

A note on the finite variance of the averaging function for polynomial system solving.

Carlos Beltrán and Michael Shub *

November 14, 2008

Abstract

In [BP08], the average complexity of linear homotopy methods to solve polynomial equations with random initial input (in a sense to be described below) was proven to be finite, and even polynomial in the size of the input. In this paper, we prove that some other higher moments are also finite. In particular, we show that the variance is polynomial in the size of the input.

1 Introduction

Let $(d) = (d_1, \dots, d_n)$ be a list of positive degrees, and $\mathcal{H}_{(d)}$ be the space of homogeneous polynomial functions $f = (f_1, \dots, f_n) : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$, where f_j is homogeneous of degree d_j , $j = 1 \dots n$. We denote $\mathcal{D} = d_1 \cdots d_n$ (Bézout's Number), $d = \max\{d_j : 1 \leq j \leq n\}$ and assume that $d \geq 2$. Note that $\mathcal{D} \leq d^n$ is the number of complex projective solutions of any non-degenerate system $f \in \mathcal{H}_{(d)}$. We let $N+1$ denote the dimension of $\mathcal{H}_{(d)}$ as a vector space. We consider the Bombieri-Weyl inner product in $\mathcal{H}_{(d)}$ (cf. [SS93]), and the associated norm $\|\cdot\|$ and Riemannian structure in the sphere $\mathbb{S} = \mathbb{S}(\mathcal{H}_{(d)}) = \{f \in \mathcal{H}_{(d)} : \|f\| = 1\}$, normalized such a way that the total volume of \mathbb{S} is 1.

*C. Beltran, M. Shub, Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 2E4 (beltranc@math.toronto.edu), (shub@math.toronto.edu). Research was partially supported by MTM2007-62799, a Spanish postdoctoral grant (FE-CyT) and an NSERC Discovery Grant.

For $g \in \mathcal{H}_{(d)}$, let $V(g) = \{\zeta \in \mathbb{P}(\mathbb{C}^{n+1}) : g(\zeta) = 0\}$. The set $V = \{(g, \zeta) \in \mathbb{S} \times \mathbb{P}(\mathbb{C}^{n+1}) : \zeta \in V(g)\}$ is called the solution variety. We consider $\Omega = V$ endowed with the probability measure

$$P(A) = \int_{g \in \mathbb{S}} \frac{1}{\mathcal{D}} \sum_{\zeta \in V(g)} \chi_A(g, \zeta) d\mathbb{S},$$

for any Borel set $A \subseteq \Omega$. Let $d\Omega$ be the associated volume element. Denote by $E(\phi)$ the expected value of any measurable function $\phi : V \rightarrow \mathbb{R} \cup \{\infty\}$, and by $\text{Var}(\phi)$ its variance.

Let Σ' be the set of critical points of the projection $\pi : \Omega \rightarrow \mathbb{S}$, $\pi(g, \zeta) = g$. Let $f \in \mathbb{S}$ and $(g, \zeta) \in \Omega$, $g \neq -f$. Let $L_{g,f} \subseteq \mathbb{S}$ be the (shortest) arc of the great circle joining g and f , and let $\Gamma(g, f, \zeta)$ be the connected component of $\pi^{-1}(L_{g,f})$ that contains (g, ζ) . The implicit function theorem guarantees that $\Gamma(g, f, \zeta)$ is a smooth arc if it does not intersect Σ' , and that in that case $\pi|_{\Gamma(g,f,\zeta)}$ is a bijection, so that for $h \in L_{g,f}$ there is a unique zero ζ_h in $\Gamma(g, f, \zeta)$. Homotopy algorithms attempt to approximate this path starting at a known pair (g, ζ) to produce an approximate zero of the input problem f (cf. for example [SS93, BP06, BPar] and references therein). Define

$$\mathcal{C}_0(g, f, \zeta) = \int_{h \in L_{g,f}} \mu(h, \zeta_h) \|(\dot{h}, \dot{\zeta}_h)\| dh,$$

or ∞ if $\Gamma(g, f, \zeta)$ intersects Σ' .

From [Shu07] we know that $\mathcal{C}_0(g, f, \zeta)$ is an upper bound for the number of steps of a path following method to approximate a zero of f starting from an approximation to ζ . Thus, for $(g, \zeta) \in \Omega$, the quantity

$$\mathcal{A}_0(g, \zeta) = \int_{f \in \mathbb{S}} \mathcal{C}_0(g, f, \zeta) d\mathbb{S}$$

is an upper bound for the average number of steps required by a path-following method starting at the pair (g, ζ) to produce an approximate zero of a system f with probability 1.

In the recent paper [BP08, Theorem 1], we have seen that

$$E(\mathcal{A}_0(g, \zeta)) \leq 16\sqrt{2}\pi nN. \tag{1.1}$$

Note that the quantities \mathcal{C} and \mathcal{A} of [BP08] are slightly different from our \mathcal{C}_0 and \mathcal{A}_0 . Despite this, the results in [BP08] are still valid up to a small

constant, for $\mathcal{C}_0 \leq \sqrt{2}\mathcal{C}$ and $\mathcal{A}_0 \leq \sqrt{2}\mathcal{A}$ as follows from equation (2.1) below. The quantities \mathcal{A} of [BP08] and \mathcal{A}_0 of this paper are versions of the integrals studied in the article [SS94].

