How many Zeros of a Random Sparse Polynomial are Real?

Gorav Jindal	Aalto University
Anurag Pandey	MPI-INF, Saarbrücken
Himanshu Shukla	MPI-INF, Saarbrücken
Charilaos Zisopoulos	Saarland University

ISSAC 2020

Zeros of Sparse Polynomials

- Univariate polynomials over $\mathbb{R}$
- $f = a_0 + a_1 x + \cdots + a_d x^d$, $a_i \in \mathbb{R}$, degree d.
- Sparsity, $k :=$ the **number of non-zero coefficients** of f.
- We say that a polynomial f is sparse, when $k \ll d$.
- A lot of polynomials that naturally occur in theory and practice are sparse.
- Real zeros (roots) of polynomials: interesting from the point of view of both theory and application in science, engineering and mathematics.

Goal: Understanding the number of real zeros of sparse polynomials and generalizations.

Does less terms imply fewer zeros?

- Descartes'1637 rules of signs: Bounds the number of non-zero real roots
  - Looks at number of sign changes in the coeff. seq. $(a_d, a_{d-1}, \ldots, a_0) := S(f(x))$.
  - Number of positive real roots of f, $Z_{>0}(f(x))$ is upper bounded by $S(f(x))$
  - $Z_{>0}(f(x)) \leq S(f(x)) \leq k-1$
  - $Z_{<0}(f(x)) \leq S(f(-x)) \leq k-1$.
- Total number of non-zero real zeros (with multiplicities), $Z(f)$ is bounded by $2k-2$.

Does less terms imply fewer zeros?

- Descartes'1637 rules of signs: Bounds the number of non-zero real roots
  - Looks at number of sign changes in the coeff. seq. $(a_d, a_{d-1}, \ldots, a_0) := S(f(x))$.
  - Number of positive real roots of f, $Z_{>0}(f(x))$ is upper bounded by $S(f(x))$
  - $Z_{>0}(f(x)) \leq S(f(x)) \leq k-1$
  - $Z_{<0}(f(x)) \leq S(f(-x)) \leq k-1$.
- Total number of non-zero real zeros (with multiplicities), $Z(f)$ is bounded by $2k-2$.

  Generalizations?

Does less terms imply fewer zeros?

- Descartes'1637 rules of signs: Bounds the number of non-zero real roots
  - Looks at number of sign changes in the coeff. seq. $(a_d, a_{d-1}, \ldots, a_0) := S(f(x))$.
  - Number of positive real roots of f, $Z_{>0}(f(x))$ is upper bounded by $S(f(x))$
  - $Z_{>0}(f(x)) \leq S(f(x)) \leq k-1$
  - $Z_{<0}(f(x)) \leq S(f(-x)) \leq k-1$.
- Total number of non-zero real zeros (with multiplicities), $Z(f)$ is bounded by $2k-2$.

  Generalizations?
- Kushnirenko'70s: topological complexity vs algebraic complexity.
  - Coined the term "fewnomials" for sparse polynomials.
  - Initiated the study of system of sparse multivariate polynomial equations.
  - Notable works by Khovanskii, Bihan and Sottile.

Real Tau Conjecture

Real Tau Conjecture [Koiran'11]

Consider Sums of Products of Sparse polynomials

$$f = \sum_{i=1}^{m} \prod_{j=1}^{t} f_{ij},$$

where sparsity of $f_{ij} \leq k$. Then $Z(f)$ is bounded by $poly(mkt)$.

Real Tau Conjecture

Real Tau Conjecture [Koiran'11]

Consider Sums of Products of Sparse polynomials

$$f = \sum_{i=1}^{m} \prod_{j=1}^{t} f_{ij},$$

where sparsity of $f_{ij} \leq k$. Then $Z(f)$ is bounded by $poly(mkt)$.

- another conjecture where algebraic complexity dictates topological complexity.
- Implies superpoly lower bound on circuit complexity of **permanent** polynomial.
- The algebraic analog of the $\mathrm{P} \neq \mathrm{NP}$.

