Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems

Pierre Lairez

RIGID CONTINUATION PATHS I. QUASILINEAR AVERAGE COMPLEXITY FOR SOLVING POLYNOMIAL SYSTEMS

PIERRE LAIREZ

Abstract. How many operations do we need on average to compute an approximate root of a random Gaussian polynomial system? Beyond Smale's 17th problem that asked whether a polynomial bound is possible, we prove a quasi-optimal bound $(\text{input size})^{1+o(1)}$. This improves upon the previously known $(\text{input size})^{\frac{3}{2}+o(1)}$ bound.

The new algorithm relies on numerical continuation along *rigid continuation paths*. The central idea is to consider rigid motions of the equations rather than line segments in the linear space of all polynomial systems. This leads to a better average condition number and allows for bigger steps. We show that on average, we can compute one approximate root of a random Gaussian polynomial system of n equations of degree at most D in $n+1$ homogeneous variables with $O(n^4D^2)$ continuation steps. This is a decisive improvement over previous bounds that prove no better than $\sqrt{2}^{\min(n,D)}$ continuation steps on average.

Contents

1. Introduction

Following a line of research opened in the 20th century by Smale (1985, 1986), Renegar (1987, 1989), Demmel (1988), Shub (1993), Malajovich (1994), and Shub and Smale (1993a,b,c, 1994, 1996) and developped in the 21st century by Armentano et al. (2016, 2018), Bates et al. (2013), Beltrán (2011), Beltrán and Pardo (2008, 2009a,b, 2011), Beltrán and Shub (2009), Briquel et al. (2014), Bürgisser and Cucker (2011, 2013), Hauenstein and Liddell (2016), Hauenstein and Sottile (2012), Lairez (2017), and Malajovich (2018), to name a few, I am interested in the number of elementary operations that one needs to compute one zero of a polynomial system in a numerical setting. On this topic, Smale's question is a landmark: "Can a zero of n complex polynomial equations in n unknowns be found approximately, on average, in polynomial time with a uniform algorithm?" (Smale 1998, 17th problem). The wording is crafted to have a positive answer in spite of two major obstacles. The first

Date: September 10, 2019.
2000 *Mathematics Subject Classification.* Primary 68Q25; Secondary 65H10, 65H20, 65Y20.

one is the NP-completeness of many problems related to deciding the feasability of a polynomial system. Here, we consider well determined systems (as many equations as unknowns), over the complex numbers, in the average (*a fortiori* generic) case, so there will always be a zero. The second obstacle is the number of zeros: it is not polynomially bounded in terms of the input size (the number of coefficients that define the input system). Here, we ask for only one zero and numerical methods can take advantage of it.

Smale's question is now solved (Beltrán and Pardo 2009b; Bürgisser and Cucker 2011; Lairez 2017); it is an achievement but not an end. The most obvious question that pops up is to improve the degree hidden behind the words "polynomial time". This article presents an optimal answer, bringing down "polynomial time", that is $N^{O(1)}$, where N is the input size, to "quasilinear time", that is $N^{1+o(1)}$. The previous state of the art was $N^{\frac{3}{2}+o(1)}$ (Armentano et al. 2016).

1.1. **State of the art.** Let n and $d_1, \dots, d_n$ be positive integers, and let $\mathcal{H}$ be the vector space of tuples $(f_1, \dots, f_n)$ of complex homogeneous polynomials of respective degrees $d_1, \dots, d_n$ in the variables $x_0, \dots, x_n$. Let D denote $\max(d_1, \dots, d_n)$.

We are interested in the average complexity of finding one zero of a polynomial system, given as an element of $\mathcal{H}$. The complexity is measured with respect to the *input size*, denoted N. This is the number of complex coefficients that describe a system, namely

$$N \doteq \dim_{\mathbb{C}} \mathcal{H} = \binom{d_1 + n}{n} + \cdots + \binom{d_n + n}{n}.$$

Note that $N \geqslant 2^{\min(n,D)}$. "Average complexity" means that $\mathcal{H}$ is endowed with a probability measure (uniform on the unit sphere for some suitably chosen Hermitian norm) and that the behaviour the algorithms is analyzed *on average*, assuming that the input is distributed according to this probability measure. We will make use of randomized algorithms, that draw random numbers during their execution. In this case, the average complexity is an average with respect to both the input's distribution and the randomness used internally by the algorithm.

1.1.1. *Classical theory.* In the Shub–Smale–Beltrán–Pardo–Bürgisser–Cucker way of doing things, we compute a zero of a homogeneous polynomial system $F \in \mathcal{H}$ by numerical continuation from a random system $G \in \mathcal{H}$ of which we happen to know a zero $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$. The continuation is performed along the deformation $F_t \doteq \frac{1}{\|tF+(1-t)G\|}(tF + (1-t)G)$. Starting from $t = 0$, we repeatedly increment the parameter t and track a zero of F_t with a projective Newton iteration applied to the previous approximation of the zero. If the increment is small enough then we can be sure not to loose the zero and to obtain, when t reaches 1, an approximate zero of the target system F. The total complexity of the algorithm depends on the number of continuation steps that are performed, which in turn depends on the size of the increment. The key issue is to specify how small is "small enough".

Smale (1986) gave a sufficient condition for the Newton iteration to converge in terms of the *gamma number* $\gamma(F, z)$, depending on a polynomial system F and a point z (see §2.5(8) for a definition). Difficulties in estimating the variations of $\gamma(F, z)$ with respect to F led Shub and Smale (1993a) to consider the *condition number* $\mu(F, z)$, which upperbounds $\gamma(F, z)$ and characterizes how much a zero z of a system F is affected by a small pertubation of F. They gave a sufficient condition for a continuation step to be small enough in terms of μ. After some refinements,

Shub (2009) proved that $K(F, G, \zeta)$, the minimal number of steps to go from G to F while tracking the zero ζ, is bounded by

$$K(F, G, \zeta) \leqslant (\text{constant}) \int_0^1 \mu(F_t, \zeta_t) \sqrt{\|\dot{F}_t\|^2 + \|\dot{\zeta}_t\|^2} \mathrm{d}t. \tag{1}$$

This is called the "μ estimate". Explicit algorithms that achieve this bound have been designed by Beltrán (2011), Dedieu et al. (2013), and Hauenstein and Liddell (2016). A simpler but weaker form, called the "μ^2 estimate", reads

$$K(F, G, \zeta) \leqslant (\text{constant}) D^{\frac{3}{2}} d_{\mathbb{S}}(F, G) \int_0^1 \mu(F_t, \zeta_t)^2 \mathrm{d}t, \tag{2}$$

where $d_{\mathbb{S}}(F, G)$ is the distance in the unit sphere $\mathbb{S}(\mathcal{H})$ from F to G, that is the length of the continuation path. It is often used in practice because it is much easier to design algorithms that achieve this bound rather than the former. In one form or the other, this kind of integral estimate for the number of steps is the first mainstay of the method.

The second mainstay is a procedure discovered by Beltrán and Pardo (2011) and simplified by Bürgisser and Cucker (2011) to sample a Gaussian random system $G \in \mathcal{H}$ together with one of its zeros without the need for solving a polynomial system: (1) sample a random Gaussian linear map $L : \mathbb{C}^{n+1} \to \mathbb{C}^n$, (2) compute a nonzero vector $\zeta \in \mathbb{C}^{n+1}$ in the kernel of L and (3) sample a random Gaussian system in the affine subspace of $\mathcal{H}$ of all systems G such that $G(\zeta) = 0$ and $\mathrm{d}_\zeta G = L$. By construction, we obtain a system G and one of its zeros ζ. Less trivially, $G/\|G\|$ is uniformly distributed in the sphere $\mathbb{S}(\mathcal{H})$. We could think of a simpler procedure that (1) samples some $\zeta \in \mathbb{C}^{n+1}$ isotropically and (2) samples a random Gaussian system in the linear subspace of $\mathcal{H}$ of all systems G such that $G(\zeta) = 0$. This also gives a random system with one of its zeros, by construction, but the system is not uniformly distributed in the sphere after normalization.

These two mainstays together give a randomized algorithm to compute a zero of a polynomial system and a way to analyze its average complexity on a random input. On input $F \in \mathbb{S}(\mathcal{H})$, the algorithm is: (1) uniformly sample a random system $G \in \mathbb{S}(\mathcal{H})$ together with a zero ζ, (2) perform the numerical continuation from G to F tracking the zero ζ. If F itself is a uniformly distributed random variable, then for any $t \in [0, 1]$, F_t is also uniformly distributed, so (F_t, ζ_t) has the same distribution as (G, ζ). Therefore, the average number of steps performed by the algorithm is bounded by

$$\begin{aligned}
\mathbb{E}\left[K(F, G, \zeta)\right] &\leqslant (\text{constant})\, \mathbb{E}\left[D^{\frac{3}{2}} d_{\mathbb{S}}(F, G) \int_0^1 \mu(F_t, \zeta_t)^2 \mathrm{d}t\right] \\
&\leqslant (\text{constant}) D^{\frac{3}{2}} \int_0^1 \mathbb{E}\left[\mu(F_t, \zeta_t)^2\right] \mathrm{d}t \\
&\leqslant (\text{constant}) D^{\frac{3}{2}}\, \mathbb{E}\left[\mu(G, \zeta)^2\right].
\end{aligned}$$

This leads us to the third mainstay: estimates for $\mathbb{E}\left[\mu(G, \zeta)^2\right]$. Beltrán and Pardo (2011, Theorem 23) proved that $\mathbb{E}\left[\mu(G, \zeta)^2\right] \leqslant nN$. Therefore, the average number of steps performed by the algorithm on a random input is

$$\mathbb{E}[K(F, G, \zeta)] \leqslant (\text{constant}) nN.$$

The cost of each continuation step (basically, the computation of μ and a Newton's iteration) can be done in $O(N)$ operations (when $D \geqslant 2$). All in all, the total

average complexity of the classical algorithm is $O(nD^{\frac{3}{2}}N^2)$ as $N \to \infty$. When $\min(n, D) \to \infty$, then this is $N^{2+o(1)}$.

1.1.2. *Improvements.* How can we improve upon this complexity bound? We cannot do much about the $O(N)$ cost of a continuation step, as it is already optimal. Concerning the number of steps, we can try to use the μ estimate instead of the μ^2 estimate. Bounding $\|\dot{\zeta}_t\|$ by $\mu(F_t, \zeta_t)\|\dot{F}_t\|$ (which turns the μ estimate into the μ^2 estimate) is optimal in the worst case, but on average, when the direction $\dot{F}_t$ is random, this is pessimistic. Building upon this idea, Armentano et al. (2016) proved that $O(nD^{\frac{3}{2}}N^{\frac{1}{2}})$ continuation steps are enough on average. This leads to a total average complexity of $N^{\frac{3}{2}+o(1)}$ operations.

Beltrán and Shub (2009) proved that there exist continuation paths that makes the μ estimate polynomially bounded in terms of n and D. The construction is explicit but it requires the knowledge of a zero of a target system. This prevents it from being used algorithmically. Yet, it was the first time that the possibility of performing numerical continuation in very few steps (polynomially many with respect to n and D, not N) was supported.

Lastly, let us mention that Hauenstein and Liddell (2016) developped a γ estimate, based on Smale's γ number. It may be used as a starting point to obtain very similar results to ours in a more traditional context. However, this direction is yet to be explored.

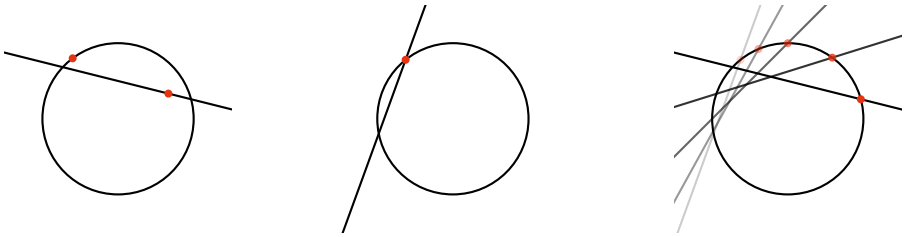
1.2. **Contribution.** I describe a randomized algorithm that uses a numerical continuation from a random start system to find one root of the input system. It performs $O(n^4D^2)$ steps on average (and each step costs $O(n^2D^2N)$ operations) for random Gaussian systems. This leads to a total average complexity of $O(n^6D^4N)$ operations as $N \to \infty$ to find one approximate root of a random Gaussian system (Theorem 40). When $\min(n, D) \to \infty$, this is $N^{1+o(1)}$. The algorithm relies on analogues of the three mainstays of the classical theory: integral estimate for the number of steps, randomization of the start system and average analysis of some condition number. However, the basic tools are thoroughly renewed.

The starting point is the observation that a typical system in $\mathcal{H}$ is poorly conditioned. As mentionned above, the expected squared condition number of a random system at a random zero is bounded by nN and it turns out that this is rather sharp. In view of Smale's question, this is satisfying, much more than bounds involving the total number of zeros, but this N is the limit of the method.

To improve the average conditioning, an idea is to define the notion of conditioning with respect to a much lower dimensional parameter space, but still big enough to be able to develop an analogue of Beltrán and Pardo's algorithm. I propose here the *rigid* setting, where the parameter space is not the whole space of polynomial systems, but the group $\mathcal{U}$ made of n copies of the unitary group $U(n+1)$, of real dimension $\sim n^3$. It acts by rigid motions on the n components of a fixed, well determined, polynomial system. Figure 1 illustrates a *rigid continuation path.*

Less parameters is less opportunities for a dramatic pertubation that will ruin the conditioning of a system. Beyond that, the continuation paths in the rigid setting preserve the geometry of the input equations. This opens a way for studying the average complexity of solving certain *structured* systems. Forthcoming work will address this topic.

A noteworthy contribution is the introduction of the *split gamma number* which tightly upper bounds Smale's gamma number (Theorem 13) and which allows for



(a) Compute one solution of each equation.

(b) Move the hypersurfaces to make the solution match.

(c) Continuously return to the original position while tracking the solution.

Figure 1. Resolution of a polynomial system with a rigid continuation path.

interesting average analysis, see §4.4. Finally, the foremost outcome of the rigid setting is Theorem 27 which gives an average bound on the necessary number of continuation steps to compute one root of a random system, with only a unitary invariance hypothesis on the probability distribution.

Ackowledgment. It is my pleasure to thank Carlos Beltrán, Peter Bürgisser and Felipe Cucker for many helpful discussions and valuable comments. I am very grateful to the referees for their conscientious work.

1.3. **Notations and basic definitions.**

n	some positive integer (used as the *number of nonhomogeneous variables*).
$\mathbb{P}^n$	complex projective space of dimension n.
$[z]$	projective class of some nonzero $z \in \mathbb{C}^{n+1}$.
$d_{\mathbb{P}}$	geodesic distance on $\mathbb{P}^n$ endowed with the Fubini-Study metric, that is $d_{\mathbb{P}}([x],[y]) = \arcsin\left(\sqrt{1-\lvert\langle x,y\rangle\rvert^2}\right)$ for any $x, y \in \mathbb{S}(\mathbb{C}^{n+1})$.
H_d	space of complex homogeneous polynomials of degree d in $x_0, \dots, x_n$.
r	some positive integer (used as the *number of equations*).
$d_1, \dots, d_r$	some positive integers (used as the *degrees of the equations*).
D	the maximum of $d_1, \dots, d_r$.
$\mathcal{H}[r]$	space of homogeneous systems of r equations of degree $d_1, \dots, d_r$, that is $H_{d_1} \times \cdots \times H_{d_r}$. Elements of $\mathcal{H}[r]$ are often considered as polynomial maps $\mathbb{C}^{n+1} \to \mathbb{C}^r$.
N	the *input size*, defined as $\dim_{\mathbb{C}} \mathcal{H}[r]$.
$U(k)$	group of unitary $k \times k$ matrices.
u^*	conjugate transpose of u.
$\mathcal{U}$	the group of r-uples of unitary matrices, $U(n+1)^r$. Elements of $\mathcal{U}$ are denoted in boldface, like $\mathbf{u}$.
$\mathbf{1}_{\mathcal{U}}$	$(\mathrm{id}, \dots, \mathrm{id})$, the neutral element in $\mathcal{U}$.
$\lVert - \rVert$	norm in a Hermitian space.
$\lVert - \rVert_W$	Weyl norm of a polynomial (see Bürgisser and Cucker 2013, §16.1).
$\lvert\lvert\lvert - \rvert\rvert\rvert$	operator norm of a map between Hermitian spaces. For a multilinear map $\varphi : E^k \to V$, this is $\sup\{\lVert\varphi(e_1, \dots, e_k)\rVert \mid \lVert e_1\rVert = \cdots = \lVert e_k\rVert = 1\}$.
$\lVert - \rVert_{\mathrm{Frob}}$	Frobenius norm of a map between Hermitian spaces.
$\lVert - \rVert_u$	$1/\sqrt{2}$ times $\lVert - \rVert_{\mathrm{Frob}}$ (used as the Riemannian metric on the tangent spaces of $U(n+1)$).

$\varphi^\dagger$ Moore-Penrose pseudo-inverse of a surjective linear map $\varphi : E \to F$, it is the unique linear map $\psi : F \to E$ such that $\varphi\psi = \mathrm{id}_F$ and $\psi\varphi$ is the orthogonal projection onto the row space of φ (the orthogonal complement of the kernel).

$\mathrm{d}_z F$ the derivative of some polynomial map $F : \mathbb{C}^{n+1} \to \mathbb{C}^r$ at $z \in \mathbb{C}^{n+1}$. We will use the same notation with $z \in \mathbb{P}^n$, which means that we choose a representative $\bar{z} \in \mathbb{C}^{n+1}$ of z such that $\|\bar{z}\| = 1$.

$\mathrm{d}_z F^\dagger$ the pseudo-inverse of the derivative.

$\mathcal{N}_F$ projective Newton's operator associated to F

$\doteq$ "is defined as"

$A = O(B)$ as $C \to \infty$
"there are $C_0 \geqslant 0$ and $k \geqslant 0$ such that $C \geqslant C_0 \Rightarrow A \leqslant kB$".

standard normal variable
a Gaussian random variable of an Euclidean space with unit covariance matrix in some orthonormal basis. The notion is relative to the underlying Euclidean inner product. For a Hermitian space, we consider the induced Euclidean structure.

2. Rigid solution varieties

The classical solution variety is the subvariety of $\mathcal{H}[r] \times \mathbb{P}^n$ of all (F, ζ) such that ζ is a zero of F. We now introduce an analogue variety in the rigid setting.

Let $X_1, \dots, X_r$ be pure-dimensional subvarieties of $\mathbb{P}^n$, with $\sum_i \operatorname{codim} X_i \leqslant n$. Let $\mathcal{U}$ denote the group $U(n+1)^r$. It acts naturally on the product $(\mathbb{P}^n)^r$ of r copies of the projective space. We denote its elements in boldface $\mathbf{u} = (u_1, \dots, u_r)$. Let $\mathcal{X}$ denote the product variety $X_1 \times \cdots \times X_r \subset (\mathbb{P}^n)^r$. For $\mathbf{u} \in \mathcal{U}$, let $\mathbf{u}\mathcal{X}$ denote the image of $\mathcal{X}$ under the action of $\mathbf{u}$, that is $\prod_{i=1}^r u_i X_i$, and let $\cap\mathbf{u}\mathcal{X}$ denote the intersection $\cap_{i=1}^r u_i X_i \subseteq \mathbb{P}^n$. The *rigid solution variety* is defined as

$$\mathcal{V} \doteq \{(\mathbf{u}, x) \in \mathcal{U} \times \mathbb{P}^n \mid x \in \cap\mathbf{u}\mathcal{X}\}.$$

There is not a single solution variety, but rather any choice of subvarieties $X_1, \dots, X_r$ leads to a solution variety. In this section, we will study the geometry of $\mathcal{V}$ with $X_1, \dots, X_r$ fixed. Later on, we will assume that $X_1, \dots, X_r$ are hypersurfaces defined by random polynomials.

Let $\mathbb{G}(k)$ denote the Grassmannian of k-dimensional projective subspaces of $\mathbb{P}^n$, that is $k+1$-dimensional linear subspaces of $\mathbb{C}^{n+1}$. For a smooth point $x \in X_i$, the projectivization of the tangent space of the cone over X_i at some representative $\bar{x} \in \mathbb{C}^{n+1}$ of x is an element of $\mathbb{G}(k_i)$ and is denoted $\mathbb{T}_x X_i$. If X_i is the zero set of some homogeneous polynomial system $F_i \in \mathcal{H}[m]$, then $\mathbb{T}_x(u_i X_i)$ is the projectivization of the kernel of $\mathrm{d}_x(F_i \circ u_i^{-1})$. Let $\mathcal{L} \doteq \mathbb{G}(\dim X_1) \times \cdots \times \mathbb{G}(\dim X_r)$, and for $\mathbf{h} = (h_1, \dots, h_r) \in \mathcal{L}$, let $\cap\mathbf{h}$ denote the intersection of the h_i in $\mathbb{P}^n$. To a generic point $(\mathbf{u}, x)$ of $\mathcal{V}$, we associate the linearization

$$L(\mathbf{u}, x) \doteq \big(\mathbb{T}_x(u_1 X_1), \dots, \mathbb{T}_x(u_r X_r)\big) \in \mathcal{L}.$$

Note that $x \in \cap L(\mathbf{u}, x)$.

This section aims at three goals: describe precisely the so-called *standard* distribution on $\mathcal{V}$ (Theorem 8), give an algorithm to sample from this distribution (Algorithm 1) and define the *split gamma number*, a variant of the gamma number well adapted to the rigid setting.

2.1. **Determinant of subspaces and incidence condition number.** Let $E_1,\dots,E_r$ be nonzero linear subspaces of a Hermitian space V. Let π_i be the orthogonal projector on E_i. We define the *multiprojection* map $\operatorname{proj}(E_1,\dots,E_r)$ by

$$\begin{aligned}\operatorname{proj}(E_1,\dots,E_r) : V &\longrightarrow E_1\times\cdots\times E_r\\ v &\longmapsto (\pi_i v,\dots,\pi_r v).\end{aligned}$$

We say that the family $E_1,\dots,E_r$ is *nondegenerate* if $\sum_i \dim E_i = \dim\left(\sum_i E_i\right)$, or, equivalently, when the multiprojection map is surjective.

