UC San Diego UC San Diego Previously Published Works

Title

Homotopy Techniques for Analytic Combinatorics in Several Variables

Permalink

https://escholarship.org/uc/item/5z93j1z6

Authors

Lee, Kisun Melczer, Stephen Smolčić, Josip

Publication Date

2022-01-15

DOI

10.1109/synasc57785.2022.00014

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

Homotopy Techniques for Analytic Combinatorics in Several Variables

Kisun Lee **Department** of Mathematics University of California San Diego La Jolla, CA, USA kil004@ucsd.edu

Stephen Melczer

Dept. of Combinatorics and Optimization Dept. of Combinatorics and Optimization University of Waterloo Waterloo, Canada smelczer@uwaterloo.ca

Josip Smolčić University of Waterloo Waterloo, Canada jsmolcic@uwaterloo.ca

Abstract-We combine tools from homotopy continuation solvers with the methods of analytic combinatorics in several variables to give the first practical algorithm and implementation for the asymptotics of multivariate rational generating functions not relying on a non-algorithmically checkable 'combinatorial' non-negativity assumption. Our homotopy implementation terminates on examples from the literature in three variables, and we additionally describe heuristic methods that terminate and correctly predict asymptotic behaviour in reasonable time on examples in even higher dimension. Our results are implemented in Julia, through the use of the HomotopyContinuation.jl package, and we provide a selection of examples and benchmarks.

Index Terms-asymptotics, analytic combinatorics in several variables, homotopy, polynomial system

Let $(f_n)_{n\in\mathbb{N}} = f_0, f_1, \ldots$ be a complex-valued sequence with generating function $F(z) = \sum_{n>0} f_n z^n$. Although F is a priori only a *formal* power series, in a wide variety of applications (in fact, whenever f_n has at most exponential growth) it represents an analytic function in a neighbourhood of the origin. The field of analytic combinatorics creates effective techniques to determine the asymptotic behaviour of f_n through a study of the analytic behaviour of F(z). Most classical methods in analytic combinatorics take as input an algebraic or differential equation satisfied by F(z) and, when successful, return the leading terms in an asymptotic expansion of f_n (see [10] or [17, Chapter 2]).

More recently, a theory of analytic combinatorics in several variables (ACSV) [17], [22] has been developed to translate the analytic behaviour of a *d*-variate generating function

$$F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} := \sum_{\mathbf{i} \in \mathbb{N}^d} f_{i_1,\dots,i_d} z_1^{i_1} \cdots z_d^{i_d}$$

into asymptotic information about its coefficient sequence $(f_i)_{i \in \mathbb{N}^d}$. In this paper we focus on the case of a power series expansion of a multivariate rational function $F(\mathbf{z}) =$ $G(\mathbf{z})/H(\mathbf{z})$ and attempt to determine asymptotics of the **r**-diagonal sequence $(f_{n\mathbf{r}})_{n\in\mathbb{N}}$ for a fixed direction vector $\mathbf{r} \in \mathbb{Z}_{>0}^{d}$. The most common situation to arise in practice is the main diagonal, when r = 1.

Remark 0.1: If r has some zero coordinates then we can reduce to the above situation by setting some of the variables equal to zero and working in a lower dimension. For instance, the $(0, r_2, r_3)$ -diagonal of any series F(x, y, z) is the (r_2, r_3) diagonal of F(0, y, z). Furthermore, our asymptotic statements continue to hold for directions $\mathbf{r} \in \mathbb{Q}_{>0}$ if they are interpreted to be valid only when $n\mathbf{r} \in \mathbb{N}^d$. In fact, the methods of ACSV show that asymptotics of the r-diagonal usually vary smoothly with r, allowing one to give a natural interpretation of asymptotics in irrational directions and derive central limit theorems [17, Section 5.3.3].

Remark 0.2: Because the methods of ACSV hold in any dimension, our requirement that F(z) be rational is less restrictive than it may seem. For instance, the r-diagonal of an algebraic function in d variables can be represented [17, Section 3.2.2] as the diagonal of a rational function in 2dvariables (and a 'skew-diagonal' of a rational function in d+1variables). The theoretical results discussed here also hold for *meromorphic* functions, when $F(\mathbf{z})$ is (locally) the ratio of analytic functions, however our restriction to rational functions allows us to stay in the realm of algebraic quantities and polynomial systems, which we use for our explicit algorithms.

There are many factors making ACSV more complicated than its univariate counterpart. Although a univariate rational function has a finite number of singularities, meaning one can determine the 'asymptotic contribution' of each and simply sum those with the fastest growth, any (non-polynomial) rational function in at least two variables must have an infinite number of singularities. In addition to obscuring which singularities contribute to asymptotics, this also means that the singular set can have non-trivial geometry, for instance by self-intersecting. The difficulties that arise mean that unlike the univariate case, which relies on standard complex-analytic results going back hundreds of years, the most advanced ACSV results rely on advanced techniques from areas of mathematics as diverse as complex analysis in several variables, the study of singular integrals, algebraic geometry, differential geometry, and topology.

The starting point of an ACSV analysis expresses the rdiagonal of $F(\mathbf{z})$ as a d-dimensional complex integral. In the simplest cases, asymptotic behaviour is determined by the behaviour of F near two types of points: *critical points*,

KL and SM acknowledge the support of the AMS Math Research Community Combinatorial Applications of Computational Geometry and Algebraic Topology, which was funded by the National Science Foundation under Grant Number DMS 1641020. SM and JS's work partially supported by NSERC Discovery Grant RGPIN-2021-02382.

defined by an explicit polynomial system, and *minimal points*, which are singularities that are coordinate-wise closest to the origin. Critical points satisfy a square polynomial system, and generically form a finite set that can be manipulated in a computer algebra system. In contrast, there are always an infinite number of minimal points, which are defined by inequalities involving the moduli of coordinates and are thus trickier and more expensive to manipulate in computations.