Understanding the real tau conjecture

Briquel-Bürgisser: Real Tau conjecture is true on average

.
$f = \sum_{i=1}^{m} \prod_{j=1}^{t} f_{ij}$, sparsity of $f_{ij} \leq k$. When the coefficients of f_{ij} are all independent gaussian variables, then the expected number of real zeros is bounded by $O(mk^2 t)$.

Understanding the real tau conjecture

Briquel-Bürgisser: Real Tau conjecture is true on average

.
$f = \sum_{i=1}^{m} \prod_{j=1}^{t} f_{ij}$, sparsity of $f_{ij} \leq k$. When the coefficients of f_{ij} are all independent gaussian variables, then the expected number of real zeros is bounded by $O(mk^2 t)$.

What about the worst case?

Understanding the real tau conjecture

Briquel-Bürgisser: Real Tau conjecture is true on average

.
$f = \sum_{i=1}^{m} \prod_{j=1}^{t} f_{ij}$, sparsity of $f_{ij} \leq k$. When the coefficients of f_{ij} are all independent gaussian variables, then the expected number of real zeros is bounded by $O(mk^2 t)$.

What about the worst case?

- Descartes' bound gives $O(mk^t)$.
- We don't completely understand the simplest non-trivial case, i.e. the case of $fg + 1$.
  - Descartes' bound gives $O(k^2)$.
  - It's not known whether it's tight. Is it $O(k)$?

Problem tackled in this work

- We aim to improve the understanding of number of real roots of sparse polynomials.
- Towards this end, we study the number of zeros of a **random** sparse polynomial.
- We solve the problem completely when the coefficients are **independent standard normal random variables**.

Problem tackled in this work

- We aim to improve the understanding of number of real roots of sparse polynomials.
- Towards this end, we study the number of zeros of a **random** sparse polynomial.
- We solve the problem completely when the coefficients are **independent standard normal random variables**.
- $S = \{e_1, \ldots, e_k\} \subseteq \mathbb{N}$. We consider the polynomial $f_S = \sum_{i=1}^{k} a_i x^{e_i}$, where a_i's are independent standard normals.
- f_S : a k-sparse random polynomial supported on $S \subseteq \mathbb{N}, |S| = k$.
- $z_S^{\mathbb{R}}$: expected number of real roots of f_S.
- Sufficient to consider $z_S := z_S^{(0,1)}$: expected number of real roots of f_S in $(0, 1)$.

Find z_S : the expected number of real zeros of f_S in $(0, 1)$

Relevant Works

Kac'43: The dense case

If $S = \{0, 1, 2, \ldots, d\}$ then

$$z_S^{\mathbb{R}^*} = \frac{2}{\pi}\log(d) + C_1 + \frac{2}{d\pi} + O\left(\frac{1}{d^2}\right), C_1 \approx 0.6257358072\ldots.$$

Edelman-Kostlan'95: A geometric derivation

For $S = \{e_1, \ldots, e_\ell\} \subseteq \mathbb{N}$, define $v_S(t) := (t^{e_1}, t^{e_2}, \ldots, t^{e_\ell})$. For $I \subseteq \mathbb{R}$, we have:

$$z_S^I = \frac{1}{\pi}\int_I \frac{\sqrt{(\|v_S(t)\|_2 \cdot \|v_S'(t)\|_2)^2 - (v_S(t) \cdot v_S'(t))^2}}{(\|v_S(t)\|_2)^2}\mathrm{d}t.$$

Bürgisser, Ergür and Tonelli-Cueto'18: Sparse case

Let $S \subseteq \mathbb{N}$ be any set as above with $|S| = k$ then we have

$$z_S \leq \frac{1}{\pi}\sqrt{k}\log(k).$$

Summary of results

Main result: Upper bound

Let $S \subseteq \mathbb{N}$ be any set as above with $\mid S \mid = k$, then we have $z_S \leq \frac{2}{\pi}\sqrt{k-1}$.

Main result: asymptotically matching lower bound

There exists a sequence of sets $S_k \subseteq \mathbb{N}$, $S_k = \{0, 1, 2^{2^1}, 2^{2^2}, \dots, 2^{2^{k-1}}\}$ with $|S_k| = k+2$, such that for $k \geq 3$, $z_{S_k} \geq \frac{\pi - \sqrt{3}}{16\pi}\sqrt{k} + \frac{1}{7}$.

