

Quantitative Singularity Theory for Random Polynomials

Paul Breiding¹, Hanieh Keneshlou^{2,*} and Antonio Lerario³

¹Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, ²Max-Planck Institute for Mathematics in Sciences, Inselstr. 22, 04103 Leipzig, Germany, and ³Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, 34136 Trieste, Italy

*Correspondence to be sent to: e-mail: keneshlo@mis.mpg.de

Motivated by Hilbert's 16th problem we discuss the probabilities of topological features of a system of random homogeneous polynomials. The distribution for the polynomials is the Kostlan distribution. The topological features we consider are *type- W singular loci*. This is a term that we introduce and that is defined by a list of equalities and inequalities on the derivatives of the polynomials. In technical terms a *type- W singular locus* is the set of points where the jet of the Kostlan polynomials belongs to a semialgebraic subset W of the jet space, which we require to be invariant under orthogonal change of variables. For instance, the zero set of polynomial functions or the set of critical points fall under this definition. We will show that, with overwhelming probability, the *type- W singular locus* of a Kostlan polynomial is ambient isotopic to that of a polynomial of lower degree. As a crucial result, this implies that complicated topological configurations are rare. Our results extend earlier results from Diatta and Lerario who considered the special case of the zero set of a single polynomial. Furthermore, for a given polynomial function p we provide a deterministic bound for the radius of the ball in the space of differentiable functions with center p , in which the W -singularity structure is constant.

Received October 23, 2019; Revised August 10, 2020; Accepted September 3, 2020
Communicated by Prof. Misha Sodin

1 Introduction

Hilbert's 16th problem was posed by David Hilbert at the Paris ICM in 1900 and, in its general form, it asks for the study of the maximal number and the possible arrangement of the components of a generic real algebraic hypersurface of degree d in real projective space. Since Hilbert had posed his question, many mathematicians have contributed to the subject: Hilbert [14], Rohn [34], Petrovsky [33], Rokhlin [35], Gudkov [9], Nikulin [29], Viro [42–44], Kharlamov [16–18], and more.

Hilbert's problem not only concerns the topology of the hypersurface but also the way it is embedded inside the projective space. The difference between these two sides of the problem can be illustrated by considering the sextic polynomials $P_1(x_0, x_1, x_2) = (x_1^4 + x_2^4 - x_0^4)(x_1^2 + x_2^2 - 2x_0^2) + x_1^5 x_2$ and $P_2(x_0, x_1, x_2) = 10(x_1^4 + x_2^4 - x_0^4)(x_1^2 + x_2^2 - 2x_0^2) + x_1^5 x_2$ and their zero sets, which are shown in Figure 1. Both of them have two connected components and so their topological types agree. But their rigid isotopy types are different, because one cannot move, in the projective plane, the inner oval on the left picture outside without crossing the outer oval. Being rigidly isotopic means that P_1 and P_2 belong to the same connected component of $\mathbb{R}[x_0, x_1, x_2]_{(6)} \setminus \Sigma$, where Σ denotes the set of singular curves.

In this article we approach this classical topic from a probabilistic point of view. That is, we assume a probability distribution on the space of polynomials and consider *statistical* properties of topological configurations. Moreover, we do not only consider the topology of zero sets. In fact, the case of the zero set of a single polynomial was already considered in [6]. Rather, we consider *type- W singular loci*. We give a rigorous definition below. Among others, those singular loci include the following:

- (1) The zero set of $p : S^n \rightarrow \mathbb{R}^k$ in the unit sphere S^n .
- (2) The set of critical points of $p : S^n \rightarrow \mathbb{R}$ on S^n .
- (3) The set of nondegenerate minima of $p : S^n \rightarrow \mathbb{R}$ on S^n .

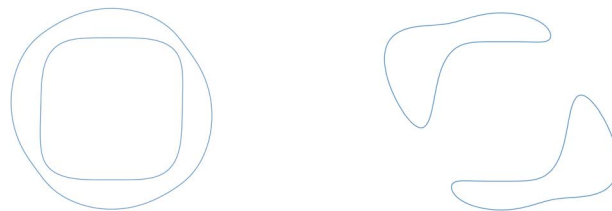


Fig. 1. The zero sets of P_1 (on the left) and P_2 (on the right) plotted on the affine chart $x_0 = 1$.

- (4) The set of points where a polynomial map $p = (p_1, p_2) : S^2 \rightarrow \mathbb{R}^2$ has a Whitney cusp.

Notice that in this list we have switched from the real projective space to the unit sphere. The reason is that polynomials define functions on the sphere, but (unless the degree is even) they do not define functions on the projective space. For this reason, in the following, we will exclusively consider loci inside the sphere S^n and not in projective space; however, we observe that, since the latter is double covered by the former, the study of the spherical case can be related to the projective one using standard algebraic topology techniques.

In this article, we follow the same philosophy of [6] and we focus on *tail probabilities*. We want to show that a system of m Kostlan polynomials (P_1, \dots, P_m) rarely has a set of point from the list above that has “complicated topology.” A Kostlan polynomial of degree d is defined as

$$P_k(x_0, \dots, x_n) = \sum_{\alpha_0 + \dots + \alpha_n = d} \xi_{\alpha_0, \dots, \alpha_n}^{(k)} \sqrt{\frac{d!}{\alpha_0! \dots \alpha_n!}} x_0^{\alpha_0} \dots x_n^{\alpha_n}, \quad (1)$$

where the $\xi_{\alpha_0, \dots, \alpha_n}^{(k)}$ are independent and identically distributed standard Gaussian random variables. By “complicated topology” we mean configurations that can’t be realized by polynomials of lower degree.

