

ON SPARSE POLYNOMIALS, TORIC VARIETIES, CONDITION LENGTH AND RANDOMNESS. EXTENDED ABSTRACT.

GREGORIO MALAJOVICH

1. SPARSE POLYNOMIAL SYSTEMS IN GENERAL

I shall consider systems of complex Laurent polynomial equations of the form:

$$(1) \quad \begin{aligned} F_1(\mathbf{Z}) &= \sum_{\mathbf{a} \in A_1} f_{i,\mathbf{a}} Z_1^{a_1} Z_2^{a_2} \cdots Z_n^{a_n} \\ &\vdots \\ F_n(\mathbf{Z}) &= \sum_{\mathbf{a} \in A_n} f_{i,\mathbf{a}} Z_1^{a_1} Z_2^{a_2} \cdots Z_n^{a_n} \end{aligned}$$

where each $A_i \subset \mathbb{Z}^n$ is a finite set. It is natural to restrict the roots of such system to points $Z \in (\mathbb{C} \setminus \{0\})^n$ or belonging to a certain toric variety, which is a compactification thereof. Systems such as (1) can be solved by homotopy methods. I intend to discuss some of the results obtained in (Malaajovich, TAb). The reason for toric varieties is the following theorem:

Theorem 1. (Bernstein, Kušnirenko, and Hovanskiĭ, 1976) *Let $A_1, \dots, A_n \subset \mathbb{Z}^n$ be finite sets. Let \mathcal{A}_i be the convex hull of A_i . Let B be the coefficient of $\lambda_1 \dots \lambda_n$ in the polynomial*

$$\text{Vol}(\lambda_1 \mathcal{A}_1 + \cdots + \lambda_n \mathcal{A}_n).$$

where $\lambda_1, \dots, \lambda_n$ are assumed to be positive real undetermined. For a generic choice of coefficients $f_{i\mathbf{a}} \in \mathbb{C}$, the system of equations (1) above has exactly B roots \mathbf{x} in $(\mathbb{C} \setminus \{0\})^n$. The number of isolated roots is never more than B .

The *scaled mixed volume* B can be exponentially sharper than Bézout or multi-homogeneous Bézout bounds for the number of solutions. The remaining *generic* solutions in \mathbb{P}^n not in $(\mathbb{C} \setminus \{0\})^n$ are actually an artifact of homogenizing and adding extra dimensions to parameter space. They are not needed!

A lot is known about *dense* random polynomial systems. Most authors now assume a unitary invariant inner product on polynomial space, which allows to introduce a Gaussian probability distribution.

Smale's 17-th problem asked about an average polynomial time algorithm for finding one approximate solution for such systems. It was solved by Beltrán and Pardo (2009, 2011) using randomized algorithms and then by Lairez (TA), who provided a deterministic algorithm.

Unfortunately, the input size is the size of the coefficient space, that is $\sum_i \binom{d_i + n}{n}$ for n homogeneous $n + 1$ -variate polynomials of degree d_1, \dots, d_n . This can be much larger than $\sum \#A_i$. Most polynomial systems coming from applications are sparse.

Talk at the conference on *Statistical Topology of Random Manifolds*, ICTP, Trieste, July 2016.

Part of these results were obtained while visiting the Simons Institute for the Theory of Computing in the University of California at Berkeley. This visit was funded by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil. Proc. BEX 2388/14-6). This research is also funded by CNPq grants 441678/2014-9 and 306673/2013-4.

Problem A. *Can a finite zero of a random sparse polynomial system as in equation (1) be found approximately, on the average, in time polynomial in $\sum_i \#A_i$ with a uniform algorithm?*

In applications we may also want to find all the roots, not just one. Nowadays all our computers and even cell phones are parallel. We should also ask

Problem B. *Can every finite zero of a random polynomial system as in equation (1) be found approximately, on the average, in time polynomial in $\sum_i \#A_i$ with a uniform algorithm running in parallel, one parallel process for every expected zero?*

One can assume that some preliminary information such as a lower mixed subdivision is given as input to a homotopy algorithm. An algorithm to find this mixed subdivision in reasonable time was given by Malaajovich (TAa). Implementation issues were also discussed, and an alternative symbolic method that can also be used to recover the mixed subdivision was given by Jensen (TA)

Then there is the challenge of producing efficient homotopy algorithms for solving a given, non-random polynomial system. As we will see, some of the mathematics involved is also related to random polynomial systems or random paths of polynomial systems. But by insisting on finding roots on toric varieties, we quickly arrive to unfamiliar mathematical objects and to a stack of open problems.

