On the bit-size of non-radical triangular sets

Dahan Xavier

Ochanomizu Univeristy, Faculty of General Educational Research

MACIS 2017, November 15th. Vienna

The problem I (Notation)

In this talk, a triangular set is:

- a lexicographic Gröbner basis (lex. G.b.) of dimension zero. (monomial order is lexicographic with $x_1 \prec \cdots \prec x_n$.)
- with as many polynomials as variables:

$$\begin{array}{rcl} T_n(x_1, x_2, x_3, \ldots, x_{n-2}, x_{n-1}, x_n) & = & x_n^{d_n} + \cdots \\ T_{n-1}(x_1, x_2, \ldots, x_{n-2}, x_{n-1}) & = & x_{n-1}^{d_{n-1}} + \cdots \\ \ddots & \vdots & \vdots \\ T_2(x_1, x_2) & = & x_2^{d_2} + \cdots \\ T_1(x_1) & = & x_1^{d_1} + \cdots \end{array}$$

- $d_i := \deg_{x_i}(T_i)$. The product $d_1 \cdots d_n$ is the (multi)degree of T.

Rmk: In general, a lex. G.b.of dimension 0 may have more polynomials than variables.

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients?

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients? in function of:

1. number of variables n

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients? in function of:

1. number of variables n
2. coefficients of input system $\mathbf{f}$ denoted $h(\mathbf{f})$

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients? in function of:

1. number of variables n
2. coefficients of input system $\mathbf{f}$ denoted $h(\mathbf{f})$
3. its total degree, multiplicity of solutions $\deg(\mathbf{f})$, $\mu(\mathbf{f})$.

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients? in function of:

1. number of variables n
2. coefficients of input system $\mathbf{f}$ denoted $h(\mathbf{f})$
3. its total degree, multiplicity of solutions $\deg(\mathbf{f})$, $\mu(\mathbf{f})$.

Rmk: Typical question in Symbolic Computation where coefficients are exact thereore often very large.

- Extended Euclidean Algorithm (subresultant)
- Mignotte's factor bound...
- Gaussian elimination, determinant (Hadamard's inequality)

The problem II (bit-size growth estimation)

Given a polynomial system $\mathbf{f} = (f_1, \ldots, f_s) \in \mathbb{Q}[x_1, \ldots, x_n]$,
such that its lex. G.b.is a triangular set,
how much grow the coefficients? in function of:

1. number of variables — n
2. coefficients of input system $\mathbf{f}$ — denoted $h(\mathbf{f})$
3. its total degree, multiplicity of solutions — $\deg(\mathbf{f})$, $\mu(\mathbf{f})$.

Rmk: Typical question in Symbolic Computation where coefficients are exact thereore often very large.

- Extended Euclidean Algorithm (subresultant)
- Mignotte's factor bound...
- Gaussian elimination, determinant (Hadamard's inequality)

. . . *quite* more difficult for polyomial systems

New results

1. Structure of non-radical triangular sets
    - Interpolation formula that extends univariate Hermite interpolation.
    - Introduce a related system denoted N based with smaller coefficients. **!** Difficult to compute from T **!**.
    - Extends known results on the radical case (D. & Schost'2004).

New results

1. Structure of non-radical triangular sets
   - Interpolation formula that extends univariate Hermite interpolation.
   - Introduce a related system denoted N based with smaller coefficients. ! Difficult to compute from T !.
   - Extends known results on the radical case (D. & Schost'2004).
2. Bit-size estimates on the family T
   - Study the growth under the interpolation process proved in 1.
   - Need the bit growth of coefficients under the inversion modulo a triangulat set.

New results

1. Structure of non-radical triangular sets
   - Interpolation formula that extends univariate Hermite interpolation.
   - Introduce a related system denoted N based with smaller coefficients. ! Difficult to compute from T !.
   - Extends known results on the radical case (D. & Schost'2004).
2. Bit-size estimates on the family T
   - Study the growth under the interpolation process proved in 1.
   - Need the bit growth of coefficients under the inversion modulo a triangulat set.

Rmk: Unable to obtain **input**-dependend bounds

Because a tool, heigh of variety is not well-defined for multiplicity.

