = I(\mathcal{V}; \mathcal{J})$, satisfying equality in (3.1).

Proof. With notation as in the proof of Theorem 3.2, the evaluation map $\mathrm{Ev} : \mathbb{F}[\mathbf{x}] \longrightarrow \mathbb{F}^{\#\mathcal{V}_{\mathcal{J}}(I) \cdot \#\mathcal{J}}$ from Definition 3.4 is $\mathbb{F}$-linear and surjective by Lemma 3.5. By definition, its kernel is

$$\mathrm{Ker}(\mathrm{Ev}) = I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}).$$

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

On the other hand, we saw in the proof of Theorem 3.2 that $I_{\mathcal{J}} \subseteq \mathrm{Ker}(\mathrm{Ev})$. This means that the evaluation map

$$\mathrm{Ev} : \mathbb{F}[\mathbf{x}]/I_{\mathcal{J}} \longrightarrow \mathbb{F}^{\#\mathcal{V}_{\mathcal{J}}(I)\cdot\#\mathcal{J}}$$

is an isomorphism if and only if $I_{\mathcal{J}} = I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J})$.

Finally, the fact that this evaluation map is an isomorphism is equivalent to (3.1) being an equality, by the proof of Theorem 3.2. Together with Proposition 3.19 and the fact that $I = I_{\mathcal{J}}$ if $I = I(\mathcal{V}; \mathcal{J})$ by the proof of Theorem 3.2, the theorem follows. □

Thanks to this result, we may establish that the bound (3.1) is the sharpest bound that is a strictly increasing function of the size of the footprint $\Delta(I_{\mathcal{J}})$, in the following sense: if equality holds for such a bound, then it holds in (3.1).

Corollary 3.21. *Let $f : \mathbb{N} \longrightarrow \mathbb{R}$ be a strictly increasing function, and assume that*

$$\#\mathcal{V}_{\mathcal{J}}(I) \leqslant f(\#\Delta(I_{\mathcal{J}})), \tag{3.9}$$

for all ideals $I \subseteq \mathbb{F}[\mathbf{x}]$. If equality holds in (3.9) *for a given ideal $I \subseteq \mathbb{F}[\mathbf{x}]$, then equality holds in* (3.1) *for such an ideal.*

Proof. First we have that $I_{\mathcal{J}} \subseteq I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J})$ as we saw in the proof of the previous theorem. Hence the reverse inclusion holds for their footprints and thus

$$f(\#\Delta(I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}))) \leqslant f(\#\Delta(I_{\mathcal{J}})). \tag{3.10}$$

Now, since $\mathcal{V}_{\mathcal{J}}(I) = \mathcal{V}_{\mathcal{J}}(I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}))$ by Proposition 3.19, and equality holds in (3.9) for I, we have that

$$f(\#\Delta(I_{\mathcal{J}})) = \#\mathcal{V}_{\mathcal{J}}(I) = \#\mathcal{V}_{\mathcal{J}}(I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J})) \leqslant f(\#\Delta(I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}))). \tag{3.11}$$

Combining (3.10) and (3.11), and using that f is strictly increasing, we conclude that

$$\#\Delta(I(\mathcal{V}_{\mathcal{J}}(I); \mathcal{J}))) = \#\Delta(I_{\mathcal{J}}),$$

which implies that equality holds in (3.1) for I by Theorem 3.20, and we are done. □

4. Applications of the footprint bound for consecutive derivatives

In this section, we present a brief collection of applications of Theorem 3.2, which are extensions to consecutive derivatives of well-known important results from the literature. Throughout the section, we will again fix finite sets $S_1, S_2, \ldots, S_m \subseteq \mathbb{F}$ and $S = S_1 \times S_2 \times \cdots \times S_m$.

4.1. Alon's combinatorial Nullstellensatz

The combinatorial Nullstellensatz is a non-vanishing theorem by Alon [1, Theorem 1.2] with many applications in combinatorics. It has been extended to non-vanishing theorems for standard multiplicities in [3, Corollary 3.2] and for multisets (sets with multiplicities) in [21, Theorem 6].

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

In this subsection, we establish and prove a combinatorial Nullstellensatz for consecutive derivatives and derive the well-known particular cases as corollaries. The formulation in [1, Theorem 1.1] is equivalent in essence. We will extend that result in the next subsection in terms of Gröbner bases.

Theorem 4.1. *Let $\mathcal{J} \subseteq \mathbb{N}^m$ be a decreasing finite set, let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a non-zero polynomial, and let $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$ for some monomial ordering. If $\mathbf{i} \in \mathcal{J}_S$, then there exist $\mathbf{s} \in S$ and $\mathbf{j} \in \mathcal{J}$ such that*

$$F^{(\mathbf{j})}(\mathbf{s}) \neq 0.$$

Proof. By Lemma 3.16, the assumptions imply that

$$\#\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}) < \#S \cdot \#\mathcal{J}.$$

On the other hand, Theorem 3.2 implies that

$$\#\mathcal{V}_{\mathcal{J}}(F(\mathbf{x})) \cdot \#\mathcal{J} \leqslant \#\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}).$$

Therefore not all points in S are zeros of $F(\mathbf{x})$ of multiplicity at least $\mathcal{J}$, and the result follows. □

We now derive the original theorem [1, Theorem 1.2]. This constitutes an alternative proof. See also [23] for another recent short proof.

Corollary 4.2 ([1]). *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$. Assume that the coefficient of $\mathbf{x}^{\mathbf{i}}$ in $F(\mathbf{x})$ is not zero and $\deg(F(\mathbf{x})) = |\mathbf{i}|$. If $\#S_j > i_j$ for all $j = 1, 2, \ldots, m$, then there exist $s_1 \in S_1, s_2 \in S_2, \ldots, s_m \in S_m$, such that*

$$F(s_1, s_2, \ldots, s_m) \neq 0.$$

Proof. There exists a graded monomial ordering such that $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$ since $\deg(F(\mathbf{x})) = |\mathbf{i}|$. Now, the assumption implies that

$$\mathbf{i} \not\geq (r_1 \# S_1, r_2 \# S_2, \ldots, r_m \# S_m),$$

for all $\mathbf{r} = (r_1, r_2, \ldots, r_m)$ such that $r_j = 1$ for some j, and the rest are zero. These are in fact all minimal multi-indices not in $\mathcal{J} = \{\mathbf{0}\}$. Thus $\mathbf{i} \in \mathcal{J}_S$, and the result follows from the previous theorem. □

The next consequence is a combinatorial Nullstellensatz for weighted multiplicities, where the particular case $w_1 = w_2 = \cdots = w_m = 1$ coincides with [3, Corollary 3.2] (recall the definition of weighted degree from Definition 2.5).

