Received: 25 July 2018 | Revised: 25 July 2019 | Accepted: 22 October 2019
DOI: 10.1002/nla.2275

RESEARCH ARTICLE

WILEY

Stably computing the multiplicity of known roots given leading coefficients

Gregory J. Clark[1] | **Joshua N. Cooper**[2]

[1]Saïd Business School, University of Oxford, Oxford, UK

[2]Department of Mathematics, University of South Carolina, Columbia, South Carolina

Correspondence
Gregory J. Clark, Saïd Business School, University of Oxford, Park End St, Oxford OX1 1HP, UK.
Email: gregory.clark@sbs.ox.ac.uk

Present Address
Gregory J. Clark, Saïd Business School, University of Oxford, Park End St, Oxford OX1 1HP, UK

Funding information
SPARC Graduate Research Grant

Summary

We show that a monic univariate polynomial over a field of characteristic zero, with k distinct nonzero known roots, is determined by precisely k of its proper leading coefficients. Furthermore, we give an explicit, numerically stable algorithm for computing the exact multiplicities of each root over $\mathbb{C}$. We provide a version of the result and accompanying algorithm when the field is not algebraically closed by considering the minimal polynomials of the roots. Then, we demonstrate how these results can be used to obtain the full homogeneous spectra of symmetric tensors—in particular, complete characteristic polynomials of hypergraphs.

KEYWORDS

polynomial root multiplicity, proper leading coefficients, spectral hypergraph theory, stable Vandermonde inversion

1 | INTRODUCTION

Győry et al.[1] present the following natural question.

Problem 1. Let K be a field of characteristic zero. Is it true that a monic polynomial $p = \sum_{i=0}^{n} c_i x^{n-i} \in K[x]$ of degree n with exactly k distinct zeros is determined up to finitely many possibilities by any k of its nonzero coefficients?

By degrees-of-freedom considerations, at least k coefficients are needed; which sets of k coefficients actually suffice, however, seems to be a delicate matter. We consider the following variation of Problem 1. The *codegree* of a monomial term Cx^k in a univariate polynomial $f(x)$ is $\deg(f) - k$. For a given polynomial $\sum_{i=0}^{n} c_i x^{n-i}$, a sequence of coefficients is *leading* if it has the form $(c_0, \ldots, c_k)$ for some k; it is *proper leading* if it has the form $(c_1, \ldots, c_k)$ for some k.

Problem 2. Let K be a field of characteristic zero. Is it true that a monic polynomial $p = \sum_{i=0}^{n} c_i x^{n-i} \in K[x]$ of unknown degree n, with exactly k distinct known zeros $r_1, r_2, \ldots, r_k$, is uniquely determined by its first k proper leading coefficients?

For present purposes, roots are "known" in an exact sense: We seek an algorithm that can determine multiplicities of roots given their exact minimal polynomials over K along with exact coefficients of p. Problem 2 is answered affirmatively with the following main result, where we exchange uncertainty in the degree of p for information about its zero root.

Theorem 1. *Let $p = \sum_{i=0}^{n} c_i x^{n-i} \in K[x]$ be a monic polynomial with known degree n, $k+1$ distinct roots, $r_0 = 0, r_1, r_2, \ldots, r_k$, with multiplicity $m_0, m_1, m_2, \ldots m_k$, respectively. Then, the multiplicities are uniquely determined by $c_0 = 1, c_1, \ldots, c_k$.*

 © 2019 John Wiley & Sons, Ltd.

Furthermore, p may be determined by fewer than k proper coefficients when K is not algebraically closed.

Theorem 2. *Let $p = \sum_{i=0}^{n} c_i x^{n-i} \in K[x]$ be a monic polynomial such that $p(0) \neq 0$. Suppose $p = \prod_{i=1}^{t} q_i^{m_i}$ for $q_i \in K[x]$. The multiplicity vector $\mathbf{m} = \langle m_1, \ldots, m_t \rangle^T$ is uniquely determined by the t proper leading coefficients if and only if $V \in \bar{K}^{t \times t}$ is nonsingular, where*

$$V_{i,j} = \sum_{r : q_j(r)=0} r^i.$$

Remark 1. Observe that, when $q_i = x - r_i$ (i.e., p splits over K), Theorem 2 provides the same conclusion as Theorem 1.

In Section 2, we prove both of the above results. In particular, we prove Theorem 1 via an algorithm that allows us to compute exactly the multiplicity of each root. In Section 3, we show that this algorithm, when performed over $\mathbb{C}$, can be made numerically stable in the sense that the requisite number of bits of precision to approximate each root in order to compute its multiplicity *exactly* is linear in k and the logarithms of (a) the ratio between the largest and smallest difference of roots, (b) the largest root, and (c) the largest coefficient of codegree at most k. We conclude by demonstrating the utility of this algorithm by obtaining previously unknown characteristic polynomials of two 3-uniform hypergraphs in Section 4. This problem from spectral hypergraph theory, which served as the original motivation for the present work, presents an unusual setting for which the tools introduced below enable heretofore impossible calculations: complete factorization of very high degree polynomials with few (known) roots, whose first several coefficients can be obtained exactly. The examples in Section 4 illustrate the regime in which our algorithm is most useful, that is, the degree n of p is large, and the number k of distinct roots is small in comparison. For such questions, Theorems 1 and 2 ensure that it suffices to check agreement among only a small number of leading coefficients upon multiplying out the factorization obtained in order to certify its correctness.

2 | DETERMINING THE MULTIPLICITIES OF KNOWN ROOTS

We begin with a proof of Theorem 1.