2. FEWNOMIAL SPACES AND TORIC VARIETIES

To be able to speak of random equations or random polynomials one should first specify an appropriate function space. Without requiring functions to be holomorphic we forego the quest for global theorems, so functions must be required to be holomorphic. For convenience we will assume that each function space is a linear space with a *given* inner product.

Definition 2. A fewnomial space \mathcal{F} of functions over a complex manifold \mathcal{M} is a Hilbert space of holomorphic functions from \mathcal{M} to \mathbb{C} , such that the evaluation form

$$V : \mathcal{M} \longrightarrow \mathbb{C} \\ f \longmapsto f(\mathbf{x})$$

satisfies:

- i. For all \mathbf{x} , $V(\mathbf{x})$ is a continuous linear form.
- ii. For all \mathbf{x} , $V(\mathbf{x})$ is not the zero form.

The fewnomial space \mathcal{F} is said to be *non-degenerate* if and only if,

- iii. For all \mathbf{x} , the projection of $DV(\mathbf{x})$ onto $V(\mathbf{x})^\perp$ has full rank.

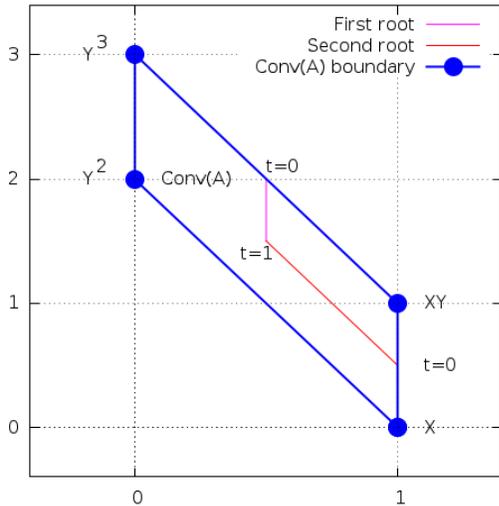


FIGURE 1. The momentum map for the two solutions of the running example between $t = 1$ (center) and $t = 0$ (on the boundary).

Fewnomial spaces are reproducing kernel spaces, with reproducing kernel $K(\mathbf{x}, \mathbf{y}) = V(\mathbf{x})(V(\mathbf{y})^*)$. The pull-back of the Fubini-Study metric in $\mathbb{P}(\mathcal{F}^*)$ defines a Hermitian structure on \mathcal{M} , denoted by $\langle \cdot, \cdot \rangle_{\mathcal{F}, \mathbf{x}}$.

A classical example is Bergman's space of holomorphic functions on a domain of \mathbb{C}^n with \mathcal{L}^2 inner product. The univariate affine and hyperbolic Gaussian analytic functions can also be understood as a random element of a suitable space of fewnomials. Other examples are the spaces \mathcal{H}_d of homogeneous polynomials in \mathbb{C}^{n+1} or spaces of sparse polynomial systems with orthonormal basis $(\dots, \rho_{\mathbf{a}} \mathbf{X}^{\mathbf{a}}, \dots)_{\mathbf{a} \in A}$ for a given $\rho : A \rightarrow (0, \infty)$. A multiplication between fewnomial spaces which is functorial in most of the invariants is described in (Malajovich, 2013).

By changing coordinates $X_i = e^{x_i}$ one can transform a sparse polynomial into a sparse exponential sum. Let $\mathcal{M} = \mathbb{C}^n \bmod 2\pi\sqrt{-1}\mathbb{Z}^n$. Let $\mathcal{F}_{\mathbf{a}}$ be the complex vector space with orthonormal basis $(\dots, \rho_{\mathbf{a}} e^{\mathbf{a}\mathbf{x}}, \dots)_{\mathbf{a} \in A}$.

Arguably, the most important tool in the theory of homotopy algorithms for homogeneous polynomial systems is the invariance by $U(n+1)$ -action, as explained by Blum et al. (1998). We cannot use this technique here. Thus we need an alternative tool.