Corollary 4.3. *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, let $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$ and let $r \in \mathbb{N}_+$. Assume that the coefficient of $\mathbf{x}^{\mathbf{i}}$ in $F(\mathbf{x})$ is not zero and $\deg_{\mathbf{w}}(F(\mathbf{x})) = |\mathbf{i}|_{\mathbf{w}}$.*

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Assume also that, for all $\mathbf{r} = (r_1, r_2, \ldots, r_m)$ *with* $|\mathbf{r}|_{\mathbf{w}} \geqslant r$, *there exists a* j *such that* $r_j \# S_j > i_j$. *Then there exist* $s_1 \in S_1, s_2 \in S_2, \ldots, s_m \in S_m$, *and some* $\mathbf{j} \in \mathbb{N}^m$ *with* $|\mathbf{j}|_{\mathbf{w}} < r$, *such that*

$$F^{(\mathbf{j})}(s_1, s_2, \ldots, s_m) \neq 0.$$

Proof. It follows from Theorem 4.1 like the previous corollary. □

We conclude with a combinatorial Nullstellensatz for multi-indices bounded on each coordinate separately.

Corollary 4.4. *Let* $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, *let* $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$, *and assume that* $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$, $\mathbf{i} = (i_1, i_2, \ldots, i_m)$, *for some monomial ordering and* $i_j < r_j \# S_j$, *for all* $j = 1, 2, \ldots, m$. *There exist* $s_1 \in S_1, s_2 \in S_2, \ldots, s_m \in S_m$, *and some* $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$ *with* $j_k < r_k$, *for all* $k = 1, 2, \ldots, m$, *such that*

$$F^{(\mathbf{j})}(s_1, s_2, \ldots, s_m) \neq 0.$$

4.2. Gröbner bases of ideals of zeros in a grid

An equivalent but more refined consequence is obtaining a Gröbner basis for ideals $I(S; \mathcal{J})$ associated with the whole grid S and to a decreasing finite set of multi-indices (recall Definition 3.18). This result is also usually referred to as combinatorial Nullstellensatz in many works in the literature (see [1, Theorem 1.1], [3, Theorem 3.1] and [21, Theorem 1]). We briefly recall the notion of Gröbner basis. We will also make repeated use of the Euclidean division on the multivariate polynomial ring and its properties. See [7, Chapter 2] for more details.

Definition 4.5 (Gröbner bases). Given a monomial ordering $\preceq_m$ and an ideal $I \subseteq \mathbb{F}[\mathbf{x}]$, we say that a finite family of polynomials $\mathcal{F} \subseteq I$ is a Gröbner basis of I with respect to $\preceq_m$ if

$$\langle \mathrm{LM}_{\preceq_m}(I) \rangle = \langle \mathrm{LM}_{\preceq_m}(\mathcal{F}) \rangle.$$

Moreover, we say that $\mathcal{F}$ is reduced if, for any two distinct $F(\mathbf{x}), G(\mathbf{x}) \in \mathcal{F}$, $\mathrm{LM}_{\preceq_m}(F(\mathbf{x}))$ does not divide any monomial in $G(\mathbf{x})$.

Recall that a Gröbner basis of an ideal generates it as an ideal. To obtain reduced Gröbner bases, we need a way to minimally generate complements of decreasing finite sets in $\mathbb{N}^m$, which is given by the following object.

Definition 4.6. For any decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$, we define

$$\mathcal{B}_{\mathcal{J}} = \{\mathbf{i} \notin \mathcal{J} : \mathbf{j} \notin \mathcal{J} \text{ and } \mathbf{j} \preceq \mathbf{i} \Longrightarrow \mathbf{i} = \mathbf{j}\}.$$

The main result of this subsection is the following.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Theorem 4.7. *For any decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$, the family*

$$\mathcal{F} = \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : (r_1, r_2, \ldots, r_m) \in \mathcal{B}_{\mathcal{J}} \right\}$$

is a reduced Gröbner basis of the ideal $I(S;\mathcal{J})$ with respect to any monomial ordering. In particular, for any $F(\mathbf{x}) \in I(S;\mathcal{J})$, there exist polynomials $H_{\mathbf{r}}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, for $\mathbf{r} = (r_1, \ldots, r_m) \in \mathcal{B}_{\mathcal{J}}$, such that

$$\deg(H_{\mathbf{r}}(\mathbf{x})) + \sum_{j=1}^{m} r_j \deg(G_j(x_j)) \leqslant \deg(F(\mathbf{x}))$$

and

$$F(\mathbf{x}) = \sum_{\mathbf{r} \in \mathcal{B}_{\mathcal{J}}} \left(H_{\mathbf{r}}(\mathbf{x}) \prod_{j=1}^{m} G_j(x_j)^{r_j} \right).$$

Proof. It suffices to prove that if $F(\mathbf{x}) \in I(S;\mathcal{J})$ and we divide it by the family $\mathcal{F}$ (in an arbitrary order), then the remainder must be the zero polynomial.

Performing such a division, we obtain $F(\mathbf{x}) = G(\mathbf{x}) + R(\mathbf{x})$, where $R(\mathbf{x})$ is the remainder of the division and $G(\mathbf{x}) \in I(S;\mathcal{J})$. Assume that $R(\mathbf{x}) \neq 0$ and let $\mathbf{x}^{\mathbf{i}}$ be the leading monomial of $R(\mathbf{x})$ with respect to the chosen monomial ordering. Since no leading monomial of the polynomials in $\mathcal{F}$ divides $\mathbf{x}^{\mathbf{i}}$, we conclude that

$$\mathbf{i} \not\succeq (r_1 \# S_1, r_2 \# S_2, \ldots, r_m \# S_m),$$

for all minimal $\mathbf{r} = (r_1, r_2, \ldots, r_m) \notin \mathcal{J}$, that is, for all $\mathbf{r} \in \mathcal{B}_{\mathcal{J}}$. Thus by Theorem 4.1, we conclude that not all points in S are zeros of $R(\mathbf{x})$ of multiplicity at least $\mathcal{J}$, which is absurd since $R(\mathbf{x}) = F(\mathbf{x}) - G(\mathbf{x}) \in I(S;\mathcal{J})$, and we are done.

The fact that $\mathcal{F}$ is reduced follows from observing that the multi-indices $\mathbf{r} \in \mathcal{B}_{\mathcal{J}}$ are minimal among those not in $\mathcal{J}$. The last part of the theorem follows by performing the Euclidean division. □

The following particular case is [1, Theorem 1.1].

