# Average minus Weak Complexity equals Adaptive Algorithms

Josué TONELLI-CUETO (BMS & TU Berlin)
based in discussions and work with
Peter BÜRGISSER, Felipe CUCKER and Alperen A. ERGÜR

August 23, 2019

*Complexity of Numerical Computation*
A conference in honor of Felipe Cucker

Slides at https://tonellicueto.xyz/pdf/felipefest2019_slides.pdf

This talk is based on discussions and work with

- · Peter Bürgisser (Technische Universität Berlin),
- · Felipe Cucker (CityU Hong Kong), and
- · Alperen Ergür (Technische Universität Berlin).

To be more developed and written out in

- · future work, and
- · chapter 5 of my PhD thesis
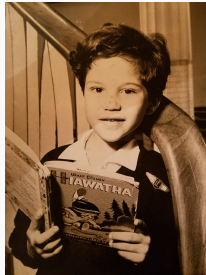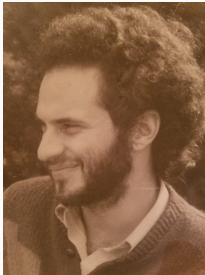  "Condition and Homology in Semialgebraic Geometry"
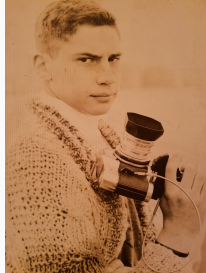
All this supported by



**Einstein Stiftung Berlin**
**Einstein Foundation Berlin,**

within the Einstein Visiting Fellowship "Complexity and accuracy of numerical algorithms in algebra and geometry" of Felipe Cucker.

Thanks to Antonia!

符号**算法**、
数值**算法**，
只要能解决未解决的问题就
是好**算法**。

## Table of contents

# Grid method until now

## Grid method

Sketch of the grid method:

- Put a grid of points approaching the space
- Check a condition in all points of the grid
- Refine if the condition is not satisfied at some of the points
    - Non-adaptive: Uniform refinement of the whole grid
    - Adaptive: Refinement only around the points where something goes wrong

Philosophy of the method:

- *divide et impera*
- Smaller operations lead to smaller errors and errors cannot sequentially accumulate

## Grid method in real algebraic geometry I

(CS98) (Cucker & Smale, 1998)
Grid method is introduced for testing feasibility of a semialgebraic set
Condition-based analysis, but no probabilistic estimate

(C99) (Cucker, 1999)
Improvements over CS99 are done
Condition number $\kappa(f)$ is introduced

(CKMW1) (Cucker, Krick, Malajovich & Wschebor, 2008)
Counting zeros of real algebraic set: Condition-based complexity

(CKMW2) (Cucker, Krick, Malajovich & Wschebor, 2009)
Counting zeros of real algebraic set: Condition number theorem and probabilistic analysis

(CKMW3) (Cucker, Krick, Malajovich & Wschebor, 2012)
Counting zeros of real algebraic set: Alternative probabilistic analysis (without using conic structure)

write the "bible of condition"
Grid method appears in Chapter 19

# Grid method in real algebraic geometry III

(CKS16)  (Cucker, Krick & Shub, 2016)
        Numerical computation of homology of real projective sets
        [using the reach together with the approximation theorem of
        (Niyogi, Smale & Weinberger, 2008)]

 (BCL17)  (Bürgisser, Cucker & Lairez, 2017)
        ...of basic semialgebraic sets

(BCTC1)  (Bürgisser, Cucker & T.-C., 2019)
        ...of closed semialgebraic sets

(BCTC2)  (Bürgisser, Cucker & T.-C., 2019)
        ...of arbitrary semialgebraic sets
        [using construction of (Grabrielov & Vorobjov, 2009)]

## Successes and defeats

Successes:

- Algorithms work in single exponential time, i.e., $\mathrm{poly}(q, D)^{\mathrm{poly}(n)}$, with high probability (weak complexity)
- Highly parallelizable algorithm

Defeats:

- Algorithm does not work for all inputs...
  (inevitable in numerical algorithms)
  We can deal with it with an hybrid approach: Under a condition threshold go numerical, and above the threshold go symbolic.
- Infinite expectation (unlike the complex case)

# Plantinga-Vegter algorithm enters the game

## Setting

What do we have?