We define the *determinant* of $E_1,\dots,E_r$ as

$$\det(E_1,\dots,E_r) \doteq \left|\det\left(\operatorname{proj}(E_1,\dots,E_r)_{|E_1+\cdots+E_r}\right)\right|,$$

in the nondegenerate case and $\det(E_1,\dots,E_r)\doteq 0$ otherwise. Note that the determinant of a map between two Hermitian spaces is well defined up to multiplication by some $e^{i\theta}$, so that the modulus is well defined. We also define the *orthogonal determinant* of $E_1,\dots,E_r$ as

$$\det{}^{\perp}(E_1,\dots,E_r) \doteq \det(E_1^{\perp},\dots,E_r^{\perp}).$$

Lastly, we define the *incidence condition number* of $E_1,\dots,E_r$ as

$$\kappa(E_1,\dots,E_r) \doteq \left\|\!\left|\operatorname{proj}(E_1,\dots,E_r)^{\dagger}\right|\!\right\|$$

when the multiprojection map is surjective, and $\kappa(E_1,\dots,E_r)\doteq\infty$ otherwise. With the appropriate distance, the incidence condition number is the inverse of the distance of the tuple $(E_1,\dots,E_r)$ to the closest $(F_1,\dots,F_r)$ such that $\dim\left(\sum_i F_i\right) < \sum_i \dim F_i$ (Breiding and Vannieuwenhoven 2018, Theorems 1.1 and 1.3).

Lemma 1. *For any subspaces $E_1,\dots,E_r\subseteq V$, $\kappa(E_1,\dots,E_r)\geqslant 1$, with equality if and only if $E_1,\dots,E_r$ are orthogonal subspaces.*

Proof. If the family $E_1,\dots,E_r$ is degenerate then $\kappa(E_1,\dots,E_r)=\infty\geqslant 1$, so we may assume it is nondegenerate. The Hermitian transpose of $\operatorname{proj}(E_1,\dots,E_r)$ is the map $\operatorname{sum}(E_1,\dots,E_r) : E_1\times\cdots\times E_r\to V$ defined by

$$\operatorname{sum}(E_1,\dots,E_r) : (u_1,\dots,u_r)\longmapsto u_1+\cdots+u_r,$$

so that $\kappa(E_1,\dots,E_r) = \left\|\!\left|\operatorname{proj}(E_1,\dots,E_r)^{\dagger}\right|\!\right\| = \left\|\!\left|\operatorname{sum}(E_1,\dots,E_r)^{\dagger}\right|\!\right\|$. In the nondegenerate case, $\operatorname{sum}(E_1,\dots,E_r)$ is injective, and $\operatorname{sum}(E_1,\dots,E_r)^{\dagger}$ is a left inverse, that is, for any $(v_1,\dots,v_r)\in\prod_i E_i$,

$$(v_1,\dots,v_r) = \operatorname{sum}(E_1,\dots,E_r)^{\dagger}(v_1+\cdots+v_r),$$

and, in particular,

$$\sum_i \|v_i\|^2 \leqslant \left\|\!\left|\operatorname{sum}(E_1,\dots,E_r)^{\dagger}\right|\!\right\|^2 \left\|\sum_i v_i\right\|^2. \tag{3}$$

Choosing $v_j\neq 0$ for some j and $v_i=0$ for $i\neq j$ shows that $\left\|\!\left|\operatorname{sum}(E_1,\dots,E_r)^{\dagger}\right|\!\right\|\geqslant 1$.

If $\left\|\!\left|\operatorname{sum}(E_1,\dots,E_r)^{\dagger}\right|\!\right\| = 1$, we have by (3), for any $v_i\in E_i$ and $v_j\in E_j$, $i\neq j$,

$$\|v_i\|^2+\|v_j\|^2 \leqslant \|v_i+v_j\|^2 = \|v_i\|^2+\|v_j\|^2+2\Re\langle v_i,v_j\rangle,$$

which implies that the real part of $\langle v_i,v_j\rangle$ is nonnegative. Since it holds for $-v_i$ and v_j too, it follows that E_i and E_j are orthogonal. Conversely, if $E_1,\dots,E_r$ are orthogonal, then the map $\operatorname{sum}(E_1,\dots,E_r)$ is an isometric embedding, and thus $\kappa(E_1,\dots,E_r)=1$. □

Lemma 2. *Let $P\doteq\pi_1+\cdots+\pi_r$, it is a self-adjoint endomorphism of the subspace $\sum_i E_i\subseteq V$. In the nondegenerate case,* $\det(E_1,\dots,E_r)^2 = \det P$.

Proof. Since $P = \mathrm{sum}(E_1, \dots, E_r) \cdot \mathrm{proj}(E_1, \dots, E_r)_{|E_1+\cdots+E_r}$ and since the map $\mathrm{sum}(E_1, \dots, E_r)$ is the Hermitian transpose of $\mathrm{proj}(E_1, \dots, E_r)_{|E_1+\cdots+E_r}$, it follows the definition of $\det(E_1, \dots, E_r)$ that $\det P = \det(E_1, \dots, E_r)^2$. □

We can check similarly that $\kappa(E_1, \dots, E_r)^2 = \left\|\left|P^{-1}\right|\right\|$, giving another interpretation of κ (which will not be used here).

Lemma 3. *For any subspaces $E_1, \dots, E_r \subseteq V$,*

$$\det{}^{\perp}(E_1, \dots, E_r) = \det{}^{\perp}(E_1, E_2) \det{}^{\perp}(E_1 \cap E_2, E_3, \dots, E_r).$$

Proof. This follows from the factorization

$$\mathrm{proj}(E_1^{\perp}, \dots, E_r^{\perp})_{\sum_i E_i^{\perp}} = \left(\mathrm{proj}(E_1^{\perp}, E_2^{\perp})_{|E_1^{\perp}+E_2^{\perp}} \times \mathrm{id}_{E_3^{\perp}} \times \cdots \times \mathrm{id}_{E_r^{\perp}}\right) \circ \mathrm{proj}(E_1^{\perp} + E_2^{\perp}, E_3^{\perp}, \dots, E_r^{\perp})_{|\sum_i E_i^{\perp}}.$$ □

2.2. **Reminders on Riemannian geometry.** We will work mainly with two Riemannian manifolds: $\mathbb{P}^n$, the n-dimensional complex projective space endowed with the Fubini–Study metric, and $U(n+1)$, the group of $(n+1) \times (n+1)$ unitary matrices. Concerning the latter, we endow $\mathbb{C}^{(n+1)\times(n+1)}$ with the norm

$$\|A\|_u \doteq \tfrac{1}{\sqrt{2}} \|A\|_{\mathrm{Frob}} \doteq \sqrt{\tfrac{1}{2} \operatorname{Tr}(AA^*)}, \quad A \in \mathbb{C}^{(n+1)\times(n+1)}, \tag{4}$$

where A^* denote the conjugate transpose, and we choose on $U(n+1)$ the Riemannian metric induced from the embedding of $U(n+1)$ in $\mathbb{C}^{(n+1)\times(n+1)}$. This metric is invariant under left and right multiplication in $U(n+1)$.

Let X and Y be Riemannian manifolds and let $f : X \to Y$ be an infinitely differentiable surjective map. For any $x \in X$, we define the *normal Jacobian* of f at x as

$$\mathrm{NJ}_x f \doteq \sqrt{\det\left(\mathrm{d}_x f \cdot \mathrm{d}_x f^*\right)}.$$

When $\mathrm{d}_x f$ is bijective, this is the absolute value of the usual Jacobian. A fundamental result is the *coarea formula* for Riemannian manifolds (Federer 1959, Theorem 3.1; Howard 1993, Appendix): for any integrable map $\Theta : X \to \mathbb{R}$,

$$\int_X \mathrm{d}x\ \Theta(x) \mathrm{NJ}_x f = \int_Y \mathrm{d}y \int_{f^{-1}(y)} \mathrm{d}x\ \Theta(x). \tag{5}$$

The special case of Riemannian submersions is important. We say that f is a *Riemannian submersion* if for any $x \in X$, the derivative $\mathrm{d}_x f$ induces an isometry from $(\ker \mathrm{d}_x f)^{\perp}$ to $T_{f(x)}Y$. In that case, we easily check that f is Lipschitz-continuous with constant 1 and that $\mathrm{NJ}_x f = 1$ for all $x \in X$. Note also that for any submanifold Z of Y, if f is a Riemannian submersion then so is $f_{|f^{-1}(Z)}$. The scaling in the definition of $\|-\|_u$ is chosen to have the following result.

Lemma 4. *For any $p \in \mathbb{P}^n$, the map $\varphi : u \in U(n+1) \mapsto up \in \mathbb{P}^n$ is a Riemannian submersion. In particular, for any variety $X \subseteq \mathbb{P}^n$, and any integrable map $\Theta : \varphi^{-1}X \to \mathbb{R}$,*

$$\int_{\varphi^{-1}X} \mathrm{d}u\ \Theta(u) = \int_X \mathrm{d}x \int_{up=x} \mathrm{d}u\ \Theta(u),$$

where $\int_{up=x} \mathrm{d}u$ denotes the integration over the variety $\varphi^{-1}(x)$.

Proof. Thanks to the invariance of the Riemannian metric of $U(n+1)$ under right multiplication, it is enough to check that the defining property of Riemannian submersion holds at id, the identity matrix. With a suitable choice of coordinates,

we may also assume that $p = [1 : 0 : \cdots : 0]$. The tangent space of $\mathbb{P}^n$ at p is canonically identified with $\{p\}^\perp$, that is $\{0\} \times \mathbb{C}^n$.

The tangent space $T_{\mathrm{id}}U(n+1)$ of $U(n+1)$ at id is the space of skew-Hermitian matrices, and for any $u \in T_{\mathrm{id}}U(n+1)$, $\mathrm{d}_{\mathrm{id}}\varphi(\dot{u}) = \dot{u}p$. Therefore,

$$(\ker \mathrm{d}_{\mathrm{id}}\varphi)^\perp = \left\{ \left(\begin{smallmatrix} 0 & v^* \\ v & \mathbf{0} \end{smallmatrix}\right) \mid v \in \mathbb{C}^n \right\},$$

and since $\left\| \left(\begin{smallmatrix} 0 & v^* \\ v & \mathbf{0} \end{smallmatrix}\right) \right\|_u = \|v\|$, the map $\dot{u} \in (\ker \mathrm{d}_{\mathrm{id}}\varphi)^\perp \to \dot{u}p$ is clearly an isometry.

The second claim follows from coarea formula (5), noting that the restriction of π to $\pi^{-1}X$ is again a Riemannian submersion. □

2.3. **Basic integral formulas.** For our problem, the manifold $\mathcal{V}$ has a natural distribution, the *standard distribution*, denoted ρ_{std}, defined as follows. Let $\mathbf{u} \in \mathcal{U}$ and $x \in \cap\mathbf{u}\mathcal{X}$ be uniformly distributed, so that the random variable $(\mathbf{u}, x)$ belongs to $\mathcal{V}$ and let ρ_{std} be its probability distribution. The uniform distribution is defined on $\mathcal{U}$ by the Riemannian metric; and on $\cap\mathbf{u}\mathcal{X}$ by the Riemannian metric on the regular locus. This section aims at describing the conditional probability distribution of $(\mathbf{u}, x)$, given x and the linearization $L(\mathbf{u}, x) \in \mathcal{L}$.

For any $x \in \mathbb{P}^n$, $\mathbf{y} \in \mathcal{X}$ and $\mathbf{h} \in \mathcal{L}$, we define

$$\begin{aligned}
\mathcal{L}_x &\doteq \{\mathbf{h} \in \mathcal{L} \mid x \in \cap\mathbf{h}\}, \\
\mathcal{U}_x &\doteq \{\mathbf{u} \in \mathcal{U} \mid x \in \cap\mathbf{u}\mathcal{X}\} \\
\mathcal{U}_{x,\mathbf{y}} &\doteq \{\mathbf{u} \in \mathcal{U}_x \mid \mathbf{u}\mathbf{y} = (x, \dots, x)\}, \\
\mathcal{U}_{x,\mathbf{y},\mathbf{h}} &\doteq \{\mathbf{u} \in \mathcal{U}_{x,\mathbf{y}} \mid L(\mathbf{u}, x) = \mathbf{h}\}.
\end{aligned}$$

Lemma 5. *For any two submanifolds X and Y of $\mathbb{P}^n$, with $\operatorname{codim} X + \operatorname{codim} Y \leqslant n$, and for any integrable function $\Theta : U(n+1) \times \mathbb{P}^n \to \mathbb{R}$,*

$$\int_{U(n+1)} \mathrm{d}u \int_{X \cap uY} \mathrm{d}z\ \Theta(u, z) = \int_X \mathrm{d}x \int_Y \mathrm{d}y \int_{uy=x} \mathrm{d}u\ \Theta(u, x) \det{}^\perp(\mathbb{T}_x X, \mathbb{T}_x uY),$$

(where $\int_{uy=x} \mathrm{d}u$ denotes the integration over the variety of all $u \in U(n+1)$ such that $uy = x$, as in Lemma 4).

Proof. It is a corollary of the "basic integral formula" of Howard (1993, §2.7). Let $p \in \mathbb{P}^n$ be some point and let φ be the Riemannian submersion $u \in U(n+1) \mapsto up \in \mathbb{P}^n$ (Lemma 4). Let $M = \varphi^{-1}X$ and $N = \varphi^{-1}Y$. Howard's basic integral formula asserts that

$$\int_{U(n+1)} \mathrm{d}u \int_{M \cap uN} \mathrm{d}v\ \Theta(u, \varphi v) = \int_M \mathrm{d}v \int_N \mathrm{d}w\ \Theta(vw^{-1}, \varphi v) \det{}^\perp(v^{-1}T_v M, w^{-1}T_w N), \tag{6}$$

where $v^{-1}T_v M$ and $w^{-1}T_w N$ are subspaces of $T_{\mathrm{id}}U(n+1)$. (The equality of our $\det^\perp$ with Howard's σ is given by Lemma 2.) On the one hand, by Lemma 4, we obtain the following expression for the left-hand side of (6):

$$\begin{aligned}
\int_{U(n+1)} \mathrm{d}u \int_{M \cap uN} \mathrm{d}v\ \Theta(u, \varphi v) &= \int_{U(n+1)} \mathrm{d}u \int_{X \cap uY} \mathrm{d}z \int_{vp=z} \mathrm{d}v\ \Theta(u, z) \\
&= \int_{U(n+1)} \mathrm{d}u \int_{X \cap uY} \mathrm{d}z\ \Theta(u, z) \operatorname{vol}\{v \in U(n+1) \mid vp = z\} \\
&= \operatorname{vol}(U(1) \times U(n)) \int_{U(n+1)} \mathrm{d}u \int_{X \cap uY} \mathrm{d}z\ \Theta(u, z),
\end{aligned}$$

where, for the last equality, we remark that $\{v \in U(n+1) \mid vp = z\}$ is isometric, by some left multiplication, to $U(1) \times U(n)$, the stabilizer of a point in $\mathbb{P}^n$. This is the

left-hand side of the claimed equality. On the other hand, regarding the right-hand side of (6), we check easily that

$$\begin{aligned}\det{}^{\perp}(v^{-1}T_vM, w^{-1}T_wN) &= \det{}^{\perp}(v^{-1}\mathbb{T}_{vp}X, w^{-1}\mathbb{T}_{wp}Y)\\ &= \det{}^{\perp}\left(\mathbb{T}_{vp}X, \mathbb{T}_{vp}(vw^{-1}Y)\right)\end{aligned}$$

and therefore (using Lemma 4 again)

$$\begin{aligned}&\int_M \mathrm{d}v \int_N \mathrm{d}w\ \Theta(vw^{-1}, \varphi v)\det{}^{\perp}(v^{-1}T_vM, w^{-1}T_wN)\\ &\qquad= \int_X \mathrm{d}x \int_{vp=x} \mathrm{d}v \int_Y \mathrm{d}y \int_{wp=y} \mathrm{d}w\ \Theta(vw^{-1}, x)\det{}^{\perp}(\mathbb{T}_xX, \mathbb{T}_xvw^{-1}Y)\\ &\qquad= \mathrm{vol}(U(1)\times U(n)) \int_X \mathrm{d}x \int_Y \mathrm{d}y \int_{uy=x} \mathrm{d}u\ \Theta(u, x)\det{}^{\perp}(\mathbb{T}_xX, \mathbb{T}_xuY),\end{aligned}$$

where the last equality is given by the change of variables $v = uw$. This gives the right-hand side of the claim. □

In our setting where we consider r varieties $X_1, \dots, X_r$, we can give the following "basic integral formula". It has been proved very similarly in the real case by Bürgisser and Lerario (2018, §7.5).

Proposition 6. *For any measurable function* $\Theta : \mathcal{V} \to [0, \infty)$

$$\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x\ \Theta(\mathbf{u}, x) = \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{U}_x} \mathrm{d}\mathbf{u}\ \Theta(\mathbf{u}, x)\det{}^{\perp}(L(\mathbf{u}, x)).$$

Proof. We proceed by induction on r. When $r = 1$,

$$\begin{aligned}\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x\ \Theta(\mathbf{u}, x) &= \int_{U(n+1)} \mathrm{d}u \int_{uX_1} \mathrm{d}x\ \Theta(u, x)\\ &= \int_{\mathbb{P}^n} \mathrm{d}x \int_{X_1} \mathrm{d}y \int_{uy=x} \mathrm{d}u\ \Theta(u, x)\det{}^{\perp}(\mathbb{T}_x\mathbb{P}^n, \mathbb{T}_xuX_1),\end{aligned}$$

by Lemma 5 with $X = \mathbb{P}^n$ and $Y = X_1$. Note that $\det^{\perp}(\mathbb{T}_x\mathbb{P}^n, \mathbb{T}_xuX_1) = \det^{\perp}(\mathbb{T}_xuX_1) = 1$, and it follows that

$$\begin{aligned}\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x\ \Theta(\mathbf{u}, x) &= \int_{\mathbb{P}^n} \mathrm{d}x \int_{X_1} \mathrm{d}y \int_{uy=x} \mathrm{d}u\ \Theta(u, x)\\ &= \int_{\mathbb{P}^n} \mathrm{d}x \int_{y=vx} \mathrm{d}v\ \Theta(v^{-1}, x), \quad \text{by the change of variable } v = u^{-1},\\ &= \int_{\mathbb{P}^n} \mathrm{d}x \int_{x\in v^{-1}X_1} \mathrm{d}v\ \Theta(v^{-1}, x), \quad \text{by Lemma 4},\\ &= \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{U}_x} \mathrm{d}u\ \Theta(u, x), \quad \text{by the change of variable } u = v^{-1}.\end{aligned}$$

This concludes the base case since $\det^{\perp}(L(\mathbf{u}, x))$ is identically 1 when $r = 1$.

Assume that the property holds for $r-1$ subvarieties $X_1, \dots, X_{r-1}$, for some $r \geqslant 1$, and let $\mathcal{U}'$, $\mathcal{X}'$, etc. denote the analogues of $\mathcal{U}$, $\mathcal{X}$, etc. for the varieties $X_1, \dots, X_{r-1}$. From the decomposition $\mathcal{U} = \mathcal{U}' \times U(n+1)$, we obtain by Lemma 5

$$\begin{aligned}&\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x\ \Theta(\mathbf{u}, x) = \int_{\mathcal{U}'} \mathrm{d}\mathbf{u}' \int_{U(n+1)} \mathrm{d}u_r \int_{(\cap \mathbf{u}'\mathcal{X}')\cap u_rX_r} \mathrm{d}x\ \Theta(\mathbf{u}', u_r, x)\\ &= \int_{\mathcal{U}'} \mathrm{d}\mathbf{u}' \int_{\cap \mathbf{u}'\mathcal{X}'} \mathrm{d}x \int_{X_r} \mathrm{d}y \int_{u_ry=x} \mathrm{d}u_r\ \Theta(\mathbf{u}', u_r, x)\det{}^{\perp}(\mathbb{T}_x(\cap \mathbf{u}'\mathcal{X}'), \mathbb{T}_xu_rX_r)\end{aligned}$$

$$= \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{U}'_x} \mathrm{d}\mathbf{u}' \int_{X_r} \mathrm{d}y \int_{u_r y = x} \mathrm{d}u_r \; \Theta(\mathbf{u}', u_r, x) \det{}^{\perp}(\mathbb{T}_x(\cap \mathbf{u}'\mathcal{X}'), \mathbb{T}_x u_r X_r) \det{}^{\perp}(L(\mathbf{u}', x)),$$

using the induction hypothesis for the last equation. Lemma 3 shows that

$$\det{}^{\perp}(\mathbb{T}_x(\cap \mathbf{u}'\mathcal{X}'), \mathbb{T}_x u_r X_r) \det{}^{\perp}(L(\mathbf{u}', x)) = \det{}^{\perp}(L(\mathbf{u}, x)).$$

Moreover, by Lemma 4 (applied as in the base case),

$$\int_{X_r} \mathrm{d}y \int_{u_r y = x} \mathrm{d}u_r \; h(u_r) = \int_{x \in u_r X_r} \mathrm{d}u_r \; h(u_r),$$

for any integrable function $h : U(n+1) \to \mathbb{R}$. This implies that

$$\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x \; \Theta(\mathbf{u}, x) = \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{U}'_x} \mathrm{d}\mathbf{u}' \int_{x \in u_r X_r} \mathrm{d}u_r \; \Theta(\mathbf{u}', u_r, x) \det{}^{\perp}_x(L(\mathbf{u}, x)).$$

To conclude, we remark that $\mathcal{U}_x = \mathcal{U}'_x \times \{u_r \in U(n+1) \mid x \in u_r X_r\}$. □

If we apply the statement above to the case where the varieties X_i are projective subspaces, that is $X_i \in \mathbb{G}(\dim X_i)$, we obtain the following corollary.

