Deformation techniques for sparse polynomial systems

Gabriela Jeronimo, Guillermo Matera,
Pablo Solernó, Ariel Waissbein

Buenos Aires, Argentina

FoCM'08, Hong Kong

Symbolic polynomial system solving

Input: $\{f_1,\ldots,f_n\} \subset \mathbb{Q}[X_1,\ldots,X_n]$ which define a variety $V := V(f_1,\ldots,f_n) \subset \mathbb{A}^n$ of dimension zero.

Output: a geometric solution.

- A generic linear form $u \in \mathbb{Q}[X_1,\ldots,X_n]$.
- The minimal polynomial $p \in \mathbb{Q}[T]$ of $\pi_u(V)$.
- The inverse map $\pi_u^{-1}(t) := (v_1(t),\ldots,v_n(t))$.

Then $V := \{(v_1(t),\ldots,v_n(t)) : p(t) = 0\}$.

Remark (Kronecker): $v_1, \ldots, v_n$ are easily computed from the minimal pol. $p(T)$ of a generic linear form u.

Projection Problem (Proj): Given

- generators $f_1, \ldots, f_n \in \mathbb{Q}[X_1, \ldots, X_n]$ of $I(V)$,
- a generic linear form u,

find the minimal polynomial $p(T) \in \mathbb{Q}[T]$.

Complexity of the projection problem

- L: complexity of evaluation of $f_1, \ldots, f_n$.
- $D := \prod_i \deg(f_i)$: the Bézout number of the system.

Theorem [Giusti, Heintz, Pardo, ...] Proj can be solved with $Ln^{O(1)}D^{\mathbf{2}}$ arithmetic operations.

Problem: for particular classes of systems, profit from better "Bézout numbers".

Symbolic homotopy methods

Let $f := (f_1, \ldots, f_n) \in \mathbb{Q}[X]^n := \mathbb{Q}[X_1, \ldots, X_n]^n$ be polynomials such that $V := V(f) \subset \mathbb{A}^n$ is 0-dimensional.

Suppose given $F := (F_1, \ldots, F_n) \in \mathbb{Q}[X, \mathcal{E}]^n$ such that:

- $V = \{F(X, 1) = 0\}$,
- $W_0 := \{F(X, 0) = 0\}$ is "known",
- $W := V(F) \subset \mathbb{A}^{n+1}$ is a curve.

Let $\pi : W \to \mathbb{A}^1$ be the projection $\pi(x, \varepsilon) := \varepsilon$,

- π is dominant and 0 is a regular value of π.

Projection problem - unramified case (ProjU): Given

- a "generic" linear form $u \in \mathbb{Q}[X]$,
- the minimal polynomial $p_0(T)$ of u in $\mathbb{Q}[W_0]$,

compute the minimal pol $p(\mathcal{E}, T)$ of u in $\mathbb{Q}(\mathcal{E}) \hookrightarrow \mathbb{Q}(W)$.

Complexity measures

- L: complexity of evaluation of $F_1, \ldots, F_n$.
- deg(π): degree of π.
- deg(W): degree of the curve W.

Theorem ([Heintz, Krick, Puddu, Sabia, Waissbein], [Schost]): ProjU can be solved with complexity $Ln^{O(1)}\deg(\pi)\deg(W)$.

An extension to the case of W_0 ramified

It's easier to produce deformations with W_0 ramified.

• Branches of W at $\mathcal{E} = 0$ are given by Puiseux series.

• We may need several terms to "separate" all the branches at $\mathcal{E} = 0$. These are called the singular parts.

Projection problem - ramified case (ProjR): Given

• a "generic" linear form $u \in \mathbb{Q}[X]$,

• the singular parts of the Puiseux expansions of the branches of W at $\mathcal{E} = 0$,

compute the minimal pol $p(\mathcal{E}, T)$ of u in $\mathbb{Q}(\mathcal{E}) \hookrightarrow \mathbb{Q}(W)$.