- An implicit curve $C$ inside $[-a, a]^2$
  given by a $C^1$ function $f : [-a, a]^2 \to \mathbb{R}$
- Interval approximations $\Box f$ of $f$ and $\Box \partial f$ of $\partial f$

What do we want?

- Piecewise-linear approximation $L$ of $C$ in $[-a, a]^2$ such that
  $([-a, a]^2, C)$ and $([-a, a]^2, L)$ are isotopic

Any assumptions?

- $C$ smooth
- $C$ Intersects the boundary of $[-a, a]^2$ transversely

**Algorithm:** PV Algorithm for curves (Plantinga, Vegter; 2004)

**Input:** $a \in (0, \infty)$ and $f : \mathbb{R}^2 \to \mathbb{R}$
with interval approximations $\square[f]$ and $\langle \square[\partial f], \square[\partial f] \rangle$

---

SUBDIVISION:
Starting with the trivial subdivision $\mathcal{S} := \{[-a, a]^n\}$, repeatedly
subdivide each $J \in \mathcal{S}$ into 4 squares until for all $J \in \mathcal{S}$,

$$0 \notin \square f(J) \text{ or } 0 \notin \langle \square \partial f(J), \square \partial f(J) \rangle$$
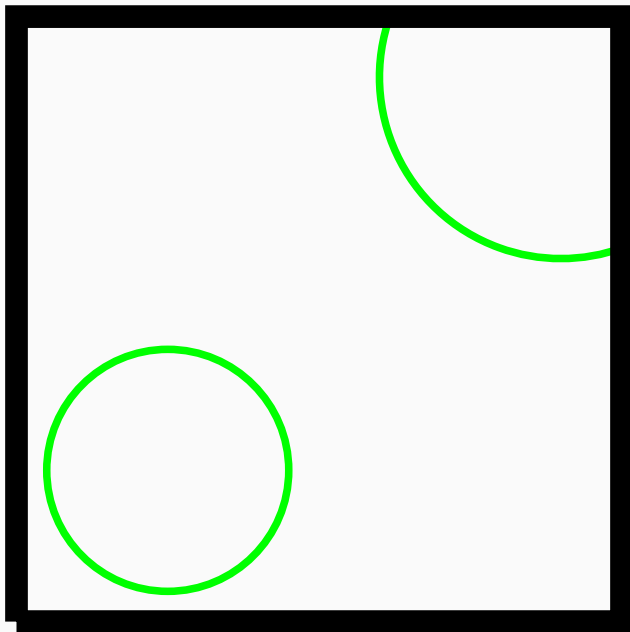
CONSTRUCTION:
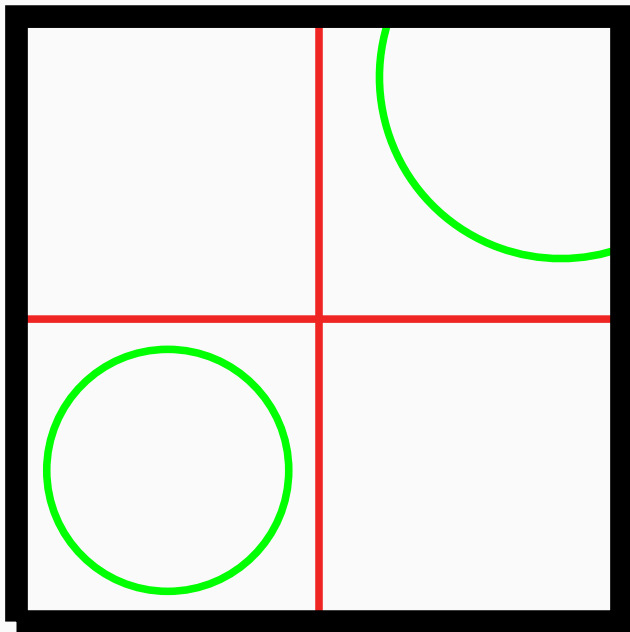Construct piecewise-linear curve $L$
joining the midpoints of "small" edges of each $J \in \mathcal{S}$ with oposite
$f$-signs at their vertices