Proof of Theorem 1. Fix such a monic polynomial p with distinct roots $r_0 = 0, r_1, \ldots, r_k$ with respective multiplicities $m_0, m_1, \ldots, m_k$. Ignoring r_0 for a moment, let $\mathbf{r} = \langle r_1, \ldots, r_k \rangle^T$ and $\mathbf{m} = \langle m_1, \ldots, m_k \rangle^T$. We denote the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ r_1 & r_2 & \ldots & r_k \\ r_1^2 & r_2^2 & \ldots & r_k^2 \\ \vdots & \vdots & \vdots & \vdots \\ r_1^{k-1} & r_2^{k-1} & \ldots & r_k^{k-1} \end{pmatrix}$$

and consider

$$V_0 = \begin{pmatrix} r_1 & r_2 & \ldots & r_k \\ r_1^2 & r_2^2 & \ldots & r_k^2 \\ \vdots & \vdots & \vdots & \vdots \\ r_1^k & r_2^k & \ldots & r_n^k \end{pmatrix}.$$

Let $\mathbf{p} \in \overline{K}^n$, where

$$\mathbf{p}_i := \sum_{j=1}^{k} r_j^i m_j,$$

then,

$$V_0 \mathbf{m} = \mathbf{p}.$$

Notice that $V_0 = V \mathrm{diag}(\mathbf{r})$ is nonsingular as it is the product of two nonsingular matrices. We have then

$$\mathbf{m} = V_0^{-1} \mathbf{p}. \tag{1}$$

We present a formula for $\mathbf{p}$, which is a function of only the leading $k+1$ coefficients of p.

Let A be the diagonal matrix where r_i occurs m_i times, and note

$$p(x) = \det(xI - A).$$

By the Faddeev–LeVerrier algorithm (a.k.a. the method of Gantmacher[2]), we have for $j \geq 1$

$$c_j = -\frac{1}{j}\sum_{i=1}^{j} c_{j-i}\operatorname{tr}(A^i) = -\frac{1}{j}\sum_{i=1}^{j} c_{j-i}\mathbf{p}_i. \tag{2}$$

Let $\mathbf{c} = \langle c_1, c_2, \ldots, c_k\rangle^T$, $\Lambda = -\operatorname{diag}(1, 1/2, \ldots, 1/k)$, and

$$C = \begin{pmatrix} c_0 = 1 & 0 & \ldots & 0 \\ c_1 & c_0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{k-1} & c_{k-2} & \ldots & c_0 \end{pmatrix}.$$

By Equation (2), $\mathbf{c} = \Lambda C\mathbf{p}$. Moreover, as Λ and C are invertible, we have $\mathbf{p} = (\Lambda C)^{-1}\mathbf{c}$. It follows from Equation (1) that

$$\mathbf{m} = (\Lambda C V_0)^{-1}\mathbf{c}. \tag{3}$$

Furthermore, $m_0 = n - \mathbf{1} \cdot \mathbf{m}$. □

We briefly remark about the proof of Theorem 1. Problem 1 has a flavor of polynomial interpolation: given k points, how many (univariate) polynomials of degree n go through each of the k points? If $n \leq k - 1$, the polynomial is known to be unique and is relatively expensive to compute (as any standard text in numerical analysis will attest). Our proof technique mimics this approach as the classical problem of determining $p = \sum_{i=0}^{k-1} c_i x^{(k-1)-i}$, which resembles k distinct points $\{(x_i, y_i)\}_{i=1}^{k}$, can be solved by computing

$$V^T\mathbf{c} = \mathbf{y},$$

where $\mathbf{c} = \langle c_{k-1}, c_{k-2}, \ldots, c_0\rangle^T$, V is as previously defined given $r_i = x_i$, and $\mathbf{y} = \langle y_1, \ldots, y_k\rangle^T$. Suppose for a moment that each root is distinct so that

$$p(x) = \prod_{i=1}^{k}(x - r_i).$$

Then, c_j, the codegree j coefficient, is precisely the jth elementary symmetric polynomial in the variables $r_1, \ldots, r_k$. In the case of repeated roots, we have that c_j can be expressed using modified symmetric polynomials in the distinct roots, where r_i is replaced with $\binom{m_i}{j} r_i^j$. The expression for each coefficient via these modified symmetric polynomials is given by Equation (2).

Note that if the roots of p are known, it is possible to determine p with fewer coefficients than the number of distinct roots (e.g., when p is a nonlinear minimal polynomial). We modify Theorem 1 to include the case when some of the roots are known to occur with the same multiplicity. We now prove Theorem 2.

Proof of Theorem 2. The proof follows similarly to that of Theorem 1. First suppose that V is nonsingular. Let $\mathbf{m} = \langle m_1, \ldots, m_t\rangle^T$ and

$$\mathbf{p}_i = \sum_{j=1}^{t}\left(\sum_{r : q_j(r)=0} r^i\right) m_j = (V\mathbf{m})_i$$

so that if V is nonsingular, $\mathbf{m} = V^{-1}\mathbf{p}$. Let A be defined as in Theorem 1: The diagonal matrix where the roots of q_i occur m_i times and note

$$p(x) = \det(xI - A).$$

We have by the Faddeev–LeVerrier algorithm, for $j \geq 1$,

$$c_j = -\frac{1}{j}\sum_{i=1}^{j} c_{j-i}\mathrm{tr}(A^i) = -\frac{1}{j}\sum_{i=1}^{j} c_{j-i}\mathbf{p}_i$$

so that, for $\mathbf{c} = \langle c_1, \ldots, c_t\rangle^T$, $\Lambda = -\mathrm{diag}(1, 1/2, \ldots, 1/t)$, we have

$$\mathbf{m} = (\Lambda CV)^{-1}\mathbf{c}.$$

If instead $\mathbf{m}$ is uniquely determined by the first t proper coefficients, then $V\mathbf{m} = (\Lambda C)^{-1}\mathbf{c}$ has exactly one solution; hence, V is nonsingular. □

