



**HAL**  
open science

# A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time

Pierre Lairez

► **To cite this version:**

Pierre Lairez. A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. *Foundations of Computational Mathematics*, 2016, 10.1007/s10208-016-9319-7. hal-01178588v3

**HAL Id: hal-01178588**

**<https://hal.science/hal-01178588v3>**

Submitted on 19 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time

*Pierre Lairez*

## Abstract

We describe a deterministic algorithm that computes an approximate root of  $n$  complex polynomial equations in  $n$  unknowns in average polynomial time with respect to the size of the input, in the Blum-Shub-Smale model with square root. It rests upon a derandomization of an algorithm of Beltrán and Pardo and gives a deterministic affirmative answer to Smale’s 17th problem. The main idea is to make use of the randomness contained in the input itself.

## Introduction

Shub and Smale provided an extensive theory of Newton’s iteration and homotopy continuation which aims at studying the complexity of computing approximate roots of complex polynomial systems of equations with as many unknowns as equations.<sup>1</sup> In their theory, an *approximate root* of a polynomial system refers to a point from which Newton’s iteration converges quadratically to an exact zero of the system—see Definition 1. This article answers with a deterministic algorithm the following question that they left open:

**Problem** (Smale<sup>2</sup>). *Can a zero of  $n$  complex polynomial equations in  $n$  unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?*

The term *algorithm* refers to a machine *à la* Blum-Shub-Smale<sup>3</sup> (BSS): a random access memory machine whose registers can store arbitrary real numbers, that can compute elementary arithmetic operations in the real field at unit cost and that can branch according

---

TECHNISCHE UNIVERSITÄT BERLIN, GERMANY — DFG RESEARCH GRANT BU 1371/2-2

Date — March 15, 2016.

DOI — 10.1007/s10208-016-9319-7

Keywords — Polynomial system, homotopy continuation, Smale’s 17th problem, derandomization.

2010 Mathematics subject classification — Primary 68Q25; Secondary 65H10, 65H20, 65Y20.

<sup>1</sup>Smale, “Newton’s method estimates from data at one point”; Shub and Smale, “Complexity of Bézout’s theorem. I. Geometric aspects”, “Complexity of Bezout’s theorem. II. Volumes and probabilities”, “Complexity of Bezout’s theorem. IV. Probability of success; extensions”, “Complexity of Bezout’s theorem. V. Polynomial time”; Shub, “Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric”.

<sup>2</sup>Smale, “Mathematical problems for the next century”, 17th problem.

<sup>3</sup>Blum, Shub, and Smale, “On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines”.

to the sign of a given register. To avoid vain technical argumentation, I consider the BSS model extended with the possibility of computing the square root of a positive real number at unit cost. The wording *uniform algorithm* emphasizes the requirement that a single finite machine should solve all the polynomial systems whatever the degree or the dimension. The complexity should be measured with respect to the size of the input, that is the number of real coefficients in a dense representation of the system to be solved. An important characteristic of a root of a polynomial system is its *conditioning*. Because of the feeling that approximating a root with arbitrarily large condition number requires arbitrarily many steps, the problem only asks for a complexity that is polynomial *on the average* when the input is supposed to be sampled from a certain probability distribution that we choose. The relevance of the average-case complexity is arguable, for the input distribution may not reflect actual inputs arising from applications. But yet, average-case complexity sets a mark with which any other result should be compared.

The problem of solving polynomial systems is a matter of numerical analysis just as much as it is a matter of symbolic computation. Nevertheless, the reaches of these approaches differ in a fundamental way. In an exact setting, having one root of a generic polynomial system is having them all because of Galois' indeterminacy, and it turns out that the number of solutions of a generic polynomial system is the product of the degrees of the equations, Bézout's bound, and is not polynomially bounded by the number of coefficients in the input. This is why achieving a polynomial complexity is only possible in a numerical setting.

The main numerical method to solve a polynomial system  $f$  is homotopy continuation. The principle is to start from another polynomial system  $g$  of which we know a root  $\eta$  and to move  $g$  toward  $f$  step by step while tracking all the way to  $f$  an approximate root of the deformed system by Newton's iteration. The choice of the step size and the complexity of this procedure is well understood in terms of the condition number along the homotopy path.<sup>4</sup> Most of the theory so far is exposed in the book *Condition*.<sup>5</sup> The main difficulty is to choose the starting pair  $(g, \eta)$ . Shub and Smale<sup>6</sup> showed that there exists good starting pairs, and even many, for some measure, but without providing a way to compute them efficiently. Beltrán and Pardo<sup>7</sup> discovered how to pick a starting pair at random and showed that, on average, this is a good choice. This led to a nondeterministic polynomial average-time algorithm which answers Smale's question. Bürgisser and Cucker<sup>8</sup> performed a smoothed analysis of the Beltrán-Pardo algorithm and described a deterministic algorithm with complexity  $N^{\mathcal{O}(\log \log N)}$ , where  $N$  is the input size. The question of the existence of a deterministic algorithm with polynomial average complexity is still considered open.

This work provides, with Theorem 23, a complete deterministic answer to Smale's problem, even though, as we will see, it enriches the theory of homotopy continuation itself only marginally. The answer is based on a derandomization of the nondeterministic Beltrán and

---

<sup>4</sup>Beltrán and Pardo, "Fast linear homotopy to find approximate zeros of polynomial systems"; Bürgisser and Cucker, "On a problem posed by Steve Smale"; Shub, "Complexity of Bezout's theorem. VI. Geodesics in the condition (number) metric".

<sup>5</sup>Bürgisser and Cucker, *Condition*.

<sup>6</sup>Shub and Smale, "Complexity of Bezout's theorem. V. Polynomial time".

<sup>7</sup>Beltrán and Pardo, "Fast linear homotopy to find approximate zeros of polynomial systems", "Smale's 17th problem: average polynomial time to compute affine and projective solutions".

<sup>8</sup>Bürgisser and Cucker, "On a problem posed by Steve Smale".

Pardo’s algorithm according to two basic observations. Firstly, an approximate root of a system  $f$  is also an approximate root of a slight perturbation of  $f$ . Therefore, to compute an approximate root of  $f$ , one can only consider the most significant digits of the coefficients of  $f$ . Secondly, the remaining least significant digits, or noise, of a continuous random variable are practically independent from the most significant digits and almost uniformly distributed. In the BSS model, where the input is given with infinite precision, this noise can be extracted and can be used in place of a genuine source of randomness. This answer shows that for Smale’s problem, the deterministic model and the nondeterministic are essentially equivalent: randomness is part of the question from its very formulation asking for an average analysis. It is worth noting that the idea that the input is subject to a random noise that does not affect the result is what makes the smoothed analysis of algorithms relevant.<sup>9</sup> Also, the study of the resolution of a system  $f$  given only the most significant digits of  $f$  is somewhat related to recent works in the setting of machines with finite precision.<sup>10</sup>

The derandomization proposed here is different in nature from the derandomization theorem  $\text{BPP}_{\mathbb{R}} = \text{P}_{\mathbb{R}}$ ,<sup>11</sup> which states that a decision problem that can be solved over the reals in polynomial time (worst-case complexity) with randomization and bounded error probability can also be solved deterministically in polynomial time. Contrary to this work, the derandomization theorem above relies on the ability of a BSS machine to hold arbitrary constants in its definition, even hardly computable ones or worse, not computable ones which may lead to unlikely statements. For example, one can decide the termination of Turing machines with a BSS machine insofar Chaitin’s  $\Omega$  constant is built in the machine.

**Acknowledgment** I am very grateful to Peter Bürgisser for his help and constant support, and to Carlos Beltrán for having carefully commented this work. I thank the two referees for their meticulous reading and their insightful suggestions.

## Contents

<b>1</b>	<b>The method of homotopy continuation</b>	<b>4</b>
1.1	Approximate root . . . . .	4
1.2	Homotopy continuation algorithm . . . . .	5
1.3	A variant of Beltrán-Pardo randomization . . . . .	9
<b>2</b>	<b>Derandomization of the Beltrán-Pardo algorithm</b>	<b>12</b>
2.1	Duplication of the uniform distribution on the sphere . . . . .	12
2.2	Homotopy continuation with precision check . . . . .	15
2.3	A deterministic algorithm . . . . .	18
2.4	Average analysis . . . . .	18
2.5	Implementation in the BSS model with square root . . . . .	21

---

<sup>9</sup>Spielman and Teng, “Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time”.

<sup>10</sup>Briquel, Cucker, Peña, and Roshchina, “Fast computation of zeros of polynomial systems with bounded degree under finite-precision”.

<sup>11</sup>Blum, Cucker, Shub, and Smale, *Complexity and real computation*, §17.6.

# 1 The method of homotopy continuation

This part exposes the principles of Newton’s iterations and homotopy continuation upon which rests Beltrán and Pardo’s algorithm. It mostly contains known results and variations of known results that will be used in the next part ; notable novelties are the inequality relating the maximum of the condition number along a homotopy path by the integral of the cube of the condition number (Proposition 7) and a variant of Beltrán and Pardo’s randomization procedure (Theorem 9). For Smale’s problem, the affine setting and the projective setting are known to be equivalent,<sup>12</sup> so we only focus on the latter.

## 1.1 Approximate root

Let  $n$  be a positive integer. The space  $\mathbb{C}^{n+1}$  is endowed with the usual Hermitian inner product. For  $d \in \mathbb{N}$ , let  $H_d$  denote the vector space of homogeneous polynomials of degree  $d$  in the variables  $x_0, \dots, x_n$ . It is endowed with an Hermitian inner product, called *Weyl’s inner product*, for which the monomial basis is an orthogonal basis and  $\|x_0^{a_0} \cdots x_n^{a_n}\|^2 = \frac{a_0! \cdots a_n!}{(a_1 + \cdots + a_n)!}$ . Let  $d_1, \dots, d_n$  be positive integers and let  $\mathcal{H}$  denote  $H_{d_1} \times \cdots \times H_{d_n}$ , the space of all systems of homogeneous equations in  $n+1$  variables and of degree  $d_1, \dots, d_n$ . This space is endowed with the Hermitian inner product induced by the inner product of each factor. The dimension  $n$  and the  $d_i$ ’s are fixed throughout this article. Let  $D$  be the maximum of all  $d_i$ ’s and let  $N$  denote the complex dimension of  $\mathcal{H}$ , namely

$$N = \binom{n+d_1}{n} + \cdots + \binom{n+d_n}{n}.$$

Elements of  $\mathcal{H}$  are polynomial systems to be solved, and  $2N$  is the *input size*. Note that  $2 \leq N$ ,  $n^2 \leq N$  and  $D \leq N$ .

For every Hermitian space  $V$ , we endow the set  $\mathbb{S}(V)$  of elements of norm 1 with the induced Riemannian metric  $d_{\mathbb{S}}$ : the distance between two points  $x, y \in \mathbb{S}(V)$  is the angle between them, namely  $\cos d_{\mathbb{S}}(x, y) = \operatorname{Re}\langle x, y \rangle$ . The projective space  $\mathbb{P}(V)$  is endowed with the quotient Riemannian metric  $d_{\mathbb{P}}$  defined by

$$d_{\mathbb{P}}([x], [y]) \stackrel{\text{def}}{=} \min_{\lambda \in \mathbb{S}(\mathbb{C})} d_{\mathbb{S}}(x, \lambda y).$$

An element of  $f \in \mathcal{H}$  is regarded as a homogeneous polynomial function  $\mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$ . A *root*—or *solution*, or *zero*—of  $f$  is a point  $\zeta \in \mathbb{P}^n$  such that  $f(\zeta) = 0$ . Let  $V$  be the *solution variety*  $\{(f, \zeta) \in \mathcal{H} \times \mathbb{P}^n \mid f(\zeta) = 0\}$ . For  $z \in \mathbb{C}^{n+1} \setminus \{0\}$ , let  $df(z) : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$  denote the differential of  $f$  at  $z$ . Let  $z^\perp$  be the orthogonal complement of  $\mathbb{C}z$  in  $\mathbb{C}^{n+1}$ . If the restriction  $df(z)|_{z^\perp} : z^\perp \rightarrow \mathbb{C}^n$  is invertible, we define the *projective Newton operator*  $\mathcal{N}$ , introduced by Shub,<sup>13</sup> by

$$\mathcal{N}(f, z) \stackrel{\text{def}}{=} z - df(z)|_{z^\perp}^{-1}(f(z)).$$

It is clear that  $\mathcal{N}(f, \lambda z) = \lambda \mathcal{N}(f, z)$ , so  $\mathcal{N}(f, -)$  defines a partial function  $\mathbb{P}^n \rightarrow \mathbb{P}^n$ .

**Definition 1.** A point  $z \in \mathbb{P}^n$  is an *approximate root* of  $f$  if the sequence defined recursively by  $z_0 = z$  and  $z_{k+1} = \mathcal{N}(f, z_k)$  is well defined and if there exists  $\zeta \in \mathbb{P}^n$  such that  $f(\zeta) = 0$

<sup>12</sup>Beltrán and Pardo, “Smale’s 17th problem: average polynomial time to compute affine and projective solutions”.

<sup>13</sup>Shub, “Some remarks on Bezout’s theorem and complexity theory”.

(Symbols in the margin mark the place where they are defined.)

$\mathcal{H}$

$D$

$N$

$V$

$\mathcal{N}(f, z)$

and  $d_{\mathbb{P}}(z_k, \zeta) \leq 2^{1-2^k} d_{\mathbb{P}}(z, \zeta)$  for all  $k \geq 0$ . The point  $\zeta$  is the *associated root* of  $z$  and we say that  $z$  *approximates*  $\zeta$  as a root of  $f$ .