Corollary 7. *For any measurable function* $\Theta : \{(\mathbf{h}, x) \in \mathcal{L} \mid x \in \cap \mathbf{h}\} \to [0, \infty)$

$$\int_{\mathcal{L}} \mathrm{d}\mathbf{h} \int_{\cap \mathbf{h}} \mathrm{d}x \; \Theta(\mathbf{h}, x) = \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{L}_x} \mathrm{d}\mathbf{h} \; \Theta(\mathbf{h}, x) \det{}^{\perp}(\mathbf{h}).$$

Proof. Let us assume that each X_i is a projective subspace of $\mathbb{P}^n$. The map

$$\mathbf{u} \in \mathcal{U} \mapsto \mathbf{u}\mathcal{X} = (u_1 X_1, \ldots, u_r X_r) \in \mathcal{L},$$

is a Riemannian submersion, thus

$$\begin{aligned} \int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x \; \Theta(\mathbf{u}\mathcal{X}, x) &= \int_{\mathcal{L}} \mathrm{d}\mathbf{h} \int_{\mathbf{u}\mathcal{X} = \mathbf{h}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{h}} \mathrm{d}x \; \Theta(\mathbf{h}, x) \\ &= \mathrm{vol}(\mathrm{Stab}_{\mathcal{U}} \mathcal{X}) \int_{\mathcal{L}} \mathrm{d}\mathbf{h} \int_{\cap \mathbf{h}} \mathrm{d}x \; \Theta(\mathbf{h}, x), \end{aligned}$$

where $\mathrm{Stab}_{\mathcal{U}} \mathcal{X} = \{\mathbf{u} \in \mathcal{U} \mid \mathbf{u}\mathcal{X} = \mathcal{X}\}$, because all the subsets $\{\mathbf{u} \in \mathcal{U} \mid \mathbf{u}\mathcal{X} = \mathbf{h}\}$ are isometric, by some left multiplication, to $\mathrm{Stab}_{\mathcal{U}} \mathcal{X}$. Similarly,

$$\int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{U}_x} \mathrm{d}\mathbf{u} \; \Theta(\mathbf{u}\mathcal{X}, x) \det{}^{\perp}(L(\mathbf{u}, x)) = \mathrm{vol}(\mathrm{Stab}_{\mathcal{U}} \mathcal{X}) \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{L}_x} \mathrm{d}\mathbf{h} \; \Theta(\mathbf{h}, x) \det{}^{\perp}(\mathbf{h}).$$

This reduces the claim to Proposition 6. □

We now have all we need to prove the main result of this section.

Theorem 8. *For any measurable function* $\Theta : \mathcal{V} \to [0, \infty)$*,*

$$\int_{\mathcal{U}} \mathrm{d}\mathbf{u} \int_{\cap \mathbf{u}\mathcal{X}} \mathrm{d}x \; \Theta(\mathbf{u}, x) = \int_{\mathcal{L}} \mathrm{d}\mathbf{h} \int_{\mathcal{X}} \mathrm{d}\mathbf{y} \int_{\cap \mathbf{h}} \mathrm{d}x \int_{\mathcal{U}_{x,\mathbf{y},\mathbf{h}}} \mathrm{d}\mathbf{u} \; \Theta(\mathbf{u}, x).$$

In other words, if $(\mathbf{u}, x) \in \mathcal{V}$ *is a* ρ_{std}*-distributed random variable distributed, then:*

- *(i) the random variable* $L(\mathbf{u}, x)$ *is uniformly distributed in* $\mathcal{L}$*;*
- *(ii) the random variable* $\mathbf{y} \doteq (u_1^{-1} x, \ldots, u_r^{-1} x)$ *is uniformly distributed in* $\mathcal{X}$ *and independent from* $L(\mathbf{u}, x)$*;*
- *(iii) conditionally on* $L(\mathbf{u}, x)$*, the random variable* x *is uniformly distributed in* $\cap L(\mathbf{u}, x)$*;*
- *(iv) conditionally on* $L(\mathbf{u}, x)$*,* x *and* $\mathbf{y}$*, the random variable* $\mathbf{u}$ *is uniformly distributed in* $\mathcal{U}_{x,\mathbf{y},L(\mathbf{u},x)}$*.*

Algorithm 1. Sampling of a unitary solution variety
Input. Varieties $X_1, \dots, X_r \subset \mathbb{P}^n$ with $\sum_i \operatorname{codim} X_i \leqslant n$.
Output. $(\mathbf{u}, \zeta) \in \mathcal{V}$, where $\mathcal{V} = \{(u_1, \dots, u_r, x) \in \mathcal{U} \times \mathbb{P}^n \mid x \in \cap_i u_i X_i\}$.
Postcondition. $(\mathbf{u}, \zeta) \sim \rho_{\mathrm{std}}$ (Theorem 8)

function Sample($X_1, \dots, X_r$)
Sample $h_1 \in \mathbb{G}(\dim X_i), \dots, h_r \in \mathbb{G}(\dim X_r)$, uniformly and independently.
Sample $\zeta \in h_1 \cap \cdots \cap h_r \subset \mathbb{P}^n$ uniformly.
Sample $y_1 \in X_1, \dots, y_r \in X_r$ uniformly and independently.
Sample $u_1, \dots, u_r \in U(n+1)$, such that $u_i y_i = \zeta$ and $u_i(\mathbb{T}_{y_i} X_i) = h_i$, uniformly and independently.
return $(u_1, \dots, u_r) \in \mathcal{U}$ and $\zeta \in \mathbb{P}^n$.
end function

Proof. By Corollary 7,

$$\int_{\mathcal{L}} \mathrm{d}\mathbf{h} \int_{\cap \mathbf{h}} \mathrm{d}x \int_{\mathcal{X}} \mathrm{d}\mathbf{y} \int_{\mathcal{U}_{x,\mathbf{y},\mathbf{h}}} \mathrm{d}v \; \Theta(\mathbf{u}, x) = \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{L}_x} \mathrm{d}\mathbf{h} \det{}^{\perp}(\mathbf{h}) \int_{\mathcal{X}} \mathrm{d}\mathbf{y} \int_{\mathcal{U}_{x,\mathbf{y},\mathbf{h}}} \mathrm{d}v \; \Theta(\mathbf{u}, x)$$
$$= \int_{\mathbb{P}^n} \mathrm{d}x \int_{\mathcal{X}} \mathrm{d}\mathbf{y} \int_{\mathcal{L}_x} \mathrm{d}\mathbf{h} \int_{\mathcal{U}_{x,\mathbf{y},\mathbf{h}}} \mathrm{d}v \; \Theta(\mathbf{u}, x) \det{}^{\perp}(L(\mathbf{u}, x)).$$

Moreover, the map $\mathbf{u} \in \mathcal{U}_{x,\mathbf{y}} \mapsto L(\mathbf{u}, x) \in \mathcal{L}_x$ is a Riemannian submersion, thus

$$\int_{\mathcal{L}_x} \mathrm{d}\mathbf{h} \int_{\mathcal{U}_{x,\mathbf{y},\mathbf{h}}} \mathrm{d}v \; \Theta(\mathbf{u}, x) \det{}^{\perp}(L(\mathbf{u}, x)) = \int_{\mathcal{U}_{x,\mathbf{y}}} \mathrm{d}v \; \Theta(\mathbf{u}, x) \det{}^{\perp}(L(\mathbf{u}, x)).$$

The map $\mathbf{u} \in \mathcal{U}_x \mapsto (u_1^{-1} x, \dots, u_m^{-1} x) \in \mathcal{X}$ is also a Riemannian submersion, thus

$$\int_{\mathcal{X}} \mathrm{d}\mathbf{y} \int_{\mathcal{U}_{x,\mathbf{y}}} \mathrm{d}v \; \Theta(\mathbf{u}, x) \det{}^{\perp}(L(\mathbf{u}, x)) = \int_{\mathcal{U}_x} \mathrm{d}v \; \Theta(\mathbf{u}, x) \det{}^{\perp}(L(\mathbf{u}, x)).$$

To conclude, we apply Proposition 6. □

Corollary 9. *Let X be subvariety of $\mathbb{P}^n$ and $L \in \mathbb{G}(\operatorname{codim} X)$ be a uniformly distributed random projective subspace. Let $\zeta \in X \cap L$ be a uniformly distributed random variable. Then ζ is uniformly distributed in X.*

Proof. Let $L_0 \in \mathbb{G}(\operatorname{codim} X)$ and let $v \in U(n+1)$ a be uniformly distributed random variable. The random subspace vL_0 is uniformly distributed in $\mathbb{G}(\operatorname{codim} X)$: indeed, the probability distribution of vL_0 is invariant under the action of $U(n+1)$ and, by transitivity of the action of $U(n+1)$ on $\mathbb{G}(\operatorname{codim} X)$, there is a unique invariant probability distribution on $\mathbb{G}(\operatorname{codim} X)$. So we may assume that $L = vL_0$.

Let $u \in U(n+1)$ be an independent uniformly distributed random variable. The random variables u and $v' \doteq uv$ are independent and uniformly distributed, because the diffeomorphism

$$(u, v) \in U(n+1) \times U(n+1) \mapsto (u, uv) \in U(n+1) \times U(n+1)$$

has constant Jacobian, and so preserves the uniform distribution. Moreover, $u\zeta$ is uniformly distributed in $uX \cap v'L_0$. Therefore, the pair $((u, v'), u\zeta)$ is ρ_{std}-distributed in the solution variety associated to X and L_0. By Theorem 8(ii), $u^{-1}(u\zeta)$ is uniformly distributed in X, which is the claim. □

2.4. **Sampling the solution variety.** Based on Theorem 8, Algorithm 1 samples a ρ_{std}-distributed random $(\mathbf{u}, \zeta) \in \mathcal{V}$. We explain briefly how to perform the four steps of the algorithm.

For each $1 \leqslant i \leqslant r$, we sample independently linear forms $\lambda_{i,1}, \dots, \lambda_{i,\operatorname{codim} X_i} \in (\mathbb{C}^{n+1})^*$ with a standard normal distribution. We define h_i as the zero locus of $\lambda_{i,1}, \dots, \lambda_{i,\operatorname{codim} X_i}$. Next, we compute a unitary basis of the linear subspace $h_1 \cap \cdots \cap h_n$ and use it to sample $\zeta \in \mathbb{P}(h_1 \cap \cdots \cap h_n)$ with a uniform distribution.

To sample uniformly a point $y_i \in X_i$, we consider a random uniformly distributed subspace $L_i \in \mathbb{G}(\operatorname{codim} X_i)$. Almost surely, the intersection $X_i \cap L_i$ is finite and we sample uniformly a point y_i in it. By Corollary 9, y_i is uniformly distributed in X_i. Since L_i is a projective subpace, the computation of $X_i \cap L_i$ requires a polynomial system solving in $\operatorname{codim} X_i + 1$ homogeneous variables. In the typical case where X_i is a hypersurface defined by a polynomial f_i and L_i is a projective line, this amounts to compute the zeros $[x : y] \in \mathbb{P}^1$ of the homogeneous equation $f_i(xp + yq) = 0$, for some basis $\{p, q\}$ of L_i.

Once we get the y_i, we compute, for each $1 \leqslant i \leqslant r$, some $v_i \in U(n+1)$ which maps y_i to ζ and $\mathbb{T}_{y_i} X_i$ to h_i and we sample uniformly a w_i in the subgroup of all $w \in U(n+1)$ such that $w\zeta = \zeta$ and $wL_i = L_i$. This subgroup is isometrically isomorphic to $U(1) \times U(\dim X_i) \times U(\operatorname{codim} X_i)$. We can sample uniformly in a unitary group $U(k)$ by considering the Q factor of the QR decomposition of a random $k \times k$ Gaussian matrix. And then, we define $u_i \doteq w_i v_i$. When $X_1, \dots, X_r$ are all hypersurfaces, we have proved the following proposition.

Proposition 10. *If $X_1, \dots, X_r$ are all hypersurfaces, Algorithm 1 samples a ρ_{std}-distributed point in the solution variety $\mathcal{V}$ with*

- *r times root-finding of bivariate homogeneous polynomials of respective degrees $\deg X_1, \dots, \deg X_r$;*
- *$O(n^3)$ samplings of the standard normal distribution on $\mathbb{R}$; and*
- *$O(n^4)$ arithmetic operations.*

2.5. **The split gamma number.** In the classical theory, the condition number μ plays two roles: first, by definition, it bounds the variation of a zero after a pertubation of the system; second, it is an upper bound for Smale's gamma number with some regularity properties (the Lipschitz properties). Each role is reflected by a factor μ in the μ^2 estimate.

In the rigid setting, the two roles are played by different numbers: the variation of a zero is bounded by the incidence condition number κ and the upper bound for γ that we use is the *split gamma number* $\hat{\gamma}$. This will lead to a $\kappa\hat{\gamma}$ estimate for the complexity of numerical continuation in the rigid setting. In this section, we introduce the split gamma number and we start with some reminders about the gamma number.

Let $F = (f_1, \dots, f_s) \in \mathcal{H}[s]$ be a homogeneous polynomial system that we regard as a polynomial map $\mathbb{C}^{n+1} \to \mathbb{C}^s$. Let $d_i \doteq \deg f_i$ and $D \doteq \max_i d_i$.

When $s = n$, the polynomial system F has generically finitely many zeros and our primary goal is to compute them numerically and approximately. A fundamental tool is Newton's operator. We use here the projective version introduced by Shub (1989). For $z \in \mathbb{P}^n$, projective class of some $\bar{z} \in \mathbb{C}^{n+1}$, we define

$$\mathcal{N}_F(z) \doteq \left[\bar{z} - \mathrm{d}_{\bar{z}} F|_{z^\perp}^{-1}\left(F(\bar{z})\right)\right] \in \mathbb{P}^n, \tag{7}$$

where $z^\perp$ is the orthogonal complement of z in $\mathbb{C}^{n+1}$. The definition does not depend on the choice of $\bar{z}$. Given a zero $\zeta \in \mathbb{P}^n$ of F, we say that $z \in \mathbb{P}^n$ *approximates ζ as a zero of F*, or that *z is an approximate zero of F with associated zero ζ*, if for

any $k \geqslant 0$,

$$d_{\mathbb{P}}\left(\mathcal{N}_F^k(z), \zeta\right) \leqslant 2^{1-2^k} d_{\mathbb{P}}(z, \zeta),$$

where $d_{\mathbb{P}}$ is the geodesic distance in $\mathbb{P}^n$, see §1.3.

The main result of the gamma theory is a sufficient condition for a point to approximate a zero. For a polynomial system F, in the general case $r \leqslant n$, we will use the following definition (Shub and Smale 1996; Dedieu 2006, §4) for the gamma number of F at $z \in \mathbb{P}^n$:

$$\gamma(F, z) \doteq \begin{cases} \sup_{k \geqslant 2} \left\| \frac{1}{k!} \mathrm{d}_z F^\dagger \cdot \mathrm{d}_z^k F \right\|^{\frac{1}{k-1}} & \text{if } \mathrm{d}_z F \text{ is surjective,} \\ \infty & \text{otherwise.} \end{cases} \tag{8}$$

(The definition does not depend on the choice of a unit representative $\bar{z}$ of z.) When $s = n$, the pseudo-inverse $\mathrm{d}_z F^\dagger$ is often replaced by $\mathrm{d}_z F|_{z^\perp}^{-1}$, as in Newton's iteration (*e.g.* Bürgisser and Cucker 2013; Shub and Smale 1994). If z is a zero of F, both definitions coincide, so the gamma theorem (Theorem 12) is equally true for both variants.

When $F = (f)$ is a single equation, that is $s = 1$, it is useful to remark that $\mathrm{d}_z f$ is a linear form and so $\mathrm{d}_z f^\dagger$ is simply $\|\mathrm{d}_z f\|^{-1}\pi$, where π is an isometric embedding of $\mathbb{C}$ in $\mathbb{C}^{n+1}$. This gives $\gamma(f, z)$ the following form:

$$\gamma(f, z) = \sup_{k \geqslant 2} \left(\tfrac{1}{k!} \|\mathrm{d}_z f\|^{-1} \|\mathrm{d}_z^k f\| \right)^{\frac{1}{k-1}}. \tag{9}$$

The following lower bound will be occasionally useful.

Lemma 11. *For any $z \in \mathbb{P}^n$, $\gamma(F, z) \geqslant \frac{D-1}{2}$.*

Proof. We may assume that $\mathrm{d}_z F$ is surjective, otherwise the bound is trivial. Let $d_i \doteq \deg f_i$. Using the homogeneity, for any $u \in \mathbb{C}^{n+1}$,

$$\mathrm{d}_z^2 F(z, u) = \begin{pmatrix} d_1 - 1 & & \\ & \ddots & \\ & & d_m - 1 \end{pmatrix} \mathrm{d}_z F(u). \tag{10}$$

We fix some $1 \leqslant i \leqslant r$ and consider $u \doteq \mathrm{d}_z F^\dagger(e_i)$, where $e_i \doteq (0, \dots, 1, \dots, 0) \in \mathbb{C}^r$ with a single one at the ith position. Becase $\mathrm{d}_z F$ is surjective, $\mathrm{d}_z F(\mathrm{d}_z F^\dagger(e_i)) = e_i$, and by (10), we have $\mathrm{d}_z^2 F(z, u) = (d_i - 1)e_i$ and then $\mathrm{d}_z F^\dagger(\mathrm{d}_z^2 F(z, u)) = (d_i - 1)u$. This implies that $\left\|\mathrm{d}_z F^\dagger \mathrm{d}_z^2 F\right\| \geqslant d_i - 1$, for any $1 \leqslant i \leqslant m$, and the claim follows. □

We now state the main result of the gamma theory in the projective setting, primarily due to Shub and Smale (1993a).

Theorem 12 (Shub, Smale)**.** *Let $F \in \mathcal{H}[n]$ be a homogeneous polynomial system. For any zero $\zeta \in \mathbb{P}^n$ of F and any $z \in \mathbb{P}^n$, if $d_{\mathbb{P}}(z, \zeta)\gamma(F, \zeta) \leqslant \frac{1}{6}$ then z approximates ζ as a zero of F.*

Proof. This is (Bürgisser and Cucker 2013, Theorem 16.38) with $r = 0.981$ (and we use that $\gamma(F, \zeta) \geqslant \frac{1}{2}$ when $D \geqslant 2$, Lemma 11). □

Let $F_1 \in \mathcal{H}[s_1], \dots, F_r \in \mathcal{H}[s_r]$ be homogeneous polynomial systems, with $s_1 + \cdots + s_r \leqslant n$. Based on the incidence condition number of linear subspaces (§2.1), we define the *incidence condition number* of $F_1, \dots, F_r$ at a point $x \in \mathbb{P}^n$ by

$$\kappa(F_1, \dots, F_r; x) \doteq \kappa\left(\ker(\mathrm{d}_x F_1)^\perp, \dots, \ker(\mathrm{d}_x F_r)^\perp\right). \tag{11}$$

Since the orthogonal projector on $\ker(\mathrm{d}_x F_i)^\perp$ is $\mathrm{d}_x F_i^\dagger \circ \mathrm{d}_x F_i$, we can write

$$\kappa(F_1,\dots,F_r;x) = \left\| \begin{pmatrix} \mathrm{d}_x F_1^\dagger \circ \mathrm{d}_x F_1 \\ \hline \vdots \\ \hline \mathrm{d}_x F_r^\dagger \circ \mathrm{d}_x F_r \end{pmatrix}^\dagger \right\|. \tag{12}$$

The *split gamma number* of $F_1,\dots,F_r$ at a point $x \in \mathbb{P}^n$ is defined by

$$\hat\gamma(F_1,\dots,F_r;x) \doteq \kappa(F_1,\dots,F_r;x)\left(\gamma(F_1,x)^2 + \cdots + \gamma(F_r,x)^2\right)^{\frac12}. \tag{13}$$

The split gamma number separates the contribution of the γ number of each block of equations from the more geometric information contained in κ: for $x \in \cap_i V(F_i)$, the κ factor only depends on the relative position of the tangent spaces of the varieties $V(F_1),\dots,V(F_r)$ at x; while the γ factor quantifies how much each $V(F_i)$ deviates from its tangent space at x. Note that when $r=1$, then $\hat\gamma(F,x) = \gamma(F,x)$.

Theorem 13. *Let $G \doteq (F_1,\dots,F_r) \in \mathcal{H}[s_1+\cdots+s_r]$ denote the concatenation of the systems. For any $x \in \mathbb{P}^n$,*

$$\gamma(G,x) \leqslant \hat\gamma(F_1,\dots,F_r;x) \leqslant r\,\kappa(F_1,\dots,F_r;x)\gamma(G,x).$$

Proof. It is easy to see that $\hat\gamma(F_1,\dots,F_r;x)$ is finite if and only if $\gamma(G,x)$ is. Thus, we may assume that $\mathrm{d}G$ and the $\mathrm{d}F_i$ are surjective (we drop the index x in d_x). We begin with the first inequality. Let $K_i \doteq \ker \mathrm{d}_x F_i$ and $P \doteq \mathrm{proj}(K_1^\perp,\dots,K_r^\perp)$, so that $\kappa(F_1,\dots,F_r;x) = \|P^\dagger\|$. We first prove that for any $k \geqslant 2$ and any $\mathbf{y} = (y_1,\dots,y_k) \in \left(\mathbb{C}^{n+1}\right)^k$,

$$\mathrm{d}G^\dagger \cdot \mathrm{d}^k G(\mathbf{y}) = P^\dagger\left(\mathrm{d}F_1^\dagger \cdot \mathrm{d}^k F_1(\mathbf{y}),\dots,\mathrm{d}F_r^\dagger \cdot \mathrm{d}^k F_r(\mathbf{y})\right). \tag{14}$$

It is clear that

$$\mathrm{d}^k G(\mathbf{y}) = \left(\mathrm{d}^k F_1(\mathbf{y}),\dots,\mathrm{d}^k F_r(\mathbf{y})\right) \in \mathbb{C}^{s_1}\times\cdots\times\mathbb{C}^{s_r}.$$

Let $v_i \doteq \mathrm{d}^k F_i(\mathbf{y})$ and $\mathbf{v} \doteq (v_1,\dots,v_r)$. Because $\mathrm{d}G$ is surjective, we have $\mathrm{d}G \cdot \mathrm{d}G^\dagger(\mathbf{v}) = \mathbf{v}$ and, equivalently, $\mathrm{d}F_i \cdot \mathrm{d}G^\dagger(\mathbf{v}) = v_i$. Therefore $\mathrm{d}F_i^\dagger \cdot \mathrm{d}F_i \cdot \mathrm{d}G^\dagger \mathbf{v} = \mathrm{d}F_i^\dagger v_i$. But $\mathrm{d}F_i^\dagger \cdot \mathrm{d}F_i$ is simply the orthogonal projection on $K_i^\perp$. This gives $P \cdot \mathrm{d}G^\dagger \cdot \mathrm{d}^k G(\mathbf{y}) = \left(\mathrm{d}F_1^\dagger v_1,\dots,\mathrm{d}F_r^\dagger v_r\right)$, and since the image of $\mathrm{d}G^\dagger$ is orthogonal to the kernel of P we have $P^\dagger \cdot P \cdot \mathrm{d}G^\dagger = \mathrm{d}G^\dagger$. This proves (14).