A. Previous Work and Our Contributions

From the beginning of its modern period in work of Pemantle and Wilson [21], the goal of ACSV has always been to develop methods explicit enough to be implemented in a computer algebra system. The 'surgery' approach of [21], which applies to generating functions with smooth singular sets that form manifolds, essentially computes a residue in one variable to obtain a (d-1)-dimensional integral that is approximated using the saddle-point method. Although this surgery method does not require much theory beyond univariate analytic combinatorics, it requires strong conditions on the locations of minimal points that can be computationally expensive to verify. Later techniques, using cones of hyperbolicity [3] and multivariate residue and homology computations [2], rely on more advanced theory but simplify the assumptions that need to be verified for the results to hold. In the simplest cases, which hold for the majority of examples encountered in combinatorial applications, it suffices to determine which of the critical points are minimal and then add explicit asymptotic contributions corresponding to the (finite number of) minimal critical points. The most expensive step in such an analysis is almost always checking minimality.

The first systematic algorithmic study of ACSV methods was conducted by Melczer and Salvy [18], who encoded critical points using a symbolic-numeric data structure known as a Kronecker or rational univariate representation and then reduced checking minimality to rigorously approximating the roots of certain univariate polynomials to sufficiently high accuracy. Those authors created a preliminary implementation of their work, which does not certify numeric computations to provide rigorous proofs and requires combinatorial rational functions, in the Maple computer algebra system. A rational function $F(\mathbf{z})$ is *combinatorial* if all of its power series coefficients are non-negative: although this condition is satisfied for any multivariate generating function, in many combinatorial examples only one diagonal of F enumerates a combinatorial class and the non-diagonal entries have negative coefficients. It is an open problem, even in the univariate case, whether it is decidable to detect when a rational function is combinatorial (see [20] for some open problems in this area). Although Melczer and Salvy [18] detail a method that, in principle, yields an algorithm for asymptotics that does not require combinatorality, in practice an implementation in Maple would not halt in reasonable time beyond low degree examples in two or three variables.

Instead of continuing with the Kronecker representation approach of Melczer and Salvy, in this paper we exploit homotopy continuation methods to certify minimality of critical points, and ultimately determine asymptotics of r-diagonals of rational functions. Using the HOMOTOPYCONTINUATION.JL Julia package [7] for polynomial system solving, we provide the first implementation of ACSV methods under assumptions that often hold in practice. Our implementation is efficient enough to work even without the assumption of combinatorality, although when the user knows a priori that their input rational function is combinatorial then the computation is greatly reduced. In addition, we describe two heuristic methods to classify minimal critical points using numerical approximations that are extremely efficient, and are the only implemented algorithms we currently know of that can aid in the search for minimal points in more than three variables.

Example 0.3: The main diagonal of the power series expansion of

$$F(x, y, z) = \frac{1}{1 - (1 + z)(x + y - xy)}$$

is related to a result of Apéry [1] on the irrationality measure of $\zeta(2)$. After importing our package we define the denominator polynomial in Julia using

$$\begin{array}{l} \text{(lpolyvar x y z)} \\ \text{H} = 1 - (1+z) * (x+y-x*y) \end{array}$$

If we know that this power series expansion is combinatorial, then we can get the (truncated for clarity) minimal critical point

and print out the leading asymptotic term of the diagonal with

It is not obvious from the definition that F is combinatorial. If we don't know our function is combinatorial then we can determine minimality by running find_min_crits(H), which returns the same point but requires approximately 15 minutes of computation. If we want to heuristically check for minimal critical points, but don't know that F is combinatorial and don't want to wait for the full algorithm, we can run the algorithms find_min_crits(H; approx_crit=true) or find_min_crits(H; monodromy=true), described below, which also find the correct point and finish in seconds.

Remark 0.4: Because we use numeric methods, asymptotic behaviour is returned with numeric approximations of constants. If the user wants to determine the algebraic quantities involved exactly, we recommend solving for the critical points (a relatively cheap operation) symbolically using another computer algebra system like Sage or Maple and then using the results of this package to filter out the minimal ones (the most expensive operation).

The rest of this paper proceeds as follows. Section I gives a quick recap of the methods of ACSV and the high-level problems that need to be decided to find asymptotics, with a description of numerical algebraic geometry methods for polynomial system solving given in Section II. Section III uses this background material to detail our ACSVHOMOTOPY.JL Julia package, while Section IV illustrates the package on a wide variety of combinatorial examples, including benchmarks between different algorithms. Although our algorithms always terminate, due to the nature of homotopy continuation methods they may not always provide a rigorous proof of asymptotics – Section V discusses this issue and describes situations in which the algorithms do give rigorous proofs. Finally, Section VI concludes with some extensions that we believe should be addressed next.

I. SMOOTH ACSV

From now on, $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ denotes a ratio of *d*-variate coprime polynomials $G, H \in \mathbb{Z}[\mathbf{z}]$ with power series expansion $F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}}$ converging around the origin, and $\mathbf{r} \in \mathbb{Z}_{>0}^d$ is a fixed direction vector.

Definition 1.1 (minimal critical points): A point $\mathbf{w} \in \mathbb{C}^d_*$ is a (simple) smooth critical point of F if $(\nabla H)(\mathbf{w}) \neq \mathbf{0}$ and

$$\begin{cases} H(\mathbf{w}) = 0\\ r_k z_1 H_{z_1}(\mathbf{w}) - r_1 z_k H_{z_k}(\mathbf{w}) = 0 \ (2 \le k \le d). \end{cases}$$
(1)

We call $\mathbf{w} \in \mathbb{C}^d_*$ a *minimal point* if $H(\mathbf{w}) = 0$ and there does not exist $\mathbf{y} \in \mathbb{C}^d$ such that $H(\mathbf{y}) = 0$ and $|y_j| < |w_j|$ for all $j = 1, \ldots, d$.

Remark 1.2: If $(\nabla H)(\mathbf{w}) = \mathbf{0}$ then (1) is trivially satisfied. If the gradient vanishes because H has a higher-order pole (for instance, if $H = P^2$ for some polynomial P) then our analysis of minimal critical points can be performed on the square-free part of H (the product of its irreducible factors) to obtain an asymptotic expansion of $f_{n\mathbf{r}}$ with minor modifications. On the other hand, if the gradient vanishes because the zero set of H self-intersects then more advanced techniques are required [17, Part III].

