Faster Algorithms via Waring Decompositions

Kevin Pratt *

Computer Science Department, Carnegie Mellon University

July 16, 2018

Abstract

We show that decompositions of certain polynomials as sums of powers of linear forms yield faster algorithms for some algebraic problems. Such a decomposition is known as a *Waring decomposition*. Our results are:

1. We give a black-box algorithm for computing the sum of the coefficients of the degree-k multilinear monomials in a polynomial over a field of characteristic zero. Our algorithm runs in time $O^*(n^{k/2})$ and space $poly(n)$. This solves an open problem of Koutis and Williams [1].
2. We give a randomized $O^*((3e/2)^k) \approx O^*(4.08^k)$ time, polynomial space, black-box algorithm for detecting multilinear monomials of degree k over a field of characteristic zero. This improves on the $O^*(4.32^k)$ time, exponential space algorithm given in [2]. We note that there is a $O^*(2^k)$ lower bound on our approach.

1 Introduction

Let $\mathbb{F}$ be a field of characteristic zero and let C be an arithmetic circuit computing a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$. In this note we focus on the following two problems, using the terminology of [1]:

1. (k, n)-MLC: Compute the sum of the coefficients of the degree-k multilinear monomials in f.
2. k-MLD: Decide whether f contains a multilinear monomial of degree k.

We prove the following.

Theorem 1. *(k, n)-MLC can be solved in $O^*(n^{k/2})$ time and $poly(|C|)$ space assuming unit-cost arithmetic operations over $\mathbb{F}$.*

*kpratt@andrew.cmu.edu

Theorem 2. *k-MLD can be solved with constant probability and one-sided error in time $O^*((3e/2)^k)$ and space poly$(|C|)$ assuming unit-cost arithmetic operations over $\mathbb{F}$.*

The previous fastest algorithm for (k, n)-MLC ran in $O^*(n^k)$ time. Improvements were only known in special cases, see e.g. Theorems 5.2 and 5.3 in [1]. Theorem 1 yields faster polynomial space algorithms for several exact counting problems such as counting k-set packings and counting paths of length k in a directed graph.

The previous fastest algorithm for k-MLD ran in time $O^*(4.32^k)$ [2]. In the case that C is monotone over $\mathbb{Z}$ (that is, C only uses positive integers), a $O^*(2^k)$ time, polynomial space algorithm was given in [3, 4]. This algorithm is at the heart of the $O^*(2^k)$ time algorithm for detecting paths of length k in a directed graph.

Our algorithms only require black-box access to C. They are essentially consequences of *Waring decompositions* for the elementary symmetric polynomials as given in [5].

In contemporaneous work, [6] have shown some related results.

In the next section we provide the necessary facts about Waring decompositions. We then prove Theorems 1 and 2.

2 Preliminaries

Here we introduce basic concepts regarding Waring decompositions. There is an extensive theory on this topic that has developed over the past 150 years. We refer the curious reader to [7, 8].

Let $\mathbb{F}$ be a field of characteristic zero. We denote by $\mathbb{F}[x_1, \ldots, x_n] = \bigoplus_{d \in \mathbb{N}} S_d$ the polynomial ring over $\mathbb{F}$, graded by degree. Recall that the d-th graded piece S_d is the vector space of n variate homogeneous degree d polynomials over $\mathbb{F}$.

Definition 3. For $f \in S_d$, a *Waring decomposition* of f of *length* r is an expression of the form

$$f = \alpha_1 l_1^d + \cdots + \alpha_r l_r^d,$$

where $l_1, \ldots, l_r$ are linear forms, and $\alpha_1, \ldots, \alpha_r \in \mathbb{F}$. The *Waring rank* of f, denoted $rk_{\mathbb{F}}(f)$, is the length of the shortest Waring decomposition of f.

The main focus of this paper is computing the following bilinear form on S_d.

Definition 4. Let $\langle \cdot, \cdot \rangle : S_d \times S_d \to \mathbb{F}$ be given by

$$\langle f, g \rangle = f(\partial_{x_1}, \ldots, \partial_{x_n}) \circ g(x_1, \ldots, x_n).$$

This is known as the *Apolar bilinear form*. It is a basic tool in the algebro-geometric study of homogeneous polynomials (see e.g. [9]).

The following facts relate this bilinear form to Waring decompositions.

Proposition 5. *$\langle \cdot, \cdot \rangle$ is symmetric and bilinear.*

Proposition 6. *If $f = (a_1 x_1 + \cdots + a_n x_n)^d$, then $\langle f, g \rangle = d! \cdot g(a_1, \ldots, a_n)$.*

Combining these, we have the following:

Theorem 7. *If $f = \alpha_1 l_1^d + \cdots + \alpha_d l_r^d$, where $l_i = c_{i,1}x_1 + \cdots + c_{i,n}x_n$ for $i = 1, \ldots, r$, then*

$$\langle f, g \rangle = d! \sum_{i=1}^{r} \alpha_i g(c_{i,1}, \ldots, c_{i,n}).$$

2.1 A Waring decomposition for elementary symmetric polynomials

Recall that the elementary symmetric polynomial of degree k in n variables is given by

$$e_k(x_1, \ldots, x_n) = e_{k,n} = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_1} \cdots x_{i_k}.$$

We make use of the following Waring decomposition for $e_{k,n}$ given in [5]. For $S \subseteq [n]$ and $i \in [n]$, define the indicator function $\delta(S, i) := -1$ if $i \in S$, and $\delta(S, i) := 1$ otherwise.

