A Numerical Algorithm for Zero Counting

IV

An Adaptive Speedup

Josué Tonelli-Cueto
Inria Paris / IMJ-PRG

Norway | Tromsø, June 7-11, 2021
Internal Story
of the Problem
THE PROBLEM

$H_d[n] \in \mathbb{R}$

Deterministic + Numerically Stable + 'Good' Probabilistic Run-time (for random $g$)

The ALGORITHM we want!

Real Homogeneous Polynomial System in $x_0, \ldots, x_n$ & with $\deg g_i = d_i$

$\# \mathbb{Z}_P^d(g)$

# projective zeros of $g$
A numerical algorithm for zero counting. I: Complexity and accuracy
Felipe Cucker\textsuperscript{1,2}, Teresa Krick\textsuperscript{3,4}, Gregorio Malajovich\textsuperscript{5}, Mario Wschebor\textsuperscript{6}

\textsuperscript{1}Department of Mathematics, City University of Hong Kong, Hong Kong
\textsuperscript{2}Department of Mathematics, U. de Buenos Aires, Argentina
\textsuperscript{3}CONICET-Argentina
\textsuperscript{4}Instituto de Matemática Aplicada del Litoral, Facultad de Ingeniería de la Universidad Nacional del Litoral, Argentina
\textsuperscript{5}Chair of Mathematics, U. Federal de Minas Gerais, Brazil

\begin{abstract}
We describe an algorithm to count the number of distinct real roots of a polynomial system \( f \) on an interval \([-1, 1] \). The main result of the paper is that the number of distinct real roots of \( f \) on \([-1, 1] \) is \( O(n^\alpha) \) given that \( |f| \) is \( O(1) \) on \([-1, 1] \) and \( |f| \) is \( O(1) \) on \([-1, 1] \) for some \( \alpha > 0 \). The main result is obtained by applying the so-called condition number theorem.
\end{abstract}

1. Introduction
In recent years considerable attention has been paid to the complexity of counting problems over the reals. The counting complexity class \#P, was introduced [20] and completeness results for \#P were established [1] for natural geometric problems, notably, for the computation of the Euler characteristic of semi-algebraic sets. As one could expect, the \#P complexity problem consists of counting the number of solutions of a system of polynomial equations.
The Original Trilozy
(by Cuker, Knick, Malaiovich, Schebeor)

FIRST MILESTONE

Part 1: Algorithm & condition-based complexity

Part 2: Condition Number Theorem

Part 3: Probabilistic analysis & integral geometry
The CKMW algorithm

\( H_d[n] \in \mathbb{E} \)

\[ \text{DETERMINISTIC} \]

\[ \text{NUMERICALLY STABLE} \]

\[ \text{'GOOD' PROBABILISTIC RUN-TIME} \]

With 'high probability', \( \text{run-time}(\text{CKMW}, 8) \leq O(n^2) \)

for \( 8 \) KSS (average/smoothed)

KSS = Kostlan-Shub-Smale Gaussian

\( D := \max d_i \)
The Spin-offs
(by Ergür, Paouris, Rojas)

Probabilistic Condition Number Estimates for Real Polynomial Systems I: A Broader Family of Distributions

Alperten A. Ergür1 · Grigoris Paouris2 · J. Maurice Rojas3

Received: 5 December 2016 / Revised: 6 June 2017 / Accepted: 30 October 2017
Published online: 1 January 2018©Springer 2018

Abstract We consider the sensitivity of real roots of polynomial systems with respect to perturbations of the coefficients. In particular, for a version of the condition number defined by Cucker and used later by Cucker, Krzakala, Malajovich, and Wachsmuth—we establish new probabilistic estimates that allow a much broader family of measures than considered earlier. We also generalize further by allowing overdetermined systems. In Part II, we study smoothed complexity and how sparsity (in the sense of restricting which terms can appear) can help further improve earlier condition number estimates.

Keywords Condition number · Equilibrium net · Probabilistic bound · Kayan · Real-sampling · Overdetermined · Subgaussian

Smoothed analysis for the Condition Number of Structured Real Polynomial Systems

Alperten A. Ergür, Grigoris Paouris, and J. Maurice Rojas

Abstract. We consider the sensitivity of real roots of structured polynomial systems to perturbations of their coefficients. In particular, we provide explicit estimates for condition numbers of structured real polynomial systems and extend these estimates to smoothed analysis setting.

1. Introduction

Efficiently finding real roots of real polynomial systems is one of the main objectives of computational algebraic geometry. There are numerous algorithms for this task, but the core steps of these algorithms are easy to criticize. They are some combination of algebraic manipulation, a discrete polynomial computation, and a numerical iterative scheme. From a computational complexity point of view, the cost of numerical iteration is much too large compared to the cost of algebraic or discrete computation. This paper constitutes a step toward understanding the complexity of numerically solving structured real polynomial systems. Our main results are Theorems 1.14, 1.15, and 1.18 below, but we will first need to give some context for our results.