We will be able to determine asymptotics in the presence of smooth minimal critical points, assuming a nondegeneracy condition on the zero set of H.

Definition 1.3 (phase Hessian matrix): If w is a smooth critical point then the phase Hessian matrix \mathcal{H} at w is the $(d-1) \times (d-1)$ matrix defined by

$$\mathcal{H}_{i,j} = \begin{cases} V_i V_j + U_{i,j} - V_j U_{i,d} - V_i U_{j,d} + V_i V_j U_{d,d} & : i \neq j \\ \\ V_i + V_i^2 + U_{i,i} - 2V_i U_{i,d} + V_i^2 U_{d,d} & : i = j \end{cases}$$

where

$$U_{i,j} = rac{w_i w_j H_{z_i z_j}(\mathbf{w})}{w_d H_{z_d}(\mathbf{w})}$$
 and $V_i = rac{r_i}{r_d}$

Theorem 1.4 (Melczer [17, Theorem 5.1]): Suppose that the system of polynomial equations (1) admits a finite number of solutions, exactly one of which, $\mathbf{w} \in \mathbb{C}^d_*$, is minimal. Suppose further that $H_{z_d}(\mathbf{w}) \neq 0$, that the phase Hessian matrix \mathcal{H} at \mathbf{w} has non-zero determinant, and that $G(\mathbf{w}) \neq 0$. Then, as $n \to \infty$,

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}} n^{(1-d)/2} \frac{(2\pi r_d)^{(1-d)/2}}{\sqrt{\det(\mathcal{H})}} \frac{-G(\mathbf{w})}{w_d H_{z_d}(\mathbf{w})} \left(1 + O\left(\frac{1}{n}\right)\right)$$

When the zero set of H contains a finite number of points with the same coordinate-wise modulus as w, all of which satisfy the same conditions as w, then an asymptotic expansion of f_{nr} is obtained by summing the right hand side of this expansion at each point.

Remark 1.5: The condition that $G(\mathbf{w}) \neq 0$ means that the leading asymptotic term in Theorem 1.4 doesn't vanish. When $G(\mathbf{w}) = 0$ asymptotics can usually still be determined by computing higher-order terms using (increasingly complicated) explicit formulas.

A. Minimality Tests

The hardest work in applying Theorem 1.4 is computing the critical points, defined implicitly by (1), and determining which, if any, are minimal.

(Combinatorial Case) Recall that a function is called *combinatorial* if its power series expansion contains only a finite number of negative coefficients. When F is combinatorial there is a simple test for minimal critical points.

Lemma 1.6 (Melczer and Salvy [18]): Suppose F has only a finite number of negative power series coefficients f_i . If $\mathbf{y} \in \mathbb{C}^d_*$ is a minimal critical point then so is $(|y_1|, \ldots, |y_d|)$. Furthermore, $\mathbf{w} \in \mathbb{R}^d_{>0}$ is a minimal critical point if and only if the system

$$H(\mathbf{z}) = H(tz_1, \dots, tz_d) = 0$$

$$z_1 H_{z_1}(\mathbf{z}) - r_1 \lambda = \dots = z_d H_{z_d}(\mathbf{z}) - r_d \lambda = 0$$
 (2)

has a solution $(\mathbf{z}, \lambda, t) \in \mathbb{R}^{d+2}$ with $\mathbf{z} = \mathbf{w}$ and t = 1 and no solution with $\mathbf{z} = \mathbf{w}$ and 0 < t < 1.

In the combinatorial case we firstly use Lemma 1.6 to characterize the minimal critical points with positive coordinates by studying the solutions to (2). From them, find the solutions to (1) with the same coordinate-wise modulus. The following algorithm summarizes this approach.

Algorithm 1 Minimal Critical Points in the Combinatorial Case

- Determine the set S of zeros of the polynomial system
 in the variables z, λ, t. If S is not finite, FAIL.
- Find ζ ∈ ℝ^d_{>0} such that there exists (ζ, λ, t) ∈ S and for all such triples, t ∉ (0, 1). If the number of such ζ's is not exactly 1 or if there are such points with λ = 0, FAIL.
- 3) Identify ζ among the set C of zeros to (1).
- 4) Return { $\mathbf{z} \in \mathbb{C}^d \mid \exists (\mathbf{z}, \lambda) \in \mathcal{C}, |z_1| = |\zeta_1|, \cdots, |z_d| = |\zeta_d|$ }.

(General Case) If F is not combinatorial, or if we don't know a priori that F is combinatorial, then it is no longer sufficient to consider only the critical points with positive real coordinates to check minimality. In order to express the moduli of coordinates as algebraic equations, we write $H(\mathbf{x} + \mathbf{x})$

 $i\mathbf{y}$) = $H^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) + iH^{\mathfrak{I}}(\mathbf{x}, \mathbf{y})$ for real variables $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and polynomials $H^{\mathfrak{R}}, H^{\mathfrak{I}} \in \mathbb{R}[\mathbf{x}, \mathbf{y}]$. Translating the smooth critical point equations (1) into these new coordinates gives that $\mathbf{z} = \mathbf{a} + i\mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ is critical if and only if

$$H^{\mathfrak{R}}(\mathbf{a}, \mathbf{b}) = H^{\mathfrak{I}}(\mathbf{a}, \mathbf{b}) = 0 \tag{3}$$

$$a_j H_{x_j}^{\mathfrak{R}}(\mathbf{a}, \mathbf{b}) + b_j H_{y_j}^{\mathfrak{R}}(\mathbf{a}, \mathbf{b}) - r_j \lambda_R = 0$$
(4)

$$a_j H_{x_j}^{\mathfrak{I}}(\mathbf{a}, \mathbf{b}) + b_j H_{y_j}^{\mathfrak{I}}(\mathbf{a}, \mathbf{b}) - r_j \lambda_I = 0$$
 (5)