As a nonexample, consider the minimal polynomial of $\alpha = \sqrt{2}$, $q_\alpha(x) = x^2 - 2 \in \mathbb{Z}[x]$ and suppose

$$p = q_\alpha^d = x^{2d} + 0x^{2d-1} - 2dx^{2d-2} + \cdots \in \mathbb{Z}[x].$$

Observe that we cannot determine d given $c_1 = 0$; moreover, this conclusion is unsurprising, given the hypotheses from Theorem 2, because

$$V = \left[\sqrt{2} - \sqrt{2}\right] = [0] \in \mathbb{C}^{1,1}$$

is singular. However, by inspection we could determine d given $c_2 = -2d$, and in fact, matrix $[\sqrt{2}^2 + (-\sqrt{2})^2] = [4] \in \mathbb{C}^{1\times 1}$ is nonsingular. Indeed, we could determine d with a simple change of variable: Apply Theorem 2 to $p_0 = (y-2)^d$, where $y = x^2$.

3 | THE STABILITY OF COMPUTING MULTIPLICITIES

We now consider the feasibility of computing $\mathbf{m}$ over $\mathbb{C}$. In general, matrix V_0 in Theorem 1 may be poorly conditioned, so this calculation is often difficult to carry out even for modest values of k. The goal of this section is to show that if each root of a monic polynomial $p(x) \in \mathbb{Z}[x]$ is approximated by a disk of radius at most ϵ, a "reasonable" precision, then the interval approximating $\mathbf{m}_i$, resulting from a particular algorithm, contains exactly one integer. That is, we provide an algorithm for exactly computing $\mathbf{m}$ via SageMath[3] with substantially improved numerical stability over simply following the calculations in Section 1.

Theorem 3. *Let $p(x) = \sum_{i=0}^{t} c_i x^{t-i} \in \mathbb{Z}[x]$ be a monic polynomial with distinct nonzero roots $r_1, \ldots, r_n$ such that $|r_1| \geq |r_2| \geq \ldots \geq |r_n| > 0$. If each root is approximated by a disk of radius ϵ such that*

$$\epsilon < \frac{m^2 r}{2^{2n+7}n^5}\left(\frac{m}{MRc}\right)^n = \left(\frac{m}{MRc}\right)^{n(1+o(1))},$$

where

- $M = \max\{\max\{|r_i - r_j|\}, 1\}$ *and* $m = \min\{\min\{|r_i - r_j| : i \neq j\}, 1\}$;
- $R = \max\{|r_1|, 1\}$ *and* $r = \min\{|r_n|, 1\}$;
- $c = \max\{\max\{|c_i|\}_{i=1}^{n}, 1\}$,

then the resulting disk approximating $\mathbf{m}_i = (\Lambda CV_0)^{-1}\mathbf{c}_i \in \mathbb{Z}$ contains exactly one integer (i.e., the computation of $\mathbf{m}$ is exact).

Notice that $MRc \geq 1$ and $MRc = 2$ for $x^n - 1$ when n is even. Roots of unity occur frequently in the spectrum of hypergraphs; see Section 3. In particular, k-cylinders—essentially k-colorable k-graphs—have a spectrum that is invariant under multiplication by the kth roots of unity. Consider now $p(x) = x^n - 1$. We have $m = \sqrt{2 - 2\cos(\frac{2\pi}{n})}$ so that

$$\epsilon < \frac{\left(2 - 2\cos\left(\frac{2\pi}{n}\right)\right)^{\frac{n+2}{2}}}{2^{3n+7}n^5}.$$

While ϵ may seem small, we are chiefly concerned with the number of bits of precision needed to approximate each root. Indeed for $x^n - 1$, we need $|\ln \epsilon| = O(n \ln n)$ bits of precision by the small-angle approximation.

Remark 2. The bound on ϵ is "reasonable," as the number of bits required to approximate each root is proportional to the number of distinct roots of p and the logarithms of the ratio of the smallest difference of the roots with the largest difference of roots, the largest root, and the largest coefficient.

For the algorithm we describe below, the computational complexity of computing $\mathbf{m}$ is quasi-linear in nb, where $b = O(\ln \epsilon)$ is the requisite number of bits of precision given. More precisely, the complexity is the maximum of $O(nb \ln nb)$ and $O(n^\delta)$, where δ is the best known exponent of matrix multiplication, currently[4] at $\delta \approx 2.373$ but widely believed to be $2 + o(1)$. As we will see, $\mathbf{m}$ can be written as a product of $n \times n$ matrices and length n vectors populated by entries with $O(nb)$ digits. The complexity follows as the number of digits needed to approximate each entry in the product of two such matrices/vectors is $O(nb)$.