As a consequence, for any $k \geqslant 2$,

$$\begin{aligned}
\left\|\tfrac{1}{k!}\mathrm{d}G^\dagger \cdot \mathrm{d}^k G\right\|^{\frac{1}{k-1}} &\leqslant \|P^\dagger\|^{\frac{1}{k-1}} \left(\sum_i \left\|\tfrac{1}{k!}\mathrm{d}F_i^\dagger \cdot \mathrm{d}^k F_i\right\|^2\right)^{\frac{1}{2k-2}} \\
&\leqslant \|P^\dagger\|^{\frac{1}{k-1}} \left(\sum_i \gamma(F_i,x)^{2k-2}\right)^{\frac{1}{2k-2}} \\
&\leqslant \|P^\dagger\| \left(\sum_i \gamma(F_i,x)^2\right)^{\frac12}
\end{aligned}$$

using $\|P^\dagger\| \geqslant 1$ (Lemma 1) and the monotonicity of p-norms with respect to p. By definition, $\kappa(F_1,\dots,F_r;x) = \|P^\dagger\|$, so we obtain the first inequality.

Concerning the second inequality, Equation (14) implies that

$$\begin{aligned}
\left\|\tfrac{1}{k!}\mathrm{d}F_i^\dagger \cdot \mathrm{d}^k F_i\right\| &\leqslant \|P\|\,\left\|\tfrac{1}{k!}\mathrm{d}G^\dagger \cdot \mathrm{d}^k G\right\| \\
&\leqslant \|P\|\,\gamma(G,x)^{k-1}.
\end{aligned}$$

We note that $|||P||| \leqslant r^{\frac{1}{2}}$ (as the direct sum of r orthogonal projectors with orthogonal images) and therefore $\gamma(F_i, x) \leqslant r^{\frac{1}{2}}\gamma(G, x)$. Hence

$$\left(\sum_{i=1}^{r} \gamma(F_i, x)^2\right)^{\frac{1}{2}} \leqslant \left(r \cdot r\gamma(G, x)^2\right)^{\frac{1}{2}} \leqslant r\gamma(G, x),$$

and the second inequality follows. □

3. Numerical continuation

In this part, we consider a more specific setting than in the previous one. We are given a polynomial system $F = (f_1, \dots, f_n) \in \mathcal{H}[n]$, with nonzero square-free polynomials $f_1, \dots, f_n$, and a $\mathbf{u} \in \mathcal{U} = U(n+1)^n$, and we look for a zero of the polynomial system $\mathbf{u} \cdot F \doteq (f_1 \circ u_1^{-1}, \dots, f_n \circ u_n^{-1})$. We will perform the average analyses in the case where $\mathbf{u}$ is uniformly distributed and F is fixed. To this end, we consider the rigid solution variety $\mathcal{V}$ relative to the projective hypersurfaces $V(f_1), \dots, V(f_n)$. In concrete terms, we have $r = n$ and

$$\mathcal{V} = \left\{(\mathbf{u}, x) \in \mathcal{U} \times \mathbb{P}^n \;\middle|\; f_1(u_1^{-1}x) = \cdots = f_n(u_n^{-1}x) = 0\right\}.$$

As the Grassmannian of hyperplanes $\mathbb{G}(n-1)$ is isomorphic to $\mathbb{P}^n$, the manifold $\mathcal{L}$ is just $(\mathbb{P}^n)^n$ and $L(\mathbf{u}, x)$ is the n-uple

$$L(\mathbf{u}, x) = \left([\mathrm{d}_x(f_1 \circ u_1^{-1})], \dots, [\mathrm{d}_x(f_n \circ u_n^{-1})]\right).$$

In particular, we can define $L(\mathbf{u}, x)$ generically on $\mathcal{U} \times \mathbb{P}^n$, not only on $\mathcal{V}$. $L(\mathbf{u}, x)$ is not defined when one of the $\mathrm{d}_x(f_i \circ u_i^{-1})$ is zero. When it is defined, we can identify $L(\mathbf{u}, x)$ with one of its preimages under the projection map $\mathbb{S}(\mathbb{C}^{n+1})^n \to (\mathbb{P}^n)^n$, that is a $n \times (n+1)$ matrix with unit rows. Namely,

$$L(\mathbf{u}, x) = \operatorname{diag}\left(\|\mathrm{d}_x(f_1 \circ u_1^{-1})\|^{-1}, \dots, \|\mathrm{d}_x(f_n \circ u_n^{-1})\|^{-1}\right) \cdot \mathrm{d}_x\left(\mathbf{u} \cdot F\right). \tag{15}$$

Under this identification, we check that $L(\mathbf{u}, x)$ is the matrix of the map

$$\operatorname{proj}\left(\ker(\mathrm{d}_x(f_1 \circ u_1^{-1})^\perp, \dots, \ker(\mathrm{d}_x(f_n \circ u_n^{-1})^\perp\right),$$

and therefore, by definition of κ (11),

$$\kappa(\mathbf{u} \cdot F, x) = \begin{cases} |||L(\mathbf{u}, x)^\dagger||| & \text{if } L(\mathbf{u}, x) \text{ is well defined and surjective,} \\ \infty & \text{otherwise.} \end{cases} \tag{16}$$

Section 3.1 studies the possibility of continuing a zero of a polynomial system $\mathbf{u} \cdot F$ when $\mathbf{u}$ varies continuously along a path. Section 3.2 introduces the condition number associated to the solution variety $\mathcal{V}$: it quantifies the extend to which a zero of a polynomial system $\mathbf{u} \cdot F$ is affected by a pertubation of $\mathbf{u}$. Section 3.3 proves some Lipschitz-continuity properties for the condition number and $\hat{\gamma}$. They are technical but crucial for designing the continuation algorithms, what does Section 3.4. Section 3.5 provides a bound on the average number of continuation steps required to compute a zero of a random system $\mathbf{u} \cdot F$ in the rigid setting. Theorem 25 is the main outcome of this part.

3.1. Path lifting.

Let $\Sigma \subset \mathcal{V}$ denote the *singular locus*:

$$\Sigma \doteq \{(\mathbf{u}, x) \in \mathcal{V} \mid \mathrm{d}_x(\mathbf{u} \cdot F) \text{ is not surjective}\}.$$

In other words, $(\mathbf{u}, x) \in \Sigma$ if and only if $\mathbf{u} \cdot F$ has a singular zero at x. Let $\pi : \mathcal{V} \to \mathcal{U}$ be the projection $\pi(\mathbf{u}, x) = \mathbf{u}$ and $\Sigma' \doteq \pi(\Sigma)$. For any $\mathbf{u} \in \mathcal{U} \backslash \Sigma'$, the polynomial system $\mathbf{u} \cdot F$ has finitely many roots that vary continuously with $\mathbf{u}$. Given a

continous path $(\mathbf{w}_t)_{0\leqslant t\leqslant 1}$ in $\mathcal{U}\backslash\Sigma'$ and a zero ζ of $\mathbf{w}_0\cdot F$, there is a unique continuous lifting $(\mathbf{w}_t,\eta_t)\in\mathcal{V}$ such that $\eta_0=\zeta$. In contrast with the classical setting, the parameter space $\mathcal{U}$ is not a complex variety, so it is not obvious anymore that a generic path in $\mathcal{U}$ does not meet Σ' and that this lifting is possible. This section aims at proving this fact.

Lemma 14. *The real codimension of Σ' in $\mathcal{U}$ is at least 2.*

Proof. We first observe that the map $\varphi:\mathcal{U}\times\mathbb{P}^n\to(\mathbb{P}^n)^n$ defined by

$$(\mathbf{u},x)\mapsto(u_1^{-1}x,\dots,u_n^{-1}x)$$

is a proper submersion, and thus a locally trivial fibration, by Ehresmann's fibration theorem. The fibers have dimension

$$\dim_{\mathbb{R}}\varphi^{-1}(*)=\dim_{\mathbb{R}}(\mathcal{U}\times\mathbb{P}^n)-n\dim_{\mathbb{R}}\mathbb{P}^n=\dim_{\mathbb{R}}\mathcal{U}-(n-1)\dim_{\mathbb{R}}\mathbb{P}^n.$$

The solution variety $\mathcal{V}$ is $\varphi^{-1}(X_1\times\cdots\times X_n)$, and so the restriction $\varphi_{|\mathcal{V}}$ induces a locally trivial fibration $\mathcal{V}\to X_1\times\cdots\times X_n$. In particular

$$\begin{aligned}\dim_{\mathbb{R}}\mathcal{V}&=\sum_{i=1}^{n}\dim_{\mathbb{R}}X_i+\dim_{\mathbb{R}}\varphi^{-1}(*)\\&=n(\dim_{\mathbb{R}}\mathbb{P}^n-2)+\dim_{\mathbb{R}}\mathcal{U}-(n-1)\dim_{\mathbb{R}}\mathbb{P}^n\\&=\dim_{\mathbb{R}}\mathcal{U}.\end{aligned}$$

Let $\Sigma_0\subseteq\Sigma$ be the subset of all $(\mathbf{u},x)$ such that x is a singular zero of one of the $f_i\circ u_i^{-1}$ (that is a singular point of u_iX_i). By definition,

$$\Sigma_0=\varphi^{-1}\left(\bigcup_{i=1}^{n}X_1\times\cdots\times\operatorname{Sing}X_i\times\cdots\times X_n\right),$$

so Σ_0 is the preimage by a locally trivial fibration of a complex subvariety of real codimension at least 2. Thus Σ_0 has codimension at least 2 in $\mathcal{V}$.

Let $\Sigma_1\subseteq\Sigma$ be the subset of all $(\mathbf{u},x)$ such that all u_iX_i are smooth at x but $\mathrm{d}_x(\mathbf{u}\cdot F)$ is not surjective, so that $\Sigma=\Sigma_0\cup\Sigma_1$. To compute the dimension of Σ_1, we observe that the map

$$\begin{aligned}\psi:\mathcal{V}\backslash\Sigma_0&\to X_1^{\mathrm{reg}}\times\cdots\times X_n^{\mathrm{reg}}\times\mathcal{L}\\(\mathbf{u},x)&\mapsto\big(u_1^{-1}x,\dots,u_n^{-1}x,L(\mathbf{u},x)\big)\end{aligned}$$

is a proper submersion and thus, by Ehresmann's fibration theorem, a locally trivial fibration. Further, Σ_1 is the preimage of the complex subvariety of all $(y_1,\dots,y_n,\mathbf{h})$ such that the intersection of the hyperplanes $h_1,\dots,h_n\subset\mathbb{C}^{n+1}$ has complex dimension $\geqslant 2$. So Σ_1 has real codimension at least 2 in $\mathcal{V}$, and so does Σ.

Since Σ' is the image of Σ by the projection map $\pi:\mathcal{V}\to\mathcal{U}$, $\dim_{\mathbb{R}}\Sigma'\leqslant\dim_{\mathbb{R}}\Sigma$, and since $\dim_{\mathbb{R}}\mathcal{U}=\dim_{\mathbb{R}}\mathcal{V}$, it follows that Σ' has codimension at least 2 in $\mathcal{U}$. □

Proposition 15. *For almost all $\mathbf{u},\mathbf{v}\in\mathcal{U}$, the shortest path in $\mathcal{U}$ from $\mathbf{u}$ to $\mathbf{v}$ does not intersect Σ'.*

Proof. The subset $S\subset\mathcal{U}\times\mathcal{U}$ of all ill-posed pairs $(\mathbf{u},\mathbf{v})$ that do not satisfy the claim is parametrized by the data of a $\mathbf{w}\in\Sigma'$, a unit vector $\dot\gamma$ in $T_{\mathbf{w}}\mathcal{U}$ and two real numbers t and s such that $\mathbf{u}=\gamma(t)$ and $\mathbf{v}=\gamma(s)$, where $\gamma:\mathbb{R}\to\mathcal{U}$ is the unique geodesic with $\gamma(0)=\mathbf{w}$ and $\gamma'(0)=\dot\gamma$. Therefore, by Lemma 14,

$$\dim_{\mathbb{R}}S\leqslant\dim_{\mathbb{R}}\Sigma'+(\dim_{\mathbb{R}}\mathcal{U}-1)+2\leqslant\dim_{\mathbb{R}}\mathcal{U}^2-1,$$

and S has real codimension at least 1 in $\mathcal{U}\times\mathcal{U}$, which proves the claim. □

The statement still holds for more general paths as long as they are unitary invariant. Let $P : \mathcal{U} \times \mathcal{U} \times [0,1] \to \mathcal{U}$ be a map such that for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{U}$ and $t \in [0,1]$, $P(\mathbf{wu}, \mathbf{wv}, t) = \mathbf{w}P(\mathbf{u}, \mathbf{v}, t)$. Let $S \subset \mathcal{U} \times \mathcal{U}$ be the set of all $(\mathbf{u}, \mathbf{v})$ such that $P(\mathbf{u}, \mathbf{v}, t) \in \Sigma'$ for some $t \in [0,1]$. We check easily that

$$S = \left\{ \left(\mathbf{w}P(\mathbf{1}_{\mathcal{U}}, \mathbf{v}, t)^{-1}, \mathbf{w}P(\mathbf{1}_{\mathcal{U}}, \mathbf{v}, t)^{-1}\mathbf{v} \right) \mid t \in [0,1], \mathbf{w} \in \Sigma' \text{ and } \mathbf{v} \in \mathcal{U} \right\},$$

so if P is regular enough to allow dimension counting, then $\dim_{\mathbb{R}} S \leqslant 1 + \dim_{\mathbb{R}} \Sigma' + \dim_{\mathbb{R}} \mathcal{U}$, and by Lemma 14, the codimension of S in $\mathcal{U} \times \mathcal{U}$ is at least 1.

3.2. **Condition number.** The rigid solution variety, considered as a manifold of pairs problem–solution has a natural condition number. We show that this is the incidence condition number κ, defined in §2.5. This is what makes the split gamma number $\hat{\gamma}$ fit nicely in the setting of the rigid solution variety. The system F being fixed, we will denote $\kappa(\mathbf{u} \cdot F, x)$ and $\hat{\gamma}(\mathbf{u} \cdot F, x)$ simply as $\kappa(\mathbf{u}, x)$ and $\hat{\gamma}(\mathbf{u}, x)$.

Lemma 16. *For any $(\mathbf{u}, x) \in \mathcal{V} \backslash \Sigma$ and any tangent vector $(\dot{\mathbf{u}}, \dot{x}) \in T_{\mathbf{u},x}\mathcal{V}$,*

$$\|\dot{x}\| \leqslant \kappa(\mathbf{u}, x) \|\dot{\mathbf{u}}\|_u.$$

Proof. Without loss of generality, we may assume that $\mathbf{u} = \mathbf{1}_{\mathcal{U}}$ which simplifies notations. The tangent space of $\mathcal{V}$ at $(\mathbf{u}, x)$ is

$$\begin{aligned} T_{\mathbf{u},x}\mathcal{V} = \; & \{(\dot{\mathbf{u}}, \dot{x}) \in T_{\mathbf{u}}\mathcal{U} \times T_x\mathbb{P}^n \mid \\ & \qquad \forall i, f_i\left((\mathbf{1}_{\mathcal{U}} + t\dot{u}_i + o(t))^{-1}(x + t\dot{x} + o(t))\right) = o(t) \text{ as } t \to 0\} \\ = \; & \{(\dot{\mathbf{u}}, \dot{x}) \in T_{\mathbf{u}}\mathcal{U} \times T_x\mathbb{P}^n \mid \forall i, f_i\left(x + t(\dot{x} - \dot{u}_i x)\right) = o(t) \text{ as } t \to 0\} \\ = \; & \{(\dot{\mathbf{u}}, \dot{x}) \in T_{\mathbf{u}}\mathcal{U} \times T_x\mathbb{P}^n \mid \forall i, \mathrm{d}_x f_i(\dot{x}) = \mathrm{d}_x f_i(\dot{u}_i x)\}. \end{aligned} \tag{17}$$

With the identification (15), we obtain that $(\dot{\mathbf{u}}, \dot{x}) \in T_{\mathbf{u},x}\mathcal{V}$ if and only if

$$L(\mathbf{u}, x)(\dot{x}) = \left(\|\mathrm{d}_x f_i\|^{-1} \mathrm{d}_x f_i(\dot{u}_i x)\right)_{1 \leqslant i \leqslant n}. \tag{18}$$

For all $1 \leqslant i \leqslant n$, $\mathrm{d}_x f_i(x) = \deg(f_i) f_i(x) = 0$ (Euler's relation), therefore

$$|\mathrm{d}_x f_i(\dot{u}_i x)| = |\mathrm{d}_x f_i(\pi_x(\dot{u}_i x))| \leqslant \|\mathrm{d}_x f_i\| \|\pi_x(\dot{u}_i x)\|, \tag{19}$$

where π_x is the orthogonal projection on $\{x\}^{\perp}$. We check that $\|\pi_x(\dot{u}_i x)\| \leqslant \|\dot{u}_i\|_u$, as in Lemma 4. If Equation (18) does hold, then

$$\begin{aligned} \|L(\mathbf{u}, x)(\dot{x})\|^2 &= \sum_i \|\mathrm{d}_x f_i\|^{-2} |\mathrm{d}_x f_i(\dot{u}_i x)|^2 \\ &\leqslant \sum_i \|\dot{u}_i\|_u^2, && \text{by (19)}, \\ &= \|\dot{\mathbf{u}}\|_u^2. \end{aligned}$$

Moreover $\dot{x} = L(\mathbf{u}, x)^{\dagger}(L(\mathbf{u}, x)(\dot{x}))$, because $\dot{x}$ is orthogonal to x and the kernel of $L(\mathbf{u}, x)$ is $\mathbb{C}x$. Therefore

$$\|\dot{x}\| \leqslant \left\|L(\mathbf{u}, x)^{\dagger}\right\| \|L(\mathbf{u}, x)(\dot{x})\| \leqslant \left\|L(\mathbf{u}, x)^{\dagger}\right\| \|\dot{\mathbf{u}}\|_u,$$

and the claim follows with $\kappa(\mathbf{u}, x) = \left\|L(\mathbf{u}, x)^{\dagger}\right\|$, by (16). □

The expected value of the incidence condition number κ is very tame and depends on the ambient dimension n only. This is one of the key points that strongly contrasts with the classical setting.

Proposition 17. *If $(\mathbf{u}, \zeta) \in \mathcal{V}$ is ρ_{std}-distributed then, $\mathbb{E}\left[\kappa(\mathbf{u}, \zeta)^2\right] \leqslant 6n^2$.*

Proof. Let M be a random $n \times (n+1)$ matrix whose rows are independent and uniformly distributed in $\mathbb{S}(\mathbb{C}^{n+1})$. It follows from (16) and Theorem 8(i) that $\kappa(\mathbf{u}, \zeta)$ has the same probability distribution as $|||M^\dagger|||$.

Let T be an $n \times n$ diagonal random matrix whose coefficients are independent χ-distributed random variables with $2n+2$ degrees of freedom, so that TM is a random Gaussian matrix (the coefficients are independent standard normal complex numbers). Obviously, $M^\dagger = (TM)^\dagger \cdot T$ and therefore $|||M^\dagger||| \leqslant |||(TM)^\dagger||| \, |||T|||$. Hölder's inequality with conjugate exponents $n' \doteq 1 + \frac{1}{n+1}$ and $n+2$, gives

$$\mathbb{E}\left[|||M^\dagger|||^2\right] \leqslant \mathbb{E}\left[|||(TM)^\dagger|||^{2n'}\right]^{\frac{1}{n'}} \mathbb{E}\left[|||T|||^{2n+4}\right]^{\frac{1}{n+2}}. \tag{20}$$

We now give upper bounds for both factors in the right-hand side. According to Beltrán and Pardo (2011, Theorem 20), whose result is derived from the work of Edelman (1989),

$$\mathbb{E}\left[|||(TM)^\dagger|||^{2n'}\right] = 2^{-n'} \sum_{k=1}^{n} \binom{n+1}{k+1} \frac{\Gamma(k-n'+1)}{\Gamma(k)} n^{-k+n'-1}. \tag{21}$$

We proceed as in (Lairez 2017, Theorem 10) and deduce that

$$\mathbb{E}\left[|||(TM)^\dagger|||^{2n'}\right]^{\frac{1}{n'}} \leqslant \left(\frac{5}{4(2-n')}\right)^{\frac{1}{n'}} \frac{n}{2} \leqslant n. \tag{22}$$

Concerning the second factor, $|||T|||^2$ is the maximum of n χ^2-distributed random variables with $2n+2$ degrees of freedom, say $Z_1, \dots, Z_n$, hence

$$\mathbb{E}\left[|||T|||^{2n+4}\right] \leqslant \sum_{i=1}^{n} \mathbb{E}[Z_i^{n+2}] = 2^{n+2} \frac{(2n+2)!}{(n-1)!}.$$

Therefore,

$$\mathbb{E}\left[|||M^\dagger|||^2\right] \leqslant n \left(2^{n+2} \frac{(2n+2)!}{(n-1)!}\right)^{\frac{1}{n+2}} \leqslant 6n^2,$$

after a few numerical computations. □

3.3. **Lipschitz properties.** We aim at bounding the variation of the numbers κ and $\hat{\gamma}$ on the Riemannian manifold $\mathcal{U} \times \mathbb{P}^n$. In particular, we will prove that $1/\hat{\gamma}$ is Lipschitz-continuous. Traditionally, such results bound directly the value at some point x with respect to the value at some other point y and the distance from x to y. For example (Dedieu 2006, Lemme 131), for any $x, y \in \mathbb{P}^n$,

$$\gamma(F, y) \leqslant \gamma(F, x)\, q\left(d_{\mathbb{P}}(x, y)\gamma(F, x)\right), \tag{23}$$

where $q(u) \doteq \frac{1}{(1-u)(1-4u+2u^2)} = 1+5u+O(u^2)$, given that $d_{\mathbb{P}}(x, y)\gamma(F, x) \leqslant 1-1/\sqrt{2}$. As much as possible, I tried to express this kind of inequalities as a bound on the derivative of the function under consideration. To first order, this is equivalent.

Proposition 18. *Let $F \in \mathcal{H}[r]$ and $\gamma_F : x \in \mathbb{P}^n \mapsto \gamma(F, x)$. For any $x \in \mathbb{P}^n$, we have $\|\mathrm{d}_x \gamma_F\| \leqslant 5\gamma_F^2$.*

Note that γ_F may not be differentiable everywhere, so the inequality $\|\mathrm{d}_x \gamma_F\| \leqslant 5\gamma_F(x)^2$ really means that

$$\limsup_{y \to x} \frac{|\gamma_F(y) - \gamma_F(x)|}{d_{\mathbb{P}}(y, x)} \leqslant 5\gamma_F(x)^2.$$

It is easy to check that Proposition 18 is equivalent to the Lipschitz continuity of the function $1/\gamma_F$, with Lipschitz constant at most 5. We give $\|\mathrm{d}\kappa\|$ and $\|\mathrm{d}\hat{\gamma}\|$ an analogue meaning.