for some $\lambda_R, \lambda_I \in R$, where $1 \le j \le d$ in each equation. To test minimality of these critical points we add the equations

$$H^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) = H^{\mathfrak{I}}(\mathbf{x}, \mathbf{y}) = 0 \tag{6}$$

$$x_i^2 + y_j^2 - t(a_j^2 + b_j^2) = 0 (7)$$

for $1 \le j \le d$, and verify there is no real solution to (3)-(7) with 0 < t < 1. Generically (3)-(7) have a finite set of *real* solutions, corresponding to the generically finite number of critical points of F, but because this system contains 3d + 4 equations in 4d + 3 variables it will never have a (non-zero) finite number of solutions over the complex numbers. By considering critical values of the projection map onto the t coordinate, Melczer and Salvy [18] proved that minimality can be tested by adding the additional equations

$$(\nu_1 y_j - \nu_2 x_j) H_{x_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) - (\nu_1 x_j + \nu_2 y_j) H_{y_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) = 0$$

for $1 \leq j \leq d$. When $\nu_1 \neq 0$ then we can scale by ν_1 and introduce the equations

$$(y_j - \nu x_j) H_{x_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) - (x_j + \nu y_j) H_{y_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) = 0 \quad (8)$$

to (3)-(7), resulting in a square system with 4d + 4 variables and equations. The case when $\nu_1 = 0$ is dealt with separately by adding the equations

$$-x_j H_{x_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) - y_j H_{y_j}^{\mathfrak{R}}(\mathbf{x}, \mathbf{y}) = 0$$
 (8')

for $1 \le j \le d$. We determine the minimal critical points by finding $\mathbf{p} + i\mathbf{q}$ such that equations (3)-(5) have a real solution with $(\mathbf{a}, \mathbf{b}) = (\mathbf{p}, \mathbf{q})$ but neither (3)-(8) nor (3)-(8') have a real solution with $(\mathbf{a}, \mathbf{b}) = (\mathbf{p}, \mathbf{q})$ and 0 < t < 1. This process provides the following algorithm.

Algorithm 2 Minimal Critical Points in the Non-Combinatorial Case

- Determine the set S of zeros of the polynomial system
 (3)-(8) in the variables a, b, x, y, λ_R, λ_I, ν, t. If S is not finite, FAIL.
- 2) Construct a set \mathcal{U} of minimal critical points $\mathbf{a} + i\mathbf{b} \in \mathbb{C}^d$ such that there exists $(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, \nu, t) \in S \cap \mathbb{R}^{4d+4}$ and for all such tuples, $t \notin (0, 1)$. If either \mathcal{U} is empty or one of its elements has $\lambda_I = \lambda_R = 0$, or if the elements of \mathcal{U} do not all belong to the same torus, FAIL.
- 3) Identify the elements of \mathcal{U} within the set \mathcal{C} of zeros to (1) and return them.
- Do the same for the polynomial system (3)-(8') in the variables a, b, x, y, λ_R, λ_I, t.

Unfortunately, to moving 4d + 4 variables makes verifying minimality much less practical than the combinatorial case. In essence, Lemma 1.6 states that to prove minimality in the combinatorial case it is sufficient to consider specific line segments in \mathbb{R}^d , while to prove minimality in the general case one must consider a much larger set of points in \mathbb{C}^d whose coordinate-wise moduli lie on specific line segments in \mathbb{R}^d .

Remark 1.7: Melczer and Salvy [18] incorrectly state that ν_1 and ν_2 must both be non-zero: at least one is non-zero at the solutions of interest, but the other may vanish. This is why we introduce (8'). Melczer and Salvy [18] also require an extra condition that a certain Jacobian matrix is non-singular, however this is mainly required for their complexity analysis. If this condition fails then the system (3)-(8) can have extra solutions that are irrelevant to detecting minimality, but the presence of such solutions does not affect correctness of the minimality test.

II. NUMERICAL ALGEBRAIC GEOMETRY

Having reduced the ACSV analysis to questions about polynomial systems, we now recall some methods in computational algebraic geometry for the study of such systems. Although the theory of Gröbner bases is, by now, the basis of much work in this area, more recently *numerical algebraic geometry* has emerged as a practical alternative. In this section, we discuss several topics in numerical algebraic geometry that will be used for our techniques.

A. Homotopy Continuation

Homotopy continuation is one method to find numerical approximations of solutions to an $n \times n$ square system $\mathcal{F} =$ (f_1, \ldots, f_n) of polynomial equations with n variables. From the system \mathcal{F} we construct an $n \times n$ polynomial system \mathcal{G} whose solutions are known a priori. The system G is called a start system and the system \mathcal{F} is called the *target system*: connecting \mathcal{F} and \mathcal{G} using a homotopy $\mathcal{H}(x,t)$ such that $\mathcal{H}(x,0) = \mathcal{G}$ and $\mathcal{H}(x,1) = \mathcal{F}$, we obtain solutions of \mathcal{F} by tracking homotopy paths from t = 0 to t = 1. To track the homotopy paths, a numerical ordinary differential equation solving technique called the Davidenko equation and Newton iteration are used. These tracking techniques are typically referred to as predictor-corrector methods. For details, see [24, Chapter 2]. Homotopy continuation is implemented in BERTINI [4], HOMOTOPYCONTINUATION.JL [7], and NAG4M2 [16].

B. Polyhedral Homotopy Continuation

The complexity of solving a polynomial system using homotopy continuation is determined by the number of homotopy paths to track. It is thus important to track a number of paths that are at least as large as the number of solutions of the system (so that all solutions can be found) but is not too much larger (to save computation). For polytopes Q_1, \ldots, Q_n the Euclidean volume $Vol(a_1Q_1 + \cdots + a_nQ_n)$ of the *Minkowski* sum $a_1Q_1 + \cdots + a_nQ_n$ is a homogeneous polynomial in the *n* variables a_1, \ldots, a_n , whose coefficient of $a_1 a_2 \cdots a_n$ is the *mixed volume* MVol (Q_1, \ldots, Q_n) of Q_1, \ldots, Q_n .