For  $f \in \mathcal{H}$  and  $z \in \mathbb{C}^{n+1} \setminus \{0\}$ , we consider the linear map

$$\Xi(f, z) : (u_1, \dots, u_n) \in \mathbb{C}^n \mapsto df(z)|_{z^\perp}^{-1} \left( \sqrt{d_1} \|z\|^{d_1-1} u_1, \dots, \sqrt{d_n} \|z\|^{d_n-1} u_n \right) \in z^\perp$$

and the *condition number*<sup>14</sup> of  $f$  at  $z$  is defined to be  $\mu(f, z) \stackrel{\text{def}}{=} \|f\| \|\Xi(f, z)\|$ , where  $\|\Xi(f, z)\|$  is the operator norm. When  $df(z)|_{z^\perp}$  is not invertible, we set  $\mu(f, z) = \infty$ . The condition number is often denoted  $\mu_{\text{norm}}$  but we stick here to the shorter notation  $\mu$ . For all  $u, v \in \mathbb{C}^\times$  we check that  $\mu(uf, vz) = \mu(f, z)$ . We note also that  $\mu(f, z) \geq \sqrt{n} \geq 1$ .<sup>15</sup> The projective  $\mu$ -theorem (a weaker form of the better known projective  $\gamma$ -theorem) relates the condition number and the notion of approximate root:

**Theorem 2** (Shub, Smale<sup>16</sup>). *For any  $(f, \zeta) \in V$  and  $z \in \mathbb{P}^n$ , if  $D^{3/2} \mu(f, \zeta) d_{\mathbb{P}}(z, \zeta) \leq \frac{1}{3}$ , then  $z$  is an approximate root of  $f$  with associated root  $\zeta$ .*

*Remark.* The classical form of the result,<sup>17</sup> requires  $D^{3/2} \mu(f, \zeta) \tan(d_{\mathbb{P}}(z, \zeta)) \leq 3 - \sqrt{7}$ . The hypothesis required here is stronger: since  $D^{3/2} \mu(f, \zeta) \geq 1$ , if  $D^{3/2} \mu(f, \zeta) d_{\mathbb{P}}(z, \zeta) \leq \frac{1}{3}$  then  $d_{\mathbb{P}}(z, \zeta) \leq \frac{1}{3}$  and then  $\tan(d_{\mathbb{P}}(z, \zeta)) \leq 3 \tan(\frac{1}{3}) d_{\mathbb{P}}(z, \zeta) \leq \frac{3-\sqrt{7}}{D^{3/2} \mu(f, \zeta)}$  because  $\tan(\frac{1}{3}) \leq 3 - \sqrt{7}$ . The symbol  $\leq_{\square}$  indicates an inequality that is easily checked using a calculator.  $\leq_{\square}$

The algorithmic use of the condition number heavily relies on this explicit Lipschitz estimate:

**Proposition 3** (Shub<sup>18</sup>). *Let  $0 \leq \varepsilon \leq \frac{1}{7}$ . For any  $f, g \in \mathbb{P}(\mathcal{H})$  and  $x, y \in \mathbb{P}^n$ , if*

$$\mu(f, x) \max \left( D^{1/2} d_{\mathbb{P}}(f, g), D^{3/2} d_{\mathbb{P}}(x, y) \right) \leq \frac{\varepsilon}{4}$$

*then  $(1 + \varepsilon)^{-1} \mu(f, x) \leq \mu(g, y) \leq (1 + \varepsilon) \mu(f, x)$ .*

## 1.2 Homotopy continuation algorithm

Let  $I \subset \mathbb{R}$  be an interval containing 0 and let  $t \in I \mapsto f_t \in \mathbb{P}(\mathcal{H})$  be a continuous function. Let  $\zeta$  be a root of  $f_0$  such that  $df_0(\zeta)|_{\zeta^\perp}$  is invertible. There is a subinterval  $J \subset I$  containing 0 and open in  $I$ , and a continuous function  $t \in J \mapsto \zeta_t \in \mathbb{P}^n$  such that  $\zeta_0 = \zeta$  and  $f_t(\zeta_t) = 0$  for all  $t \in J$ . We choose  $J$  to be the largest such interval.

**Lemma 4.** *If  $t \mapsto f_t$  is  $C^1$  on  $I$  and if  $\mu(f_t, \zeta_t)$  is bounded on  $J$ , then  $J = I$ .*

*Proof.* Without loss of generality, we may assume that  $I$  is compact, so that  $\|\dot{f}_t\|$  is bounded on  $I$ . Let  $M$  be the supremum of  $\mu(f_t, \zeta_t) \|\dot{f}_t\|$  on  $J$ . From the construction of  $\zeta_t$  with the implicit function theorem we see that  $t \in J \mapsto \zeta_t$  is  $M$ -Lipschitz continuous. Hence the map  $t \in J \mapsto \zeta_t$  extends to a continuous map on  $\bar{J}$ . Thus  $J$  is closed in  $I$ , and  $I = J$  because  $J$  is also open.  $\square$

<sup>14</sup>Shub and Smale, "Complexity of Bézout's theorem. I. Geometric aspects"; see also Bürgisser and Cucker, *Condition*, §16, for more details about the condition number.

<sup>15</sup>Bürgisser and Cucker, "On a problem posed by Steve Smale", Lemma 16.44.

<sup>16</sup>Shub and Smale, "Complexity of Bézout's theorem. I. Geometric aspects".

<sup>17</sup>Blum, Cucker, Shub, and Smale, *Complexity and real computation*, §14, Theorems 1 and 2.

<sup>18</sup>Shub, "Complexity of Bezout's theorem. VI. Geodesics in the condition (number) metric", Theorem 1; see also Bürgisser and Cucker, *Condition*, Theorem 16.2.

**Proposition 5.** *Let  $(f, \zeta) \in V$ ,  $g \in \mathbb{P}(\mathcal{H})$  and  $0 < \varepsilon \leq \frac{1}{7}$ . If  $D^{3/2}\mu(f, \zeta)^2 d_{\mathbb{P}}(f, g) \leq \frac{\varepsilon}{4(1+\varepsilon)}$ , then:*

- (i) *there exists a unique root  $\eta$  of  $g$  such that  $d_{\mathbb{P}}(\zeta, \eta) \leq (1 + \varepsilon)\mu(f, \zeta)d_{\mathbb{P}}(f, g)$ ;*
- (ii)  $(1 + \varepsilon)^{-1}\mu(f, \zeta) \leq \mu(g, \eta) \leq (1 + \varepsilon)\mu(f, \zeta)$ ;
- (iii)  $\zeta$  *approximates  $\eta$  as a root of  $g$  and  $\eta$  approximates  $\zeta$  as a root of  $f$ .*

*Proof.* Let  $t \in [0, 1] \mapsto f_t \in \mathbb{P}(\mathcal{H})$  be a geodesic path such that  $f_0 = f$ ,  $f_1 = g$  and  $\|\dot{f}_t\| = d_{\mathbb{P}}(f, g)$ . Let  $t \in J \mapsto \zeta_t$  be the homotopy continuation associated to this path starting from the root  $\zeta$  and defined as above on a maximal interval  $J \subset [0, 1]$ . Let  $\mu_t$  denote  $\mu(f_t, \zeta_t)$ .

For all  $t \in J$  we know that  $\|\dot{\zeta}_t\| \leq \mu_t \|\dot{f}_t\|$ ,<sup>19</sup> so that

$$d_{\mathbb{P}}(\zeta_0, \zeta_t) \leq \int_0^t \|\dot{\zeta}_u\| du \leq d_{\mathbb{P}}(f, g) \int_0^t \mu_u du. \quad (1)$$

Let  $J'$  be the closed subinterval of  $J$  defined by  $J' = \{t \in J \mid \forall t' \leq t, D^{3/2}\mu_0 d_{\mathbb{P}}(\zeta_0, \zeta_{t'}) \leq \frac{\varepsilon}{4}\}$ . For all  $t \in J'$  we have  $D^{3/2}\mu_0 d_{\mathbb{P}}(\zeta_0, \zeta_t) \leq \frac{\varepsilon}{4}$ , by definition, and  $D^{1/2}\mu_0 d_{\mathbb{P}}(f_0, f_t) \leq D^{3/2}\mu_0^2 d_{\mathbb{P}}(f, g) \leq \frac{\varepsilon}{4}$ , by hypothesis. Thus, Proposition 3 ensures that

$$(1 + \varepsilon)^{-1}\mu_0 \leq \mu_t \leq (1 + \varepsilon)\mu_0, \text{ for all } t \in J'. \quad (2)$$

Thanks to Inequality (1) we conclude that  $d_{\mathbb{P}}(\zeta_0, \zeta_t) \leq (1 + \varepsilon)t d_{\mathbb{P}}(f, g)\mu_0$ , for all  $t \in J'$ , so that  $D^{3/2}\mu_0 d_{\mathbb{P}}(\zeta_0, \zeta_t) \leq \frac{t\varepsilon}{4}$ , using the assumption  $D^{3/2}\mu(f, \zeta)^2 d_{\mathbb{P}}(f, g) \leq \frac{\varepsilon}{4(1+\varepsilon)}$ . This proves that  $J'$  is open in  $J$ . Since it is also closed, we have  $J' = J$ . Since  $\mu_t$  is bounded on  $J'$ , by Inequality (2), Lemma 4 implies that  $J' = J = [0, 1]$ . Now, Inequalities (1) and (2) imply that  $d_{\mathbb{P}}(\zeta_0, \zeta_1) \leq (1 + \varepsilon)d_{\mathbb{P}}(f, g)\mu_0$ . This proves (i) and (ii) follows from (2) for  $t = 1$ .

To prove that  $\eta$  approximates  $\zeta$  as a root of  $f$ , it is enough to check that

$$D^{3/2}\mu(f, \zeta)d_{\mathbb{P}}(\zeta, \eta) \leq (1 + \varepsilon)D^{3/2}\mu(f, \zeta)^2 d_{\mathbb{P}}(f, g) \leq \frac{\varepsilon}{4} \leq_{\mathbb{H}} \frac{1}{3},$$

by Theorem 2. To prove that  $\zeta$  approximates  $\eta$  as a root of  $g$ , we check that

$$D^{3/2}\mu(g, \eta)d_{\mathbb{P}}(\zeta, \eta) \leq (1 + \varepsilon)^2 D^{3/2}\mu(f, \zeta)^2 d_{\mathbb{P}}(f, g) \leq \frac{\varepsilon(1 + \varepsilon)}{4} \leq_{\mathbb{H}} \frac{1}{3}.$$

This proves (iii) and the lemma.  $\square$

Throughout this article, let  $\varepsilon = \frac{1}{13}$ ,  $A = \frac{1}{52}$ ,  $B = \frac{1}{101}$  and  $B' = \frac{1}{65}$ . The main result that allows computing a homotopy continuation with discrete jumps is the following:

**Lemma 6.** *For any  $(f, \zeta) \in V$  and  $g \in \mathcal{H}$  and for any  $z \in \mathbb{P}^n$ , if  $D^{3/2}\mu(f, z)d_{\mathbb{P}}(z, \zeta) \leq A$  and  $D^{3/2}\mu(f, z)^2 d_{\mathbb{P}}(f, g) \leq B'$  then:*

- (i)  $z$  *is an approximate root of  $g$  with some associated root  $\eta$ ;*
- (ii)  $(1 + \varepsilon)^{-2}\mu(f, z) \leq \mu(g, \eta) \leq (1 + \varepsilon)^2\mu(f, z)$ ;
- (iii)  $D^{3/2}\mu(g, \eta)d_{\mathbb{P}}(z, \eta) \leq \frac{1}{23}$ .

*If moreover  $D^{3/2}\mu(f, z)^2 d_{\mathbb{P}}(f, g) \leq B$  then:*

<sup>19</sup>Bürgisser and Cucker, *Condition*, Corollary 16.14 and Inequality (16.12).

(iv)  $D^{3/2}\mu(g, z')d_{\mathbb{P}}(z', \eta) \leq A$ , where  $z' = \mathcal{N}(g, z)$ .

*Proof.* Firstly, we bound  $\mu(f, \zeta)$ . Since  $D^{3/2}\mu(f, z)d_{\mathbb{P}}(z, \zeta) \leq A = \frac{\varepsilon}{4}$ , Proposition 3 gives

$$(1 + \varepsilon)^{-1}\mu(f, \zeta) \leq \mu(f, z) \leq (1 + \varepsilon)\mu(f, \zeta).$$

Next, we have  $D^{3/2}\mu(f, \zeta)^2d_{\mathbb{P}}(f, g) \leq (1 + \varepsilon)^2B' \leq_{\mathbb{M}} \frac{\varepsilon}{4(1 + \varepsilon)}$ , thus Proposition 5 applies and  $\zeta$  is an approximate root of  $g$  with some associated root  $\eta$  such that  $d_{\mathbb{P}}(\zeta, \eta) \leq (1 + \varepsilon)\mu(f, \zeta)d_{\mathbb{P}}(f, g)$  and  $(1 + \varepsilon)^{-1}\mu(f, \zeta) \leq \mu(g, \eta) \leq (1 + \varepsilon)\mu(f, \zeta)$  and this gives (ii).

Then, we check that  $z$  approximates  $\eta$  as a root of  $g$ . Indeed

$$d_{\mathbb{P}}(z, \eta) \leq d_{\mathbb{P}}(z, \zeta) + d_{\mathbb{P}}(\zeta, \eta) \leq \frac{A + (1 + \varepsilon)^2B'}{D^{3/2}\mu(f, z)} \leq \frac{(1 + \varepsilon)^2(A + (1 + \varepsilon)^2B')}{D^{3/2}\mu(g, \eta)}.$$

And  $(1 + \varepsilon)^2(A + (1 + \varepsilon)^2B') \leq_{\mathbb{M}} \frac{1}{23} < \frac{1}{3}$ , so Theorem 2 applies and we obtain (i) and (iii).

We assume now that  $D^{3/2}\mu(f, z)^2d_{\mathbb{P}}(f, g) \leq B$ . All the inequalities above are valid with  $B'$  replaced by  $B$ . By definition of an approximate root  $d_{\mathbb{P}}(z', \eta) \leq \frac{1}{2}d_{\mathbb{P}}(z, \eta)$ , so that

$$D^{3/2}\mu(g, \eta)d_{\mathbb{P}}(z', \eta) \leq \frac{1}{2}(1 + \varepsilon)^2(A + (1 + \varepsilon)^2B) \leq_{\mathbb{M}} \frac{\varepsilon}{4}.$$

Thus  $(1 + \varepsilon)^{-1}\mu(g, \eta) \leq \mu(g, z') \leq (1 + \varepsilon)\mu(g, \eta)$ .

To conclude, we have  $D^{3/2}\mu(g, z')d_{\mathbb{P}}(z', \eta) \leq \frac{1}{2}(1 + \varepsilon)^3(A + (1 + \varepsilon)^2B) \leq_{\mathbb{M}} A$ .  $\square$

Let  $f, g \in \mathbb{S}(\mathcal{H})$ , with  $f \neq -g$ . Let  $t \in [0, 1] \mapsto \Gamma(g, f, t)$  be the geodesic path from  $g$  to  $f$  in  $\mathbb{S}(\mathcal{H})$ . The condition  $f \neq -g$  guarantees that the geodesic path is uniquely determined, namely

$$\Gamma(g, f, t) = \frac{\sin((1 - t)\alpha)}{\sin(\alpha)}g + \frac{\sin(t\alpha)}{\sin(\alpha)}f, \quad (3)$$

where  $\alpha = d_{\mathbb{S}}(f, g) \in [0, \pi)$  is the angle between  $f$  and  $g$ .