Toy example: for $f_1 := X_1^2 - X_1 - X_2 - 1 = 0$, $f_2 := X_2^2 - X_1 - X_2 - 1 = 0$, we consider the deformation:

$$(S_{\mathcal{E}}) : \begin{cases} F_1(X, \mathcal{E}) := X_1^2 - \mathcal{E}(X_1 + X_2 - 1) = 0, \\ F_2(X, \mathcal{E}) := X_2^2 - \mathcal{E}(X_1 + X_2 - 1) = 0. \end{cases}$$

$(S_{\mathcal{E}})$ consists of 4 distinct branches $\sigma_i(\omega)$ with

- $\sigma_i(0) = (0, 0, 0)$,
- $\sigma_i'(0) = (0, \pm 1, \pm 1)$ (all tangents are distinct),

which are obtained from the solutions at $\mathcal{E} = 0$ of

$$(S_{\mathcal{E}}^*) : \begin{cases} \widehat{F}_1(X, \mathcal{E}) := \mathcal{E}^{-2} F_1(\mathcal{E} X_1, \mathcal{E} X_2, \mathcal{E}^2) = 0, \\ \widehat{F}_2(X, \mathcal{E}) := \mathcal{E}^{-2} F_2(\mathcal{E} X_1, \mathcal{E} X_2, \mathcal{E}^2) = 0. \end{cases}$$

Indeed, at $\mathcal{E} = 0$ we have

$$(S_0^*) \quad \widehat{F}_1(X, 0) = X_1^2 - 1 = 0, \, \widehat{F}_2(X, 0) = X_2^2 - 1 = 0,$$

namely, a unramified situation. For each solution $\alpha^{(i)}$ of (S_0^*), we have a branch of $(S_{\mathcal{E}}^*)$ parametrized as

$$\widehat{\sigma}_i(\omega) := \left(\alpha_1^{(i)}\omega + \mathcal{O}(\omega^2), \alpha_2^{(i)}\omega + \mathcal{O}(\omega^2), \omega^2\right).$$

Then we obtain a parametrization of a branch of $(S_{\mathcal{E}})$ by

$$\sigma_i(\mathcal{E}) = \left(\alpha_1^{(i)}\mathcal{E}^{1/2} + \mathcal{O}(\mathcal{E}), \alpha_2^{(i)}\mathcal{E}^{1/2} + \mathcal{O}(\mathcal{E}), \mathcal{E}\right).$$

We compute these parametrizations up to order 4 and the minimal polynomial $p(\mathcal{E}, T) = \prod_i (T - \sigma_i(\mathcal{E}))$ associated to $(S_{\mathcal{E}})$.

In general, we consider a finite family of deformations:

$$F^{(\gamma)}(\mathcal{E}, X) := \mathcal{E}^{\alpha_j^{(\gamma)}} F(\mathcal{E}^{e_\gamma}, (X - \sigma^{(\gamma)}(\mathcal{E}))\mathcal{E}^{R_\gamma}),$$

where $\sigma^{(\gamma)}(\mathcal{E})$ is the singular part of the branch γ.

Theorem [Bompadre, M., Wachenchauzer, Waissbein]
If $F^{(\gamma)}$ form a standard basis for every γ, then ProjR can be solved with complexity

$$L(ne)^{O(1)} \text{deg}(\pi) \, \text{deg}(W).$$

Application: Sparse systems

Let be given a 0-dimensional system

$$(S) \qquad\qquad f_1 = 0, \ldots, f_n = 0$$

with $f_i := \sum_{q \in \Delta_i} c_{i,q} X^q, \quad c_{i,q} \neq 0 \quad (i = 1, \ldots, n)$.

Let $Q_i := \mathrm{Conv}(\Delta_i) \subset \mathbb{R}^n$ be the convex hull of the set of exponents of the nonzero monomials of f_i.