Equation (1.1) implies that the average complexity of homotopy methods (when (g, ζ) is chosen at random in Ω) is finite and, indeed, polynomial in the dimension of the problem N . This result leads to an Average Las Vegas algorithm for solving polynomial equations, in the context of Smale's 17th problem (see [BP08, Cor. 2]).

In this paper, we prove that some higher moments of \mathcal{A}_0 are also finite and the central limit theorem applies. In particular the second moment is polynomial on the dimension N . Namely, we have the following

Theorem 1. *Let $2 \leq k < 3$. Then, \mathcal{A}_0 belongs to $L^k(\mathbb{S})$, that is*

$$\mathbb{E} (\mathcal{A}_0(g, \zeta)^k) < \infty.$$

Moreover, let $2 \leq k < 3 - \frac{1}{2 \ln \mathcal{D}}$. Then,

$$\mathbb{E} (\mathcal{A}_0(g, \zeta)^k) \leq 2^{2k+k/2+4} e \pi^k n^{3k-4} N^2 \mathcal{D}^{4k-8} \ln \mathcal{D}.$$

In particular, $\text{Var} (\mathcal{A}_0(g, \zeta)) \leq \mathbb{E} (\mathcal{A}_0(g, \zeta)^2) \leq 512e\pi^2 n^2 N^2 \ln \mathcal{D}$.

From [BP08, Theorem 2], the following corollary follows immediately.

Corollary 1. *The number of homotopy steps performed by the algorithm in [BP08, Corollary 2] has variance at most $O(n^2 N^2 \ln \mathcal{D})$, thus polynomial in the size of the input.*

2 Proof of Theorem 1

Let $0 < \beta < 1 - 1/k$ and let $(\dot{h}, \dot{\zeta})$ be tangent to V . Note that

$$\|(\dot{h}, \dot{\zeta}_h)\| \leq \sqrt{\|\dot{h}\|^2 + \mu(h, \zeta_h)^2 \|\dot{h}\|^2} = \sqrt{1 + \mu(h, \zeta_h)^2} \|\dot{h}\| \leq \sqrt{2} \mu(h, \zeta_h), \quad (2.1)$$

which implies

$$\mu(h, \zeta_h) \|(\dot{h}, \dot{\zeta}_h)\| \leq 2^{\frac{1-\beta}{2}} \mu(h, \zeta_h)^{2-\beta} \|(\dot{h}, \dot{\zeta}_h)\|^\beta.$$

For $\frac{1}{p} + \frac{1}{q} = 1$, the Hölder Inequality implies

$$\mathbb{E} (\mathcal{A}_0(g, \zeta)^k) = \int_{(g, \zeta) \in \Omega} (\mathcal{A}_0(g, \zeta))^k d\Omega \leq \int_{(g, \zeta) \in \Omega} \int_{f \in \mathbb{S}} \mathcal{C}_0(g, f, \zeta)^k d\mathbb{S} d\Omega \leq$$

$$2^{\frac{k(1-\beta)}{2}} \int_{(g,\zeta) \in \Omega} \int_{f \in \mathbb{S}} \left(\int_{h \in L_{g,f}} \mu(h, \zeta_h)^{p(2-\beta)} dL_{g,f} \right)^{k/p} I d\mathbb{S} d\Omega,$$

where

$$I = \left(\int_{h \in L_{g,f}} \|(\dot{h}, \dot{\zeta}_h)\|^{q\beta} dL_{g,f} \right)^{k/q}.$$

Now, let $q = \frac{1}{\beta}$ so that $p = \frac{1}{1-\beta} < k$. Then,

$$I = \left(\int_{h \in L_{g,f}} \|(\dot{h}, \dot{\zeta}_h)\| dL_{g,f} \right)^{k\beta} \leq (\pi + 2\mathcal{D}^2)^{k\beta} \leq (3\mathcal{D}^2)^{k\beta},$$

as follows from [SS94, Lemma 7.3.a]. On the other hand, again the Hölder inequality implies

$$\left(\int_{h \in L_{g,f}} \mu(h, \zeta_h)^{p(2-\beta)} dL_{g,f} \right)^{k/p} \leq \pi^{k(1-\beta)-1} \int_{h \in L_{g,f}} \mu(h, \zeta_h)^{k(2-\beta)} dL_{g,f}$$

We have proved that

$$\mathbb{E}(\mathcal{A}_0(g, \zeta)^k) \leq 2^{\frac{k(1-\beta)}{2}} \pi^{k(1-\beta)-1} (3\mathcal{D}^2)^{k\beta} J,$$

where

$$J = \int_{(g,\zeta) \in \Omega} \int_{f \in \mathbb{S}} \int_{h \in L_{g,f}} \mu(h, \zeta_h)^{k(2-\beta)} dL_{g,f} d\mathbb{S} d\Omega.$$