Let  $z \in \mathbb{P}^n$  such that  $D^{3/2}\mu(g, z)d_{\mathbb{P}}(z, \eta) \leq A$ , for some root  $\eta$  of  $g$ . By Lemma 6(i), applied with  $g = f$  and  $\eta = \zeta$ , the point  $z$  is an approximate root of  $g$ , with associated root  $\eta$ . Given  $g$  and  $z$ , we can compute an approximate root of  $f$  in the following way. Let  $g_0 = g$ ,  $t_0 = 0$  and by induction on  $k$  we define

$$\mu_k = \mu(g_k, z_k), \quad t_{k+1} = t_k + \frac{B}{D^{3/2}\mu_k^2d_{\mathbb{S}}(f, g)}, \quad g_{k+1} = \Gamma(g, f, t_{k+1}) \text{ and } z_{k+1} = \mathcal{N}(g_{k+1}, z_k).$$

Let  $K(f, g, z)$ , or simply  $K$ , be the least integer such that  $t_{K+1} > 1$ , if any, and  $K(f, g, z) = \infty$  otherwise. Let  $\tilde{M}(f, g, z)$  denote the maximum of all  $\mu_k$  with  $0 \leq k \leq K$ . Let HC be the procedure that takes as input  $f, g$  and  $z$  and outputs  $z_K$ . Algorithm 1 recapitulates the definition. It terminates if and only if  $K < \infty$ , in which case  $K$  is the number of iterations. For simplicity, we assume that we can compute exactly the square root function, the trigonometric functions and the operator norm required for the computation of  $\mu(f, z)$ . Section 2.5 shows how to implement things in the BSS model extended with the square root only.

Let  $h_t = \Gamma(f, g, t)$  and let  $t \in J \mapsto \zeta_t$  be the homotopy continuation associated to  $t \in [0, 1] \mapsto h_t$ , where  $\eta_0$  is the associated root of  $z$ , defined on a maximal subinterval  $J \subset [0, 1]$ . Let

$$M(f, g, z) \stackrel{\text{def}}{=} \sup_{t \in J} \mu(f_t, \zeta_t) \quad \text{and} \quad I_p(f, g, z) \stackrel{\text{def}}{=} \int_J \mu(h_t, \eta_t)^p dt. \quad (4)$$

$I_p(f, g, z)$ ,  
 $M(f, g, z)$

---

*Algorithm 1.* Homotopy continuation

**Input.**  $f, g \in \mathbb{S}(\mathcal{H})$  and  $z \in \mathbb{P}^n$ .

**Precondition.** There exists a root  $\eta$  of  $g$  such that  $52 D^{3/2} \mu(g, z) d_{\mathbb{P}}(z, \eta) \leq 1$ .

**Output.**  $w \in \mathbb{P}^n$

**Postcondition.**  $w$  is an approximate root of  $f$ .

---

**function** HC( $f, g, z$ )  
 $t \leftarrow 1 / (101 D^{3/2} \mu(g, z)^2 d_{\mathbb{S}}(f, g))$   
**while**  $1 > t$  **do**  
 $h \leftarrow \Gamma(g, f, t)$   
 $z \leftarrow \mathcal{N}(h, z)$   
 $t \leftarrow t + 1 / (101 D^{3/2} \mu(h, z)^2 d_{\mathbb{S}}(f, g))$   
**end while**  
**return**  $z$   
**end function**

---

The behavior of the procedure HC can be controlled in terms of the integrals  $I_p(f, g, z)$ . It is one of the corner stone of the complexity theory of homotopy continuation methods. The following estimation of the maximum of the condition number, along a homotopy path, in terms of the third moment of the condition number seems to be original. It will be important for the average complexity analysis.

**Proposition 7.** *If  $J = [0, 1]$  then  $M(f, g, z) \leq 151 D^{3/2} I_3(f, g, z)$ .*

*Proof.* Let  $\varepsilon = \frac{1}{7}$  and let  $s \in [0, 1]$  such that  $\mu(f_s, \zeta_s)$  is maximal. For all  $t \in [0, 1]$ ,  $d_{\mathbb{S}}(f_s, f_t) \leq |t - s| d_{\mathbb{S}}(f, g)$ . Thus, if

$$|t - s| \leq \frac{\varepsilon}{4(1 + \varepsilon) D^{3/2} \mu(f_s, \zeta_s)^2 d_{\mathbb{S}}(f, g)}, \quad (5)$$

then  $\mu(f_t, \zeta_t) \geq (1 + \varepsilon)^{-1} \mu(f_s, \zeta_s)$ , by Proposition 5. Since  $d_{\mathbb{S}}(f, g) \leq \pi$ , the diameter of the interval  $H$  of all  $t \in [0, 1]$  satisfying Inequality (5) is at least  $\frac{\varepsilon}{4\pi(1+\varepsilon)D^{3/2}\mu(f_s, \zeta_s)^2}$ . Thus

$$\int_0^1 \mu(f_t, \zeta_t)^3 dt \geq \int_H \frac{\mu(f_s, \zeta_s)^3}{(1 + \varepsilon)^3} dt \geq \frac{\varepsilon \mu(f_s, \zeta_s)}{4\pi(1 + \varepsilon)^4 D^{3/2}} \geq \frac{1}{151} \frac{\mu(f_s, \zeta_s)}{D^{3/2}}. \quad \square$$

**Theorem 8** (Shub<sup>20</sup>). *With the notations above, if  $D^{3/2} \mu(g, z) d_{\mathbb{P}}(z, \eta) \leq A$  then:*

(i) HC( $f, g, z$ ) terminates if and only if  $I_2(f, g, z)$  is finite, in which case  $J = [0, 1]$ ;

If moreover HC( $f, g, z$ ) terminates then:

(ii)  $(1 + \varepsilon)^{-2} M(f, g, z) \leq \tilde{M}(f, g, z) \leq (1 + \varepsilon)^2 M(f, g, z)$ .

(iii)  $K(f, g, z) \leq 136 D^{3/2} d_{\mathbb{S}}(f, g) I_2(f, g, z)$ ;

(iv) HC( $f, g, z$ ) is an approximate root of  $f$ ;

(v)  $D^{3/2} \mu(f, \zeta) d_{\mathbb{P}}(\text{HC}(f, g, z), \zeta) \leq \frac{1}{23}$ , where  $\zeta$  is the associated root of HC( $f, g, z$ ).

<sup>20</sup>Shub, ‘‘Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric’’.

*Proof.* Let  $\eta_k$  denote  $\zeta_{t_k}$ . Since  $D^{3/2}\mu_k^2 d_{\mathbb{P}}(g_k, g_{k+1}) \leq B$  for all  $k \geq 0$ , Lemma 6(iv) proves, by induction on  $k$  that  $D^{3/2}\mu_k d_{\mathbb{P}}(z_k, \eta_k) \leq A$  for any  $k \geq 0$ .

Assume that  $[0, t_k] \subset J$  for some  $k \geq 0$  and let  $t \in [t_k, t_{k+1}] \cap J$  so that

$$D^{3/2}\mu_k^2 d(g_k, h_t) \leq D^{3/2}\mu_k^2 d(g_k, g_{k+1}) \leq B.$$

Because  $D^{3/2}\mu_k d(z_k, \eta_k) \leq A$ , Lemma 6(ii) applies to  $(g_k, \eta_k)$ ,  $h_t$  and  $z_k$  and asserts that

$$(1 + \varepsilon)^{-2}\mu_k \leq \mu(h_t, \zeta_t) \leq (1 + \varepsilon)^2\mu_k. \quad (6)$$

By definition  $\mu_k^2(t_{k+1} - t_k) = \frac{B}{D^{3/2}d_{\mathbb{S}}(f, g)}$ , so integrating over  $t$  leads to

$$\int_0^{t_k} \mu(h_t, \zeta_t)^2 dt \geq (1 + \varepsilon)^{-4} \sum_{j=0}^{k-1} \mu_j^2(t_{j+1} - t_j) = \frac{kB}{(1 + \varepsilon)^4 D^{3/2}d_{\mathbb{S}}(f, g)}, \quad (7)$$

$$\text{and } \int_0^{\sup J} \mu(h_t, \zeta_t)^2 \leq (1 + \varepsilon)^4 \sum_{j=0}^k \mu_j^2(t_{j+1} - t_j) = \frac{(1 + \varepsilon)^4(k+1)B}{D^{3/2}d_{\mathbb{S}}(f, g)}. \quad (8)$$

Assume now that  $I_2(f, g, z)$  is finite. The left-hand side of Inequality (7) is finite so there exists a  $k$  such that  $t_{k+1} \notin J$ . But then Inequalities (6) shows that  $\mu_t$  is bounded on  $J$  which implies, Lemma 4 that  $J = [0, 1]$ . And since  $t_{k+1} \notin J$ , this proves that  $K$  is finite.

Conversely, assume that  $K$  is finite, *i.e.*  $\text{HC}(f, g, z)$  terminates. Then there exists a maximal  $k$  such that  $[0, t_k] \subset J$  and thus for all  $t \in J$

$$\mu(h_t, \zeta_t) \leq (1 + \varepsilon)^2 \max_{j \leq k} \mu(g_k, z_k).$$

So  $\mu(h_t, \zeta_t)$  is bounded on  $J$ , which implies that  $J = [0, 1]$ , and thus  $k = K$ . Inequality (8) then shows that  $I_2(f, g, z)$  is finite, which concludes the proof of (i). We keep assuming that  $K$  is finite. Inequality (6) shows (ii). Since  $[0, t_K] \subset [0, 1]$ , by definition, Inequalities (7) and (8) shows that

$$\frac{1}{B(1 + \varepsilon)^4} D^{3/2}d_{\mathbb{S}}(f, g)I_2(f, g, z) - 1 \leq K \leq \frac{(1 + \varepsilon)^4}{B} D^{3/2}d_{\mathbb{S}}(f, g)I_2(f, g, z).$$

We check that  $\frac{(1 + \varepsilon)^4}{B} \leq_{\mathbb{M}} 136$ , which gives (iii). Finally, Lemmas 6(i) and 6(iii) show that  $z_K$  approximates  $\zeta_1$  as a root of  $f$  and that  $D^{3/2}\mu(f, \zeta_1)d_{\mathbb{P}}(z_K, \zeta_1) \leq \frac{1}{23}$ , which gives (iv) and (v).  $\square$

### 1.3 A variant of Beltrán-Pardo randomization

An important discovery of Beltrán and Pardo is a procedure to pick a random system and one of its root simultaneously without actually solving any polynomial system. And from the complexity point of view, it turns out that a random pair  $(g, \eta) \in V$  is a good starting point to perform the homotopy continuation.

Let  $g \in \mathbb{S}(\mathcal{H})$  be a uniform random variable, where the uniform measure is relative to the Riemannian metric on  $\mathbb{S}(\mathcal{H})$ . Almost surely  $g$  has finitely many roots in  $\mathbb{P}^n$ . Let  $\eta$  be one of them, randomly chosen with the uniform distribution. The probability distribution of the random variable  $(g, \eta) \in V$  is denoted  $\rho_{\text{std}}$ . The purpose of Beltrán and Pardo's procedure<sup>21</sup>

$\rho_{\text{std}}$

<sup>21</sup>Beltrán and Pardo, "Fast linear homotopy to find approximate zeros of polynomial systems", §2.3; see also

is to generate random pairs  $(g, \eta)$ , according to the distribution  $\rho_{\text{std}}$  without solving any polynomial system. We give here a variant which requires only a uniform random variable in  $\mathbb{S}(\mathbb{C}^N) \simeq \mathbb{S}(\mathcal{H})$  as the source of randomness.

Let us assume that  $f = (f_1, \dots, f_n) \in \mathbb{S}(\mathcal{H})$  is a uniform random variable and write  $f$  as

$$f_i = c_i x_0^{d_i} + \sqrt{d_i} x_0^{d_i-1} \sum_{j=1}^n a_{i,j} x_j + f'_i(x_0, \dots, x_n),$$

for some  $c_i, a_{i,j} \in \mathbb{C}$  and  $f'_i \in H_{d_i}$  such that  $f'_i(e_0) = 0$  and  $df'_i(e_0) = 0$ . Let  $f' = (f'_1, \dots, f'_n) \in \mathcal{H}$ . By construction,  $f'(e_0) = 0$  and  $df'(e_0) = 0$ . Let

$$M \stackrel{\text{def}}{=} \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & c_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & c_n \end{pmatrix} \in \mathbb{C}^{n \times (n+1)}.$$

Almost surely,  $\ker M$  has dimension 1; Let  $\zeta \in \mathbb{P}^n$  be the point representing  $\ker M$  and let  $\zeta' \in \mathbb{S}(\mathbb{C}^{n+1})$  be the unique element of  $\ker M \cap \mathbb{S}(\mathbb{C}^{n+1})$  whose first nonzero coordinate is a real positive number. Let  $\Psi_{M, \zeta'} = (\Psi_1, \dots, \Psi_n) \in \mathcal{H}$  be defined by

$$\Psi_i \stackrel{\text{def}}{=} \sqrt{d_i} \left( \sum_{i=0}^n x_i \bar{\zeta}'_i \right)^{d_i-1} \sum_{j=0}^n m_{i,j} x_j, \quad (9)$$

where  $\bar{\zeta}'_i$  denotes the complex conjugation. By construction  $\Psi_{M, \zeta'}(\zeta) = 0$ . Let  $u \in U(n+1)$ , the unitary group of  $\mathbb{C}^{n+1}$ , such that  $u(e_0) = \zeta$ . The choice of  $u$  is arbitrary but should depend only on  $\zeta$ . For example, we can choose  $u$ , for almost all  $\zeta$ , to be the unique element of  $U(n+1)$  with determinant 1 that is the identity on the orthogonal complement of  $\{e_0, \zeta\}$  and that sends  $e_0$  to  $\zeta$ . Finally, let  $g = f' \circ u^{-1} + \Psi_{M, \zeta'} \in \mathcal{H}$ . By construction  $g(\zeta) = 0$ . We define  $\text{BP}(f) \stackrel{\text{def}}{=} (g, \zeta)$  which is a point in the solution variety  $V$ . BP(f)

**Theorem 9.** *If  $f \in \mathbb{S}(\mathcal{H})$  is a uniform random variable, then  $\text{BP}(f) \sim \rho_{\text{std}}$ .*

*Proof.* We reduce to another variant of Beltrán-Pardo procedure given by Bürgisser and Cucker<sup>22</sup> in the case of Gaussian distributions. Let  $f \in \mathbb{S}(\mathcal{H})$  be a uniform random variable, and let  $\chi \in [0, \infty)$  be an independent random variable following the *chi* distribution with  $2N$  degrees of freedom, so that  $\chi f$  is a centered Gaussian variable in  $\mathcal{H}$  with covariance matrix  $I_{2N}$  (which we call hereafter a *standard normal variable*). For  $\zeta \in \mathbb{P}^n$ , let  $R_\zeta \subset \mathcal{H}$  be the subspace of all  $g$  such that  $g(\zeta) = 0$  and  $dg(\zeta) = 0$  and let  $S_\zeta$  be the orthogonal complement of  $R_\zeta$  in  $\mathcal{H}$ . The system  $\chi f$  splits orthogonally as  $\chi f' + \chi h$ , where  $\chi f' \in R_{e_0}$  and  $\chi h \in S_{e_0}$  are independent standard normal variables.