In practice, the difficulty of computing $\mathbf{m}$ as described in Theorem 1 is in computing the inverse of the Vandermonde matrix, whose entries may vary widely in magnitude and which may be very poorly conditioned. The task of inverting Vandermonde matrices has been studied extensively. Eisenberg et al.[5] provide a brief history of the topic and results concerning the accuracy and effectiveness of several known algorithms. However, these algorithms provide good approximations for the entries of V^{-1}, whereas we seek to express them exactly as elements of the field of algebraic complex numbers, because $\mathbf{m}$ is a vector of integers. Soto-Eguibar et al.[6] showed that $V^{-1} = \Delta W L$, where Δ is the diagonal matrix

$$\Delta_{i,j} = \begin{cases} \prod_{k=1,k\neq i}^{n} \frac{1}{r_i - r_k} & : i = j \\ 0 & : i \neq j, \end{cases}$$

W is the lower triangular matrix

$$W_{i,j} = \begin{cases} 0 & : i > j \\ \prod_{k=j+1,k\neq i}^{n} (r_i - r_k) & : \text{ otherwise,} \end{cases}$$

and L is the upper triangular matrix

$$L_{i,j} = \begin{cases} 0 & : i < j \\ 1 & : i = j \\ L_{i-1,j-1} - L_{i-1,j} r_{i-1} & : i \in [2,n], j \in [2, i-1]. \end{cases}$$

Using this decomposition, it is possible to compute $\mathbf{m}$ exactly. To prove Theorem 3, we first provide an upper bound for the diameter of the disk approximating an entry of Δ, W, and L, respectively; to do so, we extensively employ the computations of Petković et al.,[7] found in Chapter 1.3. We present the necessary background here.

Let $D(z, \epsilon)$ be the open disk in the complex plane centered at z of radius ϵ. For $A = D(a, r_1)$, $B = D(b, r_2)$ complex open disks, we have the following.

1. $A \pm B = D(a \pm b, r_1 + r_2)$
2. $1/B = D\left(\frac{\bar{b}}{|b|^2 - r_2^2}, \frac{r_2}{|b|^2 - r_2^2}\right)$
3. $AB = D(ab, |a|r_2 + |b|r_2 + r_1 r_2)$

In particular, for the special case of A^n, we have

$$D(a, r_1)^n = D(a^n, (|a| - r_1)^n - |a|^n). \tag{4}$$

Moreover, given $0 < r_1 < 1 \leq |a|$,

$$(|a| - r_1)^n - |a|^n \leq r_1(2|a|)^n \tag{5}$$

because

$$(|a| - r_1)^n - |a|^n \leq \sum_{k=1}^{n} \binom{n}{k} r_1^k |a|^{n-k} \leq r_1(2|a|)^n.$$

Finally, let $d(A) = 2r_1$ denote the diameter of A and let

$$|A| = |a| + r_1$$

be the absolute value of A. Then, for $u \in \mathbb{C}$ we have the following.

1. $d(A \pm B) = d(A) + d(B)$
2. $d(uA) = |u| d(A)$
3. $d(AB) \leq |B| d(A) + |A| d(B)$

For the remainder of this paper, some numbers will be exact (e.g., rational numbers), whereas others will be approximated by a disk. The nonexact entries of a matrix $M \in \mathbb{C}^{n \times n}$ will be referred to as disks; this will be clear from the problem formulation or derived from the computations. With a slight abuse of notation, we use $d(M_{i,j})$ and $|M_{i,j}|$ to denote the diameter and absolute value of the disk approximating the entry $M_{i,j}$. Moreover, we write

$$d(M) = \max\{d(M_{i,j}) : i, j \in [n]\} \text{and} |M| = \max\{|M_{i,j}| : i, j \in [n]\}.$$

In the case when the entry is exact, the diameter is zero and the absolute value (of the disk) is simply the modulus.

Theorem 4. *Assume the notations of Theorem 2, let V denote the Vandermonde matrix from the proof of Theorem 1, and let $V^{-1} = \Delta W L$ by Soto-Eguibar et al.[6] Then,*

$$d(V^{-1}) \leq \frac{2^{2n+4} n}{m^2} \left(\frac{MR}{m} \right)^n \epsilon$$

and

$$|V^{-1}| \leq 2n \left(\frac{RM}{m} \right)^n.$$

Proof. Let

$$D_i := D(r_i, \epsilon)$$

denote the disk centered at r_i with radius ϵ. By Equation (5), we have for $s \neq t$

$$\begin{aligned} d(\Delta) &\leq d\left(\left(\frac{1}{D_s - D_t} \right)^n \right) \\ &\leq 2^n \left(\frac{2\epsilon}{m^2 - (2\epsilon)^2} \right) \left(\frac{m}{m^2 - (2\epsilon)^2} \right)^n \\ &\leq \frac{2^{2n+2}}{m^{n+2}} \cdot \epsilon, \end{aligned}$$

because $\epsilon < m/4$,

$$d(W) \leq d((D_s - D_t)^n) \leq 2^{n+1} M^n \epsilon,$$

and

$$d(L) \leq d(D_s^n) \leq (2R)^n \epsilon.$$

We first consider $d(\Delta W)$. Observe that ΔW is upper triangular and each nonzero entry of ΔW is a product of exactly one nonzero entry of Δ and W. In this way,

$$d((\Delta W)_{i,j}) \leq |W_{i,j}| d(\Delta_{i,i}) + |\Delta_{i,i}| d(W_{i,j}) \leq \frac{2^{2n+3}}{m^2} \left(\frac{M}{m} \right)^n \epsilon$$

and

$$|\Delta W| \le 2\left(\frac{M}{m}\right)^n.$$

We now determine $d(\Delta WL)$ by first computing

$$d((\Delta W)_{i,k}L_{k,j}) \le |L_{k,j}|d(\Delta W_{i,k}) + |\Delta W_{i,k}|d(L_{k,j}) \le \frac{2^{2n+4}}{m^2}\left(\frac{RM}{m}\right)^n \epsilon.$$