Recovering the bound in the dense case

If $S = \{0, 1, 2, \dots, n\}$ then $z_S \leq \frac{3}{4}\log_2(n)$.

Roots concentrated around 1

For a fixed $\epsilon > 0$ and any $S \subseteq \mathbb{N}$ with $|S| = k$, we have

$$z_S^{(0,1-\epsilon)} \leq \frac{1}{2\pi}\left(\log\left(\frac{2}{\epsilon}\right) + \frac{4}{\sqrt{\epsilon}} - 4\right).$$

Proof Technique

Main technical contribution: a reformulation of the integral by Edelman-Kostlan

Edelman-Kostlan

$$z_S^I = \frac{1}{\pi} \int_I \frac{\sqrt{(\|v_S(t)\|_2 \cdot \|v_S'(t)\|_2)^2 - (v_S(t) \cdot v_S'(t))^2}}{(\|v_S(t)\|_2)^2} \mathrm{d}t.$$

- **Key insight**: Above integral can be rewritten as a function of $(\|v_S(t)\|_2)^2$.
- For a set $S = \{e_1, e_2, \ldots, e_k\} \subseteq \mathbb{N}$, we define

$$g_S(t) := (\|v_S(t)\|_2)^2 = \sum_{i=1}^{k} t^{2e_i}$$

- Then we have the following equalities:

$v_S(t) = (t^{e_1}, t^{e_2}, \ldots, t^{e_k})$

$$v_S(t) \cdot v_S'(t) = \sum_{i=1}^{k} e_i t^{2e_i - 1} = \frac{g_S'(t)}{2}$$

$$(\|v_S'(t)\|_2)^2 = \frac{1}{4} g_S''(t) + \frac{1}{4t} g_S'(t)$$

- For all sets $S \subseteq \mathbb{N}$ let us define

$$\mathscr{I}(g_S(t)) := \frac{g_S''(t)}{g_S(t)} - \left(\frac{g_S'(t)}{g_S(t)}\right)^2 + \frac{g_S'(t)}{t g_S(t)}.$$

- Then we have the following reformulation:

Reformulaion of Edelman-Kostlan

For all sets $S \subseteq \mathbb{N}$, we have the following equality for z_S^I

$$z_S^I = \frac{1}{2\pi} \int_I \sqrt{\mathscr{I}(g_S(t))} \mathrm{d}t,$$

- **Strength:** Allows us to bound $z_{S_1 \uplus S_2}$ in terms of z_{S_1} and z_{S_2}.
- We can build the set S starting from a singleton while controlling the growth of z_S throughout.

Proposition: For singleton sets S, we have $\mathscr{I}(g_S) = 0 \implies z_S = 0$.

$$\mathscr{I}(g_{S_1 \uplus S_2}) = \mathscr{I}(g_{S_1} + g_{S_2})$$

$$= \frac{g''_{S_1} + g''_{S_2}}{g_{S_1} + g_{S_2}} - \left(\frac{g'_{S_1} + g'_{S_2}}{g_{S_1} + g_{S_2}}\right)^2 + \frac{1}{t}\left(\frac{g'_{S_1} + g'_{S_2}}{g_{S_1} + g_{S_2}}\right)$$

$$= \frac{g_{S_1}}{g_{S_1} + g_{S_2}} \cdot \mathscr{I}(g_{S_1}) + \frac{g_{S_2}}{g_{S_1} + g_{S_2}} \cdot \mathscr{I}(g_{S_2}) + \frac{1}{g_{S_1} g_{S_2}} \left(\frac{g_{S_1} {g_{S_2}}' - g_{S_2} {g_{S_1}}'}{g_{S_1} + g_{S_2}}\right)^2$$