The additive group $((\mathbb{R}^n)^*, +)$ acts on the set of **all** exponential sums by

$$\mathbf{m}, \sum_{\mathbf{a} \in A} f_{\mathbf{a}} \rho_{\mathbf{a}} e^{\mathbf{a}\mathbf{x}} \mapsto \sum_{\mathbf{a} \in A} f_{\mathbf{a}} \rho_{\mathbf{a}} e^{(\mathbf{a}-\mathbf{m})\mathbf{x}} = e^{-\mathbf{m}\mathbf{x}} \sum_{\mathbf{a} \in A} f_{\mathbf{a}} \rho_{\mathbf{a}} e^{\mathbf{a}\mathbf{x}}.$$

A particular choice of $\mathbf{m} \in (\mathbb{R}^n)^*$ plays the rôle of the canonical basis in the $U(n+1)$ -invariant homogeneous theory. This particular choice is related to an invariant of the toric action on \mathbb{C}^n : each $\theta \in (S^1)^n = \mathbb{R}^n \bmod \mathbb{Z}^n$ maps \mathbf{x} to $\mathbf{x} + 2\pi\theta\sqrt{-1}$. The reproducing kernel $K(\mathbf{x}, \mathbf{x})$ is invariant through this action, and the Hermitian metric happens to be equivariant. The *momentum map* associated to the toric action is

$$\begin{aligned} \mathbf{m} : \mathbb{C}^n &\longrightarrow \text{Conv}(A) \subseteq (\mathbb{R}^n)^* \\ \mathbf{x} &\longmapsto \mathbf{m}(\mathbf{x}) = \frac{1}{2} D \log(K(\mathbf{x}, \mathbf{x})) \end{aligned}$$

At each point \mathbf{x} , the momentum map $\mathbf{m}(\mathbf{x})$ is also a convex linear combination of the points in A . Points at toric infinity map to points on the boundary of $\text{Conv}(A)$ (Figure 1).

Now let A_1, \dots, A_n be finite sets. We consider the sparse system (1) where each f_i belongs to some \mathcal{F}_{A_i} . We have now n reproducing kernels, each one inducing a Hermitian inner product on $M = (\mathbb{C} \setminus \{0\})^n$ and with one momentum map. Each of the function spaces admits an independent $(\mathbb{R}^n)^*$ additive action.

The *toric variety* \mathcal{V} associated to the system 1 is the Zariski closure of $\{([V_1(\mathbf{x})], \dots, [V_n(\mathbf{x})]) : \mathbf{x} \in M\} \subseteq \mathbb{P}(\mathcal{F}_{A_1}) \times \dots \times \mathbb{P}(\mathcal{F}_{A_n})$. Each possible choice of f_1, \dots, f_n corresponds to a codimension- n subspace of $\mathbb{P}(\mathcal{F}_{A_1}) \times \dots \times \mathbb{P}(\mathcal{F}_{A_n})$. Not all codimension- n subspaces are possible. The generic number of intersections with \mathcal{V} is precisely the scaled mixed volume B from Theorem 1.

Each Hermitian inner form can be recovered from its imaginary part, which is a Kähler 1-1 form ω_i . Assuming that each f_i is an independent Gaussian variable with unit covariance, the root density is proportional to the outer product $\omega_1 \wedge \dots \wedge \omega_n$.

In this sense, the toric variety \mathcal{V} is a manifold with n Hermitian metrics (or equivalently n Kähler forms). At this time, we may add all of them and use product metric to measure norms on \mathcal{V} :

$$(2) \quad \langle \cdot, \cdot \rangle_{\mathbf{x}} \stackrel{\text{def}}{=} \langle \cdot, \cdot \rangle_{\mathcal{F}_{A_1}, \mathbf{x}} + \dots + \langle \cdot, \cdot \rangle_{\mathcal{F}_{A_n}, \mathbf{x}}.$$

Running example, part 1. The family

$$(3) \quad \mathbf{f}_t(X, Y) = \begin{pmatrix} tX - tXY + Y^2 - t^2Y^3 \\ X + XY - Y^2 - Y^3 \end{pmatrix}$$

admits two 'finite' solutions on the toric variety \mathcal{V} , namely (t^{-2}, t^{-1}) and $(-\frac{1+t^2}{2t}, -1)$. When $t \rightarrow 0$, both solutions converge to different points at toric 'infinity' and those can be efficiently approximated.