Hence,

$$d(V^{-1}) = d(\Delta WL) \le \max_{i,j} \sum_{k=1}^{n} d((\Delta W)_{i,k}L_{k,j}) \le \frac{2^{2n+4}n}{m^2}\left(\frac{RM}{m}\right)^n \epsilon$$

and

$$|V^{-1}| \le 2n\left(\frac{RM}{m}\right)^n.$$

□

In our computations, we are concerned with $V_0 = V \cdot \text{diag}(\mathbf{r})$, where $\text{diag}(\mathbf{r}) = \text{diag}(r_1, \ldots, r_n)$, so that

$$V_0^{-1} = \text{diag}(\mathbf{r})^{-1}V^{-1}.$$

The following corollary is immediate from the observation that

$$d(\text{diag}(\mathbf{r})^{-1}) \le \frac{2}{r}.$$

Corollary 1.

$$d(V_0^{-1}) \le \frac{2^{2n+6}n}{m^2 r}\left(\frac{MR}{m}\right)^n \epsilon$$

and

$$|V_0^{-1}| \le \frac{2n}{r}\left(\frac{RM}{m}\right)^n.$$

We are now able to prove Theorem 3.

Proof of Theorem 3. Recall $\mathbf{m} = V_0^{-1}C^{-1}\Lambda^{-1}\mathbf{c}$ as defined in the proof of Theorem 1. Fortunately, the remainder of the computations are straightforward as C^{-1}, Λ^{-1}, and $\mathbf{c}$ have integer, and thus exact, entries. As

$$C_{i,j}^{-1} = \begin{cases} 0 & : i < j \\ 1 & : i = j \\ -\sum_{k=1}^{i-1} c_{i-k}C_{k,j}^{-1} & : i > j, \end{cases}$$

we have

$$d\left(V_0^{-1}C^{-1}\right) \le n(nc^{n-1})\frac{2^{2n+6}n}{m^2 r}\left(\frac{MR}{m}\right)^n \epsilon = \frac{2^{2n+6}n^3}{m^2 rc}\left(\frac{MRc}{m}\right)^n \epsilon.$$

Furthermore, because $\Lambda^{-1} = -\text{diag}(1, 2, \ldots, n)$, we have

$$d\left(V_0^{-1}C^{-1}\Lambda^{-1}\right) \le |-n| d\left(V_0^{-1}C^{-1}\right) = \frac{2^{2n+6}n^4}{m^2 rc}\left(\frac{MRc}{m}\right)^n \epsilon$$

and, finally,

$$\begin{aligned} d\left(\left(V_0^{-1}C^{-1}\Lambda^{-1}\right)\mathbf{c}\right) &\le nc \cdot d\left(V_0^{-1}C^{-1}\Lambda^{-1}\right) \\ &\le \frac{2^{2n+6}n^5}{m^2 r}\left(\frac{MRc}{m}\right)^n \epsilon < \frac{1}{2}. \end{aligned}$$

Thus, each interval will contain at most one integer as desired. □

4 | APPLICATION TO HYPERGRAPH SPECTRA

For the present authors, Problem 2 arose organically in the context of spectral hypergraph theory. In short, the authors were concerned with determining high-degree polynomials when the roots (without multiplicity) are known and all but the lowest-codegree coefficients are too costly to compute. We briefly explain the context of spectral hypergraph theory for those interested in the origin of such questions. However, our presentation of the computations is self-contained; the reader who wishes to see Theorem 1 applied immediately may skip the next few paragraphs.

For $k \geq 2$, a *k-uniform hypergraph* is a pair $\mathcal{H} = (V, E)$, where $V = [n]$ is the set of *vertices* and $E \subseteq \binom{[n]}{k}$ is the set of *edges*. It is common to refer to such hypergraphs as *k-graphs* when $k > 2$ and as just *graphs* when $k = 2$. We are particularly interested in the computation of the characteristic polynomial of a uniform hypergraph. The characteristic polynomial of the adjacency matrix of a graph is straightforward to compute; however, the same cannot be said for hypergraphs. The *characteristic polynomial* of the *(normalized) adjacency hypermatrix* $\mathcal{A}$ of $\mathcal{H}$, denoted by $\phi_{\mathcal{H}}(\lambda)$, is the resultant of a family of $|E|$ homogeneous polynomials† of degree $k-1$ in the indeterminate λ; *order k*, *dimension n*, hypermatrix $\mathcal{A} \in \mathbb{R}^{[n]^k}$, whose rows and columns are indexed by the vertices of $\mathcal{H}$ and whose $(v_1, \ldots, v_k)$ entry is $1/(k-1)!$ times the indicator of the event that $\{v_1, \ldots, v_k\}$ is an edge of $\mathcal{H}$ is also sometimes called the *adjacency tensor* of $\mathcal{H}$. Equivalently, one can define the characteristic polynomial to be the hyperdeterminant of $\mathcal{A} - \lambda\mathcal{I}$ (as in Gelfand et al.[8]), where $\mathcal{I}$ is the identity hypermatrix, that is, $\mathcal{I}(v_1, \ldots, v_k)$ is the indicator of the event that $v_1 = \ldots = v_k$. The set $\sigma(\mathcal{H}) = \{r : \phi_{\mathcal{H}}(r) = 0\} \subset \mathbb{C}$ is the *spectrum* of $\mathcal{H}$ and each $r \in \sigma(\mathcal{H})$ is an *eigenvalue* of $\mathcal{H}$. It is known that $\phi_{\mathcal{H}}(\lambda)$ is a monic polynomial of degree $n(k-1)^{n-1}$, and many of the properties of characteristic polynomials of graphs generalize nicely to hypergraphs; we refer the interested reader to a work of Cooper et al.[9] and another by Qi[10] for further exploration of the topic.