---

**Output:** Piecewise-linear approximation $L$ of $C = f^{-1}(0) \cap [-a, a]^2$
isotopic to it
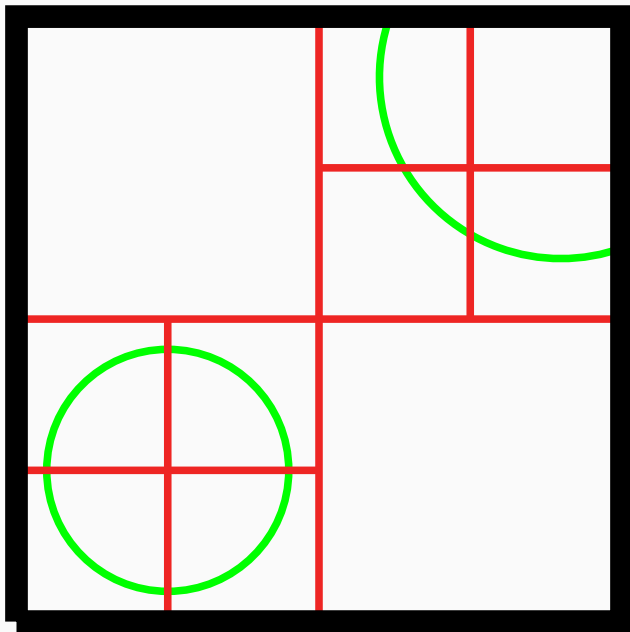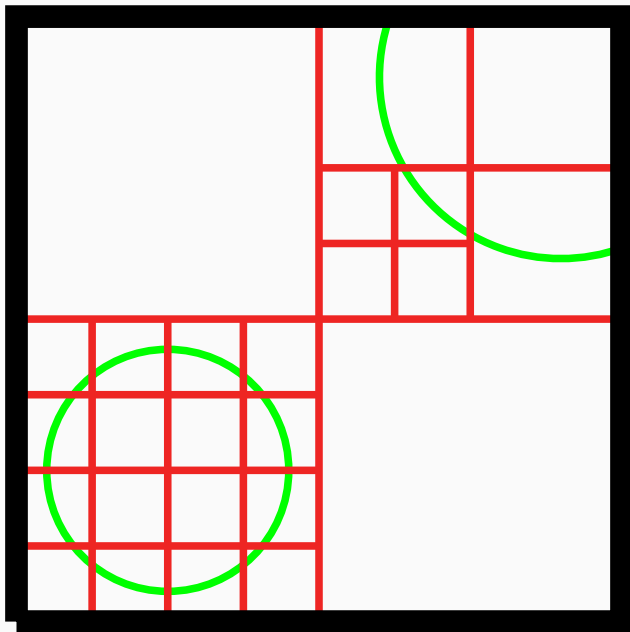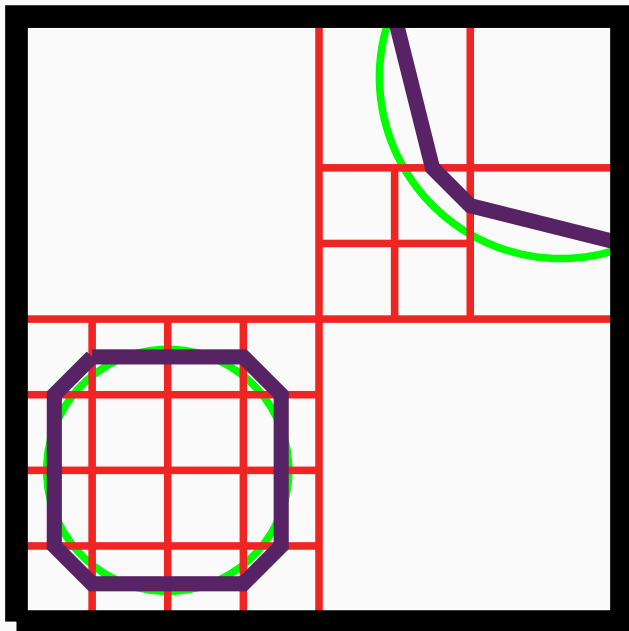
1. (Plantinga, Vegter; 2004) introduced the algorithm for curves and surfaces
   Very efficient in practice
2. (Burr, Gao & Tsigaridas, ISSAC2017) generalized the subdivision method to higher dimensions
   but no construction of the piecewise-linear approximation...