3. CONDITION NUMBERS

Two condition numbers turn out to be extremely helpful for the analysis of sparse homotopy algorithms. We start with the toric condition number:

Definition 3. The *toric condition number* of \mathbf{f} at \mathbf{x} is

$$\mu(\mathbf{f}, \mathbf{x}) = \left\| \begin{pmatrix} \frac{f_1}{\|f_1\| \|V_1(\mathbf{x})\|} (DV_1(\mathbf{x}) - V_1(\mathbf{x})m_1(\mathbf{x})) \\ \vdots \\ \frac{f_n}{\|f_n\| \|V_n(\mathbf{x})\|} (DV_n(\mathbf{x}) - V_n(\mathbf{x})m_n(\mathbf{x})) \end{pmatrix} \right\|_{\mathbf{x}}^{-1}$$

where the operator norm from \mathbb{C}^n (with canonical inner product) into $(\mathcal{M}, \|\cdot\|_{\mathbf{x}})$ is assumed.

Malajovich and Rojas (2004) provided some probabilistic estimates for the condition number in the *unmixed* case, that is when $A_1 = \dots = A_n$. There are also results that hold for non-degenerate fewnomial spaces. Little is known about the general *mixed* case.

By momentum map action, we may assume that $m_1(x) = \dots = m_n(x) = 0$. In that case, each of the rows of the matrix reads as

$$\frac{1}{\|f_i\|} f_i \cdot \frac{1}{\|V_i(x)\|} \begin{pmatrix} \vdots \\ \rho_{i,\mathbf{a}} e^{\mathbf{a}\mathbf{x}} \\ \vdots \end{pmatrix}_{\mathbf{a} \in A_i}$$

Open question 1. Assume that x is fixed. Let $\mathbf{f} = (f_1, \dots, f_n)$ be a Gaussian random variable with unit covariance. Find a sharp bound for $\text{Prob}[\mu(\mathbf{f}, \mathbf{x}) > \epsilon^{-1}]$.

It would be nice if the probability of ill-posedness could be bounded in terms of the mixed volume form, for subsequent integration. Compare with Bürgisser and Cucker (2013, Cor. 4.23)

Open question 2. Assume that \mathbf{x} is fixed. Let $\mathbf{f} = (f_1, \dots, f_n)$ be a Gaussian random variable with unit covariance. Estimate $\text{Prob}[\omega_1(\mathbf{x}) \wedge \dots \wedge \omega_n(\mathbf{x})\mu(\mathbf{f}, \mathbf{x}) > \epsilon^{-1}]$.

A really useful theorem for dense polynomial systems is the higher derivative estimate

$$\gamma(\mathbf{f}, \mathbf{x}) \leq \frac{(\max d_i)^{3/2}}{2} \mu(\mathbf{f}, \mathbf{x}).$$

as explained by Blum et al. (1998, Th. 2 Sec.14.2). Above, $\gamma(\mathbf{f}, \mathbf{x})$ is the invariant from alpha-theory and the radius of quadratic convergence of Newton iteration is $O(\gamma(\mathbf{f}, \mathbf{x})^{-1})$. This high derivative estimate is not true any more in the sparse context, but we have instead

$$\gamma(\mathbf{f}, \mathbf{x}) \leq \frac{1}{2} \mu(\mathbf{f}, \mathbf{x}) \nu(\mathbf{f}, \mathbf{x}).$$

for the invariant ν defined below:

Definition 4.

$$\nu(\mathbf{x}) = \max_i \nu_i(\mathbf{x})$$

where

$$\nu_i(\mathbf{x}) = \max_{\mathbf{a} \in A_i} \sup_{\|\mathbf{u}\|_{i, \mathbf{x}} \leq 1} |(\mathbf{a} - \mathbf{m}_i(\mathbf{x}))\mathbf{u}|.$$

Open question 3. What is $\text{Prob}(\nu(f, x) > \epsilon^{-1})$, when x is the root of a random Gaussian polynomial?