Proof of Proposition 18. The most direct way to see this is by (23):

$$\frac{\gamma_F(y)-\gamma_F(x)}{d_{\mathbb{P}}(y,x)} \leqslant \gamma_F(x)\frac{5u+O(u^2)}{d_{\mathbb{P}}(y,x)} = 5\gamma_F(x)^2\left(1+O(d_{\mathbb{P}}(x,y))\right),$$

as $y \to x$, where $u \doteq d_{\mathbb{P}}(x,y)\gamma_F(x)$, and

$$\frac{\gamma_F(x)-\gamma_F(y)}{d_{\mathbb{P}}(x,y)} \leqslant \gamma_F(y)\frac{5v+O(v^2)}{d_{\mathbb{P}}(x,y)} = 5(\gamma_F(x)+o(1))^2\left(1+O(d_{\mathbb{P}}(x,y))\right),$$

where $v \doteq d_{\mathbb{P}}(x,y)\gamma_F(y)$. □

Lemma 19. *On $\mathcal{U}\times\mathbb{P}^n$, $\|\mathrm{d}\kappa\| \leqslant \kappa^2 + 3\kappa\hat{\gamma}$. Moreover, if $D \geqslant 2$ then $\|\mathrm{d}\kappa\| \leqslant 5\kappa\hat{\gamma}$.*

Proof. The second inequality follows from the first one: If $D \geqslant 2$ then at least one $\gamma(u_i\cdot f, x)$ is greater or equal to $\frac{D-1}{2}$, by Lemma 11. It follows that $\kappa \leqslant 2\hat{\gamma}$ and then $\kappa^2+3\kappa\hat{\gamma} \leqslant 5\kappa\hat{\gamma}$.

To prove the first inequality, we first remark that $1/\kappa(\mathbf{u},x)$ is a Lipschitz-continuous function of $L(\mathbf{u},x)$ with constant 1. Indeed, $1/\kappa(\mathbf{u},x)$ is the least singular value of $L(\mathbf{u},x)$ as a matrix, see Equation (16), and the Eckart–Young theorem expresses this number as the distance to the set of singular matrices, which is a Lipschitz continuous function with constant 1. Moreover $\mathrm{d}\kappa = -\kappa^2\mathrm{d}\frac{1}{\kappa}$, so it is enough to prove that $|||\mathrm{d}L|||$ is bounded by $1+3\frac{\hat{\gamma}}{\kappa}$.

Let $L_i(u_i,z)\in\mathbb{P}^n$ be the ith component of $L(\mathbf{u},z)$, that is the projective class of the linear form $\mathrm{d}_x(u_i\cdot f_i)$. The tangent space of $\mathbb{P}^n$ at $L_i(u_i,z)$ is isometrically identified with the quotient $\mathbb{C}^{n+1}/\mathbb{C}\cdot L_i(u_i,z)$. Denoting $h_i \doteq u_i\cdot f_i$, we check that, at a point $(\mathbf{u},x)$,

$$\mathrm{d}L_i(0,\dot{x}) = \frac{1}{\|\mathrm{d}_x h_i\|}\mathrm{d}_x^2 h_i(\dot{x}) \mod L_i(u_i,z),$$

and in particular, $\|\mathrm{d}L_i(0,\dot{x})\| \leqslant 2\gamma(h_i,x)\|\dot{x}\|$ for any $\dot{x}\in T_x\mathbb{P}^n$. Besides, $L_i(u_i,u_ix) = L_i(\mathrm{id},x)\circ u_i^*$, which is a 1-Lipschitz continuous function of u_i (Lemma 4 applied to the dual projective space). This proves that $\|\mathrm{d}L_i(\dot{u}_i,\dot{u}_ix)\| \leqslant \|\dot{u}_i\|_u$. Therefore,

$$\begin{aligned}\|\mathrm{d}L_i(\dot{u}_i,\dot{x})\| &\leqslant \|\mathrm{d}L_i(\dot{u}_i,\dot{u}_ix)\| + \|\mathrm{d}L_i(0,\dot{x}-\dot{u}_ix)\|\\ &\leqslant \|\dot{u}_i\|_u + 2\gamma(h_i,x)\left(\|\dot{x}\|+\|\dot{u}_i\|_u\right)\\ &\leqslant \|\dot{u}_i\|_u + 2\sqrt{2}\gamma(h_i,x)\left(\|\dot{\mathbf{u}}\|_u^2+\|\dot{x}\|^2\right)^{\frac{1}{2}},\end{aligned}$$

and then, by the triangle inequality,

$$\begin{aligned}\|\mathrm{d}L(\dot{\mathbf{u}},\dot{x})\| &\leqslant \|\dot{\mathbf{u}}\|_u + 2\sqrt{2}\left(\sum_i \gamma(h_i,x)^2\right)^{\frac{1}{2}}\left(\|\dot{\mathbf{u}}\|_u^2+\|\dot{x}\|^2\right)^{\frac{1}{2}}\\ &\leqslant \left(1+3\frac{\hat{\gamma}(\mathbf{u};x)}{\kappa(\mathbf{u},x)}\right)\left(\|\dot{\mathbf{u}}\|_u^2+\|\dot{x}\|^2\right)^{\frac{1}{2}},\end{aligned}$$

which concludes the proof. □

We now derive a bound for $\|\mathrm{d}\hat{\gamma}\|$.

Lemma 20. *For any homonegeneous polynomial $f:\mathbb{C}^{n+1}\to\mathbb{C}$ the map*

$$(u,x)\in U(n+1)\times\mathbb{P}^n \longmapsto \frac{1}{\gamma(u\cdot f,x)}$$

is Lipschitz continuous with constant at most $5\sqrt{2}$.

Proof. The γ number is invariant under unitary transformations, that is $\gamma(u \cdot f, x) = \gamma(f, u^*x)$. Moreover, the map $(u, x) \in U(n+1) \times \mathbb{P}^n \mapsto u^*x \in \mathbb{P}^n$ is 1-Lipschitz continuous with respect to u (Lemma 4) and to x, thus it is $\sqrt{2}$-Lipschitz continuous on $U(n+1) \times \mathbb{P}^n$. Since $1/\gamma(f, x)$ is 5-Lipschitz continous with respect to x (Proposition 18), the map $1/\gamma(f, u^*x)$ is $5\sqrt{2}$-Lipschitz continuous on $U(n+1) \times \mathbb{P}^n$. □

Lemma 21. *On* $\mathcal{U} \times \mathbb{P}^n$, $\|\mathrm{d}\hat{\gamma}\| \leqslant 13\,\hat{\gamma}^2$. *Equivalently,* $1/\hat{\gamma}$ *is 13-Lipschitz continuous.*

Proof. We may assume that $D \geqslant 2$, otherwise $\hat{\gamma}$ is identically 0. Let $\gamma_i(\mathbf{u}, x) \doteq \gamma(u_i \cdot f_i, x)$ and $\eta \doteq \sum_i \gamma_i^2$, so that $\hat{\gamma} = \kappa\sqrt{\eta}$. By Lemma 20, $\|\mathrm{d}\gamma_i\| \leqslant 5\sqrt{2}\gamma_i^2$ and then

$$\|\mathrm{d}\eta\| \leqslant 2\sum_i \gamma_i \|\mathrm{d}\gamma_i\| \leqslant 10\sqrt{2}\sum_i \gamma_i^3 \leqslant 10\sqrt{2}\eta^{3/2}.$$

Therefore,

$$\begin{aligned}
\|\mathrm{d}\hat{\gamma}\| &\leqslant \|\mathrm{d}\kappa\|\sqrt{\eta} + \tfrac{1}{2}\kappa\eta^{-\frac{1}{2}}\|\mathrm{d}\eta\| \\
&\leqslant 5\hat{\gamma}^2 + 5\sqrt{2}\hat{\gamma}\sqrt{\eta}, && \text{by Lemma 19,} \\
&\leqslant 13\hat{\gamma}^2, && \text{using } \kappa \geqslant 1, \text{ Lemma 1.}
\end{aligned}$$ □

3.4. **Numerical continuation along rigid paths.** We describe a continuation algorithm in the rigid solution variety and bound its complexity in terms of the integral of $\kappa\hat{\gamma}$ along the continuation path. It is the analogue of the μ^2 estimate of the classical theory, see §1.1. The approach proposed here differs from the usual treatment only in a more systematic use of derivatives. We assume $D \geqslant 2$ as otherwise there is only a linear system of equations to solve.

As we will see in §4, it may be valuable not to compute $\hat{\gamma}$ but rather an easier to compute upper bound. That is why the algorithm is described in terms of a function $g : \mathcal{U} \times \mathbb{P}^n \to (0, \infty]$ that can be chosen freely, as long as:

(H1) $\hat{\gamma} \leqslant g$, on $\mathcal{U} \times \mathbb{P}^n$;
(H2) $\frac{1}{g}$ is C-Lipschitz continuous, for some $C \geqslant 10$.

The following proposition describes one continuation step. Observe that the conclusion (ii) concerning the triple $(\mathbf{v}, \zeta', z')$ is similar to the hypothesis (a) concerning $(\mathbf{u}, \zeta, z)$, so that we can chain the continuation steps. The geodesic distance in $\mathcal{U}$ is denoted $d_{\mathcal{U}}$.

Proposition 22. *Let* $\mathbf{u}, \mathbf{v} \in \mathcal{U}$, *let* ζ *be a zero of the polynomial system* $\mathbf{u} \cdot F$, *let* $z \in \mathbb{P}^n$ *and let* $z' \doteq \mathcal{N}_{\mathbf{v}\cdot F}(z)$. *For any positive real number* $A \leqslant \frac{1}{4C}$, *if*

(a) $d_{\mathbb{P}}(z, \zeta) g(\mathbf{u}, \zeta) \leqslant A$, *and*
(b) $d_{\mathcal{U}}(\mathbf{u}, \mathbf{v}) \kappa(\mathbf{u}, z) g(\mathbf{u}, z) \leqslant \frac{1}{4}A$,

then there exists a unique zero ζ' *of* $\mathbf{v} \cdot F$ *such that*

(i) z *is an approximate zero of* $\mathbf{v} \cdot F$ *with associated zero* ζ', *and*
(ii) $d_{\mathbb{P}}(z', \zeta') g(\mathbf{v}, \zeta') \leqslant A$.

Proof. It suffices to construct a zero ζ' of $\mathbf{v} \cdot F$ such that $d_{\mathbb{P}}(z, \zeta') g(\mathbf{v}, \zeta') \leqslant 2A$. Indeed, since $2A \leqslant \frac{1}{6}$ and $\gamma \leqslant \hat{\gamma} \leqslant g$, it would imply by Theorem 12 that: (i) z is an approximate zero of $\mathbf{v} \cdot F$; and (ii) that

$$d_{\mathbb{P}}(z', \zeta') g(\mathbf{v}, \zeta') \leqslant \tfrac{1}{2} d_{\mathbb{P}}(z, \zeta') g(\mathbf{v}, \zeta') \leqslant A.$$

Consider a geodesic $t \in [0,\infty) \mapsto \mathbf{v}_t \in \mathcal{U}$ such that $\|\dot{\mathbf{v}}_t\| = 1$, $\mathbf{v}_0 = \mathbf{u}$ and $\mathbf{v}_{d_{\mathcal{U}}(\mathbf{u},\mathbf{v})} = \mathbf{v}$. We may assume that $\kappa(\mathbf{u},\zeta) < \infty$, otherwise $\mathbf{u} = \mathbf{v}$, by Hypothesis (b), and there is nothing to prove. So $(\mathbf{u},\zeta)$ is not in the singular locus Σ and for t small enough, there is a unique lift $(\mathbf{v}_t,\eta_t)$ in $\mathcal{V}\backslash\Sigma$, see §3.1. Let $[0,\tau)$, with $0 < \tau \leqslant \infty$, be the maximal interval of definition of η_t.

For $t \in [0,\tau)$, let $p_t \doteq (\mathbf{v}_t,\eta_t)$ and let $\delta_t \doteq d_{\mathbb{P}}(z,\eta_t)$. Let moreover $g_t \doteq g(\mathbf{v}_t,\eta_t)$, $\beta_t \doteq g_t\delta_t$ and $\kappa_t \doteq \kappa(\mathbf{v}_t,\eta_t)$. The derivative with respect to t is denoted with a dot.

We first observe that $\dot{\delta}_t \leqslant \kappa_t$, by Lemma 16, and that

$$\|\dot{p}_t\| \leqslant (1+\dot{\delta}_t^2)^{\frac{1}{2}} \leqslant 2\kappa_t \tag{24}$$

(using $\kappa \geqslant 1$, by Lemma 1). By Lemma 19, $\dot{\kappa}_t \leqslant 5\kappa_t g_t \|\dot{p}_t\| \leqslant 10\kappa_t^2 g_t$ and by the Lipschitz hypothesis on $1/g$, we have $\dot{g}_t \leqslant C g_t^2 \|\dot{p}_t\| \leqslant 2C\kappa_t g_t^2$. This implies that $\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t g_t \leqslant 3C(\kappa_t g_t)^2$ (using $C \geqslant 10$) and equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{\kappa_t g_t} \geqslant -3C. \tag{25}$$

It follows, after integration, that $\frac{1}{\kappa_t g_t} \geqslant \frac{1}{\kappa_0 g_0} - 3Ct$, and thus

$$\kappa_t g_t \leqslant \frac{\kappa_0 g_0}{1 - 3Ct\kappa_0 g_0}, \tag{26}$$

for any $t \in [0,\tau)$ such that $1 - 3Ct\kappa_0 g_0 > 0$.

A first consequence of (26) is that

$$\tau \geqslant (3C\kappa_0 g_0)^{-1}. \tag{27}$$

Indeed, assuming that $\tau < (3C\kappa_0 g_0)^{-1}$, Inequalities (24), (26) and $g_t \geqslant \frac{1}{2}$ (Lemma 11) show that $\|\dot{p}_t\|$ is bounded on $[0,\tau)$. Therefore p_t has a limit as $t \to \tau$, and then η_t is defined on the interval $[0,\tau]$. But $g_\tau < \infty$, by (26) with $t \to \tau$, therefore η_τ is not a singular zero of $\mathbf{v}_\tau \cdot F$ and η_t can be continued in a neighborhood of τ contradicting that τ is maximal.

Next, we compute that

$$\begin{aligned}
\dot{\beta}_t &= g_t\dot{\delta}_t + \dot{g}_t\delta_t \\
&\leqslant g_t\kappa_t + 2C\kappa_t g_t^2\delta_t \\
&= (1+2C\beta_t)\kappa_t g_t,
\end{aligned}$$

then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\log(1+2C\beta_t) &\leqslant 2C\kappa_t g_t \\
&\leqslant \frac{2C\kappa_0 g_0}{1-3Ct\kappa_0 g_0}, && \text{by (26)}, \\
&= -\frac{2}{3}\frac{\mathrm{d}}{\mathrm{d}t}\log(1-3Ct\kappa_0 g_0),
\end{aligned}$$

and therefore, after integration with respect to t,

$$\log\left(\frac{1+2C\beta_t}{1+2C\beta_0}\right) \leqslant \frac{2}{3}\log\left(\frac{1}{1-3Ct\kappa_0 g_0}\right). \tag{28}$$

Exponentiating both sides leads to

$$\beta_t \leqslant \frac{1+2C\beta_0}{2C\left(1-3Ct\kappa_0 g_0\right)^{\frac{2}{3}}} - \frac{1}{2C}. \tag{29}$$

We now bound $\kappa_0 g_0$. As a function on $\mathcal{U} \times \mathbb{P}^n$, we compute that

$$\begin{aligned}(30)\qquad \|\mathrm{d}(\kappa g)\| &\leqslant \|\mathrm{d}\kappa\| g + \|\mathrm{d}g\|\kappa \\ &\leqslant 5\kappa\hat{\gamma} g + C g^2 \kappa, \text{ by Lemma 19 and } C\text{-Lipschitz continuity of } \frac{1}{g}, \\ &\leqslant \tfrac{3}{2} C \kappa g^2, \quad \text{because } \hat{\gamma} \leqslant g \text{ and } 10 \leqslant C, \text{ by assumption,}\end{aligned}$$

and then,

$$(31)\qquad \|\mathrm{d}\log(\kappa g)\| = (\kappa g)^{-1} \|\mathrm{d}(\kappa g)\| \leqslant \tfrac{3}{2} C g.$$

Since $1/g$ is C-Lipschitz continuous on $\mathcal{U} \times \mathbb{P}^n$, for any $w \in \mathbb{P}^n$ on a shortest path between z and ζ, $|g(\mathbf{u}, w)^{-1} - g(\mathbf{u}, \zeta)^{-1}| \leqslant C d_{\mathbb{P}}(\zeta, z)$ and then

$$(32)\qquad g(\mathbf{u}, w) \leqslant \frac{g(\mathbf{u}, \zeta)}{1 - C d_{\mathbb{P}}(\zeta, z) g(\mathbf{u}, \zeta)}.$$

After integrating the relation (31) on a path from $(\mathbf{u}, \zeta)$ to $(\mathbf{u}, z)$, and bounding the right-hand side with (32), we obtain

$$\log(\kappa_0 g_0) \leqslant \log(\kappa(\mathbf{u}, z) g(\mathbf{u}, z)) + \frac{\frac{3}{2} C d_{\mathbb{P}}(\zeta, z) g(\mathbf{u}, \zeta)}{1 - C d_{\mathbb{P}}(\zeta, z) g(\mathbf{u}, \zeta)},$$

and then

$$(33)\qquad \kappa_0 g_0 \leqslant \kappa(\mathbf{u}, z) g(\mathbf{u}, z) \exp\left(\frac{\frac{3}{2} C d_{\mathbb{P}}(z, \zeta) g(\mathbf{u}, \zeta)}{1 - C d_{\mathbb{P}}(z, \zeta) g(\mathbf{u}, \zeta)} \right).$$

We multiply by $d_{\mathcal{U}}(\mathbf{u}, \mathbf{v})$ both sides, use the hypotheses (a) and (b), and obtain

$$(34)\qquad d_{\mathcal{U}}(\mathbf{u}, \mathbf{v}) \kappa_0 g_0 \leqslant \tfrac{1}{4} A \exp\left(\frac{\frac{3}{2} C A}{1 - C A} \right) < \tfrac{5}{12} A,$$

where the last inequality follows from the hypothesis $CA \leqslant \frac{1}{4}$. Together with (27) and the hypothesis $CA \leqslant \frac{1}{4}$, we deduce $d_{\mathcal{U}}(\mathbf{u}, \mathbf{v}) < \tau$. In particular, we can define $\zeta' = \eta_{d_{\mathcal{U}}(\mathbf{u}, \mathbf{v})}$, which is a zero of $\mathbf{v} \cdot F$, and apply Inequality (29) at $t = d_{\mathcal{U}}(\mathbf{u}, \mathbf{v})$. This gives, in combination with (34) and Hypothesis (a),

$$\begin{aligned} d_{\mathbb{P}}(z, \zeta') g(\mathbf{v}, \zeta') &\leqslant \frac{1 + 2CA}{2C(1 - \frac{5}{4} CA)^{\frac{2}{3}}} - \frac{1}{2C} \\ &\leqslant \left(\frac{1 + \frac{2}{4}}{\frac{2}{4}(1 - \frac{5}{16})^{\frac{2}{3}}} - \frac{4}{2} \right) A \\ &\leqslant 2A \end{aligned}$$

using again that $CA \leqslant \frac{1}{4}$. This concludes the proof. □

Based on Proposition 22, the procedure NC (Algorithm 2) computes an approximate zero of a system $\mathbf{u} \cdot F$ given a zero of another system $\mathbf{v} \cdot F$ using a numerical continuation along a path $(\mathbf{w}_t)_{0 \leqslant t \leqslant T}$ from $\mathbf{v}$ to $\mathbf{u}$.

Theorem 23. *On input F, $\mathbf{u}$, $\mathbf{v}$ and z, assuming that z is a zero of $\mathbf{v} \cdot F$, Algorithm 2 ouputs an approximate zero of $\mathbf{u} \cdot F$ or loops forever.*

If the continuation path $(\mathbf{w}_t)_{0 \leqslant t \leqslant T}$ chosen by the algorithm lifts as a continous path $(\mathbf{w}_t, \zeta_t)$ in $\mathcal{V}$ with $\zeta_0 = z$, then the the algorithm terminates after at most

$$25 C \int_0^T \kappa(\mathbf{w}_t, \zeta_t) g(\mathbf{w}_t, \zeta_t) \mathrm{d}t$$

continuation steps.

Algorithm 2. Numerical continuation
Input. $F \in \mathcal{H}[n]$, $\mathbf{u}$, $\mathbf{v} \in \mathcal{U}$ and $z \in \mathbb{P}^n$
Precondition. z is a zero of $\mathbf{v} \cdot F$ and the function g satisfies (H1) and (H2).
Output. $w \in \mathbb{P}^n$.
Postcondition. w is an approximate zero of $\mathbf{u} \cdot F$.

```
function NC(F, u, v, z)
    (w_t)_{0≤t≤T} ← a 1-Lipschitz continuous path from v to u
    t ← 1/(16C κ(w_0, z) g(w_0, z))
    while t < T do
        z ← N_{w_t}(z)
        t ← t + 1/(16C κ(w_t, z) g(w_t, z))
    end while
    return z
end function
```

Proof. Let $t_0 \doteq 0$, let t_k be the value of t at the beginning of the kth iteration, let $z_0 \doteq z$ and let z_k be the value of z at the end of the kth iteration; namely

$$t_{k+1} \doteq t_k + \frac{1}{16C\,\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k)} \tag{35}$$

$$z_{k+1} \doteq \mathcal{N}_{\mathbf{w}_{t_{k+1}}}(z_k). \tag{36}$$

Let K be the largest integer such that $t_K \leqslant T$ (or $K \doteq \infty$ if there is no such integer). The output of the algorithm, if any, is z_K. Thanks to the Lipschitz hypothesis for $\mathbf{w}$, we have, for any $k \geqslant 0$,

$$d_{\mathcal{U}}(\mathbf{w}_{t_k}, \mathbf{w}_{t_{k+1}}) \leqslant t_{k+1} - t_k = \left(16C\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k)\right)^{-1}.$$

Repeated application of Proposition 22 with $A \doteq \frac{1}{4C}$ leads to

$$d_{\mathbb{P}}(z_k, \zeta_{t_k}) g(\mathbf{w}_{t_k}, \zeta_{t_k}) \leqslant A = \frac{1}{4C}, \tag{37}$$

for any $k \geqslant 0$ (conclusion (ii) of Proposition 22 is used for hypothesis (a) at the next step, initialization is trivial since z is a zero). By Proposition 22 again, for any $k \geqslant 0$ and any $t \in [t_k, t_{k+1}]$, z_k is an approximate zero of $\mathbf{w}_t \cdot F$. In particular, z_K is an approximate zero of $\mathbf{w}_T \cdot F$. This proves the correctness of the algorithm.