Let  $M \in \mathbb{C}^{n \times (n+1)}$ ,  $\zeta \in \mathbb{P}^n$ ,  $\zeta' \in \mathbb{S}^n$  and  $u \in U(n+1)$  be defined in the same way as in the definition of  $\text{BP}(f)$ . The map that gives  $M$  as a function of  $h$  is an isometry  $S_{e_0} \rightarrow \mathbb{C}^{n \times (n+1)}$ , so  $\chi M$  is a standard normal variable that is independent from  $f'$ . Let  $\lambda \in \mathbb{S}(\mathbb{C})$  be an independent uniform random variable, so that  $\lambda \zeta'$  is uniformly distributed in  $\ker M \cap \mathbb{S}^n$  when  $M$  has full rank, which is the case with probability 1. The composition map  $g \in R_{e_0} \mapsto g \circ u^{-1} \in R_\zeta$  is an isometry. Thus, conditionally to  $\zeta$ , the system  $\chi f' \circ u^{-1}$  is a standard

<sup>21</sup>Bürgisser and Cucker, *Condition*, Chap. 17.

<sup>22</sup>Bürgisser and Cucker, *Condition*, Theorem 17.21(a).

normal variable in  $R_\zeta$ . As a consequence, and according to Bürgisser and Cucker<sup>23</sup>, the system

$$H \stackrel{\text{def}}{=} \chi f' \circ u^{-1} + \Psi_{\chi M, \lambda \zeta'}$$

is a standard normal variable in  $\mathcal{H}$  and  $\zeta$  is uniformly distributed among its roots. Hence  $H/\|H\|$  is uniformly distributed in  $\mathbb{S}(\mathcal{H})$  and  $(H/\|H\|, \zeta) \sim \rho_{\text{std}}$ .

We check easily that  $\|\Psi_{M, \lambda \zeta'}\| = \|M\|_F = \|h\|$ , where  $\|M\|_F$  denotes the Froebenius matrix norm, that is the usual Hermitian norm on  $\mathbb{C}^{n \times (n+1)}$ . Moreover  $\|f' \circ u^{-1}\| = \|f'\|$ , this is the fundamental property of Weyl's inner product on  $\mathcal{H}$ . Thus  $\|H\| = \|f\| = \chi$ , and in turn

$$(f' \circ u^{-1} + \Psi_{M, \lambda \zeta'}, \zeta) = (H/\|H\|, \zeta) \sim \rho_{\text{std}},$$

which is almost what we want, up to the presence of  $\lambda$ . Let  $\Delta \in \mathbb{C}^{n \times n}$  be the diagonal matrix given by  $(\bar{\lambda}^{d_i - 1})_{1 \leq i \leq n}$ . It is clear that  $\Psi_{M, \lambda \zeta'} = \Psi_{\Delta M, \zeta'}$ . The map  $M \mapsto \Delta M$  is an isometry of  $\mathbb{C}^{n \times (n+1)}$  and  $\ker \Delta M = \ker M$  so  $(\chi M, u, \zeta')$  and  $(\chi \Delta M, u, \zeta')$  have the same probability distribution. Since  $\chi f'$  is independent from  $\chi M$  and  $\lambda$ , it follows that the system  $H'$  defined by

$$H' \stackrel{\text{def}}{=} \chi f' \circ u^{-1} + \Psi_{\chi M, \zeta'},$$

has the same probability distribution as  $H$ . To conclude the proof, we note that  $\|H'\| = \chi$  and that  $(H'/\chi, \zeta) = \text{BP}(f)$ .  $\square$

Given  $f \in \mathbb{S}(\mathcal{H})$ , Beltrán and Pardo's algorithm proceeds in sampling a system  $g \in \mathbb{S}(\mathcal{H})$  from the uniform distribution and then computing  $\text{HC}(f, \text{BP}(g))$ . If the input  $f$  is a uniform random variable then we can evaluate the expected number of homotopy steps  $\mathbb{E}(K(f, \text{BP}(g)))$ . Indeed, let  $\eta$  be root of  $g$ , uniformly chosen, the theorem above asserts that  $\text{BP}(g)$  has the same probability distribution as  $(g, \eta)$  so  $\mathbb{E}(K(f, \text{BP}(g))) = \mathbb{E}(K(f, g, \eta))$ . Thanks to Theorem 8(iii), it is not difficult to see that  $\mathbb{E}(K(f, g, \eta)) \leq 214 D^{3/2} \mathbb{E}(\mu(g, \eta)^2)$ . This is why the estimation of  $\mathbb{E}(\mu(g, \eta)^2)$  is another corner stone of the average complexity analysis of homotopy methods. Deriving from a identity of Beltrán and Pardo, we obtain the following:

**Theorem 10.** *If  $(g, \eta) \sim \rho_{\text{std}}$ , then  $\mathbb{E}(\mu(g, \eta)^p) \leq \frac{3}{4-p} (nN)^{p/2}$  for any  $2 \leq p < 4$ .*

*Proof.* Let  $s = p/2 - 1$ . Beltrán and Pardo<sup>24</sup> state that

$$\mathbb{E}(\mu(g, \eta)^{2+2s}) = \frac{\Gamma(N+1)}{\Gamma(N-s)} \sum_{k=1}^n \binom{n+1}{k+1} \frac{\Gamma(k-s)}{\Gamma(k)} n^{-k+s}.$$

We use the inequalities  $x^{-y}\Gamma(x) \leq \Gamma(x-y) \leq (x-1)^{-y}\Gamma(x)$ , for  $x \in [1, \infty)$  and  $y \in [0, 1]$ , which comes from the log-convexity of  $\Gamma$ . In particular, since  $0 \leq s < 1$ ,

$$\frac{\Gamma(N+1)}{\Gamma(N-s)} = \frac{N\Gamma(N)}{\Gamma(N-s)} \leq N^{1+s} \quad \text{and} \quad \frac{\Gamma(k-s)}{\Gamma(k)} \leq (k-1)^{-s}.$$

Thus

$$\mathbb{E}(\mu(g, \eta)^{2+2s}) \leq N^{1+s} \left( \binom{n+1}{2} \frac{\Gamma(1-s)}{\Gamma(1)} n^{s-1} + \sum_{k=2}^n \binom{n+1}{k+1} (k-1)^{-s} n^{-k+s} \right).$$

<sup>23</sup>Bürgisser and Cucker, *Condition*, Theorem 17.21(a).

<sup>24</sup>Beltrán and Pardo, "Fast linear homotopy to find approximate zeros of polynomial systems", Theorem 23.

On the one hand  $(1-s)\Gamma(1-s) = \Gamma(2-s) \leq \Gamma(2) = \Gamma(1)$ , so

$$\binom{n+1}{2} \frac{\Gamma(1-s)}{\Gamma(1)} n^{s-1} \leq \frac{(n+1)n}{2} \frac{1}{1-s} n^{s-1} \leq \frac{n^{1+s}}{1-s}.$$

On the other hand,

$$\begin{aligned} \sum_{k=2}^n \binom{n+1}{k+1} (k-1)^{-s} n^{-k+s} &\leq n^{1+s} \sum_{k=3}^{n+1} \binom{n+1}{k} n^{-k} \\ &= n^{1+s} \left( \left(1 + \frac{1}{n}\right)^{n+1} - 1 - \frac{n+1}{n} - \frac{1}{n^2} \binom{n+1}{2} \right) \\ &\leq \frac{n^{1+s}}{4} \leq \frac{n^{1+s}}{4(1-s)}. \end{aligned}$$

Putting together all above, we obtain the claim.  $\square$

## 2 Derandomization of the Beltrán-Pardo algorithm

### 2.1 Duplication of the uniform distribution on the sphere

An important argument of the construction is the ability to produce approximations of two independent uniform random variables in  $\mathbb{S}^{2N-1}$  from a single uniform random variable in  $\mathbb{S}^{2N-1}$  given with infinite precision. More precisely, let  $Q$  be a positive integer. This section is dedicated to the construction of two functions  $\lfloor - \rfloor_Q$  and  $\{ - \}_Q$  from the sphere  $\mathbb{S}^{2N-1}$  to itself, respectively called the *truncation* and the *fractional part* at precision  $Q$ . For  $u \in \mathbb{S}^{2N-1}$ ,  $\lfloor u \rfloor_Q$  is close to  $u$  and if  $u$  is uniform random variable, then  $\{u\}_Q$  is *nearly* uniformly distributed in  $\mathbb{S}^{2N-1}$  and *nearly* independent from  $\lfloor u \rfloor_Q$ , in the following sense:

**Lemma 11.** *For any  $u \in \mathbb{S}^{2N-1}$ ,  $d_{\mathbb{S}}(\lfloor u \rfloor_Q, u) \leq 3N^{1/2}/Q$ . Moreover, for any continuous nonnegative function  $\Theta : \mathbb{S}^{2N-1} \times \mathbb{S}^{2N-1} \rightarrow \mathbb{R}$ ,*

$$\frac{1}{\text{vol}(\mathbb{S}^{2N-1})} \int_{\mathbb{S}^{2N-1}} \Theta(\lfloor u \rfloor_Q, \{u\}_Q) du \leq \frac{\exp\left(\frac{2N^{3/2}}{Q}\right)}{\text{vol}(\mathbb{S}^{2N-1})^2} \int_{\mathbb{S}^{2N-1}} \int_{\mathbb{S}^{2N-1}} \Theta(\lfloor u \rfloor_Q, v) du dv.$$

For  $x \in \mathbb{R}$ , let  $A(x)$  denote the integral part of  $a$  and let  $A_Q(a) \stackrel{\text{def}}{=} Q^{-1}A(Qa)$  be the truncation at precision  $Q$ . For  $x \in \mathbb{R}^{2N-1}$ , let  $A_Q(x) \in \mathbb{R}^{2N-1}$  be the vector  $(A_Q(x_1), \dots, A_Q(x_{2N-1}))$  and let  $B_Q(x) \stackrel{\text{def}}{=} (x - A_Q(x))Q$ , which is a vector in  $[0, 1]^{2N-1}$ . We note that  $\|A_Q(x) - x\|^2 \leq (2N-1)/Q^2$ , because the difference is bounded componentwise by  $1/Q$ .

Let  $C$  and  $C_+$  denote  $[-1, 1]^{2N-1}$  and  $[0, 1]^{2N-1}$  respectively, and let  $F(x) = (1 + \|x\|^2)^{-N}$ . We first show that if  $x \in C$  is a random variable with probability density function  $F$  (divided by the appropriate normalization constant) then  $B_Q(x)$  is *nearly* uniformly distributed in  $C_+$  and *nearly* independent from  $A_Q(x)$ .

**Lemma 12.** *For any continuous nonnegative function  $\Theta : [-1, 1]^{2N-1} \times [0, 1]^{2N-1} \rightarrow \mathbb{R}$ ,*

$$\int_C \Theta(A_Q(x), B_Q(x)) F(x) dx \leq \exp\left(\frac{2N^{3/2}}{Q}\right) \int_{C_+} \int_C \Theta(A_Q(x), y) F(x) dx dy.$$

*Proof.* For any integers  $-Q \leq k_i < Q$ , for  $1 \leq i \leq 2N - 1$ , the function  $A_Q$  is constant on the set  $\prod_{i=1}^{2N-1} \left[ \frac{k_i}{Q}, \frac{k_i+1}{Q} \right)$ , and these sets form a partition of  $C$ . Let  $X_1, \dots, X_{(2Q)^{2N-1}}$  denote an enumeration of these sets and let  $a_k$  denote the unique value of  $A_Q$  on  $X_k$ . The diameter of  $X_k$  is  $\sqrt{2N-1}/Q$ . Since the function  $x \in [0, \infty) \mapsto -N \log(1+x^2)$  is  $N$ -Lipschitz continuous, we derive that

$$\max_{X_k} F \leq e^{N\sqrt{2N-1}/Q} \min_{X_k} F \leq e^{2N^{3/2}/Q} \min_{X_k} F. \quad (10)$$

For any  $1 \leq k \leq (2Q)^{2N-1}$ , we have

$$\int_{X_k} \Theta(A_Q(x), B_Q(x)) F(x) dx \leq \max_{X_k} F \int_{X_k} \Theta(a_k, (x - a_k)Q) dx,$$

because  $A_Q(x) = a_k$  on  $X_k$  and by definition of  $B_Q(x)$ . A simple change of variable shows that

$$\int_{X_k} \Theta(a_k, (x - a_k)Q) dx = \text{vol}(X_k) \int_{C_+} \Theta(a_k, y) dy.$$

Besides, for all  $y \in C_+$ ,

$$\Theta(a_k, y) \leq \frac{1}{\text{vol}(X_k) \min_{X_k} F} \int_{X_k} \Theta(A_Q(x), y) F(x) dx.$$

Putting together all above and summing over  $k$  gives the claim.  $\square$

Thanks to a method due to Sibuya, we may transform a uniform random variable of  $C_+$  into a uniform random variable in  $\mathbb{S}^{2N-1}$ . Let  $x = (x_1, \dots, x_{2N-1}) \in C_+$ , let  $y_1, \dots, y_{N-1}$  denote the numbers  $x_{N+1}, \dots, x_{2N-1}$  arranged in ascending order, and let  $y_0 = 0$  and  $y_N = 1$ . Let  $S(x) \in \mathbb{R}^{2N}$  denote the vector such that for any  $1 \leq i \leq N$

$$S(x)_{2i-1} \stackrel{\text{def}}{=} \sqrt{y_i - y_{i-1}} \cos(2\pi x_i) \quad \text{and} \quad S(x)_{2i} \stackrel{\text{def}}{=} \sqrt{y_i - y_{i-1}} \sin(2\pi x_i). \quad (11)$$

**Proposition 13** (Sibuya<sup>25</sup>). *If  $x$  is a uniformly distributed random variable in  $C_+$ , then  $S(x)$  is uniformly distributed in  $\mathbb{S}^{2N-1}$ .*

We now define  $[-]_Q$  and  $\{-\}_Q$ . Let  $\Sigma \subset \mathbb{R}^{2N}$  be the set of all  $x \in \mathbb{R}^{2N}$  such that  $\|x\|_\infty = 1$ . It is divided into  $4N$  faces that are isometric to  $C$ : they are the sets  $\Sigma_i^\varepsilon \stackrel{\text{def}}{=} \{x \in \Sigma \mid x_i = \varepsilon\}$ , for  $\varepsilon \in \{-1, 1\}$  and  $1 \leq i \leq 2N$  and the isometries are given by the maps

$$t_{i,\varepsilon} : \Sigma_i^\varepsilon \rightarrow C, \quad x \mapsto (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n).$$

Through these isometries, we define the functions  $A'_Q$  and  $B'_Q$  on  $\Sigma$ : for  $x \in \Sigma_i^\varepsilon$  we set

$$A'_Q(x) \stackrel{\text{def}}{=} t_{i,\varepsilon}^{-1}(A_Q(t_{i,\varepsilon}(x))) \in \Sigma_i^\varepsilon \quad \text{and} \quad B'_Q(x) \stackrel{\text{def}}{=} B_Q(t_{i,\varepsilon}(x)) \in C_+.$$

Let  $\nu_\infty : u \in \mathbb{S}^{2N-1} \mapsto u/\|u\|_\infty \in \Sigma$  and its inverse  $\nu_2 : x \in \Sigma \mapsto x/\|x\| \in \mathbb{S}^{2N-1}$ . Finally, we define, for  $u \in \mathbb{S}^{2N-1}$ , using Sibuya's function  $S$ , see Equation (11),

$$[u]_Q \stackrel{\text{def}}{=} \nu_2(A'_Q(\nu_\infty(u))) \quad \text{and} \quad \{u\}_Q \stackrel{\text{def}}{=} S(B'_Q(\nu_\infty(u))). \quad (12)$$

We may now prove Lemma 11.