Corollary 4.8 ([1]). *If $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ vanishes at all points in S, then there exist polynomials $H_j(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that $\deg(H_j(\mathbf{x})) + \deg(G_j(x_j)) \leqslant \deg(F(\mathbf{x}))$, for $j = 1, 2, \ldots, m$, and*

$$F(\mathbf{x}) = \sum_{j=1}^{m} H_j(\mathbf{x}) G_j(x_j).$$

To study the case of weighted multiplicities, we observe the following.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Remark 4.9. Given a vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, a positive integer $r \in \mathbb{N}_+$, and the set $\mathcal{J} = \{\mathbf{i} \in \mathbb{N}^m : |\mathbf{i}|_{\mathbf{w}} < r\}$, it holds that $\mathcal{B}_{\mathcal{J}} = \mathcal{B}_{\mathbf{w}}$, where

$$\mathcal{B}_{\mathbf{w}} = \left\{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : r \leqslant \sum_{j=1}^{m} i_j w_j < r + \min\{w_j : i_j \neq 0\}\right\}.$$

We then obtain the next consequence, where the particular case $w_1 = w_2 = \cdots = w_m = 1$ coincides with [3, Theorem 3.1], which in turn extends the finite-field result [5, Theorem 1.3].

Corollary 4.10. *Given a vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$ and a positive integer $r \in \mathbb{N}_+$, if $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ vanishes at all points in S with weighted multiplicity at least r, then there exist polynomials $H_{\mathbf{r}}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that*

$$\deg(H_{\mathbf{r}}(\mathbf{x})) + \sum_{j=1}^{m} r_j \deg(G_j(x_j)) \leqslant \deg(F(\mathbf{x})),$$

for all $\mathbf{r} = (r_1, r_2, \ldots, r_m) \in \mathcal{B}_{\mathbf{w}}$, and

$$F(\mathbf{x}) = \sum_{\mathbf{r} \in \mathcal{B}_{\mathbf{w}}} \left(H_{\mathbf{r}}(\mathbf{x}) \prod_{j=1}^{m} G_j(x_j)^{r_j} \right).$$

We conclude with the case of multi-indices bounded on each coordinate separately.

Corollary 4.11. *Given a vector $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$, if $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is such that $F^{(\mathbf{j})}(\mathbf{s}) = 0$, for all $\mathbf{s} \in S$ and all $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$ satisfying $j_k < r_k$, for all $k = 1, 2, \ldots, m$, then there exist polynomials $H_j(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that*

$$\deg(H_j(\mathbf{x})) + r_j \deg(G_j(x_j)) \leqslant \deg(F(\mathbf{x})),$$

for all $j = 1, 2, \ldots, m$, and

$$F(\mathbf{x}) = \sum_{j=1}^{m} H_j(\mathbf{x}) G_j(x_j)^{r_j}.$$

Proof. It follows from Theorem 4.7 observing that if

$$\mathcal{J} = \{(j_1, j_2, \ldots, j_m) \in \mathbb{N}^m : j_k < r_k, k = 1, 2, \ldots, m\},$$

then

$$\mathcal{B}_{\mathcal{J}} = \{r_j \mathbf{e}_j \in \mathbb{N}^m : j = 1, 2, \ldots, m\},$$

where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m \in \mathbb{N}^m$ are the vectors in the canonical basis. □

4.3. Hermite interpolation over grids with consecutive derivatives

In the Appendix we show that the evaluation map (Definition 3.4) is surjective. This has been used to prove Theorem 3.2. In this subsection, we see that the combinatorial

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Nullstellensatz (Theorem 4.1) implies that the evaluation map over the whole grid S, with consecutive derivatives, is an isomorphism when taking an appropriate domain. More concretely, we show the existence and uniqueness of Hermite interpolating polynomials over S with derivatives in $\mathcal{J}$ when choosing monomials in $\mathcal{J}_S$. Finding appropriate sets of points, derivatives and polynomials to guarantee existence and uniqueness of Hermite interpolating polynomials has been extensively studied [10, 19, 22]. The next result is new to the best of our knowledge.

Theorem 4.12. *Given a decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$, the evaluation map in Definition 3.4 for the finite set $S = S_1 \times S_2 \times \cdots \times S_m$, defined as*

$$\mathrm{Ev} : \langle \mathcal{J}_S \rangle_{\mathbb{F}} \longrightarrow \mathbb{F}^{\#S \cdot \#\mathcal{J}},$$

is a vector space isomorphism. In other words, for all $b_{\mathbf{j},\mathbf{a}} \in \mathbb{F}$, where $\mathbf{j} \in \mathcal{J}$ and $\mathbf{a} \in S$, there exists a unique polynomial of the form

$$F(\mathbf{x}) = \sum_{\mathbf{i} \in \mathcal{J}_S} F_{\mathbf{i}} \mathbf{x}^{\mathbf{i}} \in \mathbb{F}[\mathbf{x}],$$

where $F_{\mathbf{i}} \in \mathbb{F}$ for all $\mathbf{i} \in \mathcal{J}_S$, such that $F^{(\mathbf{j})}(\mathbf{a}) = b_{\mathbf{j},\mathbf{a}}$, for all $\mathbf{j} \in \mathcal{J}$ and all $\mathbf{a} \in S$.

Proof. The map is one-to-one by Theorem 4.1, and both vector spaces have the same dimension over $\mathbb{F}$ by Lemma 3.14, hence the map is a vector space isomorphism. □

Remark 4.13. Observe that we may similarly prove that the following two maps are vector space isomorphisms:

$$\langle \mathcal{J}_S \rangle_{\mathbb{F}} \xrightarrow{\rho} \mathbb{F}[\mathbf{x}]/I(S;\mathcal{J}) \xrightarrow{\mathrm{Ev}} \mathbb{F}^{\#S \cdot \#\mathcal{J}},$$

where ρ is the projection to the quotient ring. We may then extend the notion of *reduction* of a polynomial as follows (see [6, Section 3.1] and [10, Section 6.3], for instance). Given $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, we define its reduction over the set S with derivatives in $\mathcal{J}$ as

$$G(\mathbf{x}) = \rho^{-1}(F(\mathbf{x}) + I(S;\mathcal{J})).$$

As an immediate consequence, we obtain the following result on Hermite interpolation with weighted multiplicities.

Corollary 4.14. *For every vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, every positive integer $r \in \mathbb{N}_+$, and elements $b_{\mathbf{j},\mathbf{a}} \in \mathbb{F}$, for $\mathbf{j} \in \mathbb{N}^m$ with $|\mathbf{j}|_{\mathbf{w}} < r$ and for $\mathbf{a} \in S$, there exists a unique polynomial of the form*

$$F(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{N}^m} F_{\mathbf{i}} \mathbf{x}^{\mathbf{i}},$$

where $F_{\mathbf{i}} \in \mathbb{F}$ for all $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$, and $F_{\mathbf{i}} = 0$ whenever

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

$$\sum_{j=1}^{m} \left\lfloor \frac{i_j}{\#S_j} \right\rfloor w_j \geqslant r,$$

such that $F^{(\mathbf{j})}(\mathbf{a}) = b_{\mathbf{j},\mathbf{a}}$, *for all* $\mathbf{j} \in \mathbb{N}^m$ *with* $|\mathbf{j}|_{\mathbf{w}} < r$ *and all* $\mathbf{a} \in S$.