We show in Theorem 4 below that with high probability the type- W singular locus of a system of Kostlan polynomials (P_1, \dots, P_m) with maximal degree $d := \max\{\deg(P_1), \dots, \deg(P_m)\}$ is *ambient isotopic* to the singular locus of a system of polynomials of degree approximatively $\sqrt{d \log d}$. By this we mean that there exists a continuous family of diffeomorphisms $\varphi_t : S^n \rightarrow S^n$ with $\varphi_0 = \text{id}_{S^n}$ that at time $t = 1$ brings the singular locus of the 1st system to the singular locus of the 2nd one. The notion ambient isotopy is weaker than the notion rigid isotopy that Hilbert used. However, we can’t work with rigid isotopies in our setting, because this is not defined for pairs of polynomials that live in different spaces—on the one hand, polynomials of degree d and on the other hand polynomials of degree $\sqrt{d \log d}$. We need to compare those polynomials in the space of all C^∞ functions!

In particular, our results also imply that *maximal configurations*, that is, type- W singularities that can’t be realized by polynomials of lower degree, have exponential small probability as $d \rightarrow \infty$. They are virtually non-existent under the Kostlan distribution. This has implications for numerical experiments: for large d it is

impossible to find maximal configurations by sampling Kostlan polynomials. For zero sets this was already observed in [19].

Sarnak [36] suggested using the probabilistic point of view in 2011. Since then the research area has seen much progress [6, 8, 10–12, 23–27, 30, 31, 36, 39]. Today, the merging of algebraic geometry and probability theory goes under the name of *random algebraic geometry*. The motivation behind taking a statistical point of view is that already for curves in the plane the number of rigid isotopy types grows super-exponentially as the degree of the curve goes to infinity [32]. Therefore, a deterministic enumeration of all possibilities is a hopeless endeavor. On the other hand, it was shown in [6] that for *Kostlan* curves there are only few types that appear with significantly high probability. This result motivated the more general study in this article.

1.1 The Kostlan distribution

Our choice of the Kostlan distribution (1) has several reasons. The 1st is that Kostlan polynomials are invariant under orthogonal change of variables: if P is a Kostlan polynomial in $n + 1$ variables, then for any orthogonal $(n + 1) \times (n + 1)$ -matrix R we have $P \circ R \sim P$. We believe that a reasonable probability distribution should have this property, since we are considering topological features of *geometric sets*. Following Klein [20] those should be defined being invariant under orthogonal change of coordinates of the ambient space. However, Kostlan polynomials are not the only orthogonally invariant probability distribution. We need more reasons to justify this choice. A 1st possible one is that the Kostlan distribution is particularly suited for comparing real algebraic geometry with complex algebraic geometry: in fact if one considers complex Kostlan polynomials (defined by taking complex Gaussians in (1)), the resulting distribution is the unique gaussian distribution (up to multiples), which is invariant under unitary change of coordinates. Ultimately, this connection is why random real algebraic geometry behaves as the “square root” of complex algebraic geometry, see [27]. Moreover, up to multiples, the Kostlan distribution is the unique, among the orthogonally invariant ones, for which we can write a random polynomial as a linear combination of the monomials with independent gaussian coefficients—thus it is “simple” to write a Kostlan polynomial. Finally, another reason is that among the orthogonally invariant probability distributions on the space of polynomials the Kostlan distribution behaves well under a certain projection, which is the key part in the proof of our main result Theorem 4 on the tail probabilities.

1.2 Singularities

The examples of singular loci above can be summarized under the following technical definition involving the *jet space* $J^r(S^n, \mathbb{R}^m)$. We recall the precise definition of the jet space in Section 3.1. One may think of points in $J^r(S^n, \mathbb{R}^m)$ as lists of derivatives of functions at a point. Those lists are called *jets*. In this interpretation, each function $f \in C^r(S^n, \mathbb{R}^m)$ defines a map $j^r f : S^n \rightarrow J^r(S^n, \mathbb{R}^m)$, called the r -th jet prolongation, which maps x to the list of derivatives of f at x . The precise definition of this is given in Definition 9 below.

The key aspect is that $J^r(S^n, \mathbb{R}^m)$ has a natural semialgebraic structure and we can therefore define semialgebraic sets therein.

Definition 1 (The type- W singular locus). We call a subset $W \subseteq J^r(S^n, \mathbb{R}^m)$ a *singularity type*, if it is semialgebraic and invariant under diffeomorphisms induced by orthogonal change of variables. Given $f \in C^r(S^n, \mathbb{R}^m)$, the subset $U = j^r f^{-1}(W) \subseteq S^n$ is called the type- W singular locus of f .

The semialgebraic sets describing the above examples are as follows.

- (1) $W = S^n \times \{0\} \subset J^0(S^n, \mathbb{R}^m)$.
- (2) $W = \{df = 0\} \subset J^1(S^n, \mathbb{R})$.
- (3) $W = \{df = 0, d^2f > 0\} \subset J^2(S^n, \mathbb{R})$.
- (4) $W \subset J^3(S^2, \mathbb{R}^2)$ gives conditions on the derivatives of $f : S^2 \rightarrow \mathbb{R}^2$ up to order three, such that in some local coordinates f has the form $(x_1, x_2) \mapsto (x_1, x_2^3 - x_1 x_2)$ (see [4]).

Definition 2 (Ambient isotopic pairs). Let C_0, C_1 be stratified subcomplexes of the sphere S^n . We say that the two pairs are *ambient isotopic*, denoted

$$(S^n, C_0) \sim (S^n, C_1),$$

if there exists a family of diffeomorphisms $(\varphi_t : S^n \rightarrow S^n)_{0 \leq t \leq 1}$ with $\varphi_0 = \text{id}_{S^n}$ and $\varphi_1(C_0) = C_1$.