Theorem 8. *If $k = 2r + 1$, then*

$$2^{k-1} k! \cdot e_{k,n} = \sum_{S \subset [n], |S| \le r} (-1)^{|S|} \binom{n - r - |S| - 1}{r - |S|} (\delta(S,1)x_1 + \delta(S,2)x_2 + \cdots + \delta(S,n)x_n)^k.$$

Similarly, when $k = 2r$,

$$2^k (n-k) k! \cdot e_{k,n} = \sum_{S \subset [n], |S| \le r} (-1)^{|S|} \binom{n - r - |S| - 1}{r - |S|} (n - 2|S|)(\delta(S,1)x_1 + \cdots + \delta(S,n)x_n)^k.$$

Note that these are valid over any field of characteristic zero (or sufficiently large characteristic).

It is also shown in [5] that for k odd, $rk_{\mathbb{C}}(e_{k,n}) \ge \sum_{i=0}^{\lfloor k/2 \rfloor} \binom{n}{i}$, and for k even, $rk_{\mathbb{C}}(e_{k,n}) \ge (\sum_{i=0}^{k/2} \binom{n}{i}) - \binom{n-1}{k/2}$. Thus the decomposition for k odd is optimal, and the decomposition for k even is essentially optimal.

2.2 Degree k recovery

In the proofs of Theorems 1 and 2 we assume that the input circuit C computes a homogeneous polynomial of degree k. This is justified through the following observation, which is also used in [1]. Suppose that C computes $f = f_1 + f_2 + \cdots + f_d$, where f_i has degree i. Then we can compute $f_i(x_1, \ldots, x_n)$ by scaling all input gates in C by an indeterminate z, evaluating the resulting circuit over $\mathbb{F}[x_1, \ldots, x_n][z]/(z^{i+1})$, and returning the coefficient of z^i.

3 An algorithm for k-MLC

In this section we give a proof of Theorem 1.

Proof of Theorem 2. Suppose that d is odd; the case when d is even is analogous. First note that $\langle e_{k,n}, g\rangle$ equals the sum of the coefficients of the multilinear monomials in g. Then by Theorem 7 and Corollary 6, we have that

$$\langle e_{k,n}, g\rangle = \frac{1}{2^{d-1}} \sum_{S\subset[n],|S|\leq\lfloor k/2\rfloor} (-1)^{|S|} \binom{n-k-|S|-1}{k-|S|} g(\delta(S,1),\delta(S,2),\ldots,\delta(S,n)).$$

The result follows immediately from the right hand side. □

4 An algorithm for k-MLD

Given $f, g \in S_d$, we define $\overline{g} := g(x_1y_1,\ldots,x_ny_n) \in \mathbb{F}[x_1,\ldots,x_n][y_1,\ldots,y_n]$. Note that $\langle f, \overline{g}\rangle \in \mathbb{F}[y_1,\ldots,y_n]$.

We are now ready to prove Theorem 2.

Proof of Theorem 2. Our objective will be to sample a multilinear polynomial f such that the following holds: for all $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, the monomial $x_{i_1}\cdots x_{i_k}$ is contained in f with probability $1/N$. Then observe that if g contains a multilinear monomial, $\langle f, \overline{g}\rangle \not\equiv 0$ with probability $1/N$. If g does not contain a multilinear monomial we will always have $\langle f, \overline{g}\rangle \equiv 0$ (as f is multilinear). We will then use a Waring decomposition of f in conjunction with the Schwartz-Zippel lemma to test if $\langle f, \overline{g}\rangle \equiv 0$. To boost the probability of accepting in the case that g contains a multilinear monomial we repeat this N times.

Let $N = \Theta((3^{-1/2}e)^k)$. We sample f as follows. Partition the variables $x_1,\ldots,x_n$ into $M := \lceil 1.5k\rceil$ disjoint sets $X_1,\ldots,X_M$ by assigning each variable to a set uniformly at random. With the set X_i we associate the linear form $l_i = \sum_{x\in X_i} x$.

Now define $f(x_1,\ldots,x_n) := e_k(l_1,\ldots,l_M)$. First, note that f is multilinear since the elementary symmetric polynomials are multilinear and $l_1,\ldots,l_M$ are linear forms in disjoint sets of variables.

Next we consider the probability P_A that a given k-multilinear monomial, without loss of generality $x_1\cdots x_k$, is contained in f. This is precisely the probability that the variables $x_1,\ldots,x_k$ are assigned to distinct sets. Hence

$$P_A = \prod_{j=0}^{k-1} \frac{M-j}{M} = \frac{\lceil 1.5k\rceil!}{\lceil 0.5k\rceil!\lceil 1.5k\rceil^k}.$$

Applying Stirling's approximation (ignoring $poly(k)$ factors),

$$\begin{aligned} P_A &\approx (1.5k\cdot e^{-1})^{1.5k}(0.5k\cdot e^{-1})^{-0.5k}(1.5k)^{-k}, \\ &= \left(3^{1/2}e^{-1}\right)^k. \end{aligned}$$

Note that $P_A = \Theta(1/N)$. This shows that f has the desired properties.