We conclude with the case of multi-indices bounded on each coordinate separately.

Corollary 4.15. *Given* $(r_1, r_2, \ldots, r_m) \in \mathbb{N}_+^m$ *and elements* $b_{\mathbf{j},\mathbf{a}} \in \mathbb{F}$, *for* $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$ *with* $j_k < r_k$, *for all* $k = 1, 2, \ldots, m$, *and for* $\mathbf{a} \in S$, *there exists a unique polynomial of the form*

$$F(\mathbf{x}) = \sum_{i_1=0}^{r_1\#S_1-1} \sum_{i_2=0}^{r_2\#S_2-1} \cdots \sum_{i_m=0}^{r_m\#S_m-1} F_{\mathbf{i}} \mathbf{x}^{\mathbf{i}},$$

such that $F^{(\mathbf{j})}(\mathbf{a}) = b_{\mathbf{j},\mathbf{a}}$, *for all* $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$ *with* $j_k < r_k$, *for all* $k = 1, 2, \ldots, m$, *and all* $\mathbf{a} \in S$.

4.4. Evaluation codes with consecutive derivatives

In this subsection, we extend the notion of *evaluation code* from the theory of error-correcting codes (see [12, Section 2] and [17, Section 4.1], for instance) to evaluation codes with consecutive derivatives. By doing so, we generalize *multiplicity codes* [20], which have been shown to achieve good parameters in decoding, local decoding and list decoding [19, 20]. We compute the dimensions of the new codes and give a lower bound on their minimum Hamming distance.

Definition 4.16. Given a decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$ and a set of monomials $\mathcal{M} \subseteq \mathcal{J}_S$, we define the $\mathbb{F}$-linear code (that is, the $\mathbb{F}$-linear vector space)

$$\mathcal{C}(S, \mathcal{M}, \mathcal{J}) = \mathrm{Ev}(\langle \mathcal{M} \rangle_{\mathbb{F}}) \subseteq \mathbb{F}^{\#S \cdot \#\mathcal{J}},$$

where Ev is the evaluation map from Definition 3.4.

As in [20], we will consider these codes over the alphabet $\mathbb{F}^{\#\mathcal{J}}$, that is, each evaluation $(F^{(\mathbf{i})}(\mathbf{a}))_{\mathbf{i} \in \mathcal{J}} \in \mathbb{F}^{\#\mathcal{J}}$, for $\mathbf{a} \in S$, constitutes one symbol of the alphabet. Thus each codeword has length $\#S$ over this alphabet. This leads to the following definition of minimum Hamming distance of an $\mathbb{F}$-linear code.

Definition 4.17. Given an $\mathbb{F}$-linear code $\mathcal{C} \subseteq (\mathbb{F}^{\#\mathcal{J}})^{\#S}$, we define its minimum Hamming distance as

$$d_H(\mathcal{C}) = \min\{\mathrm{wt_H}(\mathbf{c}) : \mathbf{c} \in \mathcal{C}, \mathbf{c} \neq \mathbf{0}\},$$

where, for any $\mathbf{c} \in (\mathbb{F}^{\#\mathcal{J}})^{\#S}$, $\mathrm{wt_H}(\mathbf{c})$ denotes the number of its non-zero components over the alphabet $\mathbb{F}^{\#\mathcal{J}}$.

As a consequence of Theorem 4.12, we may exactly compute the dimensions of the codes in Definition 4.16 and give a lower bound on their minimum Hamming distance.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Corollary 4.18. *The code in Definition 4.16 satisfies that*

$$\dim_{\mathbb{F}}(\mathcal{C}(S,\mathcal{M},\mathcal{J})) = \#\mathcal{M}, \quad \text{and}$$
$$d_H(\mathcal{C}(S,\mathcal{M},\mathcal{J})) \geqslant \left\lceil \frac{\min\{\#\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}) : F(\mathbf{x}) \in \langle\mathcal{M}\rangle_{\mathbb{F}}\}}{\#\mathcal{J}} \right\rceil.$$

Remark 4.19. Given a vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, a positive integer $r \in \mathbb{N}_+$, and a set of monomials

$$\mathcal{M} \subseteq \left\{ x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m} : \sum_{j=1}^{m} \left\lfloor \frac{i_j}{\#S_j} \right\rfloor w_j < r \right\},$$

we may define, as a particular case of the codes in Definition 4.16, the corresponding weighted multiplicity code as the $\mathbb{F}$-linear code

$$\mathcal{C}(S,\mathcal{M},\mathbf{w},r) = \mathrm{Ev}(\langle\mathcal{M}\rangle_{\mathbb{F}}) \subseteq (\mathbb{F}^{\mathrm{B}(\mathbf{w};r)})^{\#S}.$$

Observe that weighted multiplicity codes contain as particular cases classical Reed–Muller codes (see [18, Section 13.2]), by choosing $\mathbf{w} = (r, r, \ldots, r)$ for a given $r \in \mathbb{N}_+$, and classical multiplicity codes [20] by choosing $\mathbf{w} = (1, 1, \ldots, 1)$ and an arbitrary $r \in \mathbb{N}_+$. Therefore, choices of $\mathbf{w} \in \mathbb{N}^m$ such that $1 \leqslant w_i \leqslant r$, for $i = 1, 2, \ldots, m$, give codes with the same length but intermediate alphabet sizes between those of Reed–Muller and multiplicity codes. This has the extra flexibility (see [20, Section 1.2]) of choosing alphabets of sizes $\#(\mathbb{F}^{\mathrm{B}(\mathbf{w};r)})$ (whenever $\mathbb{F}$ is finite), where

$$1 \leqslant \mathrm{B}(\mathbf{w};r) \leqslant \binom{m+r-1}{m}.$$

4.5. Bounds by DeMillo, Lipton, Zippel, Alon and Füredi

In this subsection, we obtain a weaker but more concise version of the bound (3.1) for a single polynomial, which has as particular cases the bounds by DeMillo and Lipton [8], Zippel [26, Theorem 1], [27, Proposition 3] and Alon and Füredi [2, Theorem 5]. We observe that Counterexample 7.4 in [4] shows that a straightforward extension of these bounds to standard multiplicities as in (1.1) is not possible, in contrast with the bound given by Schwartz in [25, Lemma 1], which has been already extended in [9, Lemma 8].