Note that...

cost(subdivision algorithm) $\sim \#$ (subdivisions) $\cdot$ cost(evaluations)

**Algorithm:** Subdivision of PV Algorithm (Burr, Gao & Tsigaridas, ISSAC2017)

**Input:** $a \in (0, \infty), f : \mathbb{R}^n \to \mathbb{R}$

with interval approximations $\square[hf]$ and $\square[h'\partial f]$

for some functions $h, h' : \mathbb{R}^n \to (0, \infty)$

Starting with the trivial subdivision $\mathcal{S} := \{[-a, a]^n\}$, repeatedly
subdivide each $J \in \mathcal{S}$ into $2^n$ cubes until the condition

$$C_f(J) : 0 \notin \square[hf](J) \text{ or } 0 \notin \langle \square[h'\partial f], \square[h'\partial f] \rangle$$

holds for all $J \in \mathcal{S}$

**Output:** Subdivision $\mathcal{S} \subseteq \mathcal{I}_n$ of $[-a, a]^n$
    such that for all $J \in \mathcal{S}$, $C_f(J)$ is true

$h, h'$ depend on the setting and the interval arithmetic one uses

17

**Idea**: Give a bound to the size of the smallest box containing a point not satisfying the condition.

**Definition (Burr, Gao & Tsigaridas, ISSAC2017)**

A *local size bound* for $f$ is a function $b_f : \mathbb{R}^n \to [0, \infty)$ such that for all $x \in \mathbb{R}^n$,

$$b_f(x) \leq \inf\{\operatorname{vol}(J) \mid x \in J \in \mathcal{I}_n \text{ and } C_f(J) \text{ false}\},$$

where $C_f(J) : 0 \notin \square[hf](J)$ or $0 \notin \langle \square[h'\partial f], \square[h'\partial f] \rangle$

**Idea**: Boxes cover the cube. The volume of the boxes should add to that of the cube. Volume of the boxes is at least number of boxes times volume of smallest box.

### Proposition (Burr, Gao & Tsigaridas, ISSAC2017)

*The number of n-cubes of the final subdivision of the subdivision of the PV algoirhtm on input $(f, a)$, regardless of how the subdivision step is done, is at most*

$$(2a)^n / \inf\{b_f(x) \mid x \in [-a, a]^n\} \qquad \Box$$

*where $b_f$ is a local size bound for f.*

(Burr, Gao & Tsigaridas, ISSAC2017) construct a local size bound for an integer polynomial $f \in \mathcal{P}_{n,d}$ and obtain the bound

$$2^{\mathcal{O}(nd^{n+1}(n\tau + nd \log{(nd)} + 9n + d) \log a)}$$

where $\tau$ is the bit-size of the coefficients.

**Refinement of the idea**: Not all boxes have the same size. "$\frac{dx}{b_f(x)}$ is, up to constant, the infinitesimal number of boxes needed to cover $x$." Formalized by (Burr, 2016) using the technique known as *continuous amortization* introduced in (Burr, Krahmer & Yap, 2009)

**Theorem (Burr, Gao & Tsigaridas, ISSAC2017)**
*The number of n-cubes of the final subdivision of the PV algorithm on input $(f, a)$ is at most*

$$\max\left\{1, \int_{[-a,a]^n} \frac{2^n}{b_f(x)} \, dx\right\}$$

*where $b_f$ is a local size bound for f. Moreover, the bound is finite if and only if the algorithm terminates.*  □

Techniques of (Burr, Gao & Tsigaridas, ISSAC2017) cannot exploit it!

## Can we exploit it?

(Burr, Gao & Tsigaridas, ISSAC2017) said...

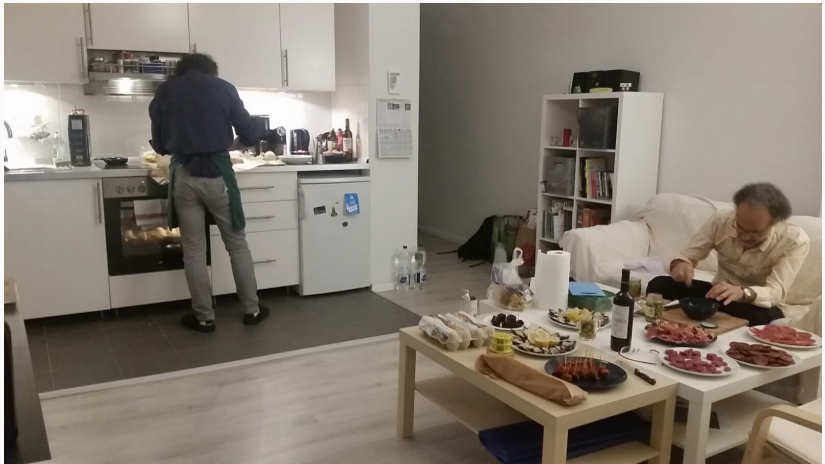> *Even though our bounds are optimal, in practice, these are quite pessimistic [...]*