Theorem 2.1 (Bernstein's theorem [5, Theorem A]): Let \mathcal{F} be a system of polynomials f_1, \ldots, f_n in $\mathbb{C}[x_1, \ldots, x_n]$. The number of isolated solutions of \mathcal{F} over \mathbb{C}^n_* is at most $MVol(Q_{f_1}, \ldots, Q_{f_n})$, where Q_{f_i} is the Newton polytope of f_i . Furthermore, for polynomials f_1, \ldots, f_n with generic coefficients, the number of solutions for \mathcal{F} over the torus is exactly $MVol(Q_{f_1}, \ldots, Q_{f_n})$.

The *polyhedral homotopy continuation* method established by Huber and Sturmfels [13] is one common way to construct a start system whose solutions form a set with the size of the mixed volume of a system. Consider a polynomial

$$f(\mathbf{x}) = \sum_{\mathbf{a} \in A} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}} \in \mathbb{C}[x_1, \dots, x_n]$$

where A is a collection of integer lattice points. Multiplying each monomial $\mathbf{x}^{\mathbf{a}}$ of f by some term $t^{w(\mathbf{a})}$ for a *lifting* function $w: A \to \mathbb{Z}$, we obtain the *lifted polynomial*

$$\overline{f}(\mathbf{x},t) = \sum_{\mathbf{a} \in A} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}} t^{w(\mathbf{a})}$$

Suppose that a target system \mathcal{F} consists of polynomials f_1, \ldots, f_n supported on A_{f_1}, \ldots, A_{f_n} , respectively. Lifting all polynomials f_1, \ldots, f_n in \mathcal{F} gives a lifted system $\overline{\mathcal{F}}(\mathbf{x}, t)$ satisfying $\overline{\mathcal{F}}(\mathbf{x}, 1) = \mathcal{F}$. The solutions of $\overline{\mathcal{F}}$ can be expressed by Puiseux series $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$ where

$$x_i(t) = t^{\alpha_i} y_i + \text{ higher order terms}$$

for some $\alpha_i \in \mathbb{Q}$ and nonzero constant y_i , and substituting $\mathbf{x}(t)$ back into our polynomials gives

$$\overline{f}_j(\mathbf{x}(t),t) = \sum_{\mathbf{a} \in A_{f_j}} c_{\mathbf{a}} \mathbf{y}^{\mathbf{a}} t^{\langle \mathbf{a}, \boldsymbol{\alpha} \rangle + w(\mathbf{a})} + \text{ higher order terms}$$

For a suitable choice of w, the constants y and exponents a can be computed at each branch of $\overline{\mathcal{F}}$, ultimately describing a start system $\mathcal{G} = \overline{\mathcal{F}}(\mathbf{x}, 0)$ with the right number of solutions. The polyhedral homotopy continuation is implemented in HOM4PS2 [15], HOMOTOPYCONTINUATION.JL [7], and PHCPACK [26].

C. Monodromy

As seen in Section I, we typically have some solutions of a polynomial system representing critical points and want to determine additional solutions to rule out those that are non-minimal. This 'bootstrapping' can be accomplished by monodromy.

For $m, n \in \mathbb{N}$, consider the complex linear space of $n \times n$ square systems $\mathcal{F}_p = (f_p^1, \ldots, f_p^n)$ depending on some coefficient parameters $p \in \mathbb{C}^m$, where the monomial support for each polynomial f_p^i is fixed. If we consider an affine linear map $\varphi : p \mapsto F_p$ for $p \in \mathbb{C}^m$ then we can write $\varphi(\mathbb{C}^m) = B$, where B is a parametrized linear variety of systems, and we define the *solution variety* $V = \{(F_p, x) \in B \times \mathbb{C}^n \mid F_p(x) = 0\}$ and *projection map* $\pi : V \to B$. Assume that the fiber $\pi^{-1}(F_p)$ only has finitely many points for a generic choice of p. The set D of systems in B with nongeneric fiber is called the *branch locus* of π . Each element in the *fundamental group* $\pi_1(B \setminus D)$ of loops in $B \setminus D$ modulo homotopy equivalence induces a permutation on the fiber $\pi^{-1}(F_p)$, which is called a *monodromy action*. To find all solutions of a system $F_p \in \pi(V)$ with generic p, one can first find a seed solution $(p_0, x_0) \in V$ and numerically compute the monodromy action to find all solutions of F_p . When the solution variety V is irreducible then the monodromy action is transitive. This method for finding solutions of polynomial systems is studied and implemented in [7], [9].

D. Certification

By construction, numerical methods return approximations, so some kind of certification is necessary for rigorous results. Specifically, a user needs a certificate that an approximation obtained by the homotopy method is properly approximating a solution of a system. A numerical approximation is called *certified* if it can be refined to an actual solution of the system to an arbitrary precision by applying iterative operators (such as Newton iteration). Software providing such certification includes ALPHACERTIFIED [12], the function CER-TIFY implemented in HOMOTOPYCONTINUATION.JL [6] and NUMERICALCERTIFICATION [14]. In our implementation, we use the function CERTIFY in HOMOTOPYCONTINUATION.JL exploiting the Krawczyk's method via interval arithmetic [19, Chapter 8].

III. THE ACSVHOMOTOPY PACKAGE

We now combine the theory of ACSV presented in Section I with the techniques described in Section II to create effective and practical algorithms for the asymptotics of multivariate rational functions. Our algorithms are implemented in the JULIA package ACSVHOMOTOPY.JL, using the HOMOTOPYCON-TINUATION.JL package for our homotopy and monodromy computations.

The package is available at

github.com/ACSVMath/ACSVHomotopy

and our example worksheet can be viewed at

github.com/ACSVMath/ACSVHomotopy/blob/main/ExampleWorksheet.ipynb

A. Combinatorial Case

For the combinatorial case we first compute the distinct solutions to (1) using a polyhedral homotopy with certification by Krawczyk's method. We then solve and certify (2) with the added equation $(1 - t)\mu - 1 = 0$ to eliminate all solutions with t = 1 (there are never any solutions with t = 0 as this would imply $H(\mathbf{0}) = 0$, contradicting F having a power series expansion). Since we no longer have solutions where t = 1, by refining the solutions to sufficient precision we can determine the solutions with positive real coordinates where 0 < t < 1, match the projection onto the z variables of each to a distinct solution of (1), and thus rule out all non-minimal critical points with positive coordinates. We then find all critical points with the same coordinate-wise moduli and return that set.