Theorem 4.20. *For any decreasing finite set $\mathcal{J} \subseteq \mathbb{N}^m$ and any polynomial $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, if $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x})) \in \mathcal{J}_S$, for some monomial ordering, then it holds that*

$$\#(S \setminus \mathcal{V}_{\mathcal{J}}(F(\mathbf{x})))\#\mathcal{J} \geqslant \#\{\mathbf{j} \in \mathcal{J}_S : \mathbf{j} \succeq \mathbf{i}\}. \tag{4.1}$$

Proof. First, from the bound (3.1) and Lemma 3.14, we obtain that

$$\#(S \setminus \mathcal{V}_{\mathcal{J}}(F(\mathbf{x})))\#\mathcal{J} \geqslant \#S\#\mathcal{J} - \#\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}) = \#(\mathcal{J}_S \setminus \Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}})), \tag{4.2}$$

where we consider $\Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}) \subseteq \mathbb{N}^m$ by abuse of notation. As explained in Section 3.3, we may lower-bound $\#(\mathcal{J}_S \setminus \Delta(\langle F(\mathbf{x})\rangle_{\mathcal{J}}))$ by the number of multi-indices $\mathbf{j} \in \mathcal{J}_S$ satisfying $\mathbf{j} \succeq \mathbf{i}$, and we are done. □

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

The following consequence summarizes the results by DeMillo and Lipton [8], and Zippel [26, Theorem 1], [27, Proposition 3].

Corollary 4.21 ([8, 26, 27]). *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be such that its degree in the jth variable is $d_j \in \mathbb{N}$, for $j = 1, 2, \ldots, m$. If $d_j < \#S_j$, for $j = 1, 2, \ldots, m$, then the number of non-zeros in S of $F(\mathbf{x})$ is at least*

$$\prod_{j=1}^{m} (\#S_j - d_j).$$

Proof. The result is the particular case $\mathcal{J} = \{\mathbf{0}\}$ of the previous theorem using any monomial ordering and the facts that $\mathcal{J}_S = S$ and $i_j \leqslant d_j$, for $j = 1, 2, \ldots, m$. □

The following is a similar bound due to Alon and Füredi [2, Theorem 5].

Corollary 4.22 ([2]). *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$. If not all points in S are zeros of $F(\mathbf{x})$, then the number of its non-zeros in S is at least*

$$\min\left\{ \prod_{j=1}^{m} y_j : 1 \leqslant y_j \leqslant \#S_j, \sum_{j=1}^{m} y_j \geqslant \sum_{j=1}^{m} \#S_j - \deg(F(\mathbf{x})) \right\}.$$

Proof. The result follows from Theorem 4.20 as in the previous corollary, taking any monomial ordering and considering $y_j = \#S_j - i_j$, for $j = 1, 2, \ldots, m$. □

Similarly, we may derive the following generalization given in [4, Theorem 1.2].

Corollary 4.23 ([4]). *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and let $1 \leqslant b_j \leqslant \#S_j$, for $j = 1, 2, \ldots, m$. If the degree of $F(\mathbf{x})$ in the jth variable is at most $\#S_j - b_j$, for $j = 1, 2, \ldots, m$, then the number of its non-zeros in S is at least*

$$\min\left\{ \prod_{j=1}^{m} y_j : b_j \leqslant y_j \leqslant \#S_j, \sum_{j=1}^{m} y_j \geqslant \sum_{j=1}^{m} \#S_j - \deg(F(\mathbf{x})) \right\}.$$

We omit the case of weighted multiplicities. In the next section, we will give an extension of the bound given by Schwartz in [25, Lemma 1] to weighted multiplicities in the sense of (1.1), which is stronger than the bound in Corollary 3.9 in some cases.

We conclude with the case of multi-indices bounded on each coordinate separately.

Corollary 4.24. *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ with $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$, $\mathbf{i} = (i_1, i_2, \ldots, i_m)$, for some monomial ordering. If $i_j < r_j \#S_j$, for $j = 1, 2, \ldots, m$, then the number N of elements $\mathbf{s} \in S$ such that $F^{(\mathbf{j})}(\mathbf{s}) \neq 0$, for some $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$ with $j_k < r_k$, for all $k = 1, 2, \ldots, m$, satisfies*

$$N \cdot \prod_{j=1}^{m} r_j \geqslant \prod_{j=1}^{m} (r_j \#S_j - i_j).$$

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

4.6. The Schwartz–Zippel bound on the whole grid

In the next section, we will give an extension of bound given by Schwartz in [25, Lemma 1] for weighted multiplicities that can be proved as the extensions to standard multiplicities given in [9, Lemma 8] and [12, Theorem 5]. In this subsection, we observe that the case where all points in S are zeros of a given weighted multiplicity follows from Corollary 4.3.

Corollary 4.25. *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, let $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, let $r \in \mathbb{N}_+$, and assume that $s = \#S_1 = \#S_2 = \cdots = \#S_m$. If all points in $S = S_1 \times S_2 \times \cdots \times S_m$ are zeros of $F(\mathbf{x})$ of weighted multiplicity at least r, then*

$$r\#S \leqslant \deg_{\mathbf{w}}(F(\mathbf{x}))s^{m-1}.$$

Proof. Assume that the bound does not hold, take $\mathbf{x}^{\mathbf{i}}$ such that $|\mathbf{i}|_{\mathbf{w}} = \deg_{\mathbf{w}}(F(\mathbf{x}))$ and whose coefficient in $F(\mathbf{x})$ is not zero, and take a vector $\mathbf{r} = (r_1, r_2, \ldots, r_m) \in \mathbb{N}^m$ with $|\mathbf{r}|_{\mathbf{w}} \geqslant r$. Then

$$sw_1r_1 + sw_2r_2 + \cdots + sw_mr_m \geqslant sr > \deg_{\mathbf{w}}(F(\mathbf{x})) = |\mathbf{i}|_{\mathbf{w}},$$

hence there exists a j such that $r_j\#S_j > i_j$. By Corollary 4.3, some element in S is not a zero of $F(\mathbf{x})$ of weighted multiplicity at least r, which contradicts the assumptions and we are done. □

5. The Schwartz–Zippel bound for weighted multiplicities

As we will see in Proposition 5.5, the bound given by Schwartz in [25, Lemma 1] can be derived by those given by DeMillo and Lipton [8], and Zippel (see [26, Theorem 1], [27, Proposition 3]), and is usually referred to as the Schwartz–Zippel bound. This bound has recently been extended to standard multiplicities in [9, Lemma 8], and further in [12, Theorem 5]. In this section, we observe that it may be easily extended to weighted multiplicities (see Definition 2.4), due to the additivity of weighted order functions. We show the sharpness of this bound and compare it with the bound (3.1) with an example, whenever it makes sense to compare the two bounds.