Given a k-graph $\mathcal{H}$, we aim to compute $\phi_{\mathcal{H}}(\lambda)$. Unfortunately, the resultant is known to be NP-hard to compute[11] despite its utility in several fields of mathematics, perhaps nowhere more so than computational algebraic geometry. Performing this computation directly using built-in routines for computing resultants has proven to be intractable even for 3-uniform hypergraphs with few vertices. Nonetheless, one can attempt to imitate classical approaches to computing characteristic polynomials of ordinary graphs. A notable example of this is the work of Harary[12] (and Sachs[13]) who showed that the coefficients of $\phi_G(\lambda)$ can be expressed as a certain weighted sum of the counts of subgraphs of G. Clark et al.[14] have established an analogous result for the coefficients of $\phi_{\mathcal{H}}(\lambda)$. This formula allows one to compute many low-codegree coefficients—that is, the coefficients of x^{d-k} for k small and $d = \deg(\phi_{\mathcal{H}})$—by a certain linear combination of subgraph counts in $\mathcal{H}$. Unfortunately, this computation becomes exponentially harder as the codegree increases, making computation of the entire (often extremely high degree) characteristic polynomial impossible for all but the simplest cases. A method of Lu et al.,[15] "α-normal labelings," is an alternative approach that can obtain all eigenvalues with relative efficiency, but it gives no information about their multiplicities. *Combining* these two techniques, however, yields a method to obtain the full characteristic polynomial: obtain a list of roots, compute a few low-codegree coefficients using subgraph counts, and then deduce the roots' multiplicities. Therefore, we arrive at the following special case of Problem 2.

Problem 3. Let K be a field of characteristic zero. Is it true that a monic polynomial $p \in K[x]$ of degree n with exactly k distinct, known roots is determined by its k proper leading coefficients?

Returning to our application of Theorem 1, we can compute $\phi_{\mathcal{H}}(\lambda)$ if we know $\sigma(\mathcal{H})$ and the first $|\sigma(\mathcal{H})|$ coefficients (note this includes coefficients that are zero, as well as the leading term). Lu et al.[15] introduced *α-consistent incidence matrices* that can be used to find the eigenvalues of $\mathcal{H}$ whose corresponding eigenvector has all nonzero entries. These eigenvalues are referred to as *totally nonzero eigenvalues* and we denote the set of totally nonzero eigenvalues of $\mathcal{H}$ as $\sigma^+(\mathcal{H})$. Clark et al.[16] showed that for $k > 2$,

$$\sigma(\mathcal{H}) \subseteq \bigcup_{H \subseteq \mathcal{H}} \sigma^+(H),$$

where $H = (V_0, E_0) \subseteq \mathcal{H}$ if $V_0 \subseteq V$ and $E_0 \subseteq E$ (cf. Cauchy interlacing theorem when $k = 2$). Computing $\sigma(\mathcal{H})$ by way of $\sigma^+(H)$ involves solving smaller multilinear systems than the one involved in computing $\phi_{\mathcal{H}}(\lambda)$. Generally speaking, $|\sigma(\mathcal{H})|$ is considerably smaller than the degree of $\phi_{\mathcal{H}}(\lambda)$. In practice, this approach has yielded $\phi_{\mathcal{H}}(\lambda)$ when other approaches of computing $\phi_{\mathcal{H}}(\lambda)$ via the resultant have failed.

†Namely, the Lagrangians of the links of all vertices minus λ times the $(k-1)$st power of the corresponding vertices' variables, or equivalently, the coordinates f_v of the gradient of the k-form naturally associated with $\mathcal{A} - \lambda\mathcal{I}$. The (symmetric) hyperdeterminant is the unique irreducible monic polynomial in the entries of $\mathcal{A}$ whose vanishing corresponds exactly to the existence of nontrivial solutions to the system $\{f_v = 0\}_{v \in V(\mathcal{H})}$.

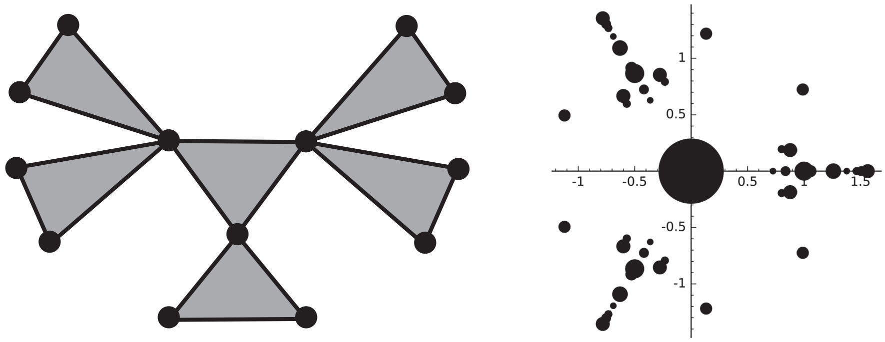

FIGURE 1 The hummingbird hypergraph and its spectrum

We present two examples demonstrating these computations, with all calculations performed in floating-point arithmetic whose precision is determined by the results above so as to ensure correct, exact integer answers. Consider the *hummingbird hypergraph* $\mathcal{B} = (\{1, 2, \ldots, 13\}, E)$, where