4. THE CONDITION LENGTH

Define the *solution variety* as

$$\mathcal{S}_0 = \left\{ (\mathbf{f}, \mathbf{x}) \in \mathbb{P}(\mathcal{F}_{A_1}) \times \dots \times \mathbb{P}(\mathcal{F}_{A_n}) \times \mathcal{M} : \mathbf{f}(\mathbf{x}) = 0 \right\}$$

Let (f_t, z_t) , $t_0 \leq t \leq t_1$ be a path in the solution variety. A typical homotopy algorithm will attempt to numerically approximate such a path, where f_t and z_{t_0} are known. Following Shub (2009), the cost of dense path-following can be bounded in terms of the condition length. Define

$$\mathcal{L}((\mathbf{f}_t, \mathbf{z}_t) : t_0, t_1) = \int_{t_0}^{t_1} \mu(\mathbf{f}_t, \mathbf{z}_t) \nu(\mathbf{z}_t) \sqrt{\|\dot{\mathbf{f}}_t\|_{\mathbf{f}_t}^2 + \|\dot{\mathbf{z}}_t\|_{\mathbf{z}_t}^2} dt.$$

The main result in (Malajovich, TAb) is:

Main Theorem. Let $(\mathbf{f}_t, \mathbf{z}_t)_{t \in [0, T]}$ be a rectifiable path in $\mathcal{S}_0 \setminus \Sigma'$. Let \mathbf{x}_0 be an approximation for \mathbf{z}_0 , satisfying

$$\frac{1}{2} \mu(\mathbf{f}_0, \mathbf{z}_0) \nu(\mathbf{z}_0) \|\mathbf{z}_0 - \mathbf{x}_0\|_{\mathbf{z}_0} \leq u_0$$

for the constant $u_0 = \frac{3-\sqrt{7}}{2} \simeq 0.090994 \dots$. Then, there is a time mesh $0 = t_0 < t_1 < \dots < t_N = T$ with

$$N \leq \left\lceil 32 \mathcal{L}((\mathbf{f}_t, \mathbf{z}_t); 0, T) \right\rceil$$

so that the approximation

$$\mathbf{x}_{i+1} = N(\mathbf{f}_{t_i}, \mathbf{x}_i)$$

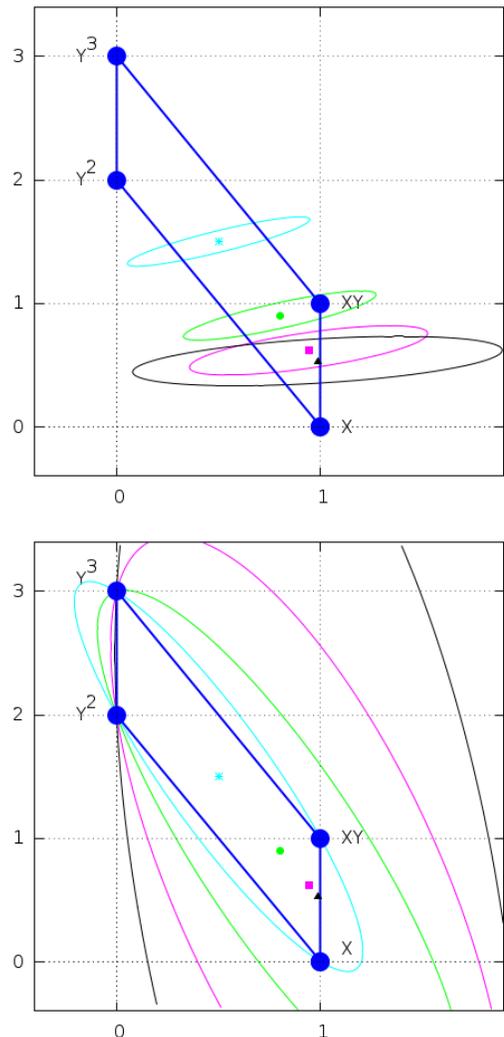


FIGURE 2. Top: Unit circles for the Hermitian metric $\langle \cdot, \cdot \rangle_{i, \mathbf{x}}$ from the running example, at several points. The circles are centered at $\mathbf{m}_i(\mathbf{x})$ and shrunk by a factor of 10 to fit in the picture. Bottom: radius ν_i circles of the dual metric. Both pictures are independent of the value of i .