5.1. The bound

Theorem 5.1. *Let $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$ be a vector of positive weights, let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and let $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$, $\mathbf{i} = (i_1, i_2, \ldots, i_m)$, with respect to the lexicographic ordering. It holds that*

$$\sum_{\mathbf{a} \in S} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) \leqslant \#S \sum_{j=1}^{m} \frac{i_j w_j}{\#S_j}. \tag{5.1}$$

When $w_1 = w_2 = \cdots = w_m = 1$, observe that [12, Theorem 5] is recovered from this theorem, and [9, Lemma 8] is recovered from the next corollary. Observe also that this corollary is stronger than Corollary 4.25.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Corollary 5.2. *Let* $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *and* $\mathbf{w} \in \mathbb{N}_+^m$. *If* $s = \#S_1 = \#S_2 = \cdots = \#S_m$, *then*

$$\sum_{\mathbf{a} \in S} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) \leqslant \deg_{\mathbf{w}}(F(\mathbf{x})) s^{m-1}.$$

To prove Theorem 5.1, we need an auxiliary lemma, whose proof can be directly translated from those of [9, Lemma 5] and [9, Corollary 7].

Lemma 5.3. *If* $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *and* $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{F}^m$, *then*

(1) $m_{\mathbf{w}}(F^{(\mathbf{i})}(\mathbf{x}), \mathbf{a}) \geqslant m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) - |\mathbf{i}|_{\mathbf{w}}$, *for all* $\mathbf{i} \in \mathbb{N}^m$, *and*
(2) $m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) \leqslant m_{w_m}(F(a_1, a_2, \ldots, a_{m-1}, x_m), a_m)$.

We may now prove Theorem 5.1. We follow closely the steps in the proof of [9, Lemma 8].

Proof of Theorem 5.1. We will prove the result by induction on m, where the case $m = 1$ follows from (1.1). Fix then $m > 1$. We may assume without loss of generality that $x_1 \prec_l x_2 \prec_l \cdots \prec_l x_m$, where $\preceq_l$ is the lexicographic ordering. Write $\mathbf{x}' = (x_1, x_2, \ldots, x_{m-1})$. There are unique polynomials $F_j(\mathbf{x}') \in \mathbb{F}[\mathbf{x}']$, for $j = 1, 2, \ldots, t$, such that

$$F(\mathbf{x}) = \sum_{j=0}^{t} F_j(\mathbf{x}') x_m^j,$$

where $\mathrm{LM}(F(\mathbf{x})) = \mathrm{LM}(F_t(\mathbf{x}')) x_m^t$. Let $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in S$ and write $\mathbf{a}' = (a_1, a_2, \ldots, a_{m-1})$ and $\mathbf{w}' = (w_1, w_2, \ldots, w_{m-1})$. Take $\mathbf{k} \in \mathbb{N}^{m-1}$ such that $|\mathbf{k}|_{\mathbf{w}'} = m_{\mathbf{w}'}(F_t(\mathbf{x}'), \mathbf{a}')$ and $F_t^{(\mathbf{k})}(\mathbf{a}') \neq 0$. By the previous lemma, we see that

$$\begin{aligned} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) &\leqslant |(\mathbf{k}, 0)|_{\mathbf{w}} + m_{\mathbf{w}}(F^{(\mathbf{k},0)}(\mathbf{x}), \mathbf{a}) \\ &\leqslant m_{\mathbf{w}'}(F_t(\mathbf{x}'), \mathbf{a}') + m_{w_m}(F^{(\mathbf{k},0)}(\mathbf{a}', x_m), a_m). \end{aligned}$$

Summing these inequalities over all $a_m \in S_m$ and applying the case $m = 1$, we obtain that

$$\sum_{a_m \in S_m} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) \leqslant m_{\mathbf{w}'}(F_t(\mathbf{x}'), \mathbf{a}') \#S_m + w_m t.$$

Using this last inequality, summing over $a_i \in S_i$, for $i = 1, 2, \ldots, m-1$, and applying the case of $m - 1$ variables, it follows that

$$\begin{aligned} \sum_{\mathbf{a} \in S} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) &\leqslant \sum_{a_1 \in S_1} \cdots \sum_{a_{m-1} \in S_{m-1}} m_{\mathbf{w}'}(F_t(\mathbf{x}'), \mathbf{a}') \#S_m + w_m t \frac{\#S}{\#S_m} \\ &\leqslant \sum_{j=1}^{m-1} w_j i_j \frac{\#S}{\#S_j} + w_m t \frac{\#S}{\#S_m}, \end{aligned}$$

and the result follows. □

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

5.2. Sharpness of the bound

In this subsection, we prove the sharpness of the bound (5.1), whose proof can be translated word by word from that of [13, Proposition 7]. Therefore, we only present a sketch of the proof.

Proposition 5.4. *For all finite sets* $S_1, S_2, \ldots, S_m \subseteq \mathbb{F}$, $S = S_1 \times S_2 \times \cdots \times S_m$, *all vectors of positive weights* $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$ *and all* $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$, *there exists a polynomial* $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *such that* $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$ *with respect to the lexicographic ordering, and such that*

$$\sum_{\mathbf{a} \in S} m_{\mathbf{w}}(F(\mathbf{x}), \mathbf{a}) = \#S \sum_{j=1}^{m} \frac{i_j w_j}{\#S_j}.$$

Sketch of proof. Denote $s_j = \#S_j$ and $S_j = \{a_1^{(j)}, a_2^{(j)}, \ldots, a_{s_j}^{(j)}\}$, and choose $r_k^{(j)} \in \mathbb{N}$ such that $i_j = r_1^{(j)} + r_2^{(j)} + \cdots + r_{s_j}^{(j)}$, for $k = 1, 2, \ldots, s_j$ and $j = 1, 2, \ldots, m$. Now define

$$F(\mathbf{x}) = \prod_{j=1}^{m} \prod_{k=1}^{s_j} (x_j - a_k^{(j)})^{r_k^{(j)}}.$$

Now, fixing integers $1 \leqslant k_j \leqslant s_j$, for $j = 1, 2, \ldots, m$, translating the point $(a_{k_1}^{(1)}, a_{k_2}^{(2)}, \ldots, a_{k_m}^{(m)})$ to the origin $\mathbf{0}$, and using the Gröbner basis from Corollary 4.10, we see that

$$m_{\mathbf{w}}(F(\mathbf{x}), (a_{k_1}^{(1)}, a_{k_2}^{(2)}, \ldots, a_{k_m}^{(m)})) = r_{k_1}^{(1)} w_1 + r_{k_2}^{(2)} w_2 + \cdots + r_{k_m}^{(m)} w_m,$$

for all $k_j = 1, 2, \ldots, s_j$ and all $j = 1, 2, \ldots, m$. The result then follows by summing these multiplicities.