and

> *Since the complexity of the algorithm can be exponential in the inputs [size], the integral must be described in terms of additional geometric and intrinsic parameters.*

What can be a solution to these issues?

### Condition numbers

# A condition-based analysis of Plantinga-Vegter

**Definition (Cucker)**

Let $f \in \mathcal{H}_{n,d}[q]$, the *local condition number* of $f$ at $y \in \mathbb{S}^n$ is

$$\kappa(f, y) := \frac{\|f\|}{\sqrt{f(y)^2 + \sigma_q(\Delta^{-1}\mathrm{D}_y f)^2}}$$

where $\|f\|$ is the Weyl norm of $f$, $\sigma_q$ the $q$th singular value and $\mathrm{D}_y f$ the tangent map with respect the sphere.

Given $f \in \mathcal{P}_{n,d}[q]$, the *local affine condition number* of $f$ at $x \in \mathbb{R}^n$ is

$$\kappa_{\text{aff}}(f, x) := \kappa(f^{\mathsf{h}}, \phi(x))$$

where $f^{\mathsf{h}}$ is the homogeneization of $f$ and $\phi : x \mapsto \frac{1}{\sqrt{1+\|x\|^2}} \begin{pmatrix} 1 \\ x \end{pmatrix}$.

Unfortunately the theory is developed for the homogenous setting. Too many translations!

## Observation

$\kappa(f, x) = \infty$ iff $V_{\mathbb{S}}(f)$ has a singularity at $x$

Ergo $\kappa(f, x)$ says how near $V_{\mathbb{S}}(f)$ is of having a singularity at $x$. Can we be more concrete? YES!

## Theorem (Condition Number Theorem)

*Let $x \in \mathbb{S}^n$ and*

$$\Sigma_x := \{g \in \mathcal{H}_{n,d}[q] \mid g(x) = 0,\ \mathrm{rank} D_x g < q\},$$

*i.e., $\Sigma_x$ is the set of $f$ having $x$ as a singular point. Then for every $f \in \mathcal{H}_{n,d}[q]$,*

$$\frac{\|f\|}{\kappa(f, x)} = \mathrm{dist}(f, \Sigma_x)$$

*where $\|\cdot\|$ is the Weyl norm of $\mathcal{H}_{n,d}[q]$ and the distance is the induced by the Weyl norm of $\mathcal{H}_{n,d}[q]$.* □

The wanted "additional geometric and intrinsic parameter"

Theorem (Cucker, Ergür & T.-C., ISSAC2019)
*Let $f \in \mathcal{P}_{n,d}[1]$. Then*

$$x \mapsto 1/ \left( 2^{5/2} dn \kappa_{\mathrm{aff}}(f, x) \right)^n$$

*is a local size bound for f with the interval approximation of (Cucker, Ergür & T.-C., ISSAC2019), and*

$$x \mapsto 1/ \left( 2^{3n} d^2 \kappa_{\mathrm{aff}}(f, x) \right)^n$$

*with the interval approximation of Remark 2.2. of (Burr, Tsigaridas, Yap; ISSAC2017)* □

## Condition-based cost (Using 1st bound)

**Theorem (Cucker, Ergür & T.-C., ISSAC2019)**

*The number of n-cubes in the final subdivision of the subdision of
the* PV *algorithm on input* $(f, a)$ *is at most*

$$d^n \max\{1, a^n\} 2^{n \log n + 9n/2} \max_{x \in [-a,a]^n} (\kappa_{\text{aff}}(f, x)^n)$$

*if the interval approximation is that of (Cucker, Ergür & T.-C.,
ISSAC2019), and at most*

$$d^{2n} \max\{1, a^n\} 2^{3n^2 + 2n} \max_{x \in [-a,a]^n} (\kappa_{\text{aff}}(f, x)^n)$$