Incrementally increasing the sparsity

Let $S \subseteq \mathbb{N}$ be a set with $0 \in S$ and $|S| = k$. If $a \in \mathbb{N}$ is such that $a > \max(S)$ then

$$z_{S \cup \{a\}} \leq z_S + \frac{1}{\pi} \arctan\left(\frac{1}{\sqrt{k}}\right)$$

$$\sqrt{\mathscr{I}(g_{S \cup \{a\}}(t))} = \sqrt{\frac{g_S \cdot \mathscr{I}(g_S)}{g_S + g_{\{a\}}} + \frac{g_{\{a\}} \cdot \mathscr{I}(g_{\{a\}})}{g_S + g_{\{a\}}} + \frac{1}{g_S g_{\{a\}}} \left(\frac{g'_{\{a\}} g_S - g'_S g_{\{a\}}}{g_S + g_{\{a\}}}\right)^2}$$

$$\leq \sqrt{\mathscr{I}(g_S(t))} + 0 + \sqrt{\frac{1}{g_S g_{\{a\}}} \left(\frac{g'_{\{a\}} g_S - g'_S g_{\{a\}}}{g_S + g_{\{a\}}}\right)^2}$$

Proof Sketch

- Since $g'_{\{a\}} g_S - g'_S g_{\{a\}} > 0$ in $(0, 1)$, we have

$$z_{S \cup \{a\}} = z_S + \frac{1}{2\pi} \int_0^1 \frac{1}{\sqrt{g_S g_{\{a\}}}} \left(\frac{g'_{\{a\}} g_S - g'_S g_{\{a\}}}{g_S + g_{\{a\}}} \right) \mathrm{d}t$$

- Now we use the substitution $u = \sqrt{\frac{g_{\{a\}}}{g_s}}$ to obtain

$$\int_0^1 \frac{1}{\sqrt{g_S g_{\{a\}}}} \left(\frac{g'_{\{a\}} g_S - g'_S g_{\{a\}}}{g_S + g_{\{a\}}} \right) \mathrm{d}t = 2 \int_\alpha^\beta \left(\frac{1}{1 + u^2} \right) \mathrm{d}u$$

- where $\alpha = \sqrt{\frac{g_{\{a\}}(0)}{g_s(0)}} = 0$ and $\beta = \sqrt{\frac{g_{\{a\}}(1)}{g_s(1)}} = \frac{1}{\sqrt{k}}$. Hence, we have

$$z_{S \cup \{a\}} \leq z_S + \frac{1}{\pi} \arctan \left(\frac{1}{\sqrt{k}} \right).$$

Finishing the proof..

Proposition: For all sets S of size two, $z_S = \frac{1}{4}$.

- Thus, starting with a set of size 2, we may always add the highest element iteratively and obtain that

$$z_S \leq \frac{1}{4} + \frac{1}{\pi}\sum_{i=2}^{k-1} \arctan\left(\frac{1}{\sqrt{i}}\right)$$

- We use the following well-known inequality

$$\arctan(x) < x \text{ for all } x > 0.$$

- This implies that

$$z_S - \frac{1}{4} \leq \frac{1}{\pi}\sum_{i=2}^{k-1}\frac{1}{\sqrt{i}} \leq \frac{1}{\pi}\int_1^{k-1}\sqrt{\frac{1}{x}}\mathrm{d}x$$
$$= \frac{2}{\pi}(\sqrt{k-1} - 1).$$

Summary of results

Upper bound

Let $S \subseteq \mathbb{N}$ be any set as above with $\mid S \mid = k$, then we have $z_S \leq \frac{2}{\pi}\sqrt{k-1}$.

Asymptotically matching lower bound

There exists a sequence of sets $S_k \subseteq \mathbb{N}$, $S_k = \{0, 1, 2^{2^1}, 2^{2^2}, \ldots, 2^{2^{k-1}}\}$ with $|S_k| = k+2$, such that for $k \geq 3$, $z_{S_k} \geq \frac{\pi - \sqrt{3}}{16\pi}\sqrt{k} + \frac{1}{7}$.

Recovering the bound in the dense case

If $S = \{0, 1, 2, \ldots, n\}$ then $z_S \leq \frac{3}{4}\log_2(n)$.

Roots concentrated around 1

For a fixed $\epsilon > 0$ and any $S \subseteq \mathbb{N}$ with $|S| = k$, we have $z_S^{(0,1-\epsilon)} \leq \frac{1}{2\pi}\left(\log\left(\frac{2}{\epsilon}\right) + \frac{4}{\sqrt{\epsilon}} - 4\right)$

Thanks for your attention :)