<sup>25</sup>Sibuya, "A method for generating uniformly distributed points on  $N$ -dimensional spheres".

*Proof of Lemma 11.* Let  $u \in \mathbb{S}^{2N-1}$ . It is well-known that  $d_{\mathbb{S}}(\lfloor u \rfloor_Q, u) \leq \frac{\pi}{2} \|\lfloor u \rfloor_Q - u\|$ . Furthermore, the map  $\nu_2$  is clearly 1-Lipschitz continuous on  $\Sigma$  so  $\|\lfloor u \rfloor_Q - u\| \leq \|A'_Q(\nu_\infty(u)) - \nu_\infty(u)\|$  and we already remarked that the latter is at most  $\sqrt{2N-1}/Q$ . With  $\frac{\pi}{2}\sqrt{2} \leq \blacksquare 3$ , this gives the first claim.

Concerning the second claim, we consider the partition of the sphere by the sets  $\nu_2(\Sigma_i^\varepsilon)$ . For any  $u \in \nu_2(\Sigma_i^\varepsilon)$ , we have

$$\nu_\infty(u) = \left( \frac{u_1}{u_i}, \dots, \frac{u_{i-1}}{u_i}, 1, \frac{u_{i+1}}{u_i}, \dots, \frac{u_n}{u_i} \right).$$

Hence, by Lemma 14 below, the absolute value of the Jacobian of the map  $\nu_\infty : \nu_2(\Sigma_i^\varepsilon) \rightarrow \Sigma_i^\varepsilon$  at some  $\nu_2(x)$ ,  $x \in \Sigma_i^\varepsilon$ , is precisely

$$(\nu_2(x)_i)^{-2N} = \|x\|^{2N} = (1 + \|t_{i,\varepsilon}(x)\|^2)^N = F(t_{i,\varepsilon}(x))^{-1}.$$

Thus, the change of variable  $\nu_\infty$  gives

$$\int_{\nu_2(\Sigma_i^\varepsilon)} \Theta(\lfloor u \rfloor_Q, \{u\}_Q) du = \int_{\Sigma_i^\varepsilon} \Theta(\nu_2(A'_Q(x)), S(B'_Q(x))) F(t_{i,\varepsilon}(x)) dx,$$

and then Lemma 12 (applied over  $\Sigma_i^\varepsilon$  with the isometry  $t_{i,\varepsilon} : \Sigma_i^\varepsilon \rightarrow C$ ) implies

$$\leq \exp\left(\frac{2N^{3/2}}{Q}\right) \int_{\Sigma_i^\varepsilon} \int_{C_+} \Theta(\nu_2(A'_Q(x)), S(y)) dy F(t_{i,\varepsilon}(x)) dx$$

and Proposition 13 gives the equality

$$= \frac{\exp\left(\frac{2N^{3/2}}{Q}\right)}{\text{vol}(\mathbb{S}^{2N-1})} \int_{\Sigma_i^\varepsilon} \int_{\mathbb{S}^{2N-1}} \Theta(\nu_2(A'_Q(x)), v) dv F(t_{i,\varepsilon}(x)) dx,$$

and applying the inverse change of variable  $\nu_2 : \Sigma_i^\varepsilon \rightarrow \nu_2(\Sigma_i^\varepsilon)$  and summing over  $i$  and  $\varepsilon$  gives the claim.  $\square$

**Lemma 14.** Let  $H = \{x \in \mathbb{R}^{2N} \mid x_1 > 0\}$  and let  $\varphi$  be the map

$$\begin{aligned} \varphi : \mathbb{S}^{2N-1} \cap H &\longrightarrow \mathbb{R}^{2N-1} \\ (u_1, \dots, u_{2N}) &\longmapsto \left( \frac{u_2}{u_1}, \dots, \frac{u_{2N}}{u_1} \right). \end{aligned}$$

Then, for any  $u \in \mathbb{S}^{2N-1} \cap H$ ,  $|\det(d\varphi(u))| = u_1^{-2N}$ , where  $\mathbb{S}^{2N-1}$  and  $\mathbb{R}^{2N-1}$  are considered with their usual Riemannian structures.

*Proof.* Let  $\psi : u \in \mathbb{R}^{2N} \cap H \rightarrow \left( \frac{u_2}{u_1}, \dots, \frac{u_{2N}}{u_1} \right)$ , so that  $\varphi$  is the restriction of  $\psi$  to the sphere  $\mathbb{S}^{2N-1}$ . Firstly, the matrix of  $d\psi(x)$ , for some  $x \in \mathbb{R}^{2N}$ , in the standard bases of  $\mathbb{R}^{2N}$  and  $\mathbb{R}^{2N-1}$ , is given by

$$\text{Mat}(d\psi(x)) = \frac{1}{x_1^2} \begin{pmatrix} -x_2 & x_1 & & 0 \\ \vdots & & \ddots & \\ -x_{2N} & 0 & & x_1 \end{pmatrix}. \quad (13)$$

Let  $u \in \mathbb{S}^{2N-1} \cap H$ . We may assume without loss of generality that  $u$  is of the form  $(u_1, u_2, 0, \dots, 0)$ , with  $u_1^2 + u_2^2 = 1$ , because  $|\det(d\varphi(u))|$  is invariant under any unitary transformation of  $u$

that preserves the first coordinate. Let  $T \subset \mathbb{R}^{2N}$  be the tangent space at  $u$  of  $\mathbb{S}^{2N-1}$ . Naturally,  $d\varphi(u) = d\psi(u)|_T$ . An orthonormal basis of  $T$  is given by  $\{f, e_3, \dots, e_{2N}\}$ , where  $f = (-u_2, u_1, 0, \dots, 0)$  and where  $e_i$  is the  $i$ th coordinate vector. Using Equation (13), we compute that  $d\varphi(u)(f) = u_1^{-2}e_1$  and that  $d\varphi(u)(e_i) = u_1^{-1}e_{i-1}$ , for  $3 \leq i \leq 2N$ . Thus  $|\det(d\varphi(u))| = u_1^{-2N}$ .  $\square$

The orthogonal monomial basis of  $\mathcal{H}$  gives an identification  $\mathcal{H} \simeq \mathbb{R}^{2N}$  and we define this way the truncation  $\lfloor f \rfloor_Q$  and the fractional part  $\{f\}_Q$  of a polynomial system  $f \in \mathbb{S}(\mathcal{H})$ . The derandomization relies on finding a approximate root of  $\lfloor f \rfloor_Q$ , for some  $Q$  large enough, and using  $\{f\}_Q$  as the source of randomness for the Beltrán-Pardo procedure. Namely, we compute  $\text{HC}(\lfloor f \rfloor_Q, \text{BP}(\{f\}_Q))$ . Almost surely, this computation produces an approximate root of  $\lfloor f \rfloor_Q$ . If  $Q$  is large enough, it is also an approximate root of  $f$ . The main technical difficulty is to choose a precision and to ensure that the result is correct while keeping the complexity under control.

## 2.2 Homotopy continuation with precision check

Let  $f, f', g \in \mathbb{S}(\mathcal{H})$  and let  $\eta \in \mathbb{P}^n$  be a root of  $g$ . Throughout this section, we assume that  $d_{\mathbb{S}}(f, f') \leq \rho$ , for some  $\rho > 0$  and that  $d_{\mathbb{S}}(f, g) \leq \pi/2$ . Up to changing  $g$  into  $-g$ , the latter is always true, since  $d_{\mathbb{S}}(f, -g) = \pi - d_{\mathbb{S}}(f, g)$ . The notations  $I_2$ ,  $M$  and  $\tilde{M}$  used in this section have been introduced in §1.2. If  $\rho$  is small enough, then  $\text{HC}(f', g, \eta)$  is an approximate root not only of  $f'$  but also of  $f$ . But if  $\rho$  fails to be small enough,  $\text{HC}(f', g, \eta)$  may not even terminate or, to say the least,  $\text{HC}(f', g, \eta)$  may take arbitrarily long to compute something that is not an approximate root of  $f$ . To control the complexity of the new algorithm, it is important to be able to recognize this situation at least as fast as  $\text{HC}(f, g, \eta)$  would terminate.

As in §1.2, let  $f_t = \Gamma(g, f, t)$  and  $f'_t = \Gamma(g, f', t)$ . Let  $t \in J \rightarrow \zeta_t \in \mathbb{P}^n$  be the homotopy continuation associated to  $f_t$ , on  $[0, 1]$ , and  $t \in J' \rightarrow \zeta'_t \in \mathbb{P}^n$  be the one associated to  $f'_t$ , defined on some maximal intervals  $J, J' \subset [0, 1]$  containing 0. Let  $\mu_t = \mu(f_t, \zeta_t)$  and  $\mu'_t = \mu(f'_t, \zeta'_t)$ .

**Lemma 15.**  $d_{\mathbb{S}}(f_t, f'_t) \leq 2d_{\mathbb{S}}(f, f')$  for any  $t \in [0, 1]$ .

*Proof.* Let  $\alpha_t = d_{\mathbb{S}}(f_t, f'_t)$ ,  $\beta = d_{\mathbb{S}}(f, g) \leq \frac{\pi}{2}$  and  $\gamma = d_{\mathbb{S}}(f', g)$ . Without loss of generality, we may assume that  $\alpha_1 < \frac{\pi}{2}$ , otherwise the inequality  $\alpha_t \leq 2\alpha_1$  that we want to check is trivial.

The spherical law of cosines applied to the spherical triangle  $\{g, f_t, f'_t\}$  gives the equality

$$\cos \alpha_t = \cos(t\beta) \cos(t\gamma) + \sin(t\beta) \sin(t\gamma) \cos A, \quad (14)$$

where  $A$  is the angle of the triangle at  $g$ . We deal with three cases. Firstly, we assume that  $\gamma \leq \frac{\pi}{2}$ . Then  $\cos \alpha_t$  decreases at  $t$  increases: Indeed, Equation (14) rewrites as

$$\cos \alpha_t = \cos(t\beta - t\gamma) - \sin(t\beta) \sin(t\gamma)(1 - \cos A) \quad (15)$$

and, as  $t$  increases,  $\cos(t\beta - t\gamma)$  decreases, because  $|\beta - \gamma| \leq \pi$ , and both  $\sin(t\beta)$  and  $\sin(t\gamma)$  increase, because  $\beta, \gamma \leq \frac{\pi}{2}$ . Thus  $\cos \alpha_t \geq \cos \alpha_1$ , for  $0 \leq t \leq 1$ , and it follows that  $\alpha_t \leq \alpha_1$ .

Second case, we assume that  $\gamma > \frac{\pi}{2}$  and  $\beta = \frac{\pi}{2}$ . For  $t \in [0, 1]$ , Equation (15) shows that

$$\cos \alpha_t \geq \cos\left(\frac{\pi}{2} - \gamma\right) - (1 - \cos A) = \sin \gamma + \cos A - 1,$$

using  $\cos(t\beta - t\gamma) \geq \cos(\beta - \gamma)$  and  $1 - \cos(A) \geq 0$ . Equation (14) shows that  $\cos \alpha_1 = \sin \gamma \cos A$ . In particular  $\cos A \geq 0$ , since  $\alpha_1 \leq \frac{\pi}{2}$  and  $\sin \gamma \geq 0$ . It follows that

$$2 \sin^2 \gamma \cos^2 A \leq \sin^4 \gamma + \cos^4 A \leq \sin \gamma + \cos A,$$

and finally that  $\cos(2\alpha_1) \leq \cos \alpha_t$ , because  $\cos(2\alpha_1) = 2 \cos^2 \alpha_1 - 1$ . Since  $2\alpha_1 \leq \pi$ , we obtain that  $2\alpha_1 \geq \alpha_t$ , which concludes in the second case.

Third case, we assume only that  $\gamma > \frac{\pi}{2}$  (and always  $\beta \leq \frac{\pi}{2}$ ). Let  $h \in \mathbb{S}(\mathcal{H})$  be the unique point on the spherical segment  $[f, f']$  such that  $d_{\mathbb{S}}(g, h) = \frac{\pi}{2}$ . In particular, we have that  $d_{\mathbb{S}}(f, f') = d_{\mathbb{S}}(f, h) + d_{\mathbb{S}}(h, f')$  and

$$\alpha_t = d_{\mathbb{S}}(f_t, f'_t) \leq d_{\mathbb{S}}(f_t, h_t) + d_{\mathbb{S}}(h_t, f'_t),$$

where  $h_t \stackrel{\text{def}}{=} \Gamma(g, h, t)$ . The first case shows that  $d_{\mathbb{S}}(f_t, h_t) \leq d_{\mathbb{S}}(f, h)$  and the second case shows that  $d_{\mathbb{S}}(h_t, f'_t) \leq 2d_{\mathbb{S}}(h, f')$ . Thus  $\alpha_t \leq d_{\mathbb{S}}(f, h) + 2d_{\mathbb{S}}(h, f') \leq 2d_{\mathbb{S}}(f, f')$ .  $\square$

Recall that  $M(f, g, \zeta)$  denotes  $\sup_{t \in J} \mu_t$ , see §1.2 and Equation (4).