However, this work 1) provides a step toward this goal.
2) understand the structure and how coefficients grow.

Motivation

- Triangular decomposition method (Wu-ritt characteristic method : Cf. talk of Dongming Wang)
  $\rightarrow$ Triangular sets are the most basic object occuring in this method.
- Modular methods: upper bounds on the running-time of the lifting/reconstruction step.
  (lifting is not yet available for non-radical triangular set)
- Understand the structure of triangular set and where the coefficients growth comes from.

Previous work (bit-size in multivariate polynomial system)

- Arithmetic Nullstellenstäze: Sombra *et al.*
- Rational Univariate Representation:
  (Rouiller 1999), (Schost-Mantzarflaris-Tsigarida '2017) . . .
- Triangular set:
  (Gallo-Mishra 1994), (Szanto 1999), (Schost & D. 2004)
- systems in two variables only:
  - General lex. G.b.in 2 variables (D. 2009)
  - RUR: (Mehrabi-Schost 2016), (Bouzidi, Lazard, Rouillier, Pouget 2013)
- Other: bi-homogeneous, multi-homoegeneous etc.

Previous work (bit-size in multivariate polynomial system)

- Arithmetic Nullstellenstäze: Sombra *et al.*
- Rational Univariate Representation:
  (Rouiller 1999), (Schost-Mantzarflaris-Tsigarida '2017) . . .
- Triangular set:
  (Gallo-Mishra 1994), (Szanto 1999), (Schost & D. 2004)
- systems in two variables only:
  - General lex. G.b.in 2 variables (D. 2009)
  - RUR: (Mehrabi-Schost 2016), (Bouzidi, Lazard, Rouillier, Pouget 2013)
- Other: bi-homogeneous, multi-homoegeneous etc.

Successfull strategy: use a universal object attached to the solution points: (independent of a polynomial system defining it).

Chow form $\rightarrow$ height of variety $\rightarrow$ Arithmetic Bézout Theorem.

unavailale yet for system with multiplicities !

Primary triangular set

Theorem (D., 2017)

All primary ideals of a triangular set are triangular sets.

Over $\mathbb{C}$, a primary triangular set is of the form:

$$\begin{aligned}
t_1(x_1) &= (x_1-\alpha_1)^{\delta_1(\alpha)},\\
t_2(x_1,x_2) &= (x_2-\alpha_2)^{\delta_2(\alpha)} + \sum_{i_1=0}^{\delta_1(\alpha)-1}\sum_{i_2=0}^{\delta_2(\alpha)-1} c[i,j](x_1-\alpha_1)^{i_1}(x_2-\alpha_2)^{i_2},\\
&\vdots\\
t_n(x_1,\dots,x_n) &= (x_n-\alpha_n)^{\delta_n(\alpha)} + \sum_{i_1=0}^{\delta_1(\alpha)-1}\sum_{i_2=0}^{\delta_2(\alpha)-1}\cdots\\
&\quad \sum_{i_{n-1}=0}^{\delta_{n-1}(\alpha)-1}\sum_{i_n=0}^{\delta_n(\alpha)-1} c[i_1,\dots,i_n]\prod_{j=1}^{n}(x_j-\alpha_j)^{i_j}
\end{aligned}$$

(i) $t_\ell(\alpha_1,\dots,\alpha_{\ell-1},x_\ell,\dots,x_n) = (x_\ell-\alpha)^{\delta_\ell(\alpha)} \Rightarrow c[0,\dots,0,i_\ell]=0$ for all i_ℓ, $\ell \geq 2$.

(ii) $c[i_1,\dots,i_n] = \frac{1}{i_1!i_2!\cdots i_{n-1}!i_n!}\frac{\partial^{i_1+\cdots+i_n}t_n}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}(\alpha_1,\dots,\alpha_n)$

Interpolating primary ideals ?

$$\vdots$$

$$t_n(x_1, \ldots, x_n) = (x_n - \alpha_n)^{\delta_n(\alpha)} + \sum_{i_1=0}^{\delta_1(\alpha)-1} \cdots \sum_{i_n=0}^{\delta_n(\alpha)-1} c[i_1, \ldots, i_n] \prod_{j=1}^{n} (x_j - \alpha_j)^{i_j}$$

(Local) multiplicity at α: $\mu(\alpha) = \delta_1(\alpha) \cdots \delta_n(\alpha)$.