1.3 Organization of the article

The rest of the article is now organized as follows. In the next section we state our main results. In Section 3 we recall the definition of jet space and use it for defining

the discriminant locus of a singularity type. Then, in Section 4, we recall the definition of harmonic polynomials and the decomposition of the space of polynomials into the harmonic basis and define several norms for polynomials, which will be used in the next sections. In Section 5 we prove our result on quantitative stability from Theorem 7 and in Section 6 we prove Theorem 4. Finally, in Section 7, we discuss what our results imply for maximal configurations.

2 Main Results

Having clarified in the previous section what we mean by type- W singular locus of a map, we can now state our main results.

2.1 Main result 1: low-degree approximation

The 1st result essentially tells that most real singularities given by polynomial equations of degree d are ambient isotopic to singularities given by polynomials of degree $O(\sqrt{d \log d})$. This means that for any singularity type W the probability of the following event goes to one as $d \rightarrow \infty$: let $p : S^n \rightarrow \mathbb{R}^m$ be given as the restriction to the sphere of a system of Kostlan polynomials. There exists a polynomial q of degree $O(\sqrt{d \log d})$ such that $(S^n, j^r q^{-1}(W)) \sim (S^n, j^r p^{-1}(W))$. The new polynomial q can be thought as a low-degree approximation of p .

The approximation procedure is constructive in the sense that one can read the approximating polynomial q from a linear projection of the given one. It is also quantitative in the sense that the approximating procedure will hold for a subset of the space of polynomials with measure increasing very quickly to full measure as the degree goes to infinity.

To be more specific, we denote by $\mathcal{P}_{n,d} = \mathbb{R}[x_0, \dots, x_n]_{(d)}$ the space of homogeneous polynomials of degree d . We recall from Section 4.1 below that this space admits a decomposition:

$$\mathcal{P}_{n,d} = \bigoplus_{d-\ell \in 2\mathbb{N}} \|x\|^{d-\ell} \mathcal{H}_{n,\ell} \quad (2)$$

where $\mathcal{H}_{n,\ell}$ denotes the space of homogeneous, harmonic polynomials of degree ℓ . Given $P \in \mathcal{P}_{n,d}$, we denote by $p : S^n \rightarrow \mathbb{R}$ its restriction to the unit sphere and for $L \in \{0, \dots, d\}$ we define $p|_L = \sum_{\ell \leq L, d-\ell \in 2\mathbb{N}} p_\ell$, where p_ℓ is the restriction to S^n of the polynomial appearing in the decomposition $P = \sum_\ell P_\ell$ given by (2). Given polynomials p_1, \dots, p_m

with $\deg(p_i) = d_i$, we can form the polynomial map $p = (p_1, \dots, p_m) : S^n \rightarrow \mathbb{R}^m$ and for $L \in \{0, \dots, d\}$ we can define

$$p|_L = (p_1|_L, \dots, p_m|_L). \tag{3}$$

We will denote by $\underline{d} = (d_1, \dots, d_m)$ and by $d = \max d_i$.

Definition 3 (Low-degree approximation event). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type. Let $p = (p_1, \dots, p_m)$ be a random polynomial map and $L \in \{0, \dots, d\}$. We can consider

$$A_L = \left\{ (S^n, j^r p|_L^{-1}(W)) \sim (S^n, j^r p^{-1}(W)) \right\} \subset \mathcal{P}_{n, \underline{d}}$$

(i.e., the type- W singularities of p and $p|_L$ are ambient isotopic). We call A_L the *low-degree approximation event* for type- W singularities.

Here is our 1st main theorem.

Theorem 4 (Low-degree approximation of type- W singular locus). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type (as defined in Definition 1). Let $P = (P_1, \dots, P_m)$ be a system of homogeneous Kostlan polynomials in $n + 1$ many variables and of degrees $\underline{d} = (d_1, \dots, d_m)$, and denote by $d = \max d_i$. Let $p = P|_{S^n}$ be the restriction of P to the sphere S^n . For an integer $0 \leq L \leq d$ let us denote by A_L the event that the type- W singular locus of p and $p|_L$ are ambient isotopic (in the sense of Definition 2). Then, we have the following behavior for three different regimes of L :

1. *There exists $b_0 > 0$ such that for all $b \geq b_0$ there exist $a_1, a_2 > 0$ with the property that, choosing $L = b\sqrt{d \log d}$ we have $\mathbb{P}(A_L) \geq 1 - a_1 d^{-a_2}$.*
2. *For every $\frac{1}{2} < b < 1$, there exist $a_1, a_2 > 0$ (with $0 < a_2 < 1$), such that choosing $L = d^b$ we have $\mathbb{P}(A_L) \geq 1 - a_1 e^{-d^{a_2}}$.*
3. *For every $0 < b < 1$ there exist $a_1, a_2 > 0$ such that choosing $L = bd$ we have that $\mathbb{P}(A_L) \geq 1 - a_1 e^{-a_2 d}$.*

Conversely:

4. *For all $a > 0$, there exists $b > 0$, such that choosing $L = b\sqrt{d \log d}$ we have $\mathbb{P}(A_L) \geq 1 - d^{-a}$ for d large enough.*
5. *For all $0 < a < 1$, there exists $0 < b < 1$, such that choosing $L = d^b$ we have for large enough d that $\mathbb{P}(A_L) \geq 1 - e^{-d^a}$.*

6. For all $a > 0$, there exists $b > 0$, such that choosing $L = bd$ we have $\mathbb{P}(E_L) \geq 1 - e^{-ad}$ for d large enough.

We call the 2nd and the 3rd regime in the theorem “exponential rarefaction of maximal configurations” because the result tells that maximal configurations, that is, polynomials of degree d whose singular loci are not ambient isotopic to that of a polynomial of smaller degree, have exponentially small probability in the space of all polynomials. In order to get an insight of these results, we will provide some applications of them for the explicit case of the singularities described in Section 1.2, in the last section of this paper.