1.1 How to control accuracy and complexity of numerical in real algebraic geometry?

In the monomial basis algebraic geometry, going back to von Neumann and Furing, condition numbers play a central role in the control of accuracy and speed of algorithms (see, e.g., [3, 6] for further background). Shaw and Smale initiated the use of condition numbers for polynomial system solving over the field of complex numbers [20, 27]. Subsequently, condition numbers played a central role in the solution of Smale’s 17th problem [2, 5, 25].

The maximum of solving polynomial systems over the reals is more subtle than complex ones. Real perturbations can cause the solution set to change drastically. One can even go from having no real root to many real roots by an arbitrarily small change in the coefficients. This behavior doesn’t appear over the complex numbers as one has theorems (such as the Fundamental Theorem of Algebra) proving that most curves are “generically” connected. Lately, a condition number theory that captures these subtleties was developed by Cucker [11]. Now we set up the notation and present Cucker’s definition.

Definition 1.1 (Bomelsini–Wojdyłło Norms). Let \( \gamma = \langle x \rangle \) where \( \gamma = (\gamma_1, \ldots, \gamma_d) \) and let \( P \equiv (p_1, \ldots, p_d) \) be a system of homogeneous polynomials with degree patterns \( d_1, \ldots, d_d \). Let \( \gamma_\alpha \) denote the coefficient of \( x^\alpha \) in \( p_\alpha \). We define the Wielandt–Bomelsini norm of \( p \) and \( P \) to be, respectively,

\[
\|p\|_W = \sqrt{\sum_{\alpha} \gamma_\alpha^2}
\]
Probabilistic Condition Number Estimates for Real Polynomial Systems: A Broader Family of Distributions

Ahmet Yildiz and Parvin Imanpour

Foundations of Computational Mathematics

Volume 16, Issue 5, November 2016

Abstract: We consider the sensitivity of real roots of polynomial systems with respect to perturbations of the coefficients. – For this purpose, we adapt the approach by Cucker and Imonpour to the case of real polynomial systems. In particular, we show that under certain conditions, all real roots of the perturbed system can be identified with high probability. This is in contrast to the case of complex polynomial systems, where such conditions are not known.

Keywords: Condition number - Perturbation analysis - Real coefficients - Polynomial systems - Perturbation theory

1 Introduction

1.1 How to control accuracy and complexity of summations in real algebraic geometry?

In the numerical linear algebra tradition, going back to von Neumann and Fuchs, condition numbers play a central role in the control of accuracy and speed of algorithms (see, e.g., [2, 6, 25] for further background). In the case of real polynomial systems, the condition number of a real polynomial system can be defined as the ratio of the maximum and minimum absolute values of its coefficients.

2. New Results

2.1 The Spin-offs (by Ergür, Paouris, Rojas)

The Spin-offs (by Ergür, Paouris, Rojas)

1st non-gaussian average complexity in Numerical Alg. Geom!

1st non-gaussian smoothed complexity in NAG!
The CKMW algorithm (after the spinoffs)

$H_d[n] \geq 8$ \rightarrow \text{CKMW} \rightarrow \#Z_p(8)

Deterministic

Numerically stable

'Good' probabilistic run-time

With 'high probability', run-time(CKMW,8) ≤ $O(n^2)$ for 8 wide class of random systems

$D := \max d_i$
For random $\mathcal{S} \in \mathcal{H}_d[\mathcal{U}]$ as before, does $E_8 \text{ run-time}(\text{CKMW}, \mathcal{S}) < \infty$? No!
Idea!

Make CKMW adaptive, then complexity should depend on

$$E_{x \sim s^n} \kappa(\delta, x)$$

which has finite expectation for a random $\delta$!

$$\kappa(\delta, x) = ||D||w/\sqrt{||s(x)||^2 + ||D_x s^T \Delta ||^2}$$

$\Delta = \text{diag}(d_i)$

Inspiration: (Cucker, Ergür, T.-C.; 2019) while studying AV algorithm
Naive adaptive version fails!

(Eckhardt, 2020) (Han, 2018)

Run-time bound in terms of

$$E \kappa(\delta, x)^{2^n}$$

which has infinite expectation for a random $\delta$!

$$\kappa(\delta, x) = \Delta \| \Psi / \sqrt{\|s(x)^2 + \|D_x s(x) \|^2 \Delta^{-1}} \Delta = \text{diag}(d_i)$$
What goes wrong?