Interpolating primary ideals ?

$$\vdots$$

$$t_n(x_1,\ldots,x_n) \quad = \quad (x_n-\alpha_n)^{\delta_n(\alpha)} + \sum_{i_1=0}^{\delta_1(\alpha)-1} \cdots \sum_{i_n=0}^{\delta_n(\alpha)-1} c[i_1,\ldots,i_n] \prod_{j=1}^{n} (x_j-\alpha_j)^{i_j}$$

(Local) multiplicity at α: $\mu(\alpha) = \delta_1(\alpha)\cdots\delta_n(\alpha)$.

Rmk: α is simple $\Leftrightarrow$ $\mu(\alpha) = 1$ and

$$\begin{cases} t_1(x_1) = x_1 - \alpha_1 \\ t_2(x_1,x_2) = x_2 - \alpha_2 \\ \quad\quad \vdots \\ t_n(x_1,\ldots,x_n) = x_n - \alpha_n. \end{cases}$$

Generalizes the standard (Lagrange) interpolation of points.

Lagrange idempotents (example)

$t_1(x_1) = (x_1 - 1)(x_1 - 2)(x_1 - 3)$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

Lagrange idempotents (example)

$t_1(x_1) = (x_1 - 1)(x_1 - 2)(x_1 - 3)$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

For $i = 1, 2, 3$, $\tilde{e}_i = \prod_{j \neq i}(x_1 - j)$. $\tilde{e}_i \tilde{e}_j = 0$ in A if $i \neq j$.

Lagrange idempotents (example)

$t_1(x_1) = (x_1 - 1)(x_1 - 2)(x_1 - 3)$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

For $i = 1, 2, 3$, $\tilde{e}_i = \prod_{j \neq i}(x_1 - j)$. $\tilde{e}_i \tilde{e}_j = 0$ in A if $i \neq j$.

Lagrange idempotent: $e_i = u_i \tilde{e}_i$ $u_i := \frac{1}{\prod_{j \neq i} j - i}$.

$$e_i^2 = e_i, \qquad e_1 + e_2 + e_3 = 1 \quad \text{in} \quad A.$$

Lagrange idempotents (example)

$t_1(x_1) = (x_1 - 1)(x_1 - 2)(x_1 - 3)$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

For $i = 1, 2, 3$, $\tilde{e}_i = \prod_{j \neq i}(x_1 - j)$. $\tilde{e}_i \tilde{e}_j = 0$ in A if $i \neq j$.

Lagrange idempotent: $e_i = u_i \tilde{e}_i$ $u_i := \frac{1}{\prod_{j \neq i} j - i}$.

$$e_i^2 = e_i, \qquad e_1 + e_2 + e_3 = 1 \quad \text{in} \quad A.$$

Lagange interpolation polynomial:

Given any function $f(x_1, x_2)$, the polynomial P_2 below is the only polynomial of degree in $x_1 < 3$ taking the values of f at $x_1 = 1, 2, 3$.

$$P(x_1, x_2) = e_1 f(1, x_2) + e_2 f(2, x_2) + e_3 f(3, x_3).$$

Idempotents occuring in Hermite interpolation

$t_1(x_1) = (x_1 - 1)(x_1 - 2)^2(x_1 - 3)^3$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

Idempotents occuring in Hermite interpolation

$t_1(x_1) = (x_1 - 1)(x_1 - 2)^2(x_1 - 3)^3$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

For $i = 1, 2, 3$, $\tilde{e}_i = \prod_{j \neq i}(x_1 - j)^j$. $\quad \tilde{e}_i \tilde{e}_j = 0$ in A if $i \neq j$.

Hermite idempotent: $e_i = \tilde{e}_i u_i$, where $u_i \tilde{e}_i + v_i (x_i - j)^j = 1$.

$$e_i^2 = e_i, \qquad e_1 + e_2 + e_3 = 1 \qquad \text{in} \quad A.$$

Idempotents occuring in Hermite interpolation

$t_1(x_1) = (x_1 - 1)(x_1 - 2)^2(x_1 - 3)^3$ Quotient ring $A = \bar{\mathbb{Q}}[x_1]/\langle t_1 \rangle$.