It remains to test if $\langle f, \overline{g}\rangle \equiv 0$. To do so we apply the Schwartz-Zippel lemma. Let S be a set of $2k$ elements in $\mathbb{F}$, and let $r_1,\ldots,r_n$ be selected uniformly from S. Then if g contains

a multilinear monomial, $\langle f, \overline{g} \rangle(r_1, \ldots, r_n) = 0$ with probability at least $1/2N$, and if g does not contain a multilinear monomial, $\langle f, \overline{g} \rangle(r_1, \ldots, r_n) = 0$. The problem is thus reduced to computing $\langle f, \overline{g} \rangle(r_1, \ldots, r_n)$.

We evaluate $\langle f, \overline{g} \rangle(r_1, \ldots, r_n) = \langle e_k(l_1, \ldots, l_M), \overline{g} \rangle(r_1, \ldots, r_n)$ by using the decomposition for $e_k(l_1, \ldots, l_M)$ given by Theorem 8. By Theorem 7 this involves evaluating g at one point for each term in the decomposition, scaling the output of g by some (signed) binomial coefficient, and adding the result to a value we maintain as we enumerate over terms in the decomposition.

The only space overhead is that required to store this value, maintain a counter, store a partition of the variables, and enumerate the terms in the Waring decomposition. Hence polynomial space suffices.

The time complexity is dominated by the total number of iterations, N, times the length of each Waring decomposition. Up to $poly(k)$ factors, the length of the decomposition for $e_{k,M}$ is

$$\begin{aligned} \sum_{i=1}^{k/2} \binom{1.5k}{i} &\approx \binom{1.5k}{0.5k} = \frac{(1.5k)!}{k!(0.5k)!}, \\ &\approx (1.5k \cdot e^{-1})^{1.5k} (k \cdot e^{-1})^{-k} (k \cdot (2e)^{-1})^{-k/2} = (3^{3/2} 2^{-1})^k. \end{aligned}$$

Hence the total runtime is

$$O^*(N \cdot (3^{3/2}2^{-1})^k) = O^*((3^{1/2}e^{-1})(3^{3/2}2^{-1}))^k = O^*((3e/2)^k). \qquad \square$$

We remark that any algorithm that depends on generating Waring decompositions of multilinear polynomials over $\mathbb{C}$ must run in time $O^*(2^k)$. This follows from the fact that the Waring rank of any multilinear degree k polynomial as at least 2^{k-1}. For suppose that $f(x_1, \ldots, x_n)$ is multilinear and contains $x_1 \cdots x_k$. Then $h(x_1, \ldots, x_k, 0, 0, \ldots, 0) = \lambda \cdot x_1 \cdots x_n$ clearly has rank at most $rk_{\mathbb{C}}(f)$. But it is known that $rk_{\mathbb{C}}(x_1 \cdots x_k) = 2^{k-1}$; see [10]. Hence we must have $rk_{\mathbb{C}}(f) \geq 2^{k-1}$.

5 Acknowledgments

I would like to thank Ryan O'Donnell for many helpful suggestions in the preparation of this work.

References

[1] I. Koutis and R. Williams, "Limits and applications of group algebras for parameterized problems," in *International Colloquium on Automata, Languages, and Programming*, pp. 653–664, Springer, 2009.

[2] C. Brand, H. Dell, and T. Husfeldt, "Extensor-coding," in *Symposium on Theory of Computing*, ACM, 2018.

[3] I. Koutis, "Faster algebraic algorithms for path and packing problems," in *International Colloquium on Automata, Languages, and Programming*, pp. 575–586, Springer, 2008.

[4] R. Williams, "Finding paths of length k in $O^*(2^k)$ time," *Information Processing Letters*, vol. 109, no. 6, pp. 315–318, 2009.

[5] H. Lee, "Power sum decompositions of elementary symmetric polynomials," *Linear Algebra and its Applications*, vol. 492, pp. 89–97, 2016.

[6] V. Arvind, A. Chatterjee, R. Datta, and P. Mukhopadhyay, "Fast Exact Algorithms Using Hadamard Product of Polynomials," *ArXiv e-prints*, July 2018.

[7] A. Iarrobino and V. Kanev, *Power sums, Gorenstein algebras, and determinantal loci*. Springer Science & Business Media, 1999.

[8] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1254–1279, 2008.

[9] R. Ehrenborg and G.-C. Rota, "Apolarity and canonical forms for homogeneous polynomials," *European Journal of Combinatorics*, vol. 14, no. 3, pp. 157–181, 1993.

[10] E. Carlini, M. V. Catalisano, and A. V. Geramita, "The solution to Waring's problem for monomials," *arXiv preprint arXiv:1110.0745*, 2011.