2.2 Main result 2: stable neighborhoods and quantitative stability

A key step in proving Theorem 4 is proving a deterministic bound on the size of the *stable neighborhood* of P . In order to state the result we first give two definitions.

Definition 5. Let $f \in C^{r+1}(S^n, \mathbb{R}^m)$, and $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type. We fix a semialgebraic stratification $W = \coprod_{i=1}^k W_i$ into smooth semialgebraic strata. The r -jet map $j^r f$ is called transversal to the stratum W_i if for all $x \in S^n$ we either have $j^r f(x) \notin W_i$, or

$$j^r f(x) \in W_i \quad \text{and} \quad d_x(j^r f)(T_x S^n) + T_{j^r f(x)} W_i = T_{j^r f(x)} J^r(S^n, \mathbb{R}^m).$$

Here, $d_x(j^r f)$ is the differential of $j^r f$ (i.e., the induced map at the level of the tangent spaces). The r -jet map $j^r f$ is called transversal to W if it is transversal to all the strata of $W = \coprod_{i=1}^k W_i$. We write

$$j^r f \pitchfork W$$

when $j^r f$ is transversal to W .

Observe that we need $f \in C^{r+1}$ in order to talk about transversality of its r -jet to W : in fact, if $f \in C^{r+s}(S^n, \mathbb{R}^m)$, the jet extension $j^r f : S^n \rightarrow J^r(S^n, \mathbb{R}^m)$ is of class C^s (see [15, page 61]). Therefore, since the transversality condition involves the differential of $j^r f$, we need this map to be at least C^1 , that is, f to be at least C^{r+1} .

Definition 6 (Stable neighborhood). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be singularity type, and let $f \in C^{r+1}(S^n, \mathbb{R}^m)$ with the property that $j^r f$ is transversal to W . The W -stable neighborhood

of f in $C^{r+1}(S^n, \mathbb{R}^m)$ is

$$\text{SN}(f, W) := \{g \in C^{r+1}(S^n, \mathbb{R}^m) \mid f \text{ and } g \text{ are } W\text{-transversely homotopic}\}$$

(i.e., there is a homotopy f_t in $C^{r+1}(S^n, \mathbb{R}^m)$ between $f_0 = f$ and $f_1 = g$ such that for every $t \in [0, 1]$ the r -th jet extension $j^r f_t$ is transversal to W).

Note the departure from polynomials to general C^{r+1} -functions in this definition. In general, it is infeasible to prove bounds on the size of a stable neighborhood. But, if f is given by a polynomial $f = P|_{S^n}$, we can measure how non-degenerate its jet with respect to W is. This measure is the distance between P and the so-called discriminant (see also (7))

$$\Sigma_{W, \underline{d}} = \{P \in \mathcal{P}_{n, \underline{d}} \mid j^r p \text{ is not transversal to } W, \text{ where } p = P|_{S^n}\}.$$

The distance measure we take is the *Bombieri–Weyl distance* $\text{dist}_{\text{BW}}(\cdot, \cdot)$ from Section 4.2.1. This distance is particularly tied to Kostlan polynomials. It is interesting to observe that both the approximation result from [6] as well as our Theorem 4 are very special of the Kostlan distribution. The reason for this is the behavior of this distribution under the projection onto the spaces of harmonic polynomials. In fact, the results from [6] are related to the more general problem: the estimation of the probability for the projection of P to the subspace of harmonic polynomials of degree L to be stable in the sense of Definition 6. In the process of proving Theorem 4 we will also face this more general problem. The following theorem, which is our 2nd main result, is the central piece in this process.

Theorem 7 (Quantitative stability). There exists a constant $c_1 > 0$ (depending on W) such that, given $P \notin \Sigma_W$, and writing $p = P|_{S^n}$, if $d_1, \dots, d_m \geq r + 1$ then we have

$$\{f \in C^{r+1}(S^n, \mathbb{R}^m) \mid \|f - p\|_{C^{r+1}} < c_1 \text{dist}_{\text{BW}}(P, \Sigma_{W, \underline{d}})\} \subseteq \text{SN}(p, W).$$

To appreciate the subtlety of this result, we remark that the f in the statement is not a polynomial but rather can be any C^{r+1} -function. The space of polynomials of a given degree is finite dimensional, and as all norms on finite-dimensional spaces are equivalent, this implies the existence of a constant $c > 0$ for which the above statement holds for all *polynomials* f ; the crucial point here is that the estimate can be made uniform over the whole infinite-dimensional space of C^{r+1} maps.

And that the bound only depends on the distance of p to the discriminant $\Sigma_{w,d}$ *within the space of polynomials*.

3 Jet Spaces and Discriminants

In this section, we briefly introduce some notations and background facts on jet spaces and discriminants. We refer the reader to the textbooks [15, 41] for more details and generalizations.

3.1 Jet spaces

We recall now the definition of jet manifolds, following [15]. Given two smooth manifolds N and M , an r -jet from N to M is an equivalence class of triples (x, f, U) where $U \subset N$ is an open set, $x \in U$ and $f : U \rightarrow M$ is a C^r map; the equivalence relation is as follows: two pairs (x, f, U) and (y, g, V) are equivalent if and only if $x = y$ and in any pair of charts adapted (Given $f : N \rightarrow M$ and $x \in N$, we say that two charts (C_x, ψ) (a chart on a neighborhood C_x of x) and $(B_{f(x)}, \varphi)$ (a chart on a neighborhood $B_{f(x)}$ of $f(x)$) are adapted to f around x if $f(C_x) \subset B_{f(x)}$.) to f around x the maps f and g have the same derivatives up to order r . The equivalence class of the triple (x, f, U) is denoted by $j^r f(x)$ and called *the r -jet of f at x* ; the point x is called the *source* of the jet and $f(x)$ the *target*. The set of all r -jets from N to M is denoted by $J^r(N, M)$ and the set of all jets with source x is denoted by $J_x^r(N, M)$.