$$E = \{\{1,2,3\},\{1,4,5\},\{1,6,7\},\{2,8,9\},\{3,10,11\},\{3,12,13\}\}.$$

We present a drawing of $\mathcal{B}$ in Figure 1, where the edges are drawn as shaded in triangles. Note that

$$\deg(\phi_{\mathcal{B}}) = n(k-1)^{n-1} = 13 \cdot 2^{12} = 53248$$

and, because $\mathcal{B}$ is a hypertree (and thus a 3-cylinder), its spectrum is invariant under multiplication by any third root of unity.[9] We compute the minimal polynomials of the totally nonzero eigenvalues of $\phi_{\mathcal{B}}$ via Macaulay2,[17]

$$\begin{aligned} \phi_{\mathcal{B}} = {} & x^{m_0}(x^9 - 6x^6 + 8x^3 - 4)^{m_1}(x^9 - 5x^6 + 5x^3 - 2)^{m_2} \\ & \cdot (x^3 - 1)^{m_3}(x^6 - 4x^3 + 2)^{m_4}(x^9 - 4x^6 + 3x^3 - 1)^{m_5} \\ & \cdot (x^6 - 3x^3 + 1)^{m_6}(x^3 - 3)^{m_7}(x^3 - 2)^{m_8}. \end{aligned}$$

With the intent of applying Theorem 1 to $\phi_{\mathcal{B}}$, we consider the change of variable $y = x^3$ and observe that we need to determine $c_3, c_6, \ldots, c_{48}$ as there are sixteen distinct nonzero cube roots. We compute

$$\begin{aligned}
c_3 &= -18432 \\
c_6 &= 169843968 \\
c_9 &= -1043209971456 \\
c_{12} &= 4804960103034624 \\
c_{15} &= -17702435302276375440 \\
c_{18} &= 54341319772238850901668 \\
c_{21} &= -142960393819753656566577552 \\
c_{24} &= 329036832924106136747171871042 \\
c_{27} &= -673063350744784559041302787109576 \\
c_{30} &= 1238925078774563882036470496247467682 \\
c_{33} &= -2072891735870949695930286542580991559916 \\
c_{36} &= 3178738418917825954994865036362341584776658 \\
c_{39} &= -4498896549573513724009044022281523093964642496 \\
c_{42} &= 5911636016042739623328802656744094043553245557890 \\
c_{45} &= -7249053168113446546908444934275696322928768819713512 \\
c_{48} &= 8332230937213678426258491158832963153453272812465162851.
\end{aligned}$$

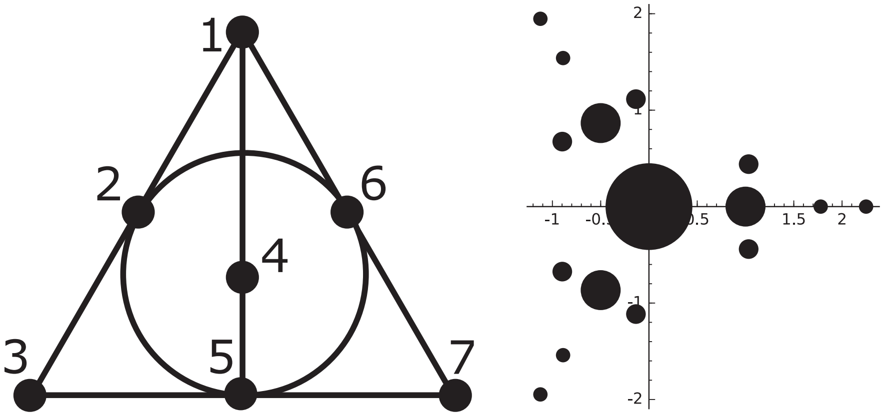

FIGURE 2 The Rowling hypergraph and its spectrum

Using Theorem 3, we have $M < 3, m > .14, R < 4.39, r > .38$, and $c = c_{48}$ so that each root of $\phi_{\mathcal{B}}$ needs to be approximated to at most 3,091 bits of precision. Using SageMath,[3] we obtain

$$\begin{aligned}\phi_{\mathcal{B}} = x^{20983}&(x^9 - 6x^6 + 8x^3 - 4)^{729}(x^9 - 5x^6 + 5x^3 - 2)^{972}\\ &\cdot (x^3 - 1)^{1782}(x^6 - 4x^3 + 2)^{486}(x^9 - 4x^6 + 3x^3 - 1)^{324}\\ &\cdot (x^6 - 3x^3 + 1)^{216}(x^3 - 3)^{54}(x^3 - 2)^{119}.\end{aligned}$$

In Figure 1, we provide a plot of $\sigma(\phi_{\mathcal{B}})$ drawn in the complex plane where a disk is centered at each root, and each disk's area is proportional to the algebraic multiplicity of the underlying root in $\phi_{\mathcal{B}}$.