The criterion to select zeros!
The CKMW algorithm

1) Refine grid \( G \subseteq \mathbb{S}^n \) until \( d_{\mathbb{S}}(G, \mathbb{S}^n) \) 'small'

2) \( \begin{cases} 
\text{Exclude points } x \in G \text{ s.t. } \|f(x)\|/\|f\|_w \text{ 'big'} \\
\text{Include points } x \in G \text{ s.t. } \|f(x)\|/\|f\|_w \text{ 'small'}
\end{cases} \)

3) Post-process the selected points to get \( \#Z_{1p}(G) \)
The CKMW algorithm

1) Refine grid $G \subseteq S^n$ until $d_S(G, S^n) \leq \frac{1}{cD^2 \chi(8)^2} : \varepsilon$

2) \begin{align*}
\begin{cases}
\text{Exclude points } x \in G \text{ s.t. } \frac{\|g(x)\|}{\|g\|_w} > \sqrt{D} \varepsilon \\
\text{Include points } x \in G \text{ s.t. } \frac{\|g(x)\|}{\|g\|_w} \leq \frac{1}{cD^2 \chi(8)^2}
\end{cases}
\end{align*}

3) Post-process the selected points to get \# $\mathbb{Z}_{1p}(8)$

Note quadratic condition in the inclusion criterion!

$$\chi(8) := \max_{x \in S^n} \frac{\|g\|_w}{\sqrt{\|g(x)\|^2 + \|D g(x)^T D^{-1}\|^{-2}}} \quad \text{condition number}$$
The adaptive CKMW algorithm

NAIVE EDITION

1) Refine adaptively \( G \subseteq \mathbb{S}^{n+1} \times (0,\infty) \) so that
   \[ \mathbb{S}^{n} \subseteq \bigcup \{ B_{g}(x,r) \mid (x,r) \in G \} \]
   & 2) \( \forall (x,r) \in G, \ r \leq \frac{1}{C} D \kappa (g) \sqrt{x^2} \)

\[
\begin{cases}
   \text{Exclude} \ (x,r) \in G \text{ if } \frac{11g(x)}{\| g \|_w} \geq \sqrt{D} r \\
   \text{Include} \ (x,r) \in G \text{ if } \frac{11g(x)}{\| g \|_w} \leq \frac{1}{C} D^2 \kappa (g) x^2
\end{cases}
\]

3) Post-process the selected points to get \( \# \mathcal{E}_{np}(g) \)

Still quadratic inclusion criterion!

\[ \kappa (g(x), x) := \frac{\| g \|_w}{\sqrt{\| g(x) \|_w^2 + \| Dg(x) \|_w^2}} \text{ local condition number} \]
Where does the square come from?

Smale's $\alpha$-criterion:
\[
\alpha(\delta, x) := \beta(\delta, x) \gamma(\delta, x) \leq \alpha^*
\]
\[
\Rightarrow \quad \# B_{\delta}(x, 1.5 \beta(\delta, x)) \cap Z_{\delta}(\delta) = 1
\]
\[
\text{and} \quad N_{\delta}^u(x) \xrightarrow{\text{quadratically}} \text{zero of} \ \delta
\]

where $\beta(\delta, x) := \| D_x \delta^{-1} \delta(x) \| \ & \gamma(\delta, x) := \sup_{k \geq 2} \| D_x \delta^{-1} \frac{1}{k!} D_x^k \delta \|

\text{Assume} \quad \sqrt{2} \kappa(\delta, x) \| g(x) \|/\|g\|_W \leq 1 \ldots$

- Higher Derivative Estimate: $\gamma(\delta, x) \leq \frac{1}{2} D^{3/2} \kappa(\delta, x)$
- A bad bound for $\beta$: $\beta(\delta, x) \leq \kappa(\delta, x) \| g(x) \|/\|g\|_W$

\[
N_{\delta}(x) := \frac{x - D_x \delta^{-1} \delta(x)}{\| x - D_x \delta^{-1} \delta(x) \|}, \quad N_{\delta}^{u+1}(x) = N_{\delta}(N_{\delta}^u(x))
\]
Where does the square come from?

Smale's α-criterion:
\[
\alpha(8, x) := \beta(8, x) \gamma(8, x) \leq \alpha
\]

\[
\implies \# B_8(x, 1.5 \beta(8, x)) \cap \mathbb{Z}_8(8) = 1
\]

\& \quad N_8(x) \xrightarrow{\text{quadratically}} \text{zero of } \&

where \(\beta(8, x) := \|D_x 8^{-1} g(x)\|\) \& \(\gamma(8, x) := \sup_{k \geq 2} \|D_x 8^{-1} D_x^k g\|\)

Assume \(\sqrt{2} \kappa(8, x) \|g(x)\|/\|g\| \leq 1\).