The most important cases in this paper are $N = \mathbb{R}^{n+1}$, $M = \mathbb{R}^m$ and $N = S^n$, $M = \mathbb{R}^m$. The next example shows how to think of the former.

Example 8. Let $N = \mathbb{R}^{n+1}$ and $M = \mathbb{R}^m$. We can represent an r -jet of $f : U \rightarrow \mathbb{R}^m$ at a point x by the list of derivatives of f at x . Therefore, the jet space $J^r(\mathbb{R}^{n+1}, \mathbb{R}^m)$ has an explicit manifold structure given by

$$J^r(\mathbb{R}^{n+1}, \mathbb{R}^m) \cong \mathbb{R}^{n+1} \times \bigoplus_{j=0}^r \mathbb{R}^{mN_j}, \quad N_j = \binom{n+1+j}{j}, \quad (4)$$

such that

$$j^r f(x) = (x, f(x), Df(x), \dots, D^k f(x), \dots, D^r f(x)),$$

where $D^k f(x) \in \mathbb{R}^{mN_j}$ is the *tensor* of the order- k partial derivatives of f at x . That is, $J^r(\mathbb{R}^{n+1}, \mathbb{R}^m)$ can be seen as the vector bundle over \mathbb{R}^{n+1} , where at each point x we attach

all derivatives of polynomials at x up to degree r . Another useful interpretation is seeing the symmetric tensors as polynomials.

$$J^r(\mathbb{R}^{n+1}, \mathbb{R}^m) \cong \mathbb{R}^{n+1} \times \bigoplus_{i=0}^r (\mathcal{P}_{n,i})^{\times m}, \tag{5}$$

such that $j^r f(x) = (x, f(x), f_1^{(1)}(x), \dots, f^{(r)}(x))$, where

$$f^{(i)}(x) = \sum_{\alpha_0 + \dots + \alpha_n = i} \left[\left(\frac{\partial}{\partial y_0} \right)^{\alpha_0} \cdots \left(\frac{\partial}{\partial y_d} \right)^{\alpha_d} f(y) \right] \Big|_{y=x} x_0^{\alpha_0} \cdots x_n^{\alpha_n}.$$

The manifold structure on $J^r(N, M)$ is defined as follows. Given open charts (U, ψ) and (W, φ) on N and M respectively we have the bijection

$$\theta : J^r(U, V) \rightarrow J^r(\psi(U), \varphi(V)), \quad j^r f(x) \mapsto j^r(\varphi \circ f \circ \psi^{-1})(\psi(x)).$$

By (4), $J^r(\psi(U), \varphi(V))$ is an open subset of a real vector space. We declare $(J^r(U, V), \theta)$ to be a chart on $J^r(N, M)$ and the set of all such charts gives an atlas, hence a differentiable structure, on $J^r(N, M)$. The map θ gives local coordinates for the r -jet of f .

When N and M are real algebraic manifolds, the manifold charts on the jet space are real algebraic as well and the jet space $J^r(N, M)$ is also a real algebraic manifold; as a consequence we can define semialgebraic subsets therein, see [3, Remark 3.2.15]: $W \subset J^r(N, M)$ is semialgebraic if and only if $\theta(W \cap J^r(U, V))$ is semialgebraic for every chart $(J^r(U, V), \theta)$.

Definition 9. Let $f : N \rightarrow M$. The r -jet prolongation $j^r f : N \rightarrow J^r(N, M)$ of f is given by $j^r f : x \mapsto j^r f(x)$.

If N and M are smooth, the jet prolongation is smooth, see [15, Chapter 2.4].

Now, we discuss how to think of the jet space $J(X, \mathbb{R}^m)$, where $X \hookrightarrow \mathbb{R}^{n+1}$ is a submanifold. Although the case $X = S^n$ is of main interest to us, it is illustrative to consider a general submanifold. First, we consider the subset of $J(\mathbb{R}^{n+1}, \mathbb{R}^m)$ where the base points are points in X :

$$J^r(\mathbb{R}^{n+1}, \mathbb{R}^m)|_X \cong X \times \bigoplus_{j=0}^r \mathbb{R}^m N_j.$$

Then, we have the commutative diagram

$$\begin{array}{ccc}
 J^r(\mathbb{R}^{n+1}, \mathbb{R}^m)|_X & \xrightarrow{\rho} & J^r(X, \mathbb{R}^m) \\
 & \searrow & \swarrow \\
 & X &
 \end{array} \tag{6}$$

where ρ projects the list of derivatives of a function f at $x \in X$ to the list of derivatives restricted to $T_x X$. In the coordinates from (5) this operator takes the following form. Let $(x, f(x), f^{(1)}, \dots, f^{(r)}) \in \mathbb{R}^{n+1} \times \bigoplus_{i=0}^r (\mathcal{P}_{n,i})^{\times m}$ be a point representing the r -jet of a function f . Then,

$$\rho(x, f(x), f^{(1)}, \dots, f^{(r)}) = (x, f(x), f^{(1)}|_{T_x X}, \dots, f^{(r)}|_{T_x X}).$$

That is, ρ restricts the polynomial functions $f^{(i)}$ to $T_x X$.

3.2 The W -discriminant

Recall from Definition 1 that we call $W \subseteq J^r(S^n, \mathbb{R}^m)$ a singularity type, if it is semialgebraic and invariant under orthogonal change of variables. Recall also the notion of transversality to W from Definition 5.