The mixed volume $MV(Q_1, \ldots, Q_n)$ is the coefficient of $\lambda_1 \cdots \lambda_n$ in $\mathrm{Vol}(\lambda_1 Q_1 + \cdots + \lambda_n Q_n) \in \mathbb{Z}_{\geq 0}$.

Theorem 1 (Bernstein): $\#_{\mathbb{C}^*}(S) \leq MV(Q_1, \ldots, Q_n)$.

Extension of Theorem 1 [Li, Wang]: If $\underline{0} \in Q_i$ for all i, then $\#_{\mathbb{C}}(S) \leq MV(Q_1, \ldots, Q_n)$.

[Huber, Sturmfels] introduces a "polyhedral" deformation of (S) aimed at a numerical continuation method.

Huber-Sturmfels deformation of (S) is given by polynomials which are obtained: for a "lifting form" $\omega = (\omega_1, \ldots, \omega_n) : \mathbb{Z}^n \to \mathbb{Z}^n$, we consider

$$(S) \quad \widehat{F}_i := \sum_{q \in \Delta_i} c_{i,q} X^q \mathcal{E}^{\omega_i(q)} \quad (1 \leq i \leq n).$$

Theorem 1: For a generic choice of the coefficients of $f_1, \ldots, f_n$, the deformation (S) is a reduction to a ramified case in our sense.

Assume that $f_1, \ldots, f_n$ are "generic".

Remark: An estimate for deg(W) is

$$\deg(W) \leq MV(\Delta_{n+1}, \widehat{Q}_1, \ldots, \widehat{Q}_n),$$

where Δ_{n+1} is the unitary simplex of $\mathbb{R}^{n+1}$.

Remark 2: Complexity depends on the degree of the morphism and the **non-archimedean height** $h(W)$, when considering $\widehat{F}_1, \ldots, \widehat{F}_n$ as elements of $\mathbb{Q}[\mathcal{E}][X]$.

Theorem 2: Let $\widehat{Q}_i \subset \mathbb{R}^{n+1}$ be the Newton polytope of $\widehat{F}_i$, and Δ be the unitary simplex of $\mathbb{R}^n \times \{0\}$. Then

$$h(W) \leq MV(\Delta, \widehat{Q}_1, \ldots, \widehat{Q}_n).$$

Theorem 3: Generic sparse systems with Newton polytopes $Q_1, \ldots, Q_n$ can be solved with complexity

$$\sum_i \#Q_i \cdot MV(Q_1, \ldots, Q_n) \cdot MV(\Delta, \widehat{Q}_1, \ldots, \widehat{Q}_n).$$

Remark: The numerical algorithm of Huber–Sturmfels also requires the input system to be generic.

For $f_1, \ldots, f_n \in \mathbb{Q}[X_1, \ldots, X_n]$ "non–generic" with supports $\Delta_1, \ldots, \Delta_n$,we consider a deformation

$$F_i(X, \mathcal{E}) := \mathcal{E} f_i(X) + (1 - \mathcal{E}) g_i(X),$$

with $g_1, \ldots, g_n$ generic with supports $\Delta_1, \ldots, \Delta_n$.

Theorem 4: A 0–dimensional sparse system with Newton polytopes $\Delta_1, \ldots, \Delta_n$ can be solved with complexity

$$L_F \cdot MV(Q_1, \ldots, Q_n) \cdot \sum_{i=1}^{n} MV(\Delta, Q_1, \ldots, Q_{i-1}, Q_{i+1}, \ldots, Q_n).$$

Conclusions:

- Complexity of 1-parameter deformations depends on:

  - the cost of the evaluation of the input polynomials,

  - the degree of the morphism (=Bézout bound)

  - the non-archimedean height of the curve.

- There exist significant families of systems with good deformations.

Questions:

- Arithmetic complexity estimates for sparse systems.

- Estimates of the non-archimedian height.

- Families of systems with good Bézout bounds.