5.3. Comparison with the footprint bound

In this subsection, we will compare the bounds (3.1) and (5.1) whenever it makes sense to do so. To that end, we will write them as follows: fix a vector of positive weights $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{N}_+^m$, a positive integer $r \in \mathbb{N}_+$, and a polynomial $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that $\mathbf{x}^{\mathbf{i}} = \mathrm{LM}(F(\mathbf{x}))$, $\mathbf{i} = (i_1, i_2, \ldots, i_m)$, with respect to the lexicographic ordering. We first consider the footprint bound as in Corollary 3.9:

$$\#\mathcal{V}_{\geqslant r, \mathbf{w}}(F(\mathbf{x})) \cdot \mathrm{B}(\mathbf{w}; r) \leqslant \#\Delta\left(\left\langle \{F(\mathbf{x})\} \bigcup \left\{ \prod_{j=1}^{m} G_j(x_j)^{r_j} : \sum_{j=1}^{m} r_j w_j \geqslant r \right\} \right\rangle\right). \tag{5.2}$$

And next we consider the bound (5.1) as follows:

$$\#\mathcal{V}_{\geqslant r, \mathbf{w}}(F(\mathbf{x})) \cdot r \leqslant \#S \sum_{j=1}^{m} \frac{i_j w_j}{\#S_j}. \tag{5.3}$$

First we observe that the bound (5.2) also holds for any other monomial ordering, and not only the lexicographic one, as is the case with (5.3). Second we observe that (5.3) gives

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Table 1. Upper bounds on the number of zeros of weighted multiplicity at least $r = 5$ when $w_1 = 2$, $w_2 = 3$ and $\#S_1 = \#S_2 = 4$, from Example 5.6

x_1^7	**15**	**15**	**15**	**15**								
x_1^6	14	**14**	**15**	**15**								
x_1^5	13	13	**14**	**15**								
x_1^4	12	13	14	15								
x_1^3	9	10	11	12	14	**14**	**14**	**14**	**15**	**15**	**15**	**15**
x_1^2	6	7	9	10	12	12	**13**	**13**	**14**	**14**	**15**	**15**
x_1	3	4	6	8	10	10	**11**	**12**	**13**	**13**	**14**	**15**
1	0	2	4	6	8	9	10	11	12	**13**	**14**	**15**
	1	x_2	x_2^2	x_2^3	x_2^4	x_2^5	x_2^6	x_2^7	x_2^8	x_2^9	x_2^{10}	x_2^{11}
x_1^7	–	–	–	–								
x_1^6	14	–	–	–								
x_1^5	**12**	13	15	–								
x_1^4	**9**	**11**	**12**	**14**								
x_1^3	**7**	**8**	**10**	12	**13**	15	–	–	–	–	–	–
x_1^2	**4**	**6**	**8**	**9**	**11**	12	14	–	–	–	–	–
x_1	**2**	4	**5**	**7**	**8**	10	12	13	15	–	–	–
1	0	**1**	**3**	**4**	**6**	**8**	**9**	11	12	14	–	–
	1	x_2	x_2^2	x_2^3	x_2^4	x_2^5	x_2^6	x_2^7	x_2^8	x_2^9	x_2^{10}	x_2^{11}

no information whereas (5.2) does, whenever

$$\sum_{j=1}^{m} \left\lfloor \frac{i_j}{\#S_j} \right\rfloor w_j < r \leqslant \sum_{j=1}^{m} \frac{i_j w_j}{\#S_j}, \tag{5.4}$$

by the discussion in Section 3.3.

Next, we observe that when we do not count multiplicities, that is, $w_1 = w_2 = \cdots = w_m = r = 1$, then (5.2) implies (5.3) via Theorem 4.20.

Proposition 5.5. *If* $w_1 = w_2 = \cdots = w_m = r = 1$, *that is,* $\mathcal{J} = \{\mathbf{0}\}$, *it holds that* $\mathrm{B}(\mathbf{w}; r) = 1$ *and*

$$\#\Delta(\langle F(\mathbf{x}), G_1(x_1), G_2(x_2), \ldots, G_m(x_m)\rangle) \leqslant \#S - \prod_{j=1}^{m}(\#S_j - i_j) \leqslant \#S \sum_{j=1}^{m} \frac{i_j}{\#S_j}.$$

In particular, (5.2) *implies* (5.3) *in this case.*

Moreover, when $m = 1$ and we count multiplicities, all bounds coincide, giving (1.2). In the following example we show that this is not the case in general. As we will see, each bound, (5.2) and (5.3), can be tighter than the other one in different cases, hence complementing each other.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Example 5.6. Consider $m = 2$, $w_1 = 2$, $w_2 = 3$, $r = 5$ and $\#S_1 = \#S_2 = 4$. Thus we have that

$$\mathcal{J} = \{(0,0),(1,0),(0,1),(2,0)\}, \quad \text{and}$$
$$\mathcal{J}_S = ([0,11] \times [0,3]) \cup ([0,3] \times [0,7]).$$

Consider all pairs $(i_1, i_2) \in \mathcal{J}_S$ and polynomials $F(x_1, x_2)$ such that $\mathrm{LM}(F(x_1, x_2)) = x_1^{i_1} x_2^{i_2}$, with respect to the lexicographic ordering. In Table 1, we show the upper bounds on the number of zeros of $F(x_1, x_2)$ of weighted multiplicity at least 5 given by (5.2) (table above) and (5.3) (table below), respectively. As is clear from the figure, in some regions of the set $\mathcal{J}_S$, the first bound is tighter than the second (bold numbers in the table above) and *vice versa* (bold numbers in the table below). Furthermore the first bound gives non-trivial information in the region given by (5.4), where the second does not (depicted by dashes).

Appendix: Proof of Lemma 3.5

In this appendix, we give the proof of Lemma 3.5. We first treat the univariate case ($m = 1$) in the classical form. The proof for Hasse derivatives can be directly translated from the result for classical derivatives.

Lemma A.1. *Let $a_1, a_2, \ldots, a_n \in \mathbb{F}$ be pairwise distinct and let $M \in \mathbb{N}_+$. There exist polynomials $F_{i,j}(x) \in \mathbb{F}[x]$ such that*

$$F_{i,j}^{(k)}(a_l) = \delta_{i,k}\delta_{j,l},$$

for all $i, k = 0, 1, 2, \ldots, M$ and all $j, l = 1, 2, \ldots, n$, where δ denotes the Kronecker delta.