Now, we are ready to introduce the W -discriminant. It is important to realize that on the one hand W is defined to be a subset of the r -th jet space, while on the other hand the associated discriminant lives in the space of C^{r+1} functions!

Definition 10 (The W -discriminant). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type.

$$\Sigma_W := \{f \in C^{r+1}(S^n, \mathbb{R}^m) \mid j^r f \text{ is not transversal to } W\} \subset C^{r+1}(S^n, \mathbb{R}^m)$$

is called the W -discriminant.

The W -discriminant for polynomial systems with degree pattern \underline{d} is defined as

$$\Sigma_{W, \underline{d}} := \Sigma_W \cap \mathcal{P}_{n, \underline{d}}. \tag{7}$$

It follows from the definition that

$$\Sigma_W = \bigcup_{x \in S^n} \Sigma_W(x), \quad \text{and} \quad \Sigma_{W, \underline{d}} = \bigcup_{x \in S^n} \Sigma_{W, \underline{d}}(x), \tag{8}$$

where $\Sigma_W(x)$ is the set of C^{r+1} functions f whose associated map $j^r f$ is not transversal to W at the point $x \in S^n$, and $\Sigma_{W, \underline{d}}(x) := \Sigma_W(x) \cap \mathcal{P}_{n, \underline{d}}$.

The discriminant $\Sigma_{W, \underline{d}}$ has the structure of a semialgebraic set with $\text{codim} \Sigma_{W, \underline{d}} \geq 1$. Moreover, considering the natural induced action of the orthogonal group $G := O(n + 1)$ on $J^r(S^n, \mathbb{R}^m)$ and Σ_W , since W is G -invariant, for every $g \in G$ we have

$$g \cdot \Sigma_W(g^{-1}x) \cong \Sigma_W(x) \quad \text{and} \quad g \cdot \Sigma_{W, \underline{d}}(g^{-1}x) \cong \Sigma_{W, \underline{d}}(x). \tag{9}$$

3.3 The degree of the W -discriminant

The next lemma estimates the degree of the W -discriminant $\Sigma_{W, \underline{d}} \subset \mathcal{P}_{n, \underline{d}}$ as a function of $d = \max d_i$. The estimate is a polynomial in d . For example, when $\Sigma \subset \mathcal{P}_{n, d}$ is the discriminant for a polynomial having degenerate zero set, then its degree is $(n + 1)(d - 1)^n = O(d^n)$. Later we will use this estimate on the degree for bounding the probability of a system of Kostlan polynomials to be close to the discriminant $\Sigma_{W, \underline{d}}$.

Proposition 11 (Degree bound). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a semialgebraic set and $\underline{d} = (d_1, \dots, d_m)$ with $d = \max d_i$. There exists a constant $u > 0$, which depends on W , and a nonzero polynomial $Q : \mathcal{P}_{n, \underline{d}} \rightarrow \mathbb{R}$ of degree bounded by $u \cdot d^u$ such that $\Sigma_{W, \underline{d}} \subset Z(Q)$, where $Z(Q)$ is the zero set of Q .

Proof. Fix a point $x_0 \in S^n$. We first show that $\Sigma_{W, \underline{d}}(x_0)$ is a semialgebraic subset of $\mathcal{P}_{n, \underline{d}}$ defined by polynomials of degree bounded by some constant $\alpha_1 > 0$ depending on W only.

Recall that $\Sigma_W(x_0) \subset C^{r+1}(S^n, \mathbb{R}^m)$ is the set of C^{r+1} functions whose r -th jet is not transversal to W at the point x_0 :

$$\Sigma_W(x_0) = \{f \in C^{r+1}(S^n, \mathbb{R}^m) \mid j^r f(x_0) \in W \text{ and } \text{im}(d_{x_0} j^r f) + T_{j^r f(x_0)} W \neq T_{j^r f(x_0)} J^r(S^n, \mathbb{R}^m)\}. \tag{15}$$

Since W is defined by semialgebraic conditions, we see that $\Sigma_W(x_0)$ is given by a semialgebraic condition on the list of the 1st $r + 1$ derivatives (i.e., the $(r + 1)$ -jet) of

$f : S^n \rightarrow \mathbb{R}^m$ at x_0 (the derivative of $j^r f$ involves $j^{r+1} f(x_0)$):

$$\Sigma_W(x_0) = \left\{ f \in C^{r+1}(S^n, \mathbb{R}^m) \mid \bigcup_{i=1}^a \bigcap_{j=1}^{b_i} \widehat{s}_{i,j}(j^{r+1} f(x_0)) *_{i,j} 0 \right\},$$

where a, b_i are some constants, $\widehat{s}_{i,j}$ are polynomials and $*_{i,j} \in \{<, =, >\}$; all these data depend on W only. Now, the map $\widehat{j} : \mathcal{P}_{n,\underline{d}} \rightarrow J_{x_0}^{r+1}(S^n, \mathbb{R}^m)$ associating to every polynomial map its $(r+1)$ -jet at x_0 is linear. Moreover, $\Sigma_{W,\underline{d}}(x_0) = \Sigma_W(x_0) \cap \mathcal{P}_{n,\underline{d}}$. Taking $s_{i,j} := \widehat{s}_{i,j} \circ \widehat{j}$ we can therefore write

$$\Sigma_{W,\underline{d}}(x_0) = \bigcup_{i=1}^a \bigcap_{j=1}^{b_i} \{s_{i,j} *_{i,j} 0\}.$$

We have $\deg(s_{i,j}) = \deg(\widehat{s}_{i,j})$, because \widehat{j} is linear. This implies that the degrees of the polynomials $s_{i,j}$ are bounded by some constant $\alpha_1 > 0$, which only depends on W (but not on d).