Concerning the bound on the number of iterations, we first note that

$$\begin{aligned}
\int_0^T \kappa_t g_t \mathrm{d}t &\geqslant \int_0^{t_K} \kappa_t g_t \mathrm{d}t, && \text{because } t_K < T, \\
&\geqslant \sum_{k=0}^{K-1} (t_{k+1} - t_k) \min_{t_k \leqslant s < t_{k+1}} \kappa_s g_s \\
&= \sum_{k=0}^{K-1} \frac{\min_{t_k \leqslant s < t_{k+1}} \kappa_s g_s}{16C\,\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k)}, && (38)
\end{aligned}$$

where we use the notations of the proof of Proposition 22 applied to the path $(\mathbf{w}_t)$.

Similarly to (25), but aiming now for lower bounds, we compute that

$$\frac{\mathrm{d}}{\mathrm{d}s} \frac{1}{\kappa_s g_s} \leqslant 3C.$$

After integration, analogously to (26), we obtain that for any $t_k \leqslant s < t_{k+1}$

$$\kappa_s g_s \geqslant \frac{\kappa_{t_k} g_{t_k}}{1 + 3C(t_{k+1} - t_k)\kappa_{t_k} g_{t_k}}.$$

We also check, similarly to (33), integrating along a shortest path from ζ_{t_k} to z_k, that

$$\begin{aligned}\kappa_{t_k} g_{t_k} &\geqslant \kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k) \exp\left(-\frac{\frac{3}{2} C d_{\mathbb{P}}(z_k, \zeta_{t_k}) g(\mathbf{w}_{t_k}, \zeta_k)}{1 + C d_{\mathbb{P}}(z_k, \zeta_{t_k}) g(\mathbf{w}_{t_k}, \zeta_{t_k})}\right) \\ &\geqslant \kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k) \exp(-\tfrac{3}{10}), \quad \text{with (37).}\end{aligned}$$

Therefore, for any $t_k \leqslant s < t_{k+1}$,

$$\begin{aligned}\kappa_s g_s &\geqslant \frac{\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k) \exp(-\frac{3}{10})}{1 + 3C(t_{k+1} - t_k)\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k) \exp(-\frac{3}{10})} \\ &\geqslant \frac{\kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k)}{\exp\left(\frac{3}{10}\right) + \frac{3}{16}}, \text{ using the value for } t_{k+1} - t_k \text{ (35)},\end{aligned}$$

and then

$$\frac{\min_{t_k \leqslant s < t_{k+1}} \kappa_s g_s}{16C\, \kappa(\mathbf{w}_{t_k}, z_k) g(\mathbf{w}_{t_k}, z_k)} \geqslant \frac{1}{16 \exp\left(\frac{3}{10}\right) + 3} \cdot \frac{1}{C} \leqslant \frac{1}{25C}.$$

Therefore, by (38), $\int_0^T \kappa_t g_t \mathrm{d}t \geqslant \frac{1}{25C} K$. □

We have some degrees of freedom but also some constraints in the choice of the path $(\mathbf{w}_t)_{0 \leqslant t \leqslant T}$ from $\mathbf{v}$ to $\mathbf{u}$:

(P1) the path is 1-Lipschitz continuous;
(P2) the path $(\mathbf{v}^{-1}\mathbf{w}_t)_{0 \leqslant t \leqslant T}$ (from $\mathbf{1}_{\mathcal{U}}$ to $\mathbf{v}^{-1}\mathbf{u}$) depends only on $\mathbf{v}^{-1}\mathbf{u}$; and
(P3) the length T of the path is at most $4n$.

The first one is required by the numerical continuation algorithm. The two others will be useful for the complexity analysis.

An obvious choice of the path between $\mathbf{v}$ and $\mathbf{u}$ is the shortest one: we write $\mathbf{v}^{-1}\mathbf{u} = (\exp(A_1), \ldots, \exp(A_n))$ for some skew-Hermitian matrices $A_1, \ldots, A_n$ and define

$$\mathbf{w}_t \doteq \mathbf{v}\left(\exp(\tfrac{t}{T} A_1), \ldots, \exp(\tfrac{t}{T} A_n)\right), \text{ for } 0 \leqslant t \leqslant T,$$

where $T \doteq (\sum_i \|A_i\|_u^2)^{\frac{1}{2}}$. We can always choose the matrices A_i such that $T \leqslant 4n$ as follows: diagonalize each A_i in a unitary basis (which preserves the Frobenius norm) as $\mathrm{diag}(\theta_{i,0}, \ldots, \theta_{i,n})\sqrt{-1}$, for some $\theta_{i,j} \in \mathbb{R}$; add integer multiples of 2π to the $\theta_{i,j}$ (which preserves the end points of the path) so that $\theta_{i,j} \in [-\pi, \pi]$; and finally, obtain

$$T^2 = \sum_{i=1}^n \|A_i\|_u^2 = \frac{1}{2} \sum_{i=1}^n \sum_{j=0}^n \theta_{i,j}^2 \leqslant \frac{1}{2} n(n+1)\pi^2 < (4n)^2.$$

Naturally, it may not be convenient to compute matrix logarithms and exponentials, especially for the complexity analysis in the BSS model. In §4.3.2 we will see paths that are cheaper to compute but still satisfy (P2) and (P3).

Remark 24. To implement Algorithm 2, the computation of the step size by which t is updated can be relaxed: it is enough to compute a number τ such that

$$M^{-1} \left(16C\, \kappa(\mathbf{w}_t, z) g(\mathbf{w}_t, z)\right)^{-1} \leqslant \tau \leqslant \left(16C\, \kappa(\mathbf{w}_t, z) g(\mathbf{w}_t, z)\right)^{-1},$$

for some constant $M \geqslant 1$, and use it as the step size. Correctness is not harmed, and the number of continuation steps is now bounded by

$$25MC \int_0^T \kappa(\mathbf{w}_t, \zeta_t) g(\mathbf{w}_t, \zeta_t) \mathrm{d}t.$$

Algorithm 3. An analogue of Beltrán-Pardo algorithm in the rigid setting
Input. Homogeneous polynomial system $F \in \mathcal{H}[n]$ and $\mathbf{u} \in \mathcal{U}$
Output. $z \in \mathbb{P}^n$.
Postcondition. z is an approximate zero of $\mathbf{u} \cdot F$.

function Solve(F, $\mathbf{u}$)
  $(\mathbf{v}, \eta) \leftarrow \text{Sample}(\mathcal{V}_F)$
  return NC $(\mathbf{u}, \mathbf{v}, \eta)$
end function

3.5. **A randomized algorithm.** We have everything we need to mimic Beltrán–Pardo algorithm in the rigid setting: a continuation algorithm with an integral estimate for its complexity and a recipe to sample points in the solution variety with the appropriate distribution. This leads to Algorithm 3 (Solve): for finding a zero of $\mathbf{u} \cdot F$, we first sample a random element $(\mathbf{v}, \eta)$ in $\mathcal{V}$, with Algorithm 1 (Sample), and then perform, with Algorithm 2, the numerical continuation along a path in $\mathcal{U}$ from $\mathbf{v}$ to $\mathbf{u}$. Based on Theorem 23, we can estimate the expected number of continuation steps performed by Solve$(F, \mathbf{u})$ when $\mathbf{u}$ is uniformly distributed.

Theorem 25. *If $\mathbf{u} \in \mathcal{U}$ is uniformly distributed then* Solve$(F, \mathbf{u})$ *terminates almost surely and outputs an approximate zero of $\mathbf{u} \cdot F$ after K continuation steps, with*

$$\mathbb{E}[K] \leqslant 100Cn\, \mathbb{E}[\kappa(\mathbf{v}, \eta) g(\mathbf{v}, \eta)].$$

Proof. Let $(\mathbf{w}_t)_{0 \leqslant t \leqslant T}$ be the path from $\mathbf{v}$ to $\mathbf{u}$ chosen in NC. By (P2), $\mathbf{v}^{-1}\mathbf{w}_t$ is a function of $\mathbf{v}^{-1}\mathbf{u}$. We first note that $\mathbf{v}$ and $\mathbf{v}^{-1}\mathbf{u}$ are independent and uniformly distributed in $\mathcal{U}$, because the Jacobian of the diffeomorphism

$$(\mathbf{u}, \mathbf{v}) \in \mathcal{U} \times \mathcal{U} \mapsto (\mathbf{v}, \mathbf{v}^{-1}\mathbf{u}) \in \mathcal{U} \times \mathcal{U}$$

is constant. Secondly, by hypothesis, for any $0 \leqslant s \leqslant 1$, the random variable $\mathbf{v}^{-1}\mathbf{w}_{Ts}$ depends only on $\mathbf{v}^{-1}\mathbf{u}$, so it is independent from $\mathbf{v}$. Therefore $\mathbf{w}_{Ts}$, which equals $\mathbf{v}(\mathbf{v}^{-1}\mathbf{w}_{Ts})$, is uniformly distributed and independent from $\mathbf{v}^{-1}\mathbf{u}$.

By Proposition 15, with probability 1, all the polynomial systems $\mathbf{w}_t \cdot F$, for $0 \leqslant t \leqslant T$, have only regular zeros. So almost surely, all the zeros of $\mathbf{v} \cdot F$ can be continued as zeros of $\mathbf{w}_t \cdot F$. Let ζ_t be the zero of $\mathbf{w}_t \cdot F$ obtained by continuation of the zero η of $\mathbf{v} \cdot F$. Since ζ is uniformly distributed among the zeros of $\mathbf{v}$, it follows that ζ_t is uniformly distributed among the zeros of $\mathbf{w}_t$, because the numerical continuation, which is almost surely well defined, induces a bijective correspondance between the two finite sets of zeros. Therefore, for any $0 \leqslant s \leqslant 1$, $(\mathbf{w}_{Ts}, \zeta_{Ts})$ is a ρ_{std}-distributed random variable independent from $\mathbf{v}^{-1}\mathbf{u}$, and in particular, independent from T.

Together with Theorem 23 (and the change of variable $t = Ts$), this implies

$$\begin{aligned}
\mathbb{E}[K] &\leqslant \mathbb{E}\left[25C \int_0^1 \kappa(\mathbf{w}_{Ts}, \zeta_{Ts}) g(\mathbf{w}_{Ts}, \zeta_{Ts}) T \mathrm{d}s\right] \\
&= 25C \int_0^1 \mathbb{E}[\kappa(\mathbf{w}_{Ts}, \zeta_{Ts}) g(\mathbf{w}_{Ts}, \zeta_{Ts}) T]\, \mathrm{d}s \\
&= 25C\, \mathbb{E}[\kappa(\mathbf{v}, \eta) g(\mathbf{v}, \eta)]\, \mathbb{E}[T],
\end{aligned}$$

which leads to the claim, with the bound $T \leqslant 4n$ given by (P3). □

4. Average complexity for random dense polynomial systems

Theorem 25 on the average complexity of computing one zero of a random system $\mathbf{u}\cdot F$, with $\mathbf{u}\in\mathcal{U}$ uniformly distributed and F fixed, is applicable to the more common question of computing one zero of a random system F, under a unitary invariance condition on the probability distribution of F.

The polynomial system $F\in\mathcal{H}[n]$ fixed in §3 is now a random variable. The probability distribution of F is assumed to be *unitary invariant*: that is, for any $\mathbf{u}\in\mathcal{U}$, the random systems F and $\mathbf{u}\cdot F$ are identically distributed.

An important example of a unitary invariant distribution is *Kostlan's distribution*: that is the standard Gaussian distribution on $\mathcal{H}[n]$ endowed with Weyl's Hermitian norm. The unitary invariance of Kostlan's distribution follows from the invariance of Weyl's norm under the action of $\mathcal{H}$. Concretely, a *Kostlan random system* $F\in\mathcal{H}[n]$ is a tuple $(f_1,\dots,f_n)$ of independent *Kostlan random polynomials*, that is

$$f_i \doteq \sum_{j_0+\cdots+j_n=d_i}\left(\frac{d_i!}{j_0!\cdots j_n!}\right)^{\frac12} c_{j_0,\dots,j_n}x_0^{j_0}\cdots x_n^{j_n},$$

where the coefficients $c_{j_0,\dots,j_n}$ are independent standard normal variables in $\mathbb{C}$.

Section 4.1 contains the average analysis of the number of continuation steps in the general setting of a unitary invariant distribution. Section 4.2 describes the function $\hat\gamma_{\mathrm{Frob}}$ that will be used for the numerical continuation (in the role of g). Next, Section 4.3 discusses the computational model and the construction of paths in $\mathcal{U}$. And lastly, Section 4.4 concludes the average analysis for Kostlan's distribution.

4.1. Unitary invariant random systems.

Let $\mathfrak{g}:\mathcal{H}[n]\to(0,\infty]$ be a function of the form (compare to Equation (13) defining $\hat\gamma$)

$$\mathfrak{g}(F,z)\doteq\kappa(F,z)\big(\mathfrak{g}_1(f_1,z)^2+\cdots+\mathfrak{g}_n(f_n,z)^2\big)^{\frac12}, \tag{39}$$

for some $\mathfrak{g}_i:H_{d_i}\times\mathbb{P}^n\to(0,\infty]$ such that for any $1\leqslant i\leqslant n$ and any $f\in H_{d_i}$,

- (H1') $x\in\mathbb{P}^n$, $\gamma(f,x)\leqslant\mathfrak{g}_i(f,x)$;
- (H2') for any $x\in\mathbb{P}^n\mapsto\mathfrak{g}_i(f,x)^{-1}$ is C'-Lipschitz continuous ($C'\geqslant 4$); and
- (H3') for any $z\in\mathbb{P}^n$ and $u\in U(n+1)$, $\mathfrak{g}_i(u\cdot f,x)=\mathfrak{g}_i(f,u^{-1}x)$.

These assumptions make it possible to use $\mathfrak{g}$ for numerical continuation in the rigid setting.

Lemma 26. *If (H1')–(H3') hold, then for any $F\in\mathcal{H}[n]$, the function*

$$g:(\mathbf{u},x)\in\mathcal{U}\times\mathbb{P}^n\mapsto\mathfrak{g}(\mathbf{u}\cdot F,x)$$

satisfies conditions (H1) and (H2), with $C=3C'$.

Proof. Condition (H1) follows directly from the definitions of $\hat\gamma$ and $\mathfrak{g}$. Condition (H2) follows exactly as the Lipschitz continuity of $\hat\gamma$ (Lemma 21). □

Theorem 27. *If the probability distribution of the system $F=(f_1,\dots,f_n)\in\mathcal{H}[n]$ is unitary invariant, the procedure* Solve$(F,\mathbf{1}_{\mathcal{U}})$*, with $g(\mathbf{u},x)\doteq\mathfrak{g}(\mathbf{u}\cdot F,x)$ for the continuation, computes an approximate root of F using K continuation steps, with*

$$\mathbb{E}[K]\leqslant 1800C'n^3\left(\sum_{i=1}^n\mathbb{E}\left[\mathfrak{g}_i(f_i,\zeta_i)^2\right]\right)^{\frac12},$$

where $\zeta_1,\dots,\zeta_n\in\mathbb{P}^n$ are random independent uniformly distributed zeros of the respective random polynomials $f_1,\dots,f_n$.

Proof. Let $K(F, \mathbf{u})$ be the (random) number of continuation steps performed by $\mathrm{Solve}(F, \mathbf{u})$. Let $\mathbf{u} \in \mathcal{U}$ be a uniformly distributed random variable independent from F. By the unitary invariance hypothesis, $F \sim \mathbf{u} \cdot F$, so $K(F, \mathbf{1}_{\mathcal{U}}) \sim K(\mathbf{u} \cdot F, \mathbf{1}_{\mathcal{U}})$, and by the assumption (P2) on the unitary invariance of the choice of the continuation path, $K(\mathbf{u} \cdot F, \mathbf{1}_{\mathcal{U}}) \sim K(F, \mathbf{u})$. In particular, $\mathbb{E}\left[K(F, \mathbf{1}_{\mathcal{U}})\right] = \mathbb{E}\left[K(F, \mathbf{u})\right]$. By Theorem 25 (with $C = 3C'$, by Lemma 26),

$$\mathbb{E}\left[K(F, \mathbf{u})\right] \leqslant 300C'n\, \mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)\mathfrak{g}(\mathbf{u} \cdot F, \zeta)\right], \tag{40}$$

where ζ is a uniformly distributed random zero of $\mathbf{u} \cdot F$. Given the form of $\mathfrak{g}$ (39), the right-hand side expands to

$$\kappa(\mathbf{u} \cdot F, \zeta)\mathfrak{g}(\mathbf{u} \cdot F, \zeta) = \kappa(\mathbf{u} \cdot F, \zeta)^2 \left(\sum_{i=1}^{n} \mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2 \right)^{\frac{1}{2}}, \tag{41}$$

using also the unitary invariance of each $\mathfrak{g}_i$ (H3').

Conditionally on F, $(\mathbf{u}, \zeta)$ is ρ_{std}-distributed in the solution variety $\mathcal{V}_F$. By Theorem 8, $u_1^{-1}\zeta, \dots, u_n^{-1}\zeta$ are independent and uniformly distributed zeros of $f_1, \dots, f_n$ respectively, and $\kappa(\mathbf{u} \cdot F, \zeta)$, which depends only on $L(\mathbf{u} \cdot F, \zeta)$, see (16), is independent with them conditionally on F. So the two factors in the right-hand side of (41) are independent conditionally on F; therefore

$$\begin{aligned} \mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)\mathfrak{g}(\mathbf{u} \cdot F, \zeta) \mid F\right] = \\ \mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)^2 \mid F\right] \mathbb{E}\left[\left(\sum_{i=1}^{n} \mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2 \right)^{\frac{1}{2}} \,\middle|\, F \right], \end{aligned} \tag{42}$$

where $\mathbb{E}\left[-|F\right]$ denotes conditional expectation. Proposition 17 gives the bound $\mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)^2 \mid F\right] \leqslant 6n^2$, and then

$$\begin{aligned} \mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)\mathfrak{g}(\mathbf{u} \cdot F, \zeta)\right] &= \mathbb{E}\Big[\mathbb{E}\left[\kappa(\mathbf{u} \cdot F, \zeta)\mathfrak{g}(\mathbf{u} \cdot F, \zeta) \mid F\right]\Big] \\ &\leqslant \mathbb{E}\left[6n^2\, \mathbb{E}\left[\left(\sum_{i=1}^{n} \mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2 \right)^{\frac{1}{2}} \,\middle|\, F \right] \right], \text{ by (42)}, \\ &= 6n^2\, \mathbb{E}\left[\left(\sum_{i=1}^{n} \mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2 \right)^{\frac{1}{2}} \right] \qquad (43) \\ &\leqslant 6n^2 \left(\sum_{i=1}^{n} \mathbb{E}\left[\mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2\right] \right)^{\frac{1}{2}}, \end{aligned}$$

where the last inequality follows from Jensen's inequality. We note as above that $u_i^{-1}\zeta$ is a uniformly distributed zero of f_i, so that $\mathbb{E}\left[\mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2\right] = \mathbb{E}\left[\mathfrak{g}_i(f_i, \zeta_i)^2\right]$. With (40), this concludes the proof. □

Remark 28. Under the same hypotheses, we also have the bound

$$\mathbb{E}[K] \leqslant 1800C'n^3 \sum_{i=1}^{n} \mathbb{E}\left[\mathfrak{g}_i(f_i, \zeta_i)\right],$$

obtained by bounding $\left(\sum_i \mathfrak{g}_i(f_i, u_i^{-1}\zeta)^2\right)^{\frac{1}{2}} \leqslant \sum_i \mathfrak{g}_i(f_i, u_i^{-1}\zeta)$ in (43).

4.2. **An efficiently computable variant of γ.** Controlling the total complexity of Algorithm 3 requires a function g that is easy to compute. The *Frobenius gamma number* is introduced here with this purpose.

4.2.1. *Norms of a multilinear map.* Let E and F be Hermitian spaces and let $h : E^k \to F$ be a multilinear map, the *operator norm* of h is defined by

$$|||h||| \doteq \max\{\|h(x_1,\dots,x_k)\| \mid x_1,\dots,x_k \in \mathbb{S}(E)\},$$

where $\mathbb{S}(E)$ is the unit sphere of E, and the *Frobenius norm* of h is defined as

$$\|h\|_{\mathrm{Frob}} \doteq \left(\sum_{1\leqslant i_1,\dots,i_k\leqslant n} \|a_{i_1,\dots,i_k}\|^2\right)^{\frac{1}{2}}, \tag{44}$$

where the $a_{i_1,\dots,i_k}$ are the coefficients of h in some unitary basis $e_1,\dots,e_{\dim E}$ of E, that is $a_{i_1,\dots,i_k} \doteq h(e_{i_1},\dots,e_{i_k}) \in F$, for $1 \leqslant i_1,\dots,i_k \leqslant n$. This definition does not depend on he choices of the unitary basis.