Example 3.1: As a simple example, we can find the minimal critical point

```
@polyvar x y
find_min_crits_comb(1-x-y)
Out:1-element Vector{Vector{ComplexF64}}:
    [0.5 + 0.0im, 0.5 + 0.0im]
```

controlling asymptotics for the central binomial coefficient $\binom{2n}{n}$ which forms the main diagonal sequence of

$$F(x,y) = \frac{1}{1-x-y}.$$

Similarly, we can compute the approximations

```
@polyvar x y z
find_min_crits_comb(1-z*(x^2*y+y+x*y^2+x))
Out:2-element Vector{Vector{ComplexF64}}:
    [1.0 + e-35im, 1.0 - e-35im, 0.25 - e-37im]
    [-1.0 - e-36im, -1.0 + e-36im, -0.25 - e-36im]
```

for the two minimal critical points $\pm(1, 1, 1/4)$ determining asymptotics for the main diagonal of

$$F(x, y, z) = \frac{(1+x)(1+y)}{1 - zxy(x+1/x+y+1/y)}$$

This diagonal enumerates walks on the cardinal directions $\{N, S, E, W\} = \{(\pm 1, 0), (0, \pm 1)\}$ that start at the origin and stay in \mathbb{N}^2 .

As in the work of Melczer and Salvy, the most expensive operation occurs when trying to group roots with the same coordinate-wise modulus as a known minimal critical point (step (4) in Algorithm 1). When using a symbolic-numeric method it is possible to compute minimal polynomials for the values of the coordinates and use this to identify points with the same coordinate-modulus by computing numerically to $O(h\delta^{3d})$ bits of precision, where h is a bound on the bitsize of the coefficients of the denominator H and δ is the degree of H (see [18, Corollary 54]). Combining past bounds in the literature [8], [23] allows us to identify an explicit precision such that if two solutions have coordinatewise moduli agreeing to this precision then their coordinatewise moduli are exactly equal. Our bound is also of $O(h\delta^{3d})$ bits, however the constant in front is worse than that when using minimal polynomials. After a minimal critical point is identified, our code continually refines precision using Newton iteration until any points with the same coordinate-wise moduli are found. In practice, any points with different coordinatewise moduli are identified at precision much lower than the worst case bound, but we must compute up to the bound when there are points with the same coordinate-wise moduli. Unfortunately, the extreme precision involved with a large number of variables means that we cannot always rigorously check to the required accuracy, and the code returns a warning to the user with its output when this occurs.

B. General Case

In general we must consider the extended systems (3)-(8) and (3)-(8'), which essentially doubles the number of variables under consideration. Mirroring the combinatorial case, we can solve (3)-(8) and (3)-(8') using a polyhedral homotopy, certify the results using Krawczyk's method, and refine to a sufficient precision to determine when 0 < t < 1 to rule out non-minimal points.

Remark 3.2: The system (3)-(8') is over determined, with 4d + 4 equations and 4d + 3 variables. In order to use HOMOTOPYCONTINUATION.JL we drop one of the equations in (8') to obtain a square system: this can introduce additional solutions which are irrelevant to determining minimality, but does not affect the correctness of our test for minimality.

Example 3.3: Straub and Zudilin [25], following Gillis, Reznick, and Zeilberger [11], study families of rational functions connected to special function theory. For instance, in three dimensions they study the constants c for which

$$F_c(x, y, z) = \frac{1}{1 - (x + y + z) + cxyz}$$

has non-negative power series coefficients on its main diagonal (which turns out to imply non-negativity of all power series coefficients). Running the code (for c = 5)

gives the minimal critical points controlling asymptotics of the main diagonal. Since this is a complex conjugate pair, the resulting asymptotic expansion implies that F_c has an infinite number of negative coefficients on its main diagonal when c = 5 (in fact, it has an infinite number of negative coefficients whenever c > 4).

C. Faster Heuristics

As seen in the examples of Section IV, the high number of variables in the extended system even in low dimensional examples means it does not terminate within reasonable time for polynomials with four or more variables. In order to speed up our solvers, we can numerically approximate the distinct solutions to the small system (3)-(5) and then substitute each of these solutions as parameters into the extended equations (6)-(8) and (6)-(8'). In the implementation, it is done by running the function find_min_crits with the flag approx_crit = true.

Remark 3.4: This approach approximates the solutions to the extended system (3)-(8) if the solutions vary smoothly with a and b, which happens whenever the Jacobian of (6)-(8) with respect to a and b is full rank at all values of a and b solving (3)-(5). Unfortunately, verifying this condition is usually about as costly as solving the extended system, so we do not do this in our computations and refer to this method only as an efficient heuristic that correctly identifies minimal critical points in a large variety of cases.

Example 3.5: To stress-test our algorithms we generate a random polynomial p(x, y, z) with six terms in four variables having coefficients in $\{1, \ldots, 100\}$ and then set H(x, y, z) = 1 - p(x, y, z). Running

```
@polyvar x y z
H=1-(72*x^3*z+97*y*z^3+53*x*z^2+47*x*y+39*z^2+71*x)
find_min_crits(H; approx_crit = true)
Out:1-element Vector{Vector{ComplexF64}}:
    [0.001+5.5e-40im, 6.2-7.5e-37im, 0.06+0.0im]
```

returns the unique minimal critical point in about three minutes. This example does not terminate without the approx_crit = true flag.