Now, consider the *Rowling hypergraph*‡

$$\mathcal{R} = ([7], \{\{1,2,3\}, \{1,4,5\}, \{1,6,7\}, \{2,5,6\}, \{3,5,7\}\}).$$

A drawing of $\mathcal{R}$ is given in Figure 2 where the edges are drawn as arcs and its spectrum is drawn similarly to that of $\phi_{\mathcal{B}}$; note that $\mathcal{R}$ is also the Fano plane minus two edges. We have

$$\deg(\phi_{\mathcal{R}}) = n(k-1)^{n-1} = 7 \cdot 2^6 = 448.$$

It is easy to verify that $\mathcal{R}$ is not a 3-cylinder; however, its spectrum, like that of 3-cylinders,[9] is invariant under multiplication by any third root of unity (see Lemma 3.11 in the work of Fan et al.[19]). By the work of Lu et al.,[15] we have

$$\begin{aligned}\phi_{\mathcal{R}} =& x^{m_0}(x^3 - 1)^{m_1}(x^{15} - 13x^{12} + 65x^9 - 147x^6 + 157x^3 - 64)^{m_2}\\ &\cdot (x^6 - x^3 + 2)^{m_3}(x^6 - 17x^3 + 64)^{m_4}.\end{aligned}$$

With the intent of applying Theorem 2 to the minimal polynomials of $x^{-m_0/3}\phi_{\mathcal{R}}(x^{1/3})$, we verify that

$$V = \begin{pmatrix} 1 & 13 & 1 & 17 \\ 1 & 39 & -3 & 161 \\ 1 & 103 & -5 & 1649 \\ 1 & 87 & 1 & 17729 \end{pmatrix}$$

‡The name was chosen for its resemblance to an important narrative device.[18]

is nonsingular. Indeed, we need to determine the first four proper leading coefficients of $x^{-m_0/3}\phi_{\mathcal{R}}(x^{1/3})$ or, equivalently, c_3, c_6, c_9, c_{12} of $\phi_{\mathcal{R}}$. We have

$$
\begin{aligned}
c_3 &= -240\\
c_6 &= 28320\\
c_9 &= -2190860\\
c_{12} &= 125012034.
\end{aligned}
$$

By Theorem 3, we have $M < 4.5, m > .69, R < 2.25$, and $r = 1$ so that at most 252 digits of precision are required to approximate each root. We compute

$$
\begin{aligned}
\phi_{\mathcal{R}} = {} & x^{133}(x^3-1)^{27}(x^{15}-13x^{12}+65x^9-147x^6+157x^3-64)^{12}\\
& \cdot (x^6-x^3+2)^6(x^6-17x^3+64)^3.
\end{aligned}
$$

ACKNOWLEDGEMENTS

Thanks to Alexander Duncan for the helpful discussions and insights, and to the anonymous referees for their rigorous attention to detail and expository clarity. This work does not have any conflicts of interest.

ORCID

Gregory J. Clark https://orcid.org/0000-0002-0960-9993

REFERENCES

1. Győry K, Hajdu L, Pintèr À, Schinzel A. Polynomials determined by a few of their coefficients. Indagationes Mathematicae. 2004;15(2):209–221.
2. Gantmacher F. The theory of matrices. Providence, RI: Chelsea Publishing; 1960.
3. SageMath. The sage mathematics software system (version 7.2). Newcastle upon Tyne, UK: The Sage Developers; 2018. Available from: http://www.sagemath.org
4. Le Gall F. Powers of tensors and fast matrix multiplication. arXiv:1401.7714.
5. Eisinberg A, Fedele G. On the inversion of the Vandermonde matrix. Appl Math Comput. 2006;174(2):1384–1397.
6. Soto-Eguibar F, Moya-Cessa H. Inverse of the Vandermonde and Vandermonde confluent matrices. Appl Math Inf Sci. 2011;5(3):361–366.
7. Petković M, Petković L. Complex interval arithmetic and its applications. Berlin, Germany: Wiley-VCH; 1998.
8. Gelfand I, Zelevinskii A, Kapranov M. Discriminants, resultants, and multidimensional determinants. Boston, MA: Birkhäuser; 1994.
9. Cooper J, Dutle A. Spectra of uniform hypergraphs. Linear Algebra Appl. 2012;436:3268–3292.
10. Qi L. Eigenvalues of a real supersymmetric tensor. J Symbolic Comput. 2005;40:1302–1324.
11. Grenet B, Koiran P, Portier N. The multivariate resultant is NP-hard in any characteristic. In: Mathematical foundations of computer science 2010. Berlin, Germany: Springer; 2010. p. 477–488. (Lecture notes in computer science; No. 6281).
12. Harary F. The determinant of the adjacency matrix of a graph. SIAM Review. 1962;4:202–210.
13. Sachs H. Über teiler, faktoren und charakteristische polynome von Graphen. Teil. 1966;I(12):7–12.
14. Clark G, Cooper G. A combinatorial description for coefficients of the characteristic polynomial of a hypergraph. In preparation.
15. Lu L, Man S. Connected hypergraphs with small spectral radius. Linear Algebra Appl. 2016;509:206–227.
16. Clark G, Cooper J. On the adjacency spectra of hypertrees. Electron J Comb. To appear.
17. Grayson D, Stillman M. Macaulay2: a software system for research in algebraic geometry. Available from: https://faculty.math.illinois.edu/Macaulay2/
18. Rowling JK. Harry Potter and the deathly hallows. New York, NY: Bloomsbury Publishing; 2007.
19. Fan Y-Z, Huang T, Bao Y-H, Zhuan-Sun C-L, Li Y-P. The spectral symmetry of weakly irreducible nonnegative tensors and connected hypergraphs. arXiv:1704.08799.

How to cite this article: Clark GJ, Cooper JN. Stably computing the multiplicity of known roots given leading coefficients. *Numer Linear Algebra Appl*. 2020;27:e2275. https://doi.org/10.1002/nla.2275