Lemma 29. *For any multilinear map $h : E^k \to F$, $|||h||| \leqslant \|h\|_{\mathrm{Frob}} \leqslant (\dim E)^{\frac{k}{2}} |||h|||$.*

Proof. This is better seen if we consider h as a map $E \to (E \to \cdots (E \to F))$. The claim follows from an induction on k and the usual comparison between the Frobenius and the operator norm. □

The following lemma relates the Weyl norm of a homogeneous polynomial with the Frobenius norm of an appropriate higher derivative. Recall that the Weyl norm of a homogeneous polynomial of degree k is defined by

$$\left\| \sum_{j_0+\cdots+j_n=k} c_{j_0,\dots,j_n} x_0^{j_0}\cdots x_n^{j_n} \right\|_W^2 \doteq \sum_{j_0+\cdots+j_n=k} \frac{j_0!\cdots j_n!}{k!} |c_{j_0,\dots,j_n}|^2. \tag{45}$$

Lemma 30. *For any homogeneous polynomial $p(x_0,\dots,x_n)$ of degree k,*

$$\|p\|_W = \left\|\tfrac{1}{k!}\mathrm{d}_0^k p\right\|_{\mathrm{Frob}}.$$

Proof. For any $0 \leqslant i_1,\dots,i_k \leqslant n$,

$$\begin{aligned}\frac{1}{k!}\mathrm{d}_0^k p(e_{i_1},\dots,e_{i_k}) &= \frac{1}{k!}\frac{\partial^k p}{\partial x_{i_1}\cdots\partial x_{i_k}}\bigg|_{x=0} \\ &= \frac{j_0!\cdots j_n!}{k!} c_{j_0,\dots,j_n},\end{aligned}$$

where j_m is the number of indices i_* that are equal to m and where $c_{j_0,\dots,j_n}$ is the coefficient of $x_0^{j_0}\cdots x_n^{j_n}$ in p. There are exactly $\frac{k!}{j_0!\cdots j_n!}$ k-uples i_* that lead to a given $(n+1)$-uple j_*. Therefore,

$$\left\|\tfrac{1}{k!}\mathrm{d}_0^k p(e_{i_1},\dots,e_{i_k})\right\|_{\mathrm{Frob}}^2 = \sum_{j_0+\cdots+j_n=k} \frac{j_0!\cdots j_n!}{k!} |c_{j_0,\dots,j_n}|^2,$$

and this is exactly $\|p\|_W^2$. □

4.2.2. *The γ number with Frobenius norms.* For a homogeneous polynomial $f : \mathbb{C}^{n+1} \to \mathbb{C}$, we define

$$\gamma_{\mathrm{Frob}}(f,z) \doteq \begin{cases} \sup_{k\geqslant 2}\left(\frac{1}{k!}\|\mathrm{d}_z f\|^{-1}\left\|\mathrm{d}_z^k f\right\|_{\mathrm{Frob}}\right)^{\frac{1}{k-1}} & \text{if } \mathrm{d}_z f \text{ is nonzero,} \\ \infty & \text{otherwise,}\end{cases} \tag{46}$$

and for a homogeneous polynomial system $F = (f_1,\dots,f_r) \in \mathcal{H}[r]$, we define

$$\hat{\gamma}_{\mathrm{Frob}}(F,z) \doteq \kappa(F,z)\left(\gamma_{\mathrm{Frob}}(f_1,z)^2 + \cdots + \gamma_{\mathrm{Frob}}(f_r,z)^2\right)^{\frac{1}{2}}.$$

Compare with the definitions of γ and $\hat{\gamma}$, §2.5.

In order to use Algorithm 2 with $\hat{\gamma}_{\mathrm{Frob}}$ as g, we must check (H1) and (H2). By Lemma 26, it is enough to check (H1'), (H2') and (H3') for γ_{Frob} of a single

polynomial. The condition (H1'), that is $\gamma \leqslant \gamma_{\mathrm{Frob}}$, is clear by Lemma 29. The condition (H3'), that is $\gamma_{\mathrm{Frob}}(u \cdot f, z) = \gamma_{\mathrm{Frob}}(f, u^{-1}z)$ follows from the unitary invariance of the Frobenius norm. The following lemma shows (H2').

Lemma 31. *For any polynomial $f \in H_d$, the function $x \in \mathbb{P}^n \mapsto 1/\gamma_{\mathrm{Frob}}(f,x)$ is 5-Lipschitz continuous.*

Proof. Let $(x_t)_{0\leqslant t\leqslant 1}$ be a differentiable path in $\mathbb{S}(\mathbb{C}^{n+1})$. Let $a_t \doteq \|\mathrm{d}_{x_t} f\|^{-1}$ and $(B_k)_t \doteq \frac{1}{k!} a_t d^k_{x_t} f$, so that $(B_k)_t$ is a multilinear map $(\mathbb{C}^{n+1})^k \to \mathbb{C}$. From now on, we drop the index t and denote the derivative with respect to t with a dot or with $\frac{\mathrm{d}}{\mathrm{d}t}$. Writing a as $\langle \mathrm{d}_x f, \mathrm{d}_x f\rangle^{-\frac{1}{2}}$, we compute that

$$\dot a = -a\,\Re\langle a\,\mathrm{d}^2_x f(\dot x), a\,\mathrm{d}_x f\rangle, \tag{47}$$

where $\Re$ denotes the real part, and the Cauchy–Schwartz inequality implies that

$$|\dot a| \leqslant a\,\|a\,\mathrm{d}^2_x f(\dot x)\|, \tag{48}$$

noting that $\|a\,\mathrm{d}_x f\| = 1$. By definition of a and $\gamma(f,x)$,

$$\begin{aligned}
\|a\,\mathrm{d}^2_x f(\dot x)\| &\leqslant \|\mathrm{d}_x f\|^{-1}\,\|\mathrm{d}^2_x f\|\,\|\dot x\| \\
&\leqslant \|\mathrm{d}_x f\|^{-1}\|\mathrm{d}^2_x f\|_{\mathrm{Frob}}\|\dot x\|, && \text{by Lemma 29,} \\
&\leqslant 2\gamma_{\mathrm{Frob}}(f,x)\|\dot x\|, && \text{by definition of } \gamma_{\mathrm{Frob}}(f,x),
\end{aligned}$$

which implies, in combination with (48), that

$$|\dot a| \leqslant 2a\gamma_{\mathrm{Frob}}(f,x)\|\dot x\|. \tag{49}$$

We note that $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}^k_x f = \mathrm{d}^{k+1}_x f(\dot x)$, therefore,

$$\dot B_k = \dot a a^{-1} B_k + (k+1)B_{k+1}(\dot x), \tag{50}$$

and it follows, since for any l, $\|B_l\|_{\mathrm{Frob}} \leqslant \gamma_{\mathrm{Frob}}(f,x)^{l-1}$, that

$$\begin{aligned}
\left|\tfrac{\mathrm{d}}{\mathrm{d}t}\|B_k\|_{\mathrm{Frob}}\right| &\leqslant \|\dot B_k\|_{\mathrm{Frob}}, \quad \text{as } \|-\|_{\mathrm{Frob}} \text{ is 1-Lipschitz continuous,} \\
&\leqslant 2\gamma_{\mathrm{Frob}}(f,x)\|B_k\|_{\mathrm{Frob}}\|\dot x\| + (k+1)\|B_{k+1}\|_{\mathrm{Frob}}\|\dot x\|, \\
&\qquad\qquad \text{by (50) and (49),} \\
&\leqslant \left(2\gamma_{\mathrm{Frob}}(f,x)^k + (k+1)\gamma_{\mathrm{Frob}}(f,x)^k\right)\|\dot x\| \\
&\leqslant (k+3)\gamma_{\mathrm{Frob}}(f,x)^k\|\dot x\|.
\end{aligned}$$

It follows that

$$\left|\tfrac{\mathrm{d}}{\mathrm{d}t}\|B_k\|_{\mathrm{Frob}}^{\frac{1}{k-1}}\right| = \frac{1}{k-1}\|B_k\|_{\mathrm{Frob}}^{\frac{1}{k-1}-1}\left|\tfrac{\mathrm{d}}{\mathrm{d}t}\|B_k\|_{\mathrm{Frob}}\right| \leqslant \frac{k+3}{k-1}\frac{\gamma^{k+1}_{\mathrm{Frob}}}{\|B_k\|_{\mathrm{Frob}}}\|\dot x\|. \tag{51}$$

By definition, γ_{Frob} is the supremum of all $\|B_k\|^{1/(k-1)}_{\mathrm{Frob}}$ $(k \geqslant 2)$, finitely many of which are nonzero, so at a given time t, there is some k such that

$$\gamma_{\mathrm{Frob}}(f,x) = \|B_k\|^{1/(k-1)}_{\mathrm{Frob}} \tag{52}$$

in a (one sided) neighborhood of t. Therefore $\dot\gamma_{\mathrm{Frob}} = \frac{\mathrm{d}}{\mathrm{d}t}\|B_k\|^{1/(k-1)}_{\mathrm{Frob}}$ and the computation above shows that

$$\begin{aligned}
|\dot\gamma_{\mathrm{Frob}}| &\leqslant \frac{k+3}{k-1}\frac{\gamma^{k+1}_{\mathrm{Frob}}}{\|B_k\|_{\mathrm{Frob}}}\|\dot x\|, && \text{by (51),} \\
&= \frac{k+3}{k-1}\frac{\gamma^{k+1}_{\mathrm{Frob}}}{\gamma^{k-1}_{\mathrm{Frob}}}\|\dot x\|, && \text{by (52),} \\
&= \frac{k+3}{k-1}\gamma^2_{\mathrm{Frob}}\|\dot x\|
\end{aligned}$$

$$\leqslant 5\gamma_{\mathrm{Frob}}^2\|\dot{x}\|, \qquad \text{because } k \geqslant 2.$$

As a function on $\mathbb{P}^n$, $\mathrm{d}_x\gamma_{\mathrm{Frob}}(f,-)(\dot{x}) = \dot{\gamma}_{\mathrm{Frob}}$, so that

$$|\mathrm{d}_x\gamma_{\mathrm{Frob}}(f,-)(\dot{x})| \leqslant 5\gamma_{\mathrm{Frob}}(f,x)^2\|\dot{x}\|.$$

Since the inequality holds for any differentiable path, covering all possible values of $\dot{x}$, it follows that $\|\mathrm{d}_x\gamma_{\mathrm{Frob}}(f,-)\| \leqslant 5\gamma_{\mathrm{Frob}}(f,x)^2$. Consequently, $\|\mathrm{d}_x\gamma_{\mathrm{Frob}}(f,-)^{-1}\| \leqslant 5$ and the claim follows. □

4.3. **Implementation details, complexity.**

4.3.1. *Computational model.* We use the Blum–Shub–Smale model (Blum et al. 1989) extended with a "6th type of node", as did Shub and Smale (1996). Unlike Shub and Smale, we will apply it to univariate (or rather homogeneous bivariate) polynomials only. A node of this type has the following behavior. If it is given as input a homogeneous polynomial $f \in \mathbb{C}[x,y]$ and an approximate zero $z \in \mathbb{P}^1$ of f, with associated zero ζ, it outputs ζ. In any other case, it fails. There is no need to specify how it fails because we will make sure that this will not happen.

From the practical point of view, given a point z which approximates a zero ζ of a homogeneous polynomial $f \in \mathbb{C}[x,y]$, one can refine the approximation to obtain $d_{\mathbb{P}}(z,\zeta) \leqslant \varepsilon$ in $\log_2\log_2\frac{\pi}{\varepsilon}$ Newton's iterations. For most practical purpose, this looks like infinite precision. In that sense, the 6th type of node does not add much power.

Do we really need a 6th type of node? The continuation method proposed here uses a start system defined in terms of the zeros of some homogeneous bivariate polynomials. Naturally, the algorithm would also work with approximate zeros only. However, if we do it this way, then the distribution of the start system is not easily described, it is only close to a nice distribution. I showed (Lairez 2017) how to deal with the complexity analysis in an analogue situation but it is too technical an argument for the little value it adds.

Interval arithmetic gives another way to remove the need for this extra type of node: wherever an exact zero is expected, we use bounding boxes instead and perform the subsequent operations with interval arithmetic. If the precision happens to be insufficient, we refine the bounding boxes with Newton's iteration and start over the computation. The convergence of Newton's iteration is so fast that even with naive estimations of the numerical stability, the number of start over will be moderate. However, this is no less technical to formalize.

For convenience, we will also assume the ability to compute fractional powers of a positive real number at unit cost. This will allow us to compute Hermitian norms and the numbers γ_{Frob} and $\hat{\gamma}_{\mathrm{Frob}}$ exactly.

4.3.2. *Continuation paths in the unitary group.* While geodesics in $\mathcal{U}$ are a natural choice for continuation paths, see §3.4, they are not easy to compute in the BSS model. We can describe more elementary continuation paths using Householder's reflections. For any 1-dimensional subspace $l \subset \mathbb{C}^{n+1}$ and any $\theta \in \mathbb{R}$, let $R(l,\theta) \in U(n+1)$ be the unique map such that $R(l,\theta)|_l = e^{i\theta}\,\mathrm{id}_l$ and $R(l,\theta)|_{l^\perp} = \mathrm{id}_{l^\perp}$. Note that for the angle π, $R(l,\pi)$ is a reflection. The computation of a matrix multiplication $AR(l,\theta)$, for any $A \in U(n+1)$, requires only $O(n^2)$ operations because $R(l,\theta) - \mathrm{id}$ has rank 1 and can be written vv^* for some $v \in \mathbb{C}^{(n+1)\times 1}$.

Given a unitary matrix $v \in U(n+1)$, the procedure of Householder (1958) (with the necessary changes in the complex case) decomposes v as $v = e^{i\alpha}R(l_1,\pi)\cdots R(l_n,\pi)$,

for some $\alpha \in [-\pi, \pi]$, with $O(n^3)$ operations. One can define the path

$$w_t \doteq e^{i\frac{t}{\tau}\alpha} R(l_1, \tfrac{t}{\tau}\pi) \cdots R(l_n, \tfrac{t}{\tau}\pi),$$

where $\tau^2 \doteq \frac{1}{2}\left(\alpha^2 + n\pi^2\right) \leqslant \frac{n+1}{2}\pi^2$. Given $\mathbf{u}, \mathbf{v} \in \mathcal{U}$, we define a 1-Lipshitz continuous path $(\mathbf{w}_t)_{t\geqslant 0}$ from $\mathbf{1}_{\mathcal{U}}$ to $\mathbf{v}^{-1}\mathbf{u}$, component-wise with the method above. It reaches $\mathbf{v}^{-1}\mathbf{u}$ at $t = \sqrt{\frac{n(n+1)}{2}}\pi < 4n$. To compute $\mathbf{w}_t$ on a BSS machine, we can replace the trigonometric functions with any other functions parametrizing the circle. This construction satisfies the three conditions (P1)–(P3) from §3.4.

For a given t, the cost of computing $\mathbf{w}_t$ is dominated by the the cost of multiplying by $R(l, \theta)$ matrices: there are n such multiplications for each of the n components of $\mathbf{w}_t$, that is $O(n^4)$ operations.

4.3.3. *Computation of $\hat{\gamma}_{\mathrm{Frob}}$, cost of a continuation step.* The reason for introducing $\hat{\gamma}_{\mathrm{Frob}}$ is that we can compute it with low complexity. By contrast, computing $\hat{\gamma}$ is NP-hard because it involves the computation of the spectral norm of symmetric multilinear maps, and there is no polynomial-time approximation scheme, unless P = NP (Hillar and Lim 2013, Theorem 10.2). Beware though, a too naive algorithm for computing γ_{Frob} requires $\Omega(N^2)$ operations, where N is the input size (see §1.3). Given a polynomial $f \in H_d$ and a point $z \in \mathbb{C}^{n+1}$, the evaluation $f(z)$ and the vector $\mathrm{d}_z f$ can be computed with $O(\dim H_d)$ operations (Bürgisser and Cucker 2013, Lemma 16.32; Baur and Strassen 1983).

Proposition 32. *Given a homogeneous polynomial $f \in H_d$ and $z \in \mathbb{P}^n$, one can compute $\gamma_{\mathrm{Frob}}(f, z)$ with $O(nd^2 \dim H_d)$ operations, as $\dim H_d \to \infty$.*

Proof. We can compute $\|\mathrm{d}_z f\|^{-1}$ in $O(\dim H_d)$ operations, so the main task is to compute $\left\|\frac{1}{k!}\mathrm{d}_z^k f\right\|_{\mathrm{Frob}}$ for all $2 \leqslant k \leqslant d$, see (46). Let g be the shifted polynomial $g : x \in \mathbb{C}^{n+1} \mapsto f(z+x)$, so that $\mathrm{d}_0^k g = \mathrm{d}_z^k f$. Let g_k denote the homogeneous component of degree k of g. According to Lemma 30, $\|\frac{1}{k!}\mathrm{d}_0^k g\|_{\mathrm{Frob}} = \|g_k\|_W$.

Let S denote the *size* of g, that is the number of coefficients in a dense nonhomogeneous polynomial of degree d in $n+1$ variables. He have $S = \binom{n+d+1}{d} \leqslant d \dim H_d$. In view of (45), the computation of $\|g_k\|_W^2$ reduces to the computation of the coefficients of g_k in the monomial basis and the multinomial coefficients $\frac{i_0!\cdots i_n!}{(i_0+\cdots+i_n)!}$. We can compute them all with $O(S)$ operations thanks to the recurrence relation

$$\frac{i_0!\cdots i_{n-1}!(i_n+1)!}{(i_0+\cdots+i_n+1)!} = \frac{i_n+1}{i_0+\cdots+i_n+1}\frac{i_0!\cdots i_n!}{(i_0+\cdots+i_n)!}.$$

Therefore, we can compute $\gamma_{\mathrm{Frob}}(f, z)$ in $O(S)$ operations given the coefficients of g in the monomial basis.

To compute g, we shift the variables one after the other. To compute the first shift $f(x_0 + z_0, x_1, \ldots, x_n)$ (subsequent shifts are identical), we write

$$f = \sum_{i_1+\cdots+i_n\leqslant d} p_{i_1,\ldots,i_n}(x_0) x_1^{i_1} \cdots x_n^{i_n},$$

where the $p_{i_1,\ldots,i_n}(x_0)$ are polynomials of degree at most d. There are at most $\dim H_d$ of them (the number of monomials of degree at most d in n variables). Moreover, computing them requires only $O(S)$ copy operations and no arithmetic operation: their coefficients in the monomial basis are directly read from the coefficients of f in the monomial basis. One can compute $p_{i_1,\ldots,i_n}(x_0 + z_0)$ with $O(d^2)$ operations with a naive algorithm. Note that we can do this with only $d^{1+o(1)}$ operations using fast evaluation and interpolation algorithms (Bostan et al. 2017; Gathen

and Gerhard 1999). All together, this is $O(d^2 \dim H_d)$ operations. We recover $f(x_0+z_0, x_1, \dots, x_n)$ in $O(S)$ operations from the $p_{i_1,\dots,i_n}(x_0+z_0)$ (and no arithmetic operation). We repeat this shift operation for each one of the $n+1$ variables and this gives the claim. □

Corollary 33. *Given $F \in \mathcal{H}[n]$, $\mathbf{u} \in \mathcal{U}$ and $z \in \mathbb{P}^n$, one can compute $\kappa(\mathbf{u} \cdot F, z)$ and $\hat{\gamma}_{\mathrm{Frob}}(\mathbf{u} \cdot F, z)$ within a factor 2 in $O(nD^2N)$ operations as $N \to \infty$.*

Proof. The evaluation of $\mathbf{u} \cdot F$ at a point $z \in \mathbb{C}^{n+1}$ can be computed as

$$(f_1(u_1^* z), \dots, f_n(u_n^* z))$$

with $O(n^3)$ operations to compute the vectors $u_i^* z$ and $\sum_i O(\dim H_{d_i}) = O(N)$ additional operations to evaluate the polynomials f_i. Therefore, the matrix $\mathrm{d}_z(\mathbf{u} \cdot F)$ be computed with $O(N + n^3)$ operations too (Baur and Strassen 1983). The computation of $L(\mathbf{u} \cdot F, z)$ requires $O(n^2)$ additional operations, following (15). Then, we can compute κ, that is the inverse of the least singular value of $L(\mathbf{u}, z)$, within a factor of 2 in $O(n^3)$ operations, using a tridiagonalization with Householder's reflections and a result by Kahan (1966). To compute $\hat{\gamma}_{\mathrm{Frob}}$, it only remains to compute the $\gamma_{\mathrm{Frob}}(u_i \cdot f, z) = \gamma_{\mathrm{Frob}}(f, u_i^* z)$, for $1 \leqslant i \leqslant n$, which requires all together $O(nD^2N + n^3)$ operations, by Proposition 32. The term n^3 is $O(nN)$, so this gives the claim. □

Since κ is defined as the operator norm of some matrix, the exact computation is difficult in the BSS model. One could use the Frobenius norm instead but computing κ within a factor 2 is satisfactory, see Remark 24.

Corollary 34. *In Algorithm 2 with $g = \hat{\gamma}_{\mathrm{Frob}}$, one continuation step can be performed in $O(n^2D^2N)$ operations.*

Proof. A step boils down to: one evaluation of $\hat{\gamma}_{\mathrm{Frob}}$ and κ, that is $O(nD^2N)$ operations by Corollary 33; one evaluation of the continuation path, that is $O(n^4) = O(n^2N^2)$ operations, see §4.3.2; and one Newton's iteration, that is $O(n^3 + N) = O(nN)$ operations (Bürgisser and Cucker 2013, Proposition 16.32). □

4.4. **Gaussian random systems.** We conclude with the study of the average complexity of $\mathrm{Solve}(F, \mathbf{1}_{\mathcal{U}})$ when F is a *Kostlan random system*, that is a standard normal variable in $\mathcal{H}[n]$ endowed with Weyl's norm. In view of Theorem 27, it only remains to study the average value of $\gamma_{\mathrm{Frob}}(f, \zeta)$ when f is a Kostlan random polynomial and ζ a uniformly distributed zero of it. To this purpose, the main tool is a corollary of a result by Beltrán and Pardo (2011).

Proposition 35. *Let $f \in H_d$ be a Kostlan random polynomial, let ζ be a random uniformly distributed point in $\{z \in \mathbb{P}^n \mid f(z) = 0\}$ and let $\bar{\zeta} \in \mathbb{C}^{n+1}$ be a random uniformly distributed vector such that $\|\bar{\zeta}\| = 1$ and $[\bar{\zeta}] = \zeta$. Then*

(i) $\frac{1}{\sqrt{d}} \mathrm{d}_\zeta f$ is a standard normal variable in $(\mathbb{C}^{n+1})^*$; and
(ii) given ζ, the orthogonal projection of f on $\{g \in H_d \mid g(\zeta) = 0 \text{ and } \mathrm{d}_\zeta g = 0\}$ is a standard normal variable and is independent from $\mathrm{d}_\zeta f$.

Proof. Let $\lambda_2, \dots, \lambda_n \in (\mathbb{C}^{n+1})^*$ be independent Gaussian linear forms, so that $F \doteq (f, \lambda_2, \dots, \lambda_n)$ is a Kostlan random system of degree $(d, 1, \dots, 1)$. Let $\eta \in \mathbb{P}^n$ be a uniformly distributed random zero of F.

The zero set L of $\lambda_2, \dots, \lambda_n$ is a uniformly distributed random line independent from f and η is uniformly distributed in $V(f) \cap L$. By Corollary 9, η is uniformly distributed in $V(f)$, so we may assume that $\zeta = \eta$.