It is also possible to use the approximations to the critical points as a start system to solve (6)-(8) using the monodromy method. More precisely, for any (a, b) solving (3)-(5) we set $(\mathbf{x}, \mathbf{y}) = (\mathbf{a}, \mathbf{b})$ and t = 1 in (6)-(8) and then compute a corresponding start value of ν by computing the left kernel of the Jacobian matrix of (6)-(7) with respect to variables \mathbf{x}, \mathbf{y} and t. From a given parameter value (\mathbf{a}, \mathbf{b}) and the initial solution $(\mathbf{x}, \mathbf{y}, \nu, t)$, we collect real solutions from the monodromy method and check if $t \in (0, 1)$: if it is then we remove the parameter value (\mathbf{a}, \mathbf{b}) as it is non-minimal. Interestingly, it appears that monodromy cannot detect solutions where $\nu = 0$ when starting with a non-zero value of ν , and vice-versa, suggesting that solution variety is the union of components corresponding to these cases. We thus repeat this process separately for the cases where $\nu = 0$ and when (8') replaces (8). Finally, we return the values of (\mathbf{a}, \mathbf{b}) that are not disregarded.

Example 3.6: Melczer and Salvy [18] introduce the rational function

$$F(x,y) = \frac{1}{(1-x-y)(20-x-40y)-1}$$

because it has two critical points with positive coordinates, one of which is smaller in the first coordinate and the other of which is smaller in the second coordinate (so it is not clear which, if any, should be minimal). Running

```
@polyvar x y
H = (1-x-y)*(20-x-40*y)-1
find_min_crits(H; monodromy=true)
Out:1-element Vector{Vector{ComplexF64}}:
    [0.54 - 9.18e-41im, 0.31 + 1.83e-40im]
```

returns the correct minimal critical point.

We believe that further study of the geometric properties of the extended system (6)-(8) could help make this monodromy approach a powerful tool for ACSV analysis.

IV. EXAMPLES AND BENCHMARKS

Tables I and II list benchmarks of our implementation against a selection of combinatorial and algebraic examples, executed on a Macbook pro, 2 GHz Quad-Core Intel Core i5, 16 GB RAM. The package supports arbitrary r-diagonal sequences, but examples in this section were done with r = 1.

Example	Comb.	Maple Comb.
1 - x - y	0.0052	0.143
Two positive CPs	0.029	0.292
square-root	0.01	0.06
Apéry $\zeta(2)$	0.025	0.06
Apéry $\zeta(3)$	0.7	0.3
random poly	0.9	840
2D Walk	0.03	0.06
3D Walk	0.08	2.7
$1 - x - y^2 - w^3 - z^4$	0.06	509

	DI	-	•
- T A	81	н	
- 1/7			

TIME, IN SECONDS, OF RUNNING OUR JULIA IMPLEMENTATIONS IN THE COMBINATORIAL CASE, COMPARED TO THE KRONECKER REPRESENTATION APPROACH OF MELCZER AND SALVY. THE TIME TO

COMPILE JULIA FUNCTIONS IS NOT INCLUDED.

Example	HSolve	HSolve Approx	Monodromy
1 - x - y	0.04	0.02	2.3
Two positive CPs	4.1	0.33	2.84
Apéry $\zeta(2)$	670	3.8	8.5
square-root	29.5	0.72	14.9
random poly	INC	189.4	583.1
2D Walk	INC	15.3	31.9
GRZ	236	3.6	3.8

TABLE II

TIME, IN SECONDS, OF RUNNING OUR JULIA IMPLEMENTATIONS THAT DO NOT ASSUME COMBINATORALITY. THE FIRST COLUMN IS THE TIME TO SOLVE THE EXTENDED CRITICAL POINT SYSTEMS, THE SECOND COLUMN IS THE TIME TO SOLVE THE SMALLER SYSTEMS AFTER SOLVING THE CRITICAL POINT SYSTEM SEPARATELY, AND THE FINAL COLUMN IS THE

TIME TO RUN THE MONODROMY METHOD. INC INDICATES THE CODE DID NOT COMPLETE AFTER RUNNING FOR AN HOUR.

See our supplementary notebook for the full details on the rational functions involved.

Remark 4.1: The HOMOTOPYCONTINUATION.JL package converts input polynomials into compiled straight-line programs for fast evaluation. In order to better see the differences between examples as they grow in degree and dimension, we have removed compilation time from our benchmarks (compilation time takes the majority of the runtime for small examples but is a small part of larger examples). This results in several seconds on small examples, and up to tens of seconds on larger examples, that are not included in the reported timings. In particular, the (non-certified) package of Melczer and Salvy beats our package in the combinatorial case on most examples in Table I when compilation time is added (except for the two high degree examples where the Maple package takes much longer).

V. RIGOR OF RESULTS

Because we certify our solutions, we never attempt to approximate a point that is not actually a solution of the polynomial systems under consideration. However, by the way they are designed, it is possible for homotopy computations to *miss* solutions, which could result in a point being deemed minimal when it is not. There are some exceptions: when the number of solutions found matches the upper bound on the number of solutions given by the mixed volume, for instance, then we can be sure we have found all solutions. Tables III and IV show a comparison between the mixed volumes of

Example	Mixed volume	# Solutions found
1 - x - y	1	1
$1 - xy - xy^2 - 2x^2y$	9	9
$1 - x - y^2 - w^3 - z^4$	96	96

TABLE III MIXED VOLUME FOR THE SYSTEM (2) AND THE ACTUAL NUMBER OF SOLUTIONS FOUND FOR SEVERAL COMBINATORIAL EXAMPLES

Example	(3)-(8)	# Sols	(3)-(8')	# Sols
1 - x - y	4	1	2	0
$1 - xy - xy^2 - 2x^2y$	3276	99	1638	126
$1 - (x + y + z) + \frac{81}{8}xyz$	13068	162	4356	216
$1 - x - y^2 - w^3 - z^4$	FAIL	N/A	442368	442368

TABLE IV

MIXED VOLUMES FOR THE SYSTEMS (3)-(8) AND (3)-(8') AND THE NUMBER OF SOLUTIONS FOUND FOR SEVERAL EXAMPLES. FAIL MEANS THAT THE CODE DID NOT COMPUTE THE MIXED VOLUME DUE TO AN OUT OF MEMORY ERROR.

several systems studied here, compared to the actual number of solutions found. It can be observed that we often reach the upper bound in the combinatorial case, but this usually does not happen in the non-combinatorial case.