**Lemma 16.** *If  $D^{3/2}M(f, g, \zeta)^2 \rho \leq \frac{1}{168}$  then  $J = J' = [0, 1]$  and for any  $t \in [0, 1]$ :*

$$(i) \quad (1 + \varepsilon)^{-1} \mu'_t \leq \mu_t \leq (1 + \varepsilon) \mu'_t;$$

$$(ii) \quad D^{3/2} \mu_t d_{\mathbb{P}}(\zeta_t, \zeta'_t) \leq \frac{1}{51}.$$

*Proof.* The assumption implies that  $M(f, g, \zeta) < \infty$ , and thus  $J = [0, 1]$ , by Lemma 4. Let  $S$  the set of all  $t \in J'$  such that  $D^{3/2} \mu_t d_{\mathbb{P}}(\zeta_t, \zeta'_t) \leq \frac{1}{51}$ . It is a nonempty closed subset of  $J'$ . Let  $t \in S$ . By Lemma 15, we have  $d_{\mathbb{P}}(f_t, f'_t) \leq 2\rho$ , so

$$D^{3/2} \mu_t^2 d_{\mathbb{P}}(f_t, f'_t) \leq \frac{2}{112} = \frac{\varepsilon}{4(1 + \varepsilon)}.$$

Proposition 5 implies that there exists a root  $\eta$  of  $f'_t$  such that  $d_{\mathbb{P}}(\eta, \zeta_t) \leq 2(1 + \varepsilon)\mu_t \rho$  and  $(1 + \varepsilon)^{-1} \mu_t \leq \mu(f'_t, \eta) \leq (1 + \varepsilon)\mu_t$ . Because  $d_{\mathbb{P}}(\eta, \zeta'_t) \leq d_{\mathbb{P}}(\eta, \zeta_t) + d_{\mathbb{P}}(\zeta_t, \zeta'_t)$  and  $t \in S$  we obtain

$$D^{3/2} \mu(f'_t, \eta) d_{\mathbb{P}}(\eta, \zeta'_t) \leq D^{3/2} (1 + \varepsilon) \mu_t \left( 2(1 + \varepsilon)\mu_t \rho + \frac{1}{51 D^{3/2} \mu_t} \right) \leq (1 + \varepsilon)^2 \frac{2}{112} + (1 + \varepsilon) \frac{1}{51} \leq_{\square} \frac{1}{3}.$$

and Theorem 2 implies that  $\zeta'_t$  approximates  $\eta$  as a root of  $f'_t$ . Since it is also an exact root of  $f'_t$ , this implies  $\zeta'_t = \eta$ . In particular  $D^{3/2} \mu_t d_{\mathbb{P}}(\zeta'_t, \zeta_t) \leq 2(1 + \varepsilon) D^{3/2} \mu_t^2 \rho <_{\square} \frac{1}{51}$ . Thus  $t$  is in the interior of  $S$ , which proves that  $S$  is open and finally that  $S = J$ . Moreover, since  $\mu'_t \leq (1 + \varepsilon)\mu_t$ ,  $\mu'_t$  is bounded on  $J'$ , thus  $J' = [0, 1]$ .  $\square$

This leads to the procedure  $\text{HC}'$ , see Algorithm 2. It modifies procedure  $\text{HC}$  (Algorithm 1) in only one respect: each iteration checks up on the failure condition  $D^{3/2} \mu(h, z)^2 \rho > \frac{1}{151}$ . If the failure condition is never met, then  $\text{HC}'$  computes exactly the same thing as  $\text{HC}$ . Recall that  $\tilde{M}(f', g, \eta)$  denotes the maximum condition number  $\mu$  that arises in the homotopy continuation  $\text{HC}(f', g, \eta)$ , and that  $I_p(f, g, \eta)$  denote the integral of  $\mu^p$  along the homotopy path from  $g$  to  $f$ , see §1.2 and Equation (4).

**Proposition 17.** *If  $d_{\mathbb{S}}(f, g) \leq \frac{\pi}{2}$  and  $d(f, f') \leq \rho$ , then the procedure  $\text{HC}'(f', g, \eta, \rho)$ :*

$$(i) \quad \textit{terminates and performs at most } 158 D^{3/2} d_{\mathbb{S}}(f, g) I_2(f, g, \eta) + 4 \textit{ steps};$$

---

*Algorithm 2.* Homotopy continuation with precision check

**Input.**  $f, g \in \mathbb{S}(\mathcal{H})$ ,  $z \in \mathbb{P}^n$  and  $\rho > 0$ .

**Output.**  $w \in \mathbb{P}^n$  or FAIL.

**Specifications.** See Proposition 17.

---

**function**  $\text{HC}'(f, g, z, \rho)$   
 $t \leftarrow 1 / (101D^{3/2}\mu(g, z)^2d_{\mathbb{S}}(f, g))$   
 $h \leftarrow g$   
**while**  $1 > t$  and  $D^{3/2}\mu(h, z)^2\rho \leq \frac{1}{151}$  **do**  
 $h \leftarrow \Gamma(g, f, t)$   
 $z \leftarrow \mathcal{N}(h, z)$   
 $t \leftarrow t + 1 / (101D^{3/2}\mu(h, z)^2d_{\mathbb{S}}(f, g))$   
**end while**  
**if**  $D^{3/2}\mu(h, z)^2\rho > \frac{1}{151}$  **then return** FAIL  
**else return**  $z$   
**end if**  
**end function**

---

(ii) outputs an approximate root of  $f$ , or fails;

(iii) succeeds (i.e. outputs some  $z \in \mathbb{P}^n$ ) if and only if  $D^{3/2}\tilde{M}(f', g, \eta)^2\rho \leq \frac{1}{151}$ ;

(iv) succeeds if  $D^{3/2}M(f, g, \eta)^2\rho \leq \frac{1}{236}$ .

*Proof.* At each iteration, the value of  $t$  increases by at least  $151\rho/(101d_{\mathbb{S}}(f', g))$ , thus there are at most  $101d_{\mathbb{S}}(f', g)/(151\rho)$  iterations before termination.

By construction, the procedure  $\text{HC}'(f', g, \eta, \rho)$  fails if and only if at some point of the procedure  $\text{HC}(f', g, \eta, \rho)$  it happens that  $D^{3/2}\mu(h, z)^2\rho > \frac{1}{151}$ . In other words, the procedure  $\text{HC}'(f', g, \eta, \rho)$  fails if and only if  $D^{3/2}\tilde{M}(f', g, \eta)^2\rho > \frac{1}{151}$ , by definition of  $\tilde{M}$ . And since the procedure terminates, it succeeds if and only if it does not fail. This proves (iii).

Let us bound the number  $K'(f', g, \eta, \rho)$  of iterations of the procedure  $\text{HC}'(f', g, \eta, \rho)$  before termination. If  $\text{HC}'(f', g, \eta, \rho)$  succeeds, then  $K'(f', g, \eta, \rho) = K(f', g, \eta)$ . Furthermore

$$K'(f', g, \eta, \rho) = \sup \{ K(f'_s, g, \eta) \mid s \in [0, 1], \text{HC}'(f'_s, g, \eta, \rho) \text{ succeeds} \}. \quad (16)$$

Let  $s \in [0, 1]$  such that  $\text{HC}'(f'_s, g, \eta, \rho)$  succeeds, that is to say  $D^{3/2}\tilde{M}(f'_s, g, \eta)^2\rho \leq \frac{1}{151}$ . Theorem 8(ii) shows that

$$(1 + \varepsilon)^{-2}M(f'_s, g, \eta) \leq \tilde{M}(f'_s, g, \eta) \leq (1 + \varepsilon)^2M(f'_s, g, \eta).$$

In particular  $D^{3/2}M(f'_s, g, \eta)^2\rho \leq \frac{(1+\varepsilon)^4}{151} \leq \frac{1}{112}$  and Lemma 16 shows that  $(1 + \varepsilon)^{-2} \leq \mu'_t \leq (1 + \varepsilon)\mu_t$  for all  $t \leq s$ . So we obtain that  $(1 + \varepsilon)^{-2}I_2(f_s, g, \eta) \leq I_2(f'_s, g, \eta) \leq (1 + \varepsilon)^2I_2(f_s, g, \eta)$  and

$$\begin{aligned} K(f'_s, g, \eta) &\leq 136 D^{3/2}d_{\mathbb{S}}(f'_s, g)I_2(f'_s, g, \eta) && \text{by Theorem 8(iii)} \\ &\leq 136(1 + \varepsilon)^2D^{3/2}(d_{\mathbb{S}}(f_s, g) + 2\rho)I_2(f_s, g, \eta) && \text{by Lemma 15.} \end{aligned}$$

Besides  $D^{3/2}I_2(f_s, g, \eta)\rho \leq (1 + \varepsilon)^2 D^{3/2}M(f'_s, g, \eta)^2\rho \leq \frac{(1+\varepsilon)^2}{112}$ , so we obtain

$$K(f'_s, g, \eta) \leq_{\blacksquare} 158D^{3/2}d_{\mathbb{S}}(f_s, g)I_2(f_s, g, \eta) + 4 \leq 158D^{3/2}d_{\mathbb{S}}(f, g)I_2(f, g, \eta) + 4.$$

Together with Equation (16), this completes the proof of (i).

Let us assume that the procedure  $\text{HC}'(f', g, \eta, \rho)$  succeeds and let  $z$  be its output, which is nothing but  $\text{HC}(f', g, \eta)$ . Theorem 8(v) shows that  $D^{3/2}\mu'_1 d_{\mathbb{P}}(z, \zeta'_1) \leq \frac{1}{23}$ , where  $\zeta'_1$  is the root of  $f'_1 = f'$  obtained by homotopy continuation. As above, with  $s = 1$ , we check that  $\mu_1 \leq (1 + \varepsilon)\mu'_1$  and  $D^{3/2}\mu'_1 d_{\mathbb{P}}(\zeta_1, \zeta'_1) \leq \frac{1}{51}$  using Lemma 16. Thus

$$D^{3/2}\mu_1 d_{\mathbb{P}}(z, \zeta_1) \leq (1 + \varepsilon) \left( \frac{1}{23} + \frac{1}{51} \right) <_{\blacksquare} \frac{1}{3}.$$

Then  $z$  approximates  $\zeta_1$  as a root of  $f_1$ , by Theorem 2. This proves (ii).

Lastly, let us assume that  $D^{3/2}M(f, g, \eta)^2\rho \leq \frac{1}{236}$ . Lemma 16 implies that  $M(f, g, \eta) \geq (1 + \varepsilon)^{-1}M(f', g, \eta)$  and Theorem 8(ii) shows that  $M(f', g, \eta) \leq (1 + \varepsilon)^2 \tilde{M}(f', g, \eta)$ . Thus

$$D^{3/2}\tilde{M}(f', g, \eta)^2\rho \leq (1 + \varepsilon)^6 D^{3/2}M(f, g, \eta)^2\rho \leq \frac{(1 + \varepsilon)^6}{236} \leq_{\blacksquare} \frac{1}{151}$$

and  $\text{HC}'(f', g, \eta, \rho)$  succeeds. This proves (iv).  $\square$

### 2.3 A deterministic algorithm

Let  $f \in \mathbb{S}(\mathcal{H})$  be the input system to be solved and let  $Q \geq 1$  be a given precision. We compute

$$f' = \lfloor f \rfloor_Q, (g, \eta) = \text{BP}(\{f\}_Q), \varepsilon = \text{sign}(\pi/2 - d_{\mathbb{S}}(f, g)) \text{ and } \rho = 3N^{1/2}/Q.$$

Lemma 11 shows that  $d_{\mathbb{S}}(f, f') \leq \rho$ . Then we run the homotopy continuation procedure with precision check  $\text{HC}'(f', \varepsilon g, \eta, \rho)$ , which may fail or output a point  $z \in \mathbb{P}^n$ . If it does succeed, then Proposition 17 ensures that  $z$  is an approximate root of  $f$ . If the homotopy continuation fails, then we replace  $Q$  by  $Q^2$  and we start again, until the call to  $\text{HC}'$  succeeds. This leads to the deterministic procedure DBP, Algorithm 3. If the computation of  $\text{DBP}(f)$  terminates then the result is an approximate root of  $f$ . Section 2.4 studies the average number of homotopy steps performed by  $\text{DBP}(f)$  while Section 2.5 studies the average total cost of an implementation of DBP in the BSS model extended with the square root.

### 2.4 Average analysis

Let  $f \in \mathbb{S}(\mathcal{H})$  be the input system, a uniform random variable, and we consider a run of the procedure  $\text{DBP}(f)$ . Let  $Q_k$  be the precision at the  $k$ th iteration, namely  $Q_k = N^{2^k}$ . We set

$$f_k = \lfloor f \rfloor_{Q_k}, (g_k, \eta_k) = \text{BP}(\{f\}_{Q_k}), \varepsilon_k = \text{sign}(\pi/2 - d_{\mathbb{S}}(f, g_k)) \text{ and } \rho_k = 3N^{1/2}/Q_k.$$

Let  $\Omega$  be the least  $k$  such that the homotopy continuation with precision check  $\text{HC}'(f_k, \varepsilon_k g_k, \eta_k, \rho_k)$  succeeds. Note that  $\Omega$  is a random variable. To perform the average analysis of the total number of homotopy steps, we first deal with each iteration separately (Lemmas 18 and 19) and then give tail bounds on the probability distribution of  $\Omega$  (Proposition 20). Even if the

$Q_k$   
 $f_k, g_k, \eta_k, \varepsilon_k,$   
 $\rho_k$

---

*Algorithm 3.* Deterministic variant of Beltrán-Pardo algorithm

---

**Input.**  $f \in \mathcal{H}$   
**Output.**  $z \in \mathbb{P}^n$   
**Postcondition.**  $z$  is an approximate root of  $f$

---

**function** DBP( $f$ )  
 $Q \leftarrow N$   
**repeat**  
 $Q \leftarrow Q^2$   
 $f' \leftarrow \lfloor f \rfloor_Q$   
 $(g, \eta) \leftarrow \text{BP}(\{f\}_Q)$   
 $\varepsilon \leftarrow \text{sign}(\text{Re}\langle f, g \rangle)$   
 $\rho \leftarrow (2N)^{1/2}/Q$   
 $z \leftarrow \text{HC}'(f', \varepsilon g, \eta, \rho)$   
**until** HC' succeeds  
**return**  $z$   
**end function**

---

number of steps in each iteration are not independent from each other and from  $\Omega$ , Hölder's inequality allows obtaining a bound on the total number of steps (Theorem 21).