Let G the orthogonal projection of F on the subspace

$$R_\zeta \doteq \{G \in \mathcal{H} \mid G(\zeta) = 0 \text{ and } \mathrm{d}_\zeta G = 0\}.$$

Beltrán and Pardo (2011, Theorem 7), and Bürgisser and Cucker (2013, Prop. 17.21) for the Gaussian case, proved that: the matrix $\mathrm{diag}(d^{-\frac{1}{2}}, 1, \dots, 1)\mathrm{d}_{\bar\zeta}F$ is a standard normal variable in $\mathbb{C}^{n\times(n+1)}$; and conditionally on ζ, G is a standard normal variable in R_ζ and is independent from $\mathrm{d}_\zeta F$. The claim follows by considering the first row of $\mathrm{d}_{\bar\zeta}F$, that is $\mathrm{d}_{\bar\zeta}f$, and the first coordinate of G. □

The following three lemmas deal with the average analysis of γ_{Frob}.

Lemma 36. *Let $f \in H_d$ and $\zeta = [1 : 0 : \cdots : 0]$. We write*

$$f = \sum_{i=0}^{d} x_0^{d-i} g_i(x_1, \dots, x_n),$$

for some uniquely determined homogeneous polynomials $g_0, \dots, g_d$ of degrees $0, \dots, d$ respectively. For any $k \geqslant 2$,

$$\left\| \tfrac{1}{k!}\mathrm{d}_\zeta^k f \right\|_{\mathrm{Frob}}^2 = \binom{d}{k} \sum_{l=0}^{k} \binom{d-l}{k-l} \left\| x_0^{d-l} g_l \right\|_W^2.$$

Proof. Let $\tilde f \doteq f(x_0 + 1, x_1, \dots, x_n)$ and let $\tilde f_{(k)}$ be the homogeneous component of degree k of $\tilde f$. We compute that

$$\tilde f_{(k)} = \sum_{l=0}^{d} \left[(x_0+1)^{d-l} g_l\right]_{(k)} = \sum_{l=0}^{k} \binom{d-l}{k-l} x_0^{k-l} g_l.$$

The terms of the sum are orthogonal for Weyl's inner product, and moreover $\left\| \frac{1}{k!}\mathrm{d}_\zeta^k f \right\|_{\mathrm{Frob}} = \left\| \tilde f_{(k)} \right\|_W$ by Lemma 30, therefore

$$\left\| \tfrac{1}{k!}\mathrm{d}_\zeta^k f \right\|_{\mathrm{Frob}}^2 = \sum_{l=0}^{k} \binom{d-l}{k-l}^2 \left\| x_0^{k-l} g_l \right\|_W^2.$$

Looking closely at the definition of Weyl's inner product reveals that

$$\left\| x_0^{k-l} g_l \right\|_W^2 = \frac{\binom{d}{d-l}}{\binom{k}{k-l}} \left\| x_0^{d-l} g_l \right\|_W^2 = \frac{\binom{d}{k}}{\binom{d-l}{k-l}} \left\| x_0^{d-l} g_l \right\|_W^2,$$

and the claim follows. □

Lemma 37. *Let $f \in H_d$ be a Kostlan random polynomial and $\zeta \in \mathbb{P}^n$ be a uniformly distributed zero of f. For any $k \geqslant 2$,*

$$\mathbb{E}\left[\|\mathrm{d}_\zeta f\|^{-2} \left\| \tfrac{1}{k!}\mathrm{d}_\zeta^k f_i \right\|_{\mathrm{Frob}}^2 \right] \leqslant \frac{1}{nd} \binom{d}{k} \binom{d+n}{k} \leqslant \left(\tfrac{1}{4} d^2 (d+n)\right)^{k-1}.$$

Proof. We can choose (random) coordinates such that $\zeta = [1 : 0 : \cdots : 0]$. Let $g_0, \dots, g_d$ be the polynomials defined in Lemma 36. To begin with, we describe the distribution of the random polynomials $g_0, \dots, g_d$. The polynomial g_0 (of degree 0) is simply zero because ζ is a zero of f. The second one g_1 has degree 1, and as a linear form, it is equal to $\mathrm{d}_\zeta f$. According to Proposition 35(i), it is a standard normal variable after multiplication by $d^{-\frac{1}{2}}$. In particular $d\|g_1\|_W^{-2}$ is the inverse of a χ^2-distributed variable with $2n+2$ degrees of freedom, therefore

$$\mathbb{E}\left[\|g_1\|_W^{-2}\right] = \frac{1}{2nd}.$$

For $l \geqslant 2$, the polynomial $x_0^{d-l} g_l$ is the orthogonal projection of f on the subspace of all polynomials of the form $x_0^{d-l} p(x_1, \dots, x_n)$, a which is a subspace of

$$\{g \in H_d \mid g(\zeta) = 0 \text{ and } \mathrm{d}_\zeta g = 0\} \subset H_d.$$

Thus, by Proposition 35(ii), $x_0^{d-l} g_l$ is a standard normal variable in the appropriate subspace and is independent from g_1. In particular, for any $l \geqslant 2$,

$$\mathbb{E}\left[\left\|x_0^{d-l} g_l\right\|_W^2\right] = \dim_{\mathbb{R}} \left\{x_0^{d-l} p(x_1, \dots, x_n) \in H_d\right\} = 2\binom{n-1+l}{l}.$$

Note also that $\|g_1\|_W^{-2} \left\|x_0^{d-l} g_1\right\|_W^2 = \frac{1}{d}$. It follows, by Lemma 36, that

$$\begin{aligned}
\mathbb{E}\left[\|\mathrm{d}_\zeta f\|^{-2} \left\|\tfrac{1}{k!}\mathrm{d}_\zeta^k f\right\|_{\mathrm{Frob}}^2\right] &= \sum_{l=0}^{k} \mathbb{E}\left[\binom{d}{k} \sum_{l=0}^{k} \binom{d-l}{k-l} \|g_1\|_W^{-2} \left\|x_0^{d-l} g_l\right\|_W^2\right] \\
&\leqslant \binom{d}{k}\left(\binom{d-1}{k-1}\frac{1}{d} + \sum_{l=2}^{k}\binom{d-l}{k-l}\mathbb{E}\left[\|g_1\|_W^{-2}\right]\mathbb{E}\left[\left\|x_0^{d-l} g_l\right\|_W^2\right]\right) \\
&\leqslant \binom{d}{k}\left(\binom{d-1}{k-1}\frac{1}{d} + \frac{1}{nd}\sum_{l=2}^{k}\binom{d-l}{k-l}\binom{n-1+l}{l}\right) \\
&\leqslant \frac{1}{nd}\binom{d}{k}\sum_{l=0}^{k}\binom{d-l}{k-l}\binom{n-1+l}{l} \\
&= \frac{1}{nd}\binom{d}{k}\binom{d+n}{k}.
\end{aligned}$$

To check the binomial identity $\sum_{l=0}^{k}\binom{d-l}{k-l}\binom{n-1+l}{l} = \binom{d+n}{k}$, we remark that $\binom{d-l}{k-l}$ counts the number of monomials of degree $k-l$ in $d-k+1$ variables while $\binom{n+l}{l}$ counts the number of monomials of degree l in $n+1$ variables. Therefore, the sum over l counts the number of monomials of degree k in $(d-k+1)+(n+1)$ variables, that is $\binom{d+n}{k}$.

Concerning the second inequality, the maximum value of $\left(\frac{1}{nd}\binom{d}{k}\binom{d+n}{k}\right)^{\frac{1}{k-1}}$, with $k \geqslant 2$, is reached for $k = 2$. That is, for any $k \geqslant 2$,

$$\begin{aligned}
\frac{1}{nd}\binom{d}{k}\binom{d+n}{k} &\leqslant \left(\frac{1}{nd}\binom{d}{2}\binom{d+n}{2}\right)^{k-1} \\
&\leqslant \left(\tfrac{1}{4}(d-1)(d+n)\left(\frac{d-1}{n}+1\right)\right)^{\frac{1}{k-1}},
\end{aligned}$$

which leads to the claim. □

Lemma 38. *With the same notations as Lemma 37,*

$$\mathbb{E}\left[\gamma_{\mathrm{Frob}}(f,\zeta)^2\right] \leqslant \tfrac{1}{4}d^3(d+n).$$

Proof. We bound the supremum in the definition of γ_{Frob} by a sum:

$$\begin{aligned}
\mathbb{E}\left[\gamma_{\mathrm{Frob}}(f,\zeta)^2\right] &\leqslant \sum_{k=2}^{d}\mathbb{E}\left[\left(\|\mathrm{d}_\zeta f\|^{-1}\left\|\tfrac{1}{k!}\mathrm{d}_\zeta^k f_i\right\|_{\mathrm{Frob}}\right)^{\frac{2}{k-1}}\right] \\
&\leqslant \sum_{k=2}^{d}\mathbb{E}\left[\|\mathrm{d}_\zeta f\|^{-2}\left\|\tfrac{1}{k!}\mathrm{d}_\zeta^k f_i\right\|_{\mathrm{Frob}}^2\right]^{\frac{1}{k-1}}, \text{ by Jensen's inequality,} \\
&\leqslant \sum_{k=2}^{d}\tfrac{1}{4}d^2(d+n), \text{ by Lemma 37,}
\end{aligned}$$

$$\leqslant \tfrac{1}{4} d^3(d+n),$$

and this is the claim. □

Proposition 39. *If $F \in \mathcal{H}[n]$ is a Kostlan random system, then* Sample(F) *(Algorithm 1) outputs a ρ_{std}-distributed $(\mathbf{u}, \zeta) \in \mathcal{V}_F$ with $O(n^3)$ samplings of the normal distribution on $\mathbb{R}$ and $O(n^4 + nD^4)$ operations on average.*

Proof. By Proposition 10, Sample(F) requires $O(n^3)$ samplings, $O(n^4)$ operations and n times root-finding of bivariate homogeneous polynomials of degree at most D. These polynomials are the restrictions of the polynomials $f_1, \dots, f_n$ on independent uniformly distributed random projective line $L_1, \dots, L_n$ respectively. Choosing random coordinates so that L_i is spanned by $[1:0:\cdots]$ and $[0:1:0:\cdots]$, the restriction of f_i to L_i is just $f_i(x_0, x_1, 0, \dots, 0)$. So the restriction map $f \in H_{d_i} \to f_{|L_i}$ is an orthogonal projector (or equivalently, the adjoint map $f \in \mathbb{C}[x_0, x_1]_d \to H_d$ is an isometric embedding, which follows from the definition (45) of Weyl's norm). In particular, $f_i|_{L_i}$ is a Kostlan random polynomial in $\mathbb{C}[x_0, x_1]_d$, and the algorithm of Beltrán and Pardo (2011) computes an approximate root of it with $O(D^4)$ operations on average. The 6th type of node recover the exact root. □

Theorem 40 (Main result)**.** *If $F \in \mathcal{H}[n]$ is a Kostlan random system, then* Solve$(F, \mathbf{1}_{\mathcal{U}})$ *(with $g(\mathbf{u}, x) \doteq \hat{\gamma}_{\mathrm{Frob}}(\mathbf{u} \cdot F, x)$ for the continuation) outputs an approximate zero of F almost surely with $O(n^4 D^2)$ continuation steps on average and $O(n^6 D^4 N)$ operations on average, when $N \to \infty$. When $\min(n, D) \to \infty$, this is $N^{1+o(1)}$.*

Proof. We first note that the conditions (H1')–(H3') are statisfied for $\hat{\gamma}_{\mathrm{Frob}}$, see §4.2.2, and that the probability distribution of F is unitary invariant (because $\mathcal{U}$ acts isometrically on $\mathcal{H}[n]$), therefore Theorem 27 applies.

For $1 \leqslant i \leqslant n$, let $\zeta_i \in \mathbb{P}^n$ be a uniformly distributed random zero of f_i. Concerning the number of continuation steps K, Theorem 27 gives (with $C' = 5$ given by Lemma 31)

$$\begin{aligned}
\mathbb{E}[K] &\leqslant 9000\, n^3 \left(\sum_{i=1}^{n} \mathbb{E}\left[\gamma_{\mathrm{Frob}}(f_i, \zeta_i)^2\right] \right)^{\frac{1}{2}} \\
&\leqslant 9000\, n^3 \left(\sum_{i=1}^{n} \tfrac{1}{4} d_i^3 (d_i + n) \right)^{\frac{1}{2}}, && \text{by Lemma 38,} \\
&\leqslant 9000\, n^3 \left(\tfrac{1}{4} n D^3 (D+n) \right)^{\frac{1}{2}} \\
&\leqslant 9000\, n^4 D^2, && \text{with } D + n \leqslant 2Dn.
\end{aligned}$$

Concerning the total number of operations, the cost of Algorithm "Solve" splits into the cost of the sampling and the cost of the numerical continuation. The former is $O(n^4 + nD^4)$ on average, by Proposition 39. As for the latter, the cost of a step is $O(n^2 D^2 N)$ (Corollary 34), and we need $O(n^4 D^2)$ continuation steps on average; as $N \to \infty$, that is $O(n^6 D^4 N)$ operations, by Corollary 34. When $\min(n, D) \to \infty$, then both n and D are $\binom{n+D}{n}^{o(1)} = N^{o(1)}$. □

References

D. Armentano, C. Beltrán, P. Bürgisser, F. Cucker, and M. Shub (2016). "Condition Length and Complexity for the Solution of Polynomial Systems". In: *Foundations of Computational Mathematics*. DOI: 10.1007/s10208-016-9309-9.

– (2018). “A Stable, Polynomial-Time Algorithm for the Eigenpair Problem”. In: *Journal of the European Mathematical Society* 20.6, pp. 1375–1437. DOI: 10.4171/JEMS/789.

D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler (2013). *Numerically Solving Polynomial Systems with Bertini*. Vol. 25. Software, Environments, and Tools. SIAM, Philadelphia, PA.

W. Baur and V. Strassen (1983). “The Complexity of Partial Derivatives”. In: *Theoretical Computer Science* 22.3, pp. 317–330. DOI: 10.1016/0304-3975(83)90110-X.

C. Beltrán (2011). “A Continuation Method to Solve Polynomial Systems and Its Complexity”. In: *Numerische Mathematik* 117.1, pp. 89–113. DOI: 10.1007/s00211-010-0334-3.

C. Beltrán and L. M. Pardo (2008). “On Smale’s 17th Problem: A Probabilistic Positive Solution”. In: *Foundations of Computational Mathematics* 8.1, pp. 1–43. DOI: 10.1007/s10208-005-0211-0.

– (2009a). “Efficient Polynomial System-Solving by Numerical Methods”. In: *Journal of Fixed Point Theory and Applications* 6.1, pp. 63–85. DOI: 10.1007/s11784-009-0113-x.

– (2009b). “Smale’s 17th Problem: Average Polynomial Time to Compute Affine and Projective Solutions”. In: *Journal of the American Mathematical Society* 22.2, pp. 363–385. DOI: 10.1090/S0894-0347-08-00630-9.

– (2011). “Fast Linear Homotopy to Find Approximate Zeros of Polynomial Systems”. In: *Foundations of Computational Mathematics* 11.1, pp. 95–129. DOI: 10.1007/s10208-010-9078-9.

C. Beltrán and M. Shub (2009). “Complexity of Bezout’s Theorem. VII. Distance Estimates in the Condition Metric”. In: *Foundations of Computational Mathematics* 9.2, pp. 179–195. DOI: 10.1007/s10208-007-9018-5.

L. Blum, M. Shub, and S. Smale (1989). “On a Theory of Computation and Complexity over the Real Numbers: NP-Completeness, Recursive Functions and Universal Machines”. In: *Bulletin of the American Mathematical Society*. N.S. 21.1, pp. 1–46. DOI: 10.1090/S0273-0979-1989-15750-9.

A. Bostan et al. (2017). *Algorithmes Efficaces En Calcul Formel*. 1st ed. Palaiseau: Frédéric Chyzak (self-pub.)

P. Breiding and N. Vannieuwenhoven (2018). “The Condition Number of Join Decompositions”. In: *SIAM Journal on Matrix Analysis and Applications* 39.1, pp. 287–309. DOI: 10.1137/17M1142880.

I. Briquel, F. Cucker, J. Peña, and V. Roshchina (2014). “Fast Computation of Zeros of Polynomial Systems with Bounded Degree under Finite-Precision”. In: *Mathematics of Computation* 83.287, pp. 1279–1317. DOI: 10.1090/S0025-5718-2013-02765-2.

P. Bürgisser and F. Cucker (2011). “On a Problem Posed by Steve Smale”. In: *Annals of Mathematics. Second Series* 174.3, pp. 1785–1836. DOI: 10.4007/annals.2011.174.3.8.

– (2013). *Condition: The Geometry of Numerical Algorithms*. Vol. 349. Grundlehren Der Mathematischen Wissenschaften. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-38896-5.

P. Bürgisser and A. Lerario (2018). “Probabilistic Schubert Calculus”. In: *Journal für die reine und angewandte Mathematik (Crelles Journal)*. DOI: 10.1515/crelle-2018-0009.

J.-P. Dedieu (2006). *Points Fixes, Zéros et La Méthode de Newton*. Vol. 54. Mathématiques & Applications. Springer. DOI: 10.1007/3-540-37660-7.

J.-P. Dedieu, G. Malajovich, and M. Shub (2013). "Adaptive Step-Size Selection for Homotopy Methods to Solve Polynomial Equations". In: *IMA Journal of Numerical Analysis* 33.1, pp. 1–29. DOI: 10.1093/imanum/drs007.

J. W. Demmel (1988). "The Probability That a Numerical Analysis Problem Is Difficult". In: *Mathematics of Computation* 50.182, pp. 449–480. DOI: 10.1090/S0025-5718-1988-0929546-7.

A. Edelman (1989). "Eigenvalues and Condition Numbers of Random Matrices". USA: Massachusetts Institute of Technology.

H. Federer (1959). "Curvature Measures". In: *Transactions of the American Mathematical Society* 93.3, pp. 418–491.

J. von zur Gathen and J. Gerhard (1999). *Modern Computer Algebra*. New York: Cambridge University Press.

J. D. Hauenstein and A. C. Liddell (2016). "Certified Predictor–Corrector Tracking for Newton Homotopies". In: *Journal of Symbolic Computation* 74, pp. 239–254. DOI: 10.1016/j.jsc.2015.07.001.

J. D. Hauenstein and F. Sottile (2012). "Algorithm 921: alphaCertified: Certifying Solutions to Polynomial Systems". In: *ACM Transactions on Mathematical Software* 38.4, pp. 1–20. DOI: 10.1145/2331130.2331136.

C. J. Hillar and L.-H. Lim (2013). "Most Tensor Problems Are NP-Hard". In: *Journal of the ACM* 60.6, pp. 1–39. DOI: 10.1145/2512329.

A. S. Householder (1958). "Unitary Triangularization of a Nonsymmetric Matrix". In: *Journal of the ACM* 5.4, pp. 339–342. DOI: 10.1145/320941.320947.

R. Howard (1993). "The Kinematic Formula in Riemannian Homogeneous Spaces". In: *Memoirs of the American Mathematical Society* 106.509. DOI: 10.1090/memo/0509.

W. Kahan (1966). *Accurate Eigenvalues of a Symmetric Tri-Diagonal Matrix*. CS41. Stanford University.

P. Lairez (2017). "A Deterministic Algorithm to Compute Approximate Roots of Polynomial Systems in Polynomial Average Time". In: *Foundations of Computational Mathematics*. DOI: 10.1007/s10208-016-9319-7.

G. Malajovich (1994). "On Generalized Newton Algorithms: Quadratic Convergence, Path-Following and Error Analysis". In: *Theoretical Computer Science* 133.1, pp. 65–84. DOI: 10.1016/0304-3975(94)00065-4.

– (2018). "Complexity of Sparse Polynomial Solving: Homotopy on Toric Varieties and the Condition Metric". In: *Foundations of Computational Mathematics*. DOI: 10.1007/s10208-018-9375-2.

J. Renegar (1987). "On the Efficiency of Newton's Method in Approximating All Zeros of a System of Complex Polynomials". In: *Mathematics of Operations Research* 12.1, pp. 121–148. DOI: 10.2307/3689676.

– (1989). "On the Worst-Case Arithmetic Complexity of Approximating Zeros of Systems of Polynomials". In: *SIAM Journal on Computing* 18.2, pp. 350–370. DOI: 10.1137/0218024.

M. Shub (1989). "On the Distance to the Zero Set of a Homogeneous Polynomial". In: *Journal of Complexity* 5.3, pp. 303–305. DOI: 10.1016/0885-064X(89)90027-7.

– (1993). “Some Remarks on Bezout’s Theorem and Complexity Theory”. In: *From Topology to Computation: Proceedings of the Smalefest*. Springer, New York, pp. 443–455.
– (2009). “Complexity of Bezout’s Theorem. VI. Geodesics in the Condition (Number) Metric”. In: *Foundations of Computational Mathematics* 9.2, pp. 171–178. DOI: 10.1007/s10208-007-9017-6.
M. Shub and S. Smale (1993a). “Complexity of Bézout’s Theorem. I. Geometric Aspects”. In: *Journal of the American Mathematical Society* 6.2, pp. 459–501. DOI: 10.2307/2152805.
– (1993b). “Complexity of Bezout’s Theorem. II. Volumes and Probabilities”. In: *Computational Algebraic Geometry (Nice, 1992)*. Vol. 109. Progr. Math. Boston: Birkhäuser, pp. 267–285.
– (1993c). “Complexity of Bezout’s Theorem. III. Condition Number and Packing”. In: *Journal of Complexity* 9.1, pp. 4–14. DOI: 10.1006/jcom.1993.1002.
– (1994). “Complexity of Bezout’s Theorem. V. Polynomial Time”. In: *Theoretical Computer Science* 133.1. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993), pp. 141–164. DOI: 10.1016/0304-3975(94)90122-8.
– (1996). “Complexity of Bezout’s Theorem. IV. Probability of Success; Extensions”. In: *SIAM Journal on Numerical Analysis* 33.1, pp. 128–148. DOI: 10.1137/0733008.
S. Smale (1985). “On the Efficiency of Algorithms of Analysis”. In: *Bulletin of The American Mathematical Society, New Series* 13.2, pp. 87–121. DOI: 10.1090/S0273-0979-1985-15391-1.
– (1986). “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, pp. 185–196.
– (1998). “Mathematical Problems for the next Century”. In: *The Mathematical Intelligencer* 20.2, pp. 7–15. DOI: 10.1007/BF03025291.

Inria Saclay Île-de-France, 91120 Palaiseau, France