Now, since $\mathcal{J}$ is finite, we may fix an integer M such that $\mathcal{J} \subseteq [0, M]^m$. Similarly, we may find a finite set $S \subseteq \mathbb{F}$ such that $T \subseteq S^m$. Then denote $s = \#S$ and $S = \{a_1, a_2, \ldots, a_s\}$, and let $F_{i,j,k}(x_k) \in \mathbb{F}[x_k]$ be polynomials as in the previous lemma in each variable x_k, for $i = 0, 1, 2, \ldots, M$, $j = 1, 2, \ldots, s$ and $k = 1, 2, \ldots, m$. Now define

$$F_{\mathbf{i},\mathbf{j}}(\mathbf{x}) = F_{i_1,j_1,1}(x_1)F_{i_2,j_2,2}(x_2)\cdots F_{i_m,j_m,m}(x_m) \in \mathbb{F}[\mathbf{x}],$$

for $\mathbf{i} = (i_1, i_2, \ldots, i_m) \in [0, M]^m$ and $\mathbf{j} = (j_1, j_2, \ldots, j_m) \in [1, s]^m$. By the previous lemma and Lemma 3.3, we see that

$$F_{\mathbf{i},\mathbf{j}}^{(\mathbf{k})} l(a_{l_1}, a_{l_2}, \ldots, a_{l_m}) = (\delta_{i_1,k_1}\delta_{i_2,k_2}\cdots\delta_{i_m,k_m})(\delta_{j_1,l_1}\delta_{j_2,l_2}\cdots\delta_{j_m,l_m}) = \delta_{\mathbf{i},\mathbf{k}}\delta_{\mathbf{j},\mathbf{l}},$$

for all $\mathbf{i}, \mathbf{k} \in [0, M]^m$ and all $\mathbf{j}, \mathbf{l} \in [1, s]^m$. Finally, given values $b_{\mathbf{i},\mathbf{j}} \in \mathbb{F}$, for $\mathbf{i} \in \mathcal{J}$ and $\mathbf{j} \in T$, define

$$F(\mathbf{x}) = \sum_{\mathbf{i} \in \mathcal{J}} \sum_{\mathbf{j} \in T} b_{\mathbf{i},\mathbf{j}} F_{\mathbf{i},\mathbf{j}}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}].$$

We see that $\mathrm{Ev}(F(\mathbf{x})) = ((b_{\mathbf{i},\mathbf{j}})_{\mathbf{i} \in \mathcal{J}})_{\mathbf{j} \in T}$, and we are done.

https://doi.org/10.1017/S0963548318000342 Published online by Cambridge University Press

Acknowledgement

The authors gratefully acknowledge the support from the Danish Council for Independent Research (grant no. DFF-4002-00367). The second author also gratefully acknowledges the support from the Danish Council for Independent Research via an EliteForsk-Rejsestipendium (grant no. DFF-5137-00076B).

References

[1] Alon, N. (1999) Combinatorial Nullstellensatz. *Combin. Probab. Comput.* **8** 7–29.
[2] Alon, N. and Füredi, Z. (1993) Covering the cube by affine hyperplanes. *European J. Combin.* **14** 79–83.
[3] Ball, S. and Serra, O. (2009) Punctured combinatorial Nullstellensätze. *Combinatorica* **29** 511–522.
[4] Bishnoi, A., Clark, P. L., Potukuchi, A. and Schmitt, J. R. (2018) On zeros of a polynomial in a finite grid. *Combin. Probab. Comput.* **27** 310–333.
[5] Bruen, A. A. (1992) Polynomial multiplicities over finite fields and intersection sets. *J. Combin. Theory Ser. A* **60** 19–33.
[6] Clark, P. L. (2014) The combinatorial Nullstellensätze revisited. *Electron. J. Combin.* **21** 1–17.
[7] Cox, D., Little, J. and O'Shea, D. (2007) *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer.
[8] DeMillo, R. A. and Lipton, R. J. (1978) A probabilistic remark on algebraic program testing. *Inform. Process. Lett.* **7** 193–195.
[9] Dvir, Z., Kopparty, S., Saraf, S. and Sudan, M. (2013) Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. *SIAM J. Comput.* **42** 2305–2328.
[10] Gasca, M. and Sauer, T. (2000) Polynomial interpolation in several variables. *Adv. Comput. Math.* **12** 1–377.
[11] Geil, O. and Høholdt, T. (2000) Footprints or generalized Bezout's theorem. *IEEE Trans. Inform. Theory* **46** 635–641.
[12] Geil, O. and Thomsen, C. (2013) Weighted Reed–Muller codes revisited. *Des. Codes Cryptogr.* **66** 195–220.
[13] Geil, O. and Thomsen, C. (2017) More results on the number of zeros of multiplicity at least r. *Discrete Math.* **340** 1028–1038.
[14] Hasse, H. (1936) Theorie der höheren Differentiale in einem algebraischen Funktionenkörper mit vollkommenem Konstantenkörper bei beliebiger Charakteristik. *J. Reine Angew. Math.* **175** 50–54.
[15] Hirschfeld, J. W. P., Korchmaros, G. and Torres, F. (2008) *Algebraic Curves over a Finite Field*, Princeton University Press.
[16] Høholdt, T. (1998) On (or in) the Blahut footprint. In *Codes, Curves, and Signals: Common Threads in Communications* (A. Vardy, ed.), Springer, pp. 3–7.
[17] Høholdt, T., van Lint, J. H. and Pellikaan, R. (1998) Algebraic geometry codes. In *Handbook of Coding Theory* (V. S. Pless and W. C. Huffman, eds), Elsevier, Vol. 1, pp. 871–961.
[18] Huffman, W. C. and Pless, V. (2003) *Fundamentals of Error-Correcting Codes*, Cambridge University Press.
[19] Kopparty, S. (2015) List-decoding multiplicity codes. *Theory Comput.* **11** 149–182.
[20] Kopparty, S., Saraf, S. and Yekhanin, S. (2014) High-rate codes with sublinear-time decoding. *J. Assoc. Comput. Mach.* **61** 28.
[21] Kós, G. and Rónyai, L. (2012) Alon's Nullstellensatz for multisets. *Combinatorica* **32** 589–605.
[22] Lorentz, R. A. (2000) Multivariate Hermite interpolation by algebraic polynomials: A survey. *J. Comput. Appl. Math.* **122** 167–201.

[23] Michałek, M. (2010) A short proof of combinatorial Nullstellensatz. *Amer. Math. Monthly* **117** 821–823.
[24] Pellikaan, R. and Wu, X.-W. (2004) List decoding of q-ary Reed–Muller codes. *IEEE Trans. Inform. Theory* **50** 679–682. Extended version: http://www.win.tue.nl/~ruudp/paper/43-exp.pdf
[25] Schwartz, J. T. (1980) Fast probabilistic algorithms for verification of polynomial identities. *J. Assoc. Comput. Mach.* **27** 701–717.
[26] Zippel, R. (1979) Probabilistic algorithms for sparse polynomials. In *Proc. International Symposium on Symbolic and Algebraic Computation (EUROSAM '79)*, Springer, pp. 216–226.
[27] Zippel, R. (1989) An explicit separation of relativised random and polynomial time and relativised deterministic polynomial time. Technical report, Cornell University, Ithaca, NY, USA.