produces $\mathbf{y}_0 = \mathbf{x}_N$ with

$$\frac{1}{2} \mu(\mathbf{f}_T, \mathbf{z}_T) \nu(\mathbf{z}_T) \|\mathbf{z}_T - \mathbf{y}_0\|_{\mathbf{z}_T} \leq u_0$$

for the same constant u_0 . Moreover, the sequence $\mathbf{y}_{i+1} = N(\mathbf{f}_T, \mathbf{y}_i)$ is well-defined and satisfies

$$\|\mathbf{y}_i - \mathbf{z}_T\|_{\mathbf{z}_T} \leq 2^{-2^i + 1} \|\mathbf{y}_0 - \mathbf{z}_T\|_{\mathbf{z}_T}.$$

Running example, part 2. I show in (Malajovich, TAb) that

$$\mathcal{L}((\mathbf{f}_t, (x_t, y_t)); \epsilon, 1) \in \Theta(\log(1/\epsilon))$$

where $(x, y) = (\log(X), \log(Y))$. In comparison, the condition length for the homogeneous setting as in (Shub, 2009) satisfies

$$L((\mathbf{f}_t, [X_t : Y_t : 1]); \epsilon, 1) \in \Theta(1/\epsilon).$$

This amounts to an exponentially worse bound on the number of homotopy steps, due to the fact that in projective space the

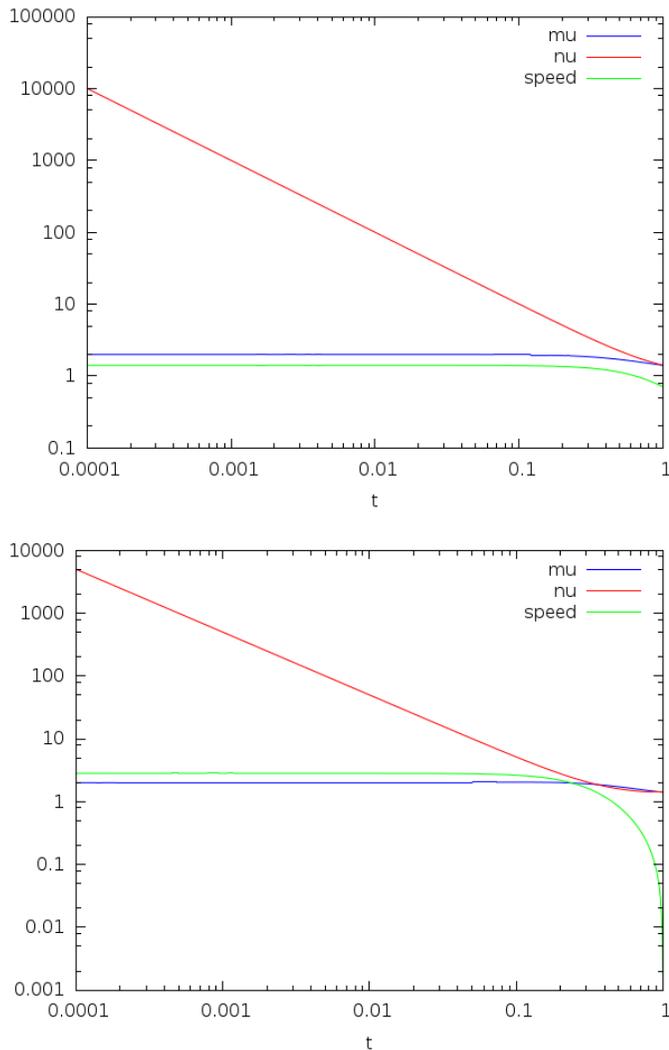


FIGURE 3. Logarithmic plot for the invariants associated to each of the solution paths, in the toric setting.

two solutions are the undistinguishable on the limit. Indeed, $\lim_{t \rightarrow 0} [X_t : Y_t : 1] = [1 : 0 : 0]$ for both curves.