From [BP08, Theorem 2 and Theorem 3],

$$J = \frac{1}{\mathcal{D}} \int_{g \in \mathbb{S}} \int_{f \in \mathbb{S}} \int_{h \in L_{g,f}} \sum_{\eta \in V(h)} \mu(h, \eta)^{k(2-\beta)} dL_{g,f} d\mathbb{S} d\mathbb{S} \leq \frac{2\pi}{\mathcal{D}} \int_{h \in \mathbb{S}} \sum_{\eta \in V(h)} \mu(h, \eta)^{k(2-\beta)} d\mathbb{S}.$$

Note that [BP08, Theorem 3] is stated for the exponent of $\mu(f, \eta)$ being 2, but the same proof holds for $k(2 - \beta)$ as used here. Finally, from [BP08, Corollary 5] we conclude that,

$$J \leq \frac{2\pi\Gamma(N+1)\Gamma(n^2+n-k(2-\beta)/2)}{\Gamma(N+1-k(2-\beta)/2)\Gamma(n^2+n)} \frac{2^{k(2-\beta)+2}}{4-k(2-\beta)} n^{3k(2-\beta)/2}, \quad (2.2)$$

is valid while $k(2 - \beta) < 4$, that is $\beta > 2 - \frac{4}{k}$. If $n \geq 2$, we may bound the last expression by

$$J \leq \frac{2^{2k+3} \pi n^{3k(2-\beta)/2-4} N^2}{2^{k\beta}(4 - k(2 - \beta))}.$$

Hence,

$$\begin{aligned} \mathbb{E}(\mathcal{A}_0(g, \zeta)^k) &\leq 2^{\frac{k(1-\beta)}{2}} \pi^{k(1-\beta)-1} (3\mathcal{D}^2)^{k\beta} \frac{2^{2k+3} \pi n^{3k(2-\beta)/2-4} N^2}{2^{k\beta}(4 - k(2 - \beta))} \leq \\ &2^{2k+k/2+3} \pi^k n^{3k-4} N^2 \frac{\mathcal{D}^{2k\beta}}{4 - k(2 - \beta)}, \end{aligned}$$

valid for $\beta \in (2 - \frac{4}{k}, 1 - 1/k)$. The minimum of this function is obtained when

$$\beta = \frac{1 + (4k - 8) \ln \mathcal{D}}{2k \ln \mathcal{D}} \in (2 - \frac{4}{k}, 1 - 1/k), \quad \text{if } 2 \leq k < 3 - \frac{1}{2 \ln \mathcal{D}}, \quad (2.3)$$

and yields the bound claimed by the theorem. If $3 - \frac{1}{2 \ln \mathcal{D}} \leq k < 3$, then the valid interval for β , $\beta \in (2 - \frac{4}{k}, 1 - 1/k)$, is non-empty and hence $\mathbb{E}(\mathcal{A}_0(g, \zeta)^k)$ is finite.

Finally, note that from [BP08, Proposition 1], equation (2.2) can be more precisely stated in the case that $n = 1$, namely in that case we have

$$J \leq \frac{2\pi\Gamma(N+1)\Gamma(2 - k(2 - \beta)/2)}{\Gamma(N+1 - k(2 - \beta)/2)\Gamma(2)} \leq \frac{4\pi N^2}{4 - k(2 - \beta)},$$

and then we have

$$\mathbb{E}(\mathcal{A}_0(g, \zeta)^k) \leq 2^{\frac{k(1-\beta)}{2}} \pi^{k(1-\beta)-1} (3\mathcal{D}^2)^{k\beta} \frac{4\pi N^2}{4 - k(2 - \beta)} \leq \frac{2^{2+k/2} \pi^k N^2 \mathcal{D}^{2k\beta}}{4 - k(2 - \beta)},$$

valid for $\beta \in (2 - \frac{4}{k}, 1 - 1/k)$. Hence, equation (2.3) still holds and so does the theorem, with an improvement in this case of 2^{2k+1} . \square

References

- [BP06] C. Beltrán and L.M. Pardo, *On Smale's 17th problem: A probabilistic positive answer.*, Found. Comput. Math. **Online First DOI 10.1007/s10208-005-0211-0** (2006).

- [BP08] ———, *Fast linear homotopy to find approximate zeros of polynomial systems*, To appear (2008).
- [BPar] ———, *Smale's 17th problem: Average polynomial time to compute affine and projective solutions*, Journal of the American Mathematical Society (To appear).
- [Shu07] M. Shub, *Complexity of Bézout's theorem. VI: Geodesics in the condition (number) metric*, J. Foundations of Computational Mathematics DOI [10.1007/s10208-007-9017-6](https://doi.org/10.1007/s10208-007-9017-6) (2007).
- [SS93] M. Shub and S. Smale, *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc. **6** (1993), no. 2, 459–501.
- [SS94] ———, *Complexity of Bezout's theorem. V. Polynomial time*, Theoret. Comput. Sci. **133** (1994), no. 1, 141–164, Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).