*if the interval approximation is that of (Burr, Tsigaridas, Yap;
ISSAC2017).*  □

Note that

$$\max_{x \in [-a,a]^n} (\kappa_{\text{aff}}(f, x)^n) \leq \kappa(f^\natural) := \max_{x \in \mathbb{S}^n} \kappa(f^\natural, x).$$

**Theorem (Cucker, Ergür & T.-C., ISSAC2019)**

*The number of n-cubes in the final subdivision of the subdision of the* PV *algorithm on input* $(f, a)$ *is at most*

$$d^n \max\{1, a^n\} 2^{n \log n + 9n/2} \, \mathbb{E}_{x \in [-a,a]^n} \left( \kappa_{\text{aff}}(f, x)^n \right)$$

*if the interval approximation is that of (Cucker, Ergür & T.-C., ISSAC2019), and at most*

$$d^{2n} \max\{1, a^n\} 2^{3n^2 + 2n} \, \mathbb{E}_{x \in [-a,a]^n} \left( \kappa_{\text{aff}}(f, x)^n \right)$$

*if the interval approximation is that of (Burr, Tsigaridas, Yap; ISSAC2017).* □

Note that

$$\mathbb{E}_{x \in [-a,a]^n} \left( \kappa_{\text{aff}}(f, x)^n \right) \leq \mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f^{\pitchfork}, x)^n \right)$$

Can

$$\mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f, x)^n \right)$$

be better than

$$\kappa(f) := \max_{x \in \mathbb{S}^n} \kappa(f, x)?$$

Yes!

### Definition (Cucker, Ergür, T.C.; ISSAC2019)

A *dobro random polynomial* $f \in \mathcal{H}_{n,d}[1]$ with parameters $K$ and $\rho$ is a polynomial

$$f := \sum_{|\alpha|=d} \binom{d}{\alpha}^{1/2} c_\alpha X^\alpha$$

such that the $c_\alpha$ are independent random variables such that

P1 $\mathbb{E}c_\alpha = 0$ (centered),

P2 $(\mathbb{E}|c_\alpha|^p)^{\frac{1}{p}} \leq K\sqrt{p}$ for $p \geq 1$ (subgaussian with $\Psi_2$-norm $\leq K$), and

P3 $\max_{u\in\mathbb{R}}\{\mathbb{P}(|c_\alpha - u| \leq \varepsilon)\}$ (anti-concentration with constant $\rho$).

A *dobro random polynomial* $f \in \mathcal{P}_{n,d}[1]$ is a polynomial $f$ such that its homogenization $f^h$ is so.

## Dobro random polynomials II

**Examples of dobro random polynomials**:

N *KSS random polynomial*:  <span style="color:red">(KSS=Kostlan-Smale-Shub)</span>
- $c_\alpha$ is Gaussian with unit variance
- $K\rho = 1/\sqrt{2\pi}$

U *Weyl random polynomial*:
- $c_\alpha$ is uniform distribution in $[-1, 1]$
- $K\rho \leq 1$

E A *p-random polynomial*:
- $c_\alpha$ has density function $\delta_p e^{-|t|^p}$ where $\delta_p$ being the appropriate constant and $p \geq 2$

**Note**: Dobro (добро) is a Russian adjective derived from добрый (dóbryj) which means kind, good, genial, gentle, soft, etc.

## Theorem (Cucker, Ergür, T.C.; ISSAC2019)

*Let $f \in \mathcal{H}_{n,d}[1]$ be a dobro random polynomial with parameters $K$ and $\rho$. Then*

$$\mathbb{E}_f \mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f^h, x)^n \right) \leq d^{\frac{n^2+n}{2}} 2^{\frac{n^2+3 \log n+9}{2}} (c_1 c_2 K \rho)^{n+1}$$

*where $c_1, c_2$ are universal constants.* □

### Corollary

*Plantinga-Vegter algorithm has average polynomial time in the degree.*

With an improvement of our condition-based techniques (i.e. changing to the max norm) we can eliminate the $n^2$ of the exponent! Also available in smoothed form!