We now proceed with proving the claim of the proposition. Let I be the set of all the pairs (i, j) such that $*_{i,j}$ equals “=” and consider the algebraic set

$$Z_{W,\underline{d}}(x_0) = \bigcup_{(i,j) \in I} \{s_{i,j} = 0\} \subset \mathcal{P}_{n,\underline{d}}.$$

By construction, $\Sigma_{W,\underline{d}}(x_0) \subset Z_{W,\underline{d}}(x_0)$. Since the $s_{i,j}$ are defined in terms of W only, the cardinality of I is bounded by some constant ℓ depending only on W . By (8) we have $\Sigma_{W,\underline{d}} = \bigcup_{x \in S^n} \Sigma_{W,\underline{d}}(x)$ and therefore, by (9), we have

$$\Sigma_{W,\underline{d}} = \bigcup_{g \in O(n+1)} g \cdot \Sigma_{W,\underline{d}}(x_0) \subseteq \bigcup_{g \in O(n+1)} g \cdot Z_{W,\underline{d}}(x_0).$$

Let us denote $Z := \bigcup_{g \in O(n+1)} g \cdot Z_{W,\underline{d}}(x_0)$. To finish the proof it is enough to find a polynomial Q that vanishes on Z and to bound its degree.

Let $G = \text{GL}(\mathbb{R}^{n+1})$ be the general linear group, and $\rho : G \rightarrow \text{GL}(\mathcal{P}_{n,\underline{d}})$ be its representation in the space of polynomial maps given by change of variables. That is, $\rho(g)P(x) = P(g(x))$. The representation ρ extends to a map between spaces of endomorphisms $\rho : \mathbb{R}^{(n+1) \times (n+1)} \rightarrow \mathbb{R}^{N \times N}$, simply by declaring $\rho(g)$ to be the linear map

that to a polynomial P associates the new polynomial $\rho(g) \cdot P = P(gx)$. We denote by $h_{i,j} : \mathbb{R}^{(n+1) \times (n+1)} \times \mathbb{R}^N \rightarrow \mathbb{R}$ the polynomial defined by

$$h_{i,j}(g, P) = s_{i,j}(\rho(g) \cdot P),$$

and we define the incidence set

$$\hat{Z} = \{(g, P) \in \mathbb{R}^{(n+1) \times (n+1)} \times \mathcal{P}_{n,d} \mid \forall (i, j) \in I : h_{i,j}(g, P) = 0\}.$$

Since the components of the representation ρ have degree at most d , and the degrees of the polynomials $s_{i,j}$ are all bounded by α_1 , the degree of each $h_{i,j}$ is bounded by $\alpha_1(d+1) \leq \alpha_2 d$, for some $\alpha_2 > 0$ (depending on W only and not on d). Therefore, \hat{Z} is defined by at most ℓ equations of degree bounded by $\alpha_2 d$.

In order to produce our polynomial Q , we move first to the complex numbers and consider the algebraic set $\hat{Z}^{\mathbb{C}}$ defined by the same equations as \hat{Z} :

$$\hat{Z}^{\mathbb{C}} = \{(g, P) \in \mathbb{C}^{(n+1) \times (n+1)} \times \mathcal{P}_{n,d}^{\mathbb{C}} \mid \forall (i, j) \in I : h_{i,j}(g, P) = 0\}.$$

Here, $\mathcal{P}_{n,d}^{\mathbb{C}}$ denotes the space of *complex* polynomial systems with degree pattern \underline{d} . Denoting by $\pi : \mathbb{C}^{(n+1) \times (n+1)} \times \mathcal{P}_{n,d}^{\mathbb{C}} \rightarrow \mathcal{P}_{n,d}^{\mathbb{C}}$ the projection on the 2nd factor, note that

$$Z \subseteq \pi(\hat{Z}^{\mathbb{C}}) \cap \mathcal{P}_{n,d}.$$

Therefore, in order to get a polynomial vanishing on Z , we can sufficiently find a real polynomial Q vanishing on $\pi(\hat{Z}^{\mathbb{C}})$.

Write the algebraic set $\hat{Z}^{\mathbb{C}}$ as

$$\hat{Z}^{\mathbb{C}} = \bigcup_{k \geq 0} \hat{Z}_k^{\mathbb{C}} \tag{10}$$

where each $\hat{Z}_k^{\mathbb{C}}$ is the union of all irreducible components of $\hat{Z}^{\mathbb{C}}$ of dimension k , namely

$$\hat{Z}_k^{\mathbb{C}} = \bigcup_{i=1}^{\gamma_k} \hat{Z}_{k,i}^{\mathbb{C}} \quad \text{with } \hat{Z}_{k,i}^{\mathbb{C}} \text{ irreducible.}$$

Observe that $\mathbb{C}^{(n+1) \times (n+1)} \times \mathcal{P}_{n,d}^{\mathbb{C}}$ is irreducible and therefore, by [38, Corollary 2, page 75], the dimension of each component of $\hat{Z}^{\mathbb{C}}$ is bounded below by $(n+1)^2 + N - \ell$, where

$N = \dim_{\mathbb{C}} \mathcal{P}_{n,d}^{\mathbb{C}}$. Therefore, the previous union (10) can be written as

$$\hat{Z}^{\mathbb{C}} = \bigcup_{k \geq (n+1)^2 + N - \ell} \hat{Z}_k^{\mathbb{C}}.$$

Fix a number $(n+1)^2 + N - \ell \leq k \leq (n+1)^2 + N$ (the number k is in range of the possible dimensions for the components of $\hat{Z}_k^{\mathbb{C}}$) and observe that

$$\sum_{i=1}^{\gamma_k} \deg(\hat{Z}_{k,i}^{\mathbb{C}}) = \deg(\hat{Z}_k^{\mathbb{C}}) \leq (\alpha_2 d)^\ell.$$

The reason for this is that, when defining the degree of $\hat{Z}_k^{\mathbb{C}}$, we need to intersect it with a generic linear space L of dimension $t = (n+1)^2 + N - k$ and this dimension is bounded above by ℓ , because $k \geq (n+1)^2 + N - \ell$. Therefore, $L \cap \hat{Z}_k^{\mathbb{C}}$ consists of finitely many points in $L \simeq \mathbb{C}^t$, and these points are defined by at most ℓ equations. Each of those equations has degree at most $\alpha_2 d$, because the defining equations in (3.3) have degree at most $\alpha_2 d$. Consequently, the number of such points is bounded by $(\alpha_2 d)^t \leq (\alpha_2 d)^\ell$.