An effective version of the Main Theorem allowed for a rigorous implementation of sparse homotopy algorithms for polyhedral and linear homotopy. This is work in progress on the software pss 5.0 (Polynomial System Solver), <https://sourceforge.net/projects/pss5/>. The development version is due to be updated to benefit from those improvements. At the time of this talk, the following was still open

Open question 4. *Decide with high probability if a homotopy path is converging to toric infinity, at a reasonable cost.*

5. THE FINSLER METRIC

The definition of the product metric in (2) is unnatural. It loses information. I expect all meaningful invariants of the theory to be invariant or linear in λ_i when we replace one of the n function spaces $F_{\mathcal{A}_i}$ by its λ_i -th power $(F_{\mathcal{A}_i})^{\lambda_i} = F_{\lambda_i \mathcal{A}_i}$. We may define instead a Finsler structure on \mathcal{V} :

$$\|u\|_x = \max_i \sqrt{\langle u, u \rangle_{\mathcal{A}_i, x}}.$$

The results in (Malajovich, TAb) also hold when the Hermitian metric is replaced by the Finsler structure in the condition numbers. This may be sharper, yet unwieldy.

Open question 5. *Let $C : T_x M \rightarrow \mathbb{C}^n$. Compute $\|C^{-1}\|_x$ in time $O(n^4)$ or better.*

Questions 1 and 2 can also be restated in terms of the Finsler structure.

Another ambitious project would be to reproduce the state of the art in probabilistic analysis of homotopy algorithms for sparse polynomials.

Open question 6. *Reproduce the results from Armentano et al. (TA) in the sparse context.*

Open question 7. *What about the results by Beltrán and Shub (2009)? Is there a sparse version?*

REFERENCES

- Armentano, Diego, Carlos Beltrán, Peter Bürgisser, Felipe Cucker, and Michael Shub. TA. *Condition Length and Complexity for the Solution of Polynomial Systems*, Found. Comput. Math., DOI 10.1007/s10208-016-9309-9.
- Beltrán, Carlos and Luis Miguel Pardo. 2009. *Smale's 17th problem: average polynomial time to compute affine and projective solutions*, J. Amer. Math. Soc. **22**, no. 2, 363–385, DOI 10.1090/S0894-0347-08-00630-9.
- . 2011. *Fast linear homotopy to find approximate zeros of polynomial systems*, Found. Comput. Math. **11**, no. 1, 95–129, DOI 10.1007/s10208-010-9078-9.
- Beltrán, Carlos and Michael Shub. 2009. *Complexity of Bezout's theorem. VII. Distance estimates in the condition metric*, Found. Comput. Math. **9**, no. 2, 179–195, DOI 10.1007/s10208-007-9018-5.
- Bernstein, D. N., A. G. Kušnirenko, and A. G. Hovanskiĭ. 1976. *Newton polyhedra*, Uspehi Mat. Nauk **31**, no. 3(189), 201–202 (Russian).
- Blum, Lenore, Felipe Cucker, Michael Shub, and Steve Smale. 1998. *Complexity and real computation*, Springer-Verlag, New York. With a foreword by Richard M. Karp.
- Bürgisser, Peter and Felipe Cucker. 2013. *Condition*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 349, Springer, Heidelberg. The geometry of numerical algorithms.
- Jensen, Anders. TA. *Tropical Homotopy Continuation*, available at <http://arxiv.org/abs/1601.02818>.
- Lairez, Pierre. TA. *A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time*, available at <http://arxiv.org/abs/1507.05485>.
- Malajovich, Gregorio. 2013. *On the expected number of zeros of nonlinear equations*, Found. Comput. Math. **13**, no. 6, 867–884, DOI 10.1007/s10208-013-9171-y.
- . TAa. *Computing mixed volume and all mixed cells in quermassintegral time*, Found. Comput. Math., DOI 10.1007/s10208-016-9320-1.
- . TAb. *Complexity of sparse polynomial solving: homotopy on toric varieties and the condition metric*, available at <https://arxiv.org/abs/1606.03410>. Preprint, ArXiv, June 2016.
- Malajovich, Gregorio and J. Maurice Rojas. 2004. *High probability analysis of the condition number of sparse polynomial systems*, Theoret. Comput. Sci. **315**, no. 2-3, 524–555, DOI 10.1016/j.tcs.2004.01.006.
- Shub, Michael. 2009. *Complexity of Bezout's theorem. VI. Geodesics in the condition (number) metric*, Found. Comput. Math. **9**, no. 2, 171–178, DOI 10.1007/s10208-007-9017-6.
- DEPARTAMENTO DE MATEMÁTICA APLICADA, INSTITUTO DE MATEMÁTICA, UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. CAIXA POSTAL 68530, RIO DE JANEIRO RJ 21941-909, BRASIL.
E-mail address: gregorio.malajovich@gmail.com