Let  $(g, \eta) \in V$  be a random variable with distribution  $\rho_{\text{std}}$  and independent of  $f$ .  $g, \eta$

**Lemma 18.** *Let  $\Theta : \mathcal{H} \times V \rightarrow \mathbb{R}$  be any nonnegative measurable function. For any  $k \geq 1$ ,*

$$\mathbb{E}(\Theta(f_k, \varepsilon_k g_k, \eta_k)) \leq 10\mathbb{E}(\Theta(f_k, g, \eta)).$$

*Proof.* It is an application of Lemma 11. We first remark that  $\varepsilon_k \in \{-1, 1\}$  so

$$\Theta(f_k, \varepsilon_k g_k, \eta_k) \leq \Theta(f_k, g_k, \eta_k) + \Theta(f_k, -g_k, \eta_k).$$

Then

$$\begin{aligned} \mathbb{E}(\Theta(f_k, g_k, \eta_k)) &= \frac{1}{\text{vol}(\mathbb{S}(\mathcal{H}))} \int_{\mathbb{S}(\mathcal{H})} \Theta(\lfloor f \rfloor_{Q_k}, \text{BP}(\{f\}_{Q_k})) \, df \\ &\leq \frac{\exp\left(\frac{2N^{3/2}}{Q_k}\right)}{\text{vol}(\mathbb{S}(\mathcal{H}))^2} \int_{\mathbb{S}(\mathcal{H}) \times \mathbb{S}(\mathcal{H})} \Theta(\lfloor f \rfloor_{Q_k}, \text{BP}(g)) \, df \, dg \quad \text{by Lemma 11} \\ &= \frac{\exp\left(\frac{2N^{3/2}}{Q_k}\right)}{\text{vol}(\mathbb{S}(\mathcal{H}))} \int_{\mathcal{H}} \int_V \Theta(\lfloor f \rfloor_{Q_k}, g, \eta) \, df \, d\rho_{\text{std}}(g, \eta) \quad \text{by Theorem 9} \\ &= \exp\left(\frac{2N^{3/2}}{Q_k}\right) \mathbb{E}(\Theta(f_k, g, \eta)). \end{aligned}$$

Similarly,  $\mathbb{E}(\Theta(f_k, -g_k, \eta_k)) \leq \exp\left(\frac{2N^{3/2}}{Q_k}\right) \mathbb{E}(\Theta(f_k, -g, \eta))$ , and since  $g$  and  $-g$  have the same probability distribution,  $\mathbb{E}(\Theta(f_k, -g, \eta)) = \mathbb{E}(\Theta(f_k, g, \eta))$ . To conclude, we remark that  $Q_k \geq N^2$  and that  $e^{\sqrt{2}} \leq 5$ . □

**Lemma 19.**  $\mathbb{E}(I_p(f, g, \eta)) = \mathbb{E}(\mu(g, \eta)^p)$  for any  $p \geq 1$  and  $k \geq 1$ .

*Proof.* Let  $h_t = \Gamma(g, f, t)$ , for  $t \in [0, 1]$ , and let  $\zeta_t$  be the associated homotopy continuation. Let  $\tau \in [0, 1]$  be a uniform random variable independent from  $f$  and  $(g, \eta)$ . Clearly  $\mathbb{E}(I_p(f, g, \eta)) = \mathbb{E}(\mu(h_\tau, \zeta_\tau)^p)$ , so it is enough to prove that  $(h_\tau, \zeta_\tau) \sim \rho_{\text{std}}$ . The systems  $f$  and  $g$  are independent and uniformly distributed on  $\mathbb{S}(\mathcal{H})$ . So their probability distributions is invariant under any unitary transformation of  $\mathcal{H}$ . Then so is the probability distribution of  $h_t$  for any  $t \in [0, 1]$ , and there is a unique such probability distribution: the uniform distribution on  $\mathbb{S}(\mathcal{H})$ . The homotopy continuation makes a bijection between the roots of  $g$  and those of  $h_t$ . Since  $\eta$  is uniformly chosen among the roots of  $g$ , so is  $\zeta_t$  among the roots of  $h_t$ . That is,  $(h_t, \zeta_t) \sim \rho_{\text{std}}$  for all  $t \in [0, 1]$ , and then  $(h_\tau, \zeta_\tau) \sim \rho_{\text{std}}$ .  $\square$

**Proposition 20.**  $\mathbb{P}(\Omega > k) \leq 2^{17} D^{9/4} n^{3/2} N^{7/4} Q_k^{-1/2}$ .

*Proof.* The probability that  $\Omega > k$  is no more than the probability that  $\text{HC}'(f_k, g_k, \eta_k, \rho_k)$  fails. By Lemma 18,  $\mathbb{P}(\text{HC}'(f_k, \varepsilon_k g_k, \eta_k, \rho_k) \text{ fails}) \leq 10 \mathbb{P}(\text{HC}'(f_k, g, \eta, \rho_k) \text{ fails})$ . Given that  $d_{\mathbb{S}}(f, f_k) \leq \rho_k$ ,

$$\begin{aligned} \mathbb{P}(\text{HC}'(f_k, g, \eta, \rho_k) \text{ fails}) &\leq \mathbb{P}\left(D^{3/2} M(f, g, \eta)^2 \rho_k \geq \frac{1}{236}\right) && \text{by Proposition 17(iv)} \\ &\leq \mathbb{P}\left(D^{9/2} I_3(f, g, \eta)^2 \rho_k \geq \frac{1}{236 \cdot 151^2}\right) && \text{by Proposition 7} \\ &\leq 151 \cdot 236^{1/2} D^{9/4} \rho_k^{1/2} \mathbb{E}(I_3(f, g, \eta)) && \text{by Markov's inequality.} \end{aligned}$$

Lemma 19 and Theorem 10 imply then

$$\mathbb{E}(I_3(f, g, \eta)) \leq \mathbb{E}(\mu(g, \eta)^3) \leq 3(nN)^{3/2}.$$

All in all, and since  $\rho_k = 3N^{1/2}/Q_k$ ,

$$\mathbb{P}(\Omega > k) \leq 10 \cdot (151 \cdot 236^{1/2} D^{9/4}) \cdot (3N^{1/2}/Q_k)^{1/2} \cdot 3(nN)^{3/2} \leq_{\blacksquare} 2^{17} D^{9/4} n^{3/2} N^{7/4} Q_k^{-1/2} \quad \square$$

Let  $K(f)$  be the total number of homotopy steps performed by procedure  $\text{DBP}(f)$  and let  $K'(f_k, \varepsilon_k g_k, \eta_k, \rho_k)$  be denoted by  $K'(f_k, \varepsilon_k g_k, \eta_k, \rho_k)$ , so that

$$K(f) = \sum_{k=1}^{\Omega} K'(f_k, \varepsilon_k g_k, \eta_k, \rho_k),$$

**Theorem 21.** *If  $N \geq 21$  then  $\mathbb{E}(K(f)) \leq 2^{17} n D^{3/2} N$ .*

*Proof.* Let  $X_k = K'(f_k, \varepsilon_k g_k, \eta_k, \rho_k)$  and let  $0 < p \leq \frac{3}{2}$ . By Lemma 18 and Proposition 17(i),

$$\mathbb{E}(X_k^p)^{1/p} \leq 10 \mathbb{E}\left(\left(158 D^{3/2} d_{\mathbb{S}}(f, g) I_2(f, g, \eta) + 4\right)^p\right)^{1/p},$$

and because  $d_{\mathbb{S}}(f, g) \leq \pi$  and by Minkowski's inequality, we obtain

$$\leq 10 \left(158 D^{3/2} \pi \mathbb{E}(I_2(f, g, \eta)^p)^{1/p} + 4\right).$$

Jensen's inequality implies that  $I_2(f, g, \eta)^p \leq I_{2p}(f, g, \eta)$ . Then  $\mathbb{E}(I_{2p}(f, g, \eta)) \leq \frac{3}{4-2p} (nN)^p \leq 3(nN)^p$ , by Lemma 19 and Theorem 10. In the end,

$$\mathbb{E}(X_k^p)^{1/p} \leq_{\blacksquare} 15000 n D^{3/2} N. \quad (17)$$

Now, let  $p = \log N / (\log N - 1)$ . If  $N \geq 21$  then  $p \leq \frac{3}{2}$ . We write the expectation of  $K(f)$  as

$$\mathbb{E}(K(f)) = \mathbb{E}\left(\sum_{k=1}^{\Omega} X_k\right) = \sum_{k=1}^{\infty} \mathbb{E}(X_k \mathbf{1}_{\Omega \geq k}).$$

Let  $q = 1/\log N$ , so that  $\frac{1}{p} + \frac{1}{q} = 1$ . From Hölder's inequality,  $\mathbb{E}(X_k \mathbf{1}_{\Omega \geq k}) \leq \mathbb{E}(X_k^p)^{1/p} \mathbb{P}(\Omega \geq k)^{1/q}$  and thus

$$\mathbb{E}(K(f)) \leq \max_{k \geq 1} \mathbb{E}(X_k^p)^{1/p} \sum_{k=1}^{\infty} \mathbb{P}(\Omega \geq k)^{1/q}.$$

Lemma 22 below, with  $C = 1$ ,  $L = 4$  and  $\delta = 1/q$ , shows that

$$\sum_{k=1}^{\infty} \mathbb{P}(\Omega \geq k)^{1/q} \leq L + 1 + \frac{2^{17/\log 21} e^5}{e^{2^L} - 1} \leq 6.$$

The claim follows then from Equation (17) and  $6 \cdot 15000 \leq 2^{17}$ .  $\square$

**Lemma 22.** *For any  $C, \delta > 0$  and any integer  $L \geq 2$  such that  $C < N^{\delta 2^L}$ ,*

$$\sum_{k=1}^{\infty} C^k \mathbb{P}(\Omega \geq k)^{\delta} \leq \sum_{k=1}^{L+1} C^k + \frac{(2^{17} N^5)^{\delta} C^{L+2}}{N^{\delta 2^L} - C}.$$

*Proof.* For any  $k$ , Proposition 20 implies that

$$\mathbb{P}(\Omega \geq k) = \mathbb{P}(\Omega > k - 1) \leq 2^{17} D^{9/4} n^{3/2} N^{7/4} Q_{k-1}^{-1/2} \leq 2^{17} N^5 N^{-2^{k-2}},$$

using  $D \leq N$  and  $n^2 \leq N$ . Moreover,  $2^{p-1} \geq p$ , for any integer  $p$ , so that  $N^{-2^{k-2}} \leq N^{-2^L(k-L-1)}$ . Of course, it also holds that  $\mathbb{P}(\Omega \geq k) \leq 1$ . Thus

$$\sum_{k \geq 1} C^k \mathbb{P}(\Omega \geq k)^{\delta} \leq \sum_{k=1}^{L+1} C^k + (2^{17} N^5)^{\delta} \sum_{k=L+2}^{\infty} C^k N^{-\delta 2^L(k-L-1)},$$

and the latter sum is a geometric sum which evaluates to  $C^{L+2}/(N^{\delta 2^L} - C)$ .  $\square$

## 2.5 Implementation in the BSS model with square root

Algorithms HC' and DBP (Algorithms 2 and 3 respectively) have been described assuming the possibility to compute exactly certain nonrational functions: the square root, the trigonometric functions sine and cosine and the operator norm of a linear map. A BSS machine can only approximate them, but it can do it efficiently. I propose here an implementation in the BSS model extended with the ability of computing the square root of a positive real number at unit cost. We could reduce further to the plain BSS model at the cost of some lengthy and nearly irrelevant technical argumentation. We now prove the main result of this article:

**Theorem 23.** *There exists a BSS machine  $A$  with square root and a constant  $c > 0$  such that for any positive integer  $n$  and any positive integers  $d_1, \dots, d_n$ :*

- (i)  $A(f)$  computes an approximate root of  $f$  for almost all  $f \in \mathcal{H}$ ;

(ii) if  $f \in \mathbb{S}(\mathcal{H})$  is a uniform random variable, then the average number of operations performed by  $A(f)$  is at most  $cnD^{3/2}N(N+n^3)$ .

Firstly, we describe an implementation of Algorithms HC' and DBP in the extended BSS model. The first difficulty is the condition number  $\mu(f, z)$ : it rests upon the operator norm for the Euclidean distance which is not computable with rational operations. While there are efficient numerical algorithms to compute such an operator norm in practice, it is not so easy to give an algorithm that approximates it in good complexity in the BSS model.<sup>26</sup> Fortunately, we can easily compute the operator norm of a matrix  $M \in \mathbb{C}^{n \times n}$  within a factor 2 as follows:<sup>27</sup> we first compute a tridiagonalization  $T$  of the Hermitian matrix  $\overline{M}^t M$  with  $\mathcal{O}(n^3)$  operations, using Householder's reduction, and then<sup>28</sup>

$$\frac{1}{\sqrt{3}}\|T\|_1 \leq \|M\|^2 \leq \|T\|_1,$$

where  $\|T\|_1$  is the operator  $\ell_1$ -norm of  $T$ , that is the maximum  $\ell_1$ -norm of a column. Therefore, up to a few modifications in the constants, we may assume that  $\mu(f, z)$  is computable in  $\mathcal{O}(n^3)$  operations, given  $\text{df}(z)$ .

The second difficulty lies in the use of the trigonometric functions sine and cosine. They first appear in the definition of the geodesic path  $\Gamma$ , Equation (3), which is used in Algorithm 2. In the case where  $d_{\mathbb{S}}(f, g) \leq \pi/2$ , it is good enough to replace  $\Gamma(g, f, \delta)$  by

$$\frac{\delta f + (1 - \delta)g}{\|\delta f + (1 - \delta)g\|}.$$

This is classical and implies modifications in the constants only.<sup>29</sup> The trigonometric functions also appear in Sibuya's function  $S$ , see Equation (11). This issue can be handled with power series approximations:

**Lemma 24.** *There is a BSS machine with square root that computes, for any  $N$  and any  $x \in [0, 1]^{2N-1}$ , with  $\mathcal{O}(N \log N)$  operations, a point  $\tilde{S}(x) \in \mathbb{S}^{2N-1}$  such that*

$$\int_{[0,1]^{2N-1}} \Theta(\tilde{S}(x)) dx \leq \frac{2}{\text{vol}(\mathbb{S}^{2N-1})} \int_{\mathbb{S}^{2N-1}} \Theta(y) dy,$$

for any nonnegative measurable function  $\Theta : \mathbb{S}^{2N-1} \rightarrow \mathbb{R}$ .