For $i = 1, 2, 3$, $\tilde{e}_i = \prod_{j \neq i}(x_1 - j)^j$. $\tilde{e}_i \tilde{e}_j = 0$ in A if $i \neq j$.

Hermite idempotent: $e_i = \tilde{e}_i u_i$, where $u_i \tilde{e}_i + v_i(x_i - j)^j = 1$.

$$e_i^2 = e_i, \qquad e_1 + e_2 + e_3 = 1 \qquad \text{in} \quad A.$$

Hermite interpolation polynomial:

$$P(x_1, x_2) = e_1 f(1, x_2) + e_2(f(2, x_2) + (x_2 - 2)\partial_{x_1} f(2, x_2)) +$$
$$e_3(f(3, x_2) + (x_1 - 3)\partial_{x_2} f(3, x_2) + \frac{1}{2}(x_1 - 3)^2 \partial_{x_3}^2 f(3, x_2)).$$

Idempotents occuring in triangular sets

- Main idea: See $t_{n+1}(x_1, \ldots, x_n, x_{n+1})$ as univariate in x_{n+1} over $\bar{\mathbb{Q}}[x_1, \ldots, x_n]/\langle t_1, \ldots, t_n \rangle$.
- Possible to iterate univariate idempotents.
- Radical triangular set (with Lagrange):
  - $\tilde{E}_n(x_1, \ldots, x_n) = \tilde{e}_1(x_1)\tilde{e}_2(x_2) \cdots \tilde{e}_n(x_n)$.
  - $E_n(x_1, \ldots, x_n) = e_1(x_1)e_2(x_2) \cdots e_n(x_n)$.

Idempotents occuring in triangular sets

- Main idea: See $t_{n+1}(x_1, \ldots, x_n, x_{n+1})$ as univariate in x_{n+1} over $\bar{\mathbb{Q}}[x_1, \ldots, x_n]/\langle t_1, \ldots, t_n\rangle$.
- Possible to iterate univariate idempotents.
- Radical triangular set (with Lagrange):
  - $\tilde{E}_n(x_1, \ldots, x_n) = \tilde{e}_1(x_1)\tilde{e}_2(x_2)\cdots\tilde{e}_n(x_n)$.
  - $E_n(x_1, \ldots, x_n) = e_1(x_1)e_2(x_2)\cdots e_n(x_n)$.
- Non-radical triangular set (with Hermite idempotents built for each primary triangular set)
  - $\tilde{E}_n(x_1, \ldots, x_n) \equiv$ $\tilde{e}_1(x_1)\tilde{e}_2(x_1, x_2)\cdots\tilde{e}_{n-1}(x_1, \ldots, x_{n-1}) \bmod \langle t_1, \ldots, t_{n-1}\rangle$
  - $E_n(x_1, \ldots, x_n) \equiv$ $e_1(x_1)e_2(x_1, x_2)\cdots e_{n-1}(x_1, \ldots, x_{n-1}) \bmod \langle t_1, \ldots, t_{n-1}\rangle$

Key tool: unique factorization over a primary triangular set

Interpolation of primary triangular sets

Theorem (This work)

Let α be a solution of T, and let $t^{(\alpha)}$ be its primary triangular set over $\bar{\mathbb{Q}}$.
To each α we can construct an idempotent $E_n(\alpha)$, and its barycentric form $\tilde{E}_n(\alpha)$.
Write $T_{n+1}[\alpha] \equiv T_{n+1} \bmod \langle t_1^{(\alpha)}, \dots, t_n^{(\alpha)} \rangle$

$$\begin{aligned} T_{n+1} &= \sum_{\alpha} E_n(\alpha) T_{n+1}[\alpha] \bmod \langle T_1, \dots, T_n \rangle. \\ N_{n+1} &= \sum_{\alpha} \tilde{E}_n(\alpha) T_{n+1}[\alpha] \bmod \langle T_1, \dots, T_n \rangle. \end{aligned}$$

Bit-size estimates $\leftarrow$ how the coefficients grow when processing these formula.