\cdot \text{Higher Derivative Estimate: } \gamma(8, x) \leq \frac{1}{2} D^{3/2} \kappa(8, x)

\cdot \text{A bad bound for } \beta:\quad \beta(8, x) \leq \kappa(8, x) \|g(x)\|/\|g\| \leq 1

\[
N_8(x) := \frac{x - D_x 8^{-1} g(x)}{\|x - D_x 8^{-1} g(x)\|}, \quad N_8^{n+1}(x) = N_8(N_8^n(x))
\]

\[
\text{This creates the square!}
\]
We should use $\beta$ directly!

Converse Smale's $\alpha$-theorem:

\[ \gamma(\beta, x) \text{dist}_S(x, \mathcal{Z}_S(\beta)) < 1 \]

\[ \Rightarrow \quad \alpha(\beta, x) \leq \frac{\gamma(\beta, x) \text{dist}_S(x, \mathcal{Z}_S(\beta))}{1 - \gamma(\beta, x) \text{dist}_S(x, \mathcal{Z}_S(\beta))} \]

'If $x$ is sufficiently near $\mathcal{Z}_S(\beta)$, then Smale's $\alpha$-criterion at $x$ holds.'

Corollary. If $\sqrt{2} \gamma(\beta, x) \| h(x) \| / \| h \|_W < 1$, then $\alpha(\beta, x) < \alpha^*$

or $B_S(x, c/D^2 \gamma(\beta, x)) \cap \mathcal{Z}_S(\beta) = \emptyset$
The adaptive CKMW algorithm

1) Refine adaptively \( C_0 \subseteq S^n \times (0, \infty) \) so that
   1) \( S^n \subseteq \bigcup \{ B_{S}(x, r) \mid (x, r) \in C_0 \} \)
   2) \( \forall (x, r) \in C_0, \ r \leq \frac{1}{\sqrt{cD^2 \lambda(x)}} \)

Excluding \( (x, r) \in C_0 \) if \( \|g(x)\|_1 / \|g\|_w \geq \sqrt{D} r \)

Include \( (x, r) \in C_0 \) if \( \beta(x) \leq \frac{1}{\sqrt{cD^2 \lambda(x)}} \)

3) Post-process the selected points to get \( \# \mathcal{Z}_{1p}(x) \)

Using \( \beta \) gives the desired \( \mathbb{E}_{x \in S^n} \lambda(x, x)^n \) bound!

\( \lambda(x, x) := \|g\|_w / \sqrt{\|g(x)\|^2 + \|Dg(x)\|^2 + \|g\|_w^2} \) local condition number
Some extra tricks

\[ \chi(8,x) := \frac{\|\delta\|_W}{\sqrt{\|\delta(x)\|^2 + \|D_x \delta \Delta^2\|^2}} \]

\[ C(\delta,x) := \frac{\|\delta\|_\infty}{\max \{\|\Delta^{-1} \delta(x)\|_\infty, \|D_x \delta \Delta^{-1}\|^2 \}} \]

where \( \Delta := \text{diag}(d_i) \)

Row-normalization

\[ \hat{\delta} := (\delta / \|\delta\|_\infty) \]
Main Result
Probabilistic Model

\[ g \in \mathcal{H}_d[n] \text{ dobro random system} \]

\[ g_i = \sum_{\alpha} \sqrt{\alpha} c_{i,\alpha} x^\alpha \]

with \( c_{i,\alpha} \) independent, centered

i.e. \( \mathbb{E} c_{i,\alpha} = 0 \)

subgaussian with cte. \( \leq K \)

i.e. \( \mathbb{E} |c_{i,\alpha}|^e \leq K_i e \epsilon^{e/2} \) for \( e \geq 1 \)

anticoncentration cte. \( \leq \rho \)

i.e. \( \mathbb{P}(|c_{i,\alpha} - t| \leq \epsilon) \leq 2 \rho e \) for \( t \in \mathbb{R} \)

It generalizes KSS random systems where \( c_{i,\alpha} \sim\mathcal{N}(0,1) \)

We also have a smoothed version!
Main Theorem

There is a deterministic, numerically stable algorithm aCKMW that given $\mathcal{G} \in \mathcal{H}_d[\nu]$ computes $\# \varepsilon_p(\mathcal{G})$ and such that

$$E_{\mathcal{G}} \text{ run-time} (a\text{CKMW}, \mathcal{G}) \leq 2^{O(n \log n)} \, D \, N + 2^{O(n \log n)} \, 2.5 \, D \, (N+D)$$

for $\mathcal{G}$ dobro random (with bounded parameters) 'GOOD' PROBABILISTIC RUN-TIME

where $D := \max d_i$, $D := \prod d_i$,

$$N := \sum \left( \frac{n+d_i}{n} \right) = \# \text{ of zero & non-zero coeff. of } \mathcal{G}$$

+ PARALLELIZABLE
Future Work

Homology computation of semialgebraic sets

Post-processing step has to be improved!

We can produce correct adaptive samples!
Tusen Takk!

Norway | Tromsø, June 7-11, 2021