*Sketch of the proof.* For any positive integer  $Q$ , let  $F_Q(x)$  be the Taylor series expansion at 0, truncated at  $x^Q$ , of the entire function  $(\exp(2i\pi x) - 1)/(x - 1)$ . It is a polynomial of degree  $Q$  that can be computed with  $\mathcal{O}(Q)$  operations, assuming that  $\pi$  is a constant of the machine, by using the linear recurrence  $(n + 2)u_{n+2} = (2i\pi + n + 2)u_{n+1} + 2i\pi u_n$  satisfied by the coefficients of  $F_Q$ . Let  $\text{Cos}_Q(x)$  and  $\text{Sin}_Q(x)$  be the real and imaginary parts of  $(1 + (x - 1)F_Q(x))/(1 + (x - 1)F_Q(x))$  respectively.

The function  $x \in [0, 1] \rightarrow (\text{Cos}_Q(x), \text{Sin}_Q(x))$  gives a parametrization of the circle  $\mathbb{S}^1$  whose

<sup>26</sup>See for example Armentano, Beltrán, Bürgisser, Cucker, and Shub, *A stable, polynomial-time algorithm for the eigenpair problem* or Armentano and Cucker, "A randomized homotopy for the Hermitian eigenpair problem"; unfortunately the Gaussian distribution that they assume does not fit the situation here.

<sup>27</sup>I thank one of the referees for having communicated this method to me.

<sup>28</sup>Kahan, *Accurate eigenvalues of a symmetric tri-diagonal matrix*.

<sup>29</sup>See for example Bürgisser and Cucker, *Condition*, §17.1.

Jacobian is almost constant: we can check that there is a universal constant  $C > 0$  such that

$$|\text{Cos}'_Q(x)^2 + \text{Sin}'_Q(x)^2 - (2\pi)^2| \leq Ce^{-Q}.$$

Thus for any continuous function  $\theta : \mathbb{S}^1 \rightarrow \mathbb{R}$

$$\int_0^1 \theta(\text{Cos}_Q(x), \text{Sin}_Q(x)) dx \leq \frac{1 + Ce^{-Q}}{2\pi} \int_{\mathbb{S}^1} \theta(y) dy.$$

Let  $\tilde{S}$  be the function  $[0, 1]^{2N-1} \rightarrow \mathbb{S}^{2N-1}$  defined in the same way as  $S$ , Equation (11), but with  $\text{Cos}_Q$  and  $\text{Sin}_Q$  in place of  $\sin$  and  $\cos$  respectively, with some  $Q \sim \log N$  such that  $(1 + Ce^{-Q})^N \leq 2$ . It is easy to check that  $\tilde{S}$  satisfies the desired properties.  $\square$

In Algorithm DBP, there is no harm in using  $\tilde{S}$  in place of  $S$ . We obtain this way variants of Algorithms HC' and DBP that fit in the BSS model with square root. It remains to evaluate the overall number of operations. It is well known that  $f(z)$  and  $df(z)$  can be computed at a point  $z \in \mathbb{C}^{n+1}$  in  $\mathcal{O}(N)$  operations—the latter as a consequence of a theorem of Baur and Strassen.<sup>30</sup> Together with the approximate computation of the operator norm discussed above, this implies the following:

**Lemma 25.** *There exists a BSS machine with square root that compute  $\mu(f, z)$  (within a factor 2) and  $\mathcal{N}(f, z)$ , for any  $f \in \mathcal{H}$  and  $z \in \mathbb{P}^n$ , in  $\mathcal{O}(N + n^3)$  operations.*

The cost of the  $k$ th iteration in Algorithm DBP is dominated by the cost of computing  $\lfloor f \rfloor_{Q_k}$  and  $\text{BP}(\{f\}_{Q_k})$  and the cost of the call to HC'. The cost of the call to HC' is dominated by the cost of the homotopy steps. Each homotopy step costs  $\mathcal{O}(N + n^3)$  operations, by Lemma 25.

We now evaluate the cost of computing  $\lfloor f \rfloor_Q$  and  $\text{BP}(\{f\}_Q)$ . Naturally, the integral part  $A(x)$  of a real number  $x$  is not a rational function of  $x$  but it can be computed in the BSS model in  $\mathcal{O}(\log(1 + |x|))$  operations using the recursive formula, say for  $x \geq 0$ ,

$$A(x) = \begin{cases} 0 & \text{if } x < 1 \\ 2A(x/2) & \text{if } x < 2A(x/2) + 1 \\ 2A(x/2) + 1 & \text{otherwise.} \end{cases}$$

Hence we can compute  $Q^{-1}A(Qx)$  in  $\mathcal{O}(\log Q)$  operations, for any positive integer  $Q$  and  $x \in [0, 1]$ . It follows that one can compute  $\lfloor f \rfloor_Q$  in  $\mathcal{O}(N \log Q)$  operations. The computation of  $\{f\}_Q$  is similar and it is done in  $\mathcal{O}(N \log Q + N \log N)$  operations, by Lemma 24, using  $\tilde{S}$  in place of Sibuya's function  $S$ . Finally, given  $\{f\}_Q$ , one compute  $\text{BP}(\{f\}_Q)$  in  $\mathcal{O}(N^2 + N \log Q_k)$  operations. In the end, the cost of the  $k$ th operation, excluding the call to HC', is thus  $\mathcal{O}(N^2 + N \log Q_k)$  operations.

Hence, the overall cost of the algorithm is

$$\mathcal{O} \left( (N + n^3)K(f) + \sum_{k=1}^{\Omega} (N^2 + N \log Q_k) \right) = \mathcal{O} \left( (N + n^3)K(f) + N^2\Omega + N \sum_{k=1}^{\Omega} \log Q_k \right),$$

where  $K(f)$  is the total number of homotopy steps. By Theorem 21,  $\mathbb{E}(K(f)) = \mathcal{O}(nD^{3/2}N)$ , so it only remains to bound the expectations of  $\Omega$  and  $\sum_{k=1}^{\Omega} \log Q_k$ .

<sup>30</sup>Baur and Strassen, "The complexity of partial derivatives".

**Lemma 26.**  $\mathbb{E}(\Omega) \leq 7$

*Proof.* By Lemma 22, with  $C = 1$ ,  $\delta = 1$  and  $L = 5$ ,

$$\mathbb{E}(\Omega) = \sum_{k=1}^{\Omega} \mathbb{P}(\Omega \geq k) \leq L + 1 + \frac{2^{L+17} N^5}{N^{2^L} - 1} \leq_{\blacksquare} 7. \quad \square$$

**Lemma 27.**  $\mathbb{E}\left(\sum_{k=1}^{\Omega} \log Q_k\right) \leq 129 \log N$ .

*Proof.* Because  $Q_k = N^{2^k}$ ,

$$\mathbb{E}\left(\sum_{k=1}^{\Omega} \log Q_k\right) = \sum_{k=1}^{\infty} \log Q_k \mathbb{P}(\Omega \geq k) = \log N \sum_{k=1}^{\infty} 2^k \mathbb{P}(\Omega \geq k).$$

Lemma 22, with  $\delta = 1$ ,  $C = 2$  and  $L = 5$  gives that

$$\sum_{k=1}^{\infty} 2^k \mathbb{P}(\Omega \geq k) \leq 2^{L+2} + \frac{2^{L+19} N^5}{N^{2^L} - 2} \leq_{\blacksquare} 129,$$

where we used that  $N \geq 2$ .  $\square$

This concludes the proof of Theorem 23.

## Conclusion

The derandomization proposed here relies on extracting randomness from the input itself, which is made possible by the BSS model and the infinite precision it provides. Actually, this might also work under finite, and rather moderate, precision. Indeed, in the  $k$ th iteration of Algorithm 3, we need, very loosely speaking, about  $\log Q_k$  digits of precision but not much more, because by construction, the homotopy continuation procedure  $\text{HC}'$  aborts when more precision would be required for the result to be relevant. And then, Lemma 27 shows that  $\log Q_k$ , is typically no more than  $129 \log N$ . So it is reasonable to think that, extending the work of Briquel et al.,<sup>31</sup> we may run a variant of Algorithm 3 on a finite precision machine and obtain a significant probability of success as long as we work with  $C \log N$  digits of precision, for some constant  $C > 0$ .

Besides, Armentano et al.<sup>32</sup> proposed recently a new complexity analysis of Beltrán-Pardo algorithm which relies on a refined homotopy continuation algorithm.<sup>33</sup> They obtained a randomized algorithm that terminates on the average on a random input after  $\mathcal{O}(nD^{3/2}N^{1/2})$  homotopy steps. This is a significant improvement on the previously known  $\mathcal{O}(nD^{3/2}N)$  bound. The derandomization method should also apply to this refined algorithm, in all likelihood, but this is not immediate: to devise the homotopy continuation with precision check, we had to look deep inside the continuation process. Adapting the method to a refined homotopy continuation process will inevitably lead to further difficulties.

<sup>31</sup>Briquel, Cucker, Peña, and Roshchina, “Fast computation of zeros of polynomial systems with bounded degree under finite-precision”.

<sup>32</sup>Armentano, Beltrán, Bürgisser, Cucker, and Shub, “Condition length and complexity for the solution of polynomial systems”.

<sup>33</sup>Shub, “Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric”; Beltrán, “A continuation method to solve polynomial systems and its complexity”.

## References

- Diego Armentano, Carlos Beltrán, Peter Bürgisser, Felipe Cucker, and Michael Shub. *A stable, polynomial-time algorithm for the eigenpair problem*. 2015. arXiv: 1505.03290.
- “Condition length and complexity for the solution of polynomial systems”. In: *Found. Comput. Math.* (2016). DOI: 10.1007/s10208-016-9309-9.
- Diego Armentano and Felipe Cucker. “A randomized homotopy for the Hermitian eigenpair problem”. In: *Found. Comput. Math.* 15.1 (2015), pp. 281–312. DOI: 10.1007/s10208-014-9217-9.
- Walter Baur and Volker Strassen. “The complexity of partial derivatives”. In: *Theoretical Computer Science* 22.3 (1983), pp. 317–330. DOI: [http://dx.doi.org/10.1016/0304-3975\(83\)90110-X](http://dx.doi.org/10.1016/0304-3975(83)90110-X).
- Carlos Beltrán. “A continuation method to solve polynomial systems and its complexity”. In: *Numer. Math.* 117.1 (2011), pp. 89–113. DOI: 10.1007/s00211-010-0334-3.
- Carlos Beltrán and Luis Miguel Pardo. “Fast linear homotopy to find approximate zeros of polynomial systems”. In: *Found. Comput. Math.* 11.1 (2011), pp. 95–129. DOI: 10.1007/s10208-010-9078-9.
- “Smale’s 17th problem: average polynomial time to compute affine and projective solutions”. In: *J. Amer. Math. Soc.* 22.2 (2009), pp. 363–385. DOI: 10.1090/S0894-0347-08-00630-9.
- Lenore Blum, Felipe Cucker, Michael Shub, and Steve Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. DOI: 10.1007/978-1-4612-0701-6.
- Lenore Blum, Michael Shub, and Steve Smale. “On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines”. In: *Bull. Amer. Math. Soc. N.S.* 21.1 (1989), pp. 1–46. DOI: 10.1090/S0273-0979-1989-15750-9.
- Irénée Briquel, Felipe Cucker, Javier Peña, and Vera Roshchina. “Fast computation of zeros of polynomial systems with bounded degree under finite-precision”. In: *Math. Comp.* 83.287 (2014), pp. 1279–1317. DOI: 10.1090/S0025-5718-2013-02765-2.
- Peter Bürgisser and Felipe Cucker. *Condition. The geometry of numerical algorithms*. Vol. 349. Grundlehren der Mathematischen Wissenschaften. Springer Berlin Heidelberg, 2013. DOI: 10.1007/978-3-642-38896-5.
- “On a problem posed by Steve Smale”. In: *Ann. of Math. (2)* 174.3 (2011), pp. 1785–1836. DOI: 10.4007/annals.2011.174.3.8.
- William Kahan. *Accurate eigenvalues of a symmetric tri-diagonal matrix*. Tech. rep. CS41. Stanford University, 1966.
- Michael Shub. “Complexity of Bezout’s theorem. VI. Geodesics in the condition (number) metric”. In: *Found. Comput. Math.* 9.2 (2009), pp. 171–178. DOI: 10.1007/s10208-007-9017-6.

- Michael Shub. “Some remarks on Bezout’s theorem and complexity theory”. In: *From Topology to Computation: Proceedings of the Smalefest*. Springer, New York, 1993. Chap. 40, pp. 443–455. DOI: 10.1007/978-1-4612-2740-3\_40.
- Michael Shub and Steve Smale. “Complexity of Bézout’s theorem. I. Geometric aspects”. In: *J. Amer. Math. Soc.* 6.2 (1993), pp. 459–501. DOI: 10.2307/2152805.
- “Complexity of Bezout’s theorem. II. Volumes and probabilities”. In: *Computational algebraic geometry (Nice, 1992)*. Vol. 109. Progr. Math. Birkhäuser Boston, Boston, MA, 1993, pp. 267–285.
  - “Complexity of Bezout’s theorem. IV. Probability of success; extensions”. In: *SIAM J. Numer. Anal.* 33.1 (1996), pp. 128–148. DOI: 10.1137/0733008.
  - “Complexity of Bezout’s theorem. V. Polynomial time”. In: *Theoret. Comput. Sci.* 133.1 (1994). Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993), pp. 141–164. DOI: 10.1016/0304-3975(94)90122-8.
- Masaaki Sibuya. “A method for generating uniformly distributed points on  $N$ -dimensional spheres”. In: *Ann. Inst. Statist. Math.* 14 (1962), pp. 81–85. DOI: 10.1007/BF02868626.
- Steve Smale. “Mathematical problems for the next century”. In: *The Mathematical Intelligencer* 20.2 (1998), pp. 7–15. DOI: 10.1007/BF03025291.
- “Newton’s method estimates from data at one point”. In: *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*. Springer, New York, 1986, pp. 185–196. DOI: 10.1007/978-1-4612-4984-9\_13.
- Daniel Spielman and Shang-Hua Teng. “Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time”. In: *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*. ACM, New York, 2001, 296–305 (electronic). DOI: 10.1145/380752.380813.

Technische Universität Berlin, Institut für Mathematik, Sekretariat MA 3-2  
 Straße des 17. Juni 136, 10623 Berlin, Deutschland

*E-mail address:* pierre@lairez.fr

*URL:* pierre.lairez.fr