# Interlude: Back to curves

$$\mathcal{O}\left(d^5\right) \text{ and } \mathcal{O}\left(d^6\right)$$

with the interval arithmetic of, respectively, (Cucker, Ergür & T.-C., ISSAC2019) and (Burr, Gao, & Tsigaridas, ISSAC2017)

$$\mathcal{O}\left(d^{5/2}\log^{5/2}d\right) \text{ and } \mathcal{O}\left(d^3\log^3 d\right)$$

with the interval arithmetic of, respectively, (Cucker, Ergür & T.-C., ISSAC2019) and (Burr, Gao & Tsigaridas, ISSAC2017) With the new techniques! Compare to bit-complexity

$$\tilde{\mathcal{O}}(d^5\tau + d^6)$$

of the deterministic algorithm of (Kobel & Sagraloff, 2015) and (Diatta, Diatta, Rouillier, Roy & Sagraloff; 2018)

*Some comments*:

- Difference should be careful, complexity measured in different ways! But still, why so efficient?!
- PV algorithm does not work with singular curves, although there is work in this direction (Burr, Choi, Galehouse & Yap, 2012)
- Can we develop an hybrid approach working on all inputs with the same worst case, but faster on average?

# A condition for adaptive grids?

## A new hope

**Proposition (T.-C., Thesis)**
*Let $f \in \mathcal{H}_{n,d}[q]$. Then*

$$\kappa(f) < \infty \text{ iff } \mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f, x)^n \right) < \infty.$$

$\square$

**Theorem (T.-C., Thesis)**
*Let $f \in \mathcal{H}_{n,d}[q]$ be KSS and $\alpha < 1 + 1/n$. Then*

$$\mathbb{E}_f \left( \mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f, x)^n \right) \right)^{\alpha} = \mathcal{O} \left( \frac{qD}{\frac{n+1}{n} - \alpha} \right)^{\mathcal{O}(n^2)}$$

$\square$

Dependence on $\alpha$ cannot be improved. This can be shown for hypersurfaces...
Using the max norm, we can get rid of the square in the exponent

### Theorem (Han, 2019)

*Let $\Lambda : \mathbb{S}^n \to (0, 1]$ be a L-Lipschitz map. Then there is an adaptive grid algorithm computing a grid $\mathcal{G} \subseteq \mathbb{S}^n$ and $r \in \mathbb{R}_>^{\mathcal{G}}$ such that*

(1) $\{B(x, r_x) \mid x \in \mathcal{G}\}$ *covers* $\mathbb{S}^n$

(2) *For all $x \in \mathcal{G}$, $r_x \leq \Lambda(x)$*

*and whose complexity is*

$$\mathcal{O}\left(2 + L\right)^n \mathbb{E}_{x \in \mathbb{S}^n}\left(\Lambda(x)^{-n}\right).$$

*In particular, the $\min_{x \in \mathbb{S}^n} \Lambda(x)$ can be computed in the above time.*

Note that

1. It's easy to see that the above bound is optimal
2. Note similarity with adaptive homotopy continuation

35

# Fast computation of $\kappa(f)$

Easy observation: $x \mapsto \kappa(f, x)^{-1}$ is $D$-Lipschitz

### Theorem (T.-C., Thesis)
*There is an adaptive grid algorithm that computes $\kappa(f)$ in average single exponential time, when $f \in \mathcal{H}_{n,d}[q]$ be KSS.*

Observe that funnily $\kappa(f)$ does not have finite expectation.

### Corollary (T.-C., Thesis)
*There is an adaptive grid algorithm that computes a lower bound of the reach of an spherical algebraic set $\mathcal{Z}(f)$ in average single exponential time, when $f \in \mathcal{H}_{n,d}[q]$ be KSS.*

The condition used is quadratic on $\kappa(f, x)$, because of the use of Smale's $\alpha$-theory. Assuming all TDA techniques work, this gives a condition-based bound with

$$\mathbb{E}_{x \in \mathbb{S}^n} \left( \kappa(f, x)^{2n} \right)$$

for basic semialgebraic sets. This does not give finite expectation!

For general semialgebraic sets, the above cannot even be obtained yet, due to even more technical problems to solve.

## More work to do!
Is there a numerical algorithm
computing homology of semialgebraic sets
in average single exponential time?

Can we develop a hybrid symbolic-numerical algorithm computing homology of semialgebraic sets in average single exponential time and worst-case doubly exponential time?

Guztiagatik eskerrik asko!