We can also conclude we know minimality rigorously when there is another way to determine that a minimal critical point must exist, and all but one point is ruled out by our algorithms. For instance, in the combinatorial case it can be shown that any polynomial whose support contains the terms $1, z_1, z_2, \ldots, z_d$ must have at least one minimal critical point with positive coordinates.

Finally, as already mentioned, in the combinatorial case the precision required to certify all minimal critical points with the same coordinate-wise modulus may not be practical. Our algorithm returns a warning with its output when it cannot verify equality between moduli to the precision required for rigorous certification.

VI. CONCLUSION

Despite the high computational cost associated to many of computations required to determine asymptotics using the methods of ACSV, the continued development of efficient computer algebra packages in Julia and other languages has made it feasible to automate the analysis beyond the simplest cases. There are many natural extensions still to be made, perhaps chiefly among them extending to the non-smooth case by incorporating algorithms for the Whitney stratification of algebraic varieties. Other interesting avenues for exploration include the development of better start systems for homotopy computations, to better match the number of critical points, and a theoretical study of the solution variety and its irreducible components for the monodromy approach (which could help the monodromy approach be competitive with or even surpass the polyhedral homotopy approach).

REFERENCES

 R. Apéry, "Sur certaines séries entières arithmétiques," in *Study group* on ultrametric analysis, 9th year: 1981/82, No. 1. Inst. Henri Poincaré, Paris, 1983, pp. Exp. No. 16, 2.

- [2] Y. Baryshnikov, S. Melczer, and R. Pemantle, "Stationary points at infinity for analytic combinatorics," *Foundations of Computational Mathematics*, 2022.
- [3] Y. Baryshnikov and R. Pemantle, "Asymptotics of multivariate sequences, part III: Quadratic points," *Adv. Math.*, vol. 228, no. 6, pp. 3127–3206, 2011.
- [4] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler, "Bertini: Software for numerical algebraic geometry," Available at bertini.nd.edu with permanent doi: dx.doi.org/10.7274/R0H41PB5.
- [5] D. N. Bernshtein, "The number of roots of a system of equations," *Functional Analysis and its applications*, vol. 9, no. 3, pp. 183–185, 1975.
- [6] P. Breiding, K. Rose, and S. Timme, "Certifying zeros of polynomial systems using interval arithmetic," arXiv preprint arXiv:2011.05000, 2020.
- [7] P. Breiding and S. Timme, "HomotopyContinuation.jl: A package for homotopy continuation in julia," in *International Congress on Mathematical Software*. Springer, 2018, pp. 458–465.
- [8] A. Dubickas and M. Sha, "Counting and testing dominant polynomials," *Exp. Math.*, vol. 24, no. 3, pp. 312–325, 2015.
- [9] T. Duff, C. Hill, A. Jensen, K. Lee, A. Leykin, and J. Sommars, "Solving polynomial systems via homotopy continuation and monodromy," *IMA Journal of Numerical Analysis*, vol. 39, no. 3, pp. 1421–1446, 2019.
- [10] P. Flajolet and R. Sedgewick, *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [11] J. Gillis, B. Reznick, and D. Zeilberger, "On elementary methods in positivity theory," *SIAM J. Math. Anal.*, vol. 14, no. 2, pp. 396–398, 1983.
- [12] J. D. and F. Sottile, "alphaCertified: Hauenstein Softfor certifying numerical solutions ware to polvnoequations," Available mial at math.tamu.edu/ sottile/research/stories/alphaCertified, 2011.
- [13] B. Huber and B. Sturmfels, "A polyhedral method for solving sparse polynomial systems," *Mathematics of computation*, vol. 64, no. 212, pp. 1541–1555, 1995.
- [14] K. Lee, "The NumericalCertification package in Macaulay2," arXiv preprint arXiv: 2208.01784, 2022.
- [15] T.-L. Lee, T.-Y. Li, and C.-H. Tsai, "HOM4PS-2.0: a software package for solving polynomial systems by the polyhedral homotopy continuation method," *Computing*, vol. 83, no. 2, pp. 109–133, 2008.
- [16] A. Leykin, "Numerical algebraic geometry," Journal of Software for Algebra and Geometry, vol. 3, no. 1, pp. 5-10, 2011.
- [17] S. Melczer, An Invitation to Analytic Combinatorics: From One to Several Variables, ser. Texts and Monographs in Symbolic Computation. Springer International Publishing, 2021.
- [18] S. Melczer and B. Salvy, "Effective coefficient asymptotics of multivariate rational functions via semi-numerical algorithms for polynomial systems," J. Symbolic Comput., vol. 103, pp. 234–279, 2021.
- [19] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to interval analysis*. Siam, 2009, vol. 110.
- [20] J. Ouaknine and J. Worrell, "Ultimate positivity is decidable for simple linear recurrence sequences," in *Automata, languages, and programming. Part II*, ser. Lecture Notes in Comput. Sci. Springer, Heidelberg, 2014, vol. 8573, pp. 330–341.
- [21] R. Pemantle and M. C. Wilson, "Asymptotics of multivariate sequences. I. Smooth points of the singular variety," J. Combin. Theory Ser. A, vol. 97, no. 1, pp. 129–161, 2002.
- [22] —, Analytic combinatorics in several variables, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2013, vol. 140.
- [23] M. Safey El Din and É. Schost, "Bit complexity for multi-homogeneous polynomial system solving—Application to polynomial minimization," *J. Symbolic Comput.*, vol. 87, pp. 176–206, 2018.
- [24] A. Sommese and C. Wampler, The Numerical Solution of Systems of Polynomials Arising in Engineering and Science. World Scientific, 2005.
- [25] A. Straub and W. Zudilin, "Positivity of rational functions and their diagonals," J. Approx. Theory, vol. 195, pp. 57–69, 2015.
- [26] J. Verschelde, "Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation," ACM Transactions on Mathematical Software (TOMS), vol. 25, no. 2, pp. 251–276, 1999.