Bit-size bounds: input data

Bounds depend on:

1. (max) bit-size $H(\alpha)$ of the coefficients in each primary triangular set $t^{(\alpha)}$

This is a natural extension of interpolating simple points. (D. & Schost 2004).

1. corresponds to bit-size of the coordinates of each point.

Bit-size bounds: input data

Bounds depend on:

1. (max) bit-size $H(\alpha)$ of the coefficients in each primary triangular set $t^{(\alpha)}$
2. degree $\delta_1(\alpha), \ldots, \delta_n(\alpha)$ of $t^{(\alpha)} \rightarrow$ degrees $d_1, \ldots, d_n$ of output T.

This is a natural extension of interpolating simple points. (D. & Schost 2004).

1. corresponds to bit-size of the coordinates of each point.
2. corresponds to the total degree (number of points)

Bit-size bounds: input data

Bounds depend on:

1. (max) bit-size $H(\alpha)$ of the coefficients in each primary triangular set $t^{(\alpha)}$
2. degree $\delta_1(\alpha), \dots, \delta_n(\alpha)$ of $t^{(\alpha)} \to$ degrees $d_1, \dots, d_n$ of output T.
3. sum over α of the bit-size $H(T) = \sum_\alpha H(\alpha)$

This is a natural extension of interpolating simple points. (D. & Schost 2004).

1. corresponds to bit-size of the coordinates of each point.
2. corresponds to the total degree (number of points)
3. corresponds to the height of variety
   (but there is no clear notion of height of variety with multiplicity)

Statement

Theorem (this work)

The bit-size of the contructed set triangular T (from its primary components) and the related system N is bounded respectively by:

$$n\,D\,H(T) + \tilde{O}(\ n\,L(T)\,D^2\,\mu(T)\ ) \qquad H(T) + \tilde{O}(\ L(T)\,D\,\mu(T)\ )$$

- $D = d_1 + d_2 + \cdots + d_n$
  (commonly much smaller than the total degree $d_1 \cdots d_n$).
- $\mu(T) := \max_\alpha \mu(\alpha)$ is the maximal multiplicty.
- $\mathrm{L}(T) := \max_\alpha H(\alpha)$ is the maximal height of the primary components.

Rmk: They are "natural" generalization of the upper-bounds obtained in the case of a radical ideal in (D. & Schost, 2004).

Conclusion (A safe conjecture)

No precise notion of height of varieties having multiplicities,
$\leftarrow$ difficult to obtain *a priori* estimates. But:

Conjecture

If a polynomial system $\mathbf{f} \subset \mathbb{Q}[x_1, \ldots, x_n]$

- *has lex. G.b.* $\mathcal{G}$ *which is a triangular set, and:*
- *its coefficients have max-bit size* $h(\mathbf{f})$
- *its maximal total degree is* d*,*

then the maximal bit-size of the coefficients of $\mathcal{G}$ *is smaller than:*

$$n^2\, h(\mathbf{f})\, d^{2n} + \tilde{O}(\; n\, d^{2n}\, \mu(T)\, L(T)\;)$$

Conclusion (A safe conjecture)

No precise notion of height of varieties having multiplicities,
$\leftarrow$ difficult to obtain *a priori* estimates. But:

Conjecture

If a polynomial system $\mathbf{f} \subset \mathbb{Q}[x_1, \dots, x_n]$

- *has lex. G.b.* $\mathcal{G}$ *which is a triangular set, and:*
- *its coefficients have max-bit size* $h(\mathbf{f})$
- *its maximal total degree is* d,

then the maximal bit-size of the coefficients of $\mathcal{G}$ *is smaller than:*

$$n^2\, h(\mathbf{f})\, d^{2n} + \tilde{O}(\; n\, d^{2n}\, \mu(T)\, L(T)\;)$$

Rmk: $\mu(T)$ and $L(T)$ are not *a priori* estimates but are local quantities.
in certain worst-case pathological situations they can be large (up to d^n yileding cubic estiamtes in d^n, otherwise should be small.)