We use now [13, Lemma 2]: since each $\hat{Z}_{k,i}^{\mathbb{C}}$ is irreducible and the projection π is linear we have $\deg(\pi(\hat{Z}_{k,i}^{\mathbb{C}})) \leq \deg(\hat{Z}_{k,i}^{\mathbb{C}})$. In particular, this implies

$$\deg(\pi(\hat{Z}_k^{\mathbb{C}})) = \sum_{i=1}^{\gamma_k} \deg(\pi(\hat{Z}_{k,i}^{\mathbb{C}})) \leq \sum_{i=1}^{\gamma_k} \deg(\hat{Z}_{k,i}^{\mathbb{C}}) \leq (\alpha_2 d)^\ell.$$

Let us denote by $\delta_{i,k}$ the degree of $\pi(\hat{Z}_{k,i}^{\mathbb{C}})$; since $\hat{Z}_{k,i}^{\mathbb{C}}$ is irreducible, then $\pi(\hat{Z}_{k,i}^{\mathbb{C}})$ is irreducible as well and we can apply [13, Proposition 3] to find a polynomial $F_{k,i}$ of degree bounded by $\delta_{k,i}$ vanishing on $\pi(\hat{Z}_{k,i}^{\mathbb{C}})$. Set now

$$F_k := F_{k,1} \cdots F_{k,\gamma_k}.$$

The polynomial F_k vanishes on $\pi(\hat{Z}_k^{\mathbb{C}})$ and has degree bounded by

$$\deg(F_k) = \sum_{i=1}^{\gamma_k} \deg(F_{k,i}) \leq \sum_{i=1}^{\gamma_k} \delta_{k,i} \leq (\alpha_2 d)^\ell.$$

Define the polynomial

$$F := \prod_{k \geq (n+1)^2 + N - \ell} F_k \tag{11}$$

and observe that, because there are at most ℓ factors in (11), then the degree of F is bounded by $\ell(\alpha_2 d)^\ell$. The polynomial F is not yet what we want, because it might not be real. To fix this, write $F = \Re(F) + i\Im(F)$ and define

$$Q := (\Re(F))^2 + (\Im(F))^2.$$

Then Q vanishes on $Z^{\mathbb{C}}$ (therefore on Z and on $\Sigma \subset Z$) and its degree is bounded by

$$\deg(Q) \leq 2\ell(\alpha_2 d)^\ell \leq ud^u$$

for some constant u , which depends on W only. This finishes the proof of the proposition. \blacksquare

4 Norms and Polynomials

In this section, we first introduce the decomposition of the space of homogeneous polynomials into the so-called harmonic basis. Then we define several norms on the space of polynomials, which will be used in the proofs later.

4.1 Harmonic polynomials

The switch from the monomial basis to the harmonic basis will be the key to obtain a low-degree approximation of singular loci.

Definition 12. Let $0 \leq \ell \leq d$ the space of homogeneous *harmonic* polynomials is

$$\mathcal{H}_{n,\ell} := \left\{ P \in \mathcal{P}_{n,\ell} \mid \sum_{i=0}^n \frac{\partial^2}{\partial x_i^2} P = 0 \right\}.$$

The space $\mathcal{P}_{n,d}$ can be decomposed as

$$\mathcal{P}_{n,d} = \bigoplus_{d-\ell \in 2\mathbb{N}} \|x\|^{d-\ell} \mathcal{H}_{n,\ell}. \quad (12)$$

The decomposition (12) has two important properties (see [21]):

- (i) Given a scalar product that is invariant under the action of $O(n+1)$ on $\mathcal{P}_{n,d}$ by change of variables, the decomposition (12) is orthogonal for this scalar product.

- (ii) The action of $O(n + 1)$ on $\mathcal{P}_{n,d}$ preserves each $\mathcal{H}_{n,\ell}$ and the induced representation on the space of harmonic polynomials is irreducible. In particular, there exists a unique, up to multiples, scalar product on $\mathcal{H}_{n,\ell}$ which is $O(n + 1)$ -invariant.

4.2 Norms of polynomials

We continue this section by defining several norms on the space of polynomials $\mathcal{P}_{n,d}$ realized as a subspace of $C^r(S^n, \mathbb{R})$. In general, we give $\mathcal{P}_{n,d}$ the structure of a normed space by endowing each space \mathcal{P}_{n,d_i} with a norm $\|\cdot\|_{d_i}$. The norm on $\mathcal{P}_{n,d}$ is

$$\|P\| := \left(\sum_{i=1}^m \|P_i\|_{d_i}^2 \right)^{1/2} \quad \text{for } P = (P_1, \dots, P_m) \in \mathcal{P}_{n,d}.$$

We identify $\mathcal{P}_{n,d}$ with its image in $C^r(S^n, \mathbb{R})$ given by $\mathcal{S}_{n,d} := \{p : S^n \rightarrow \mathbb{R} \mid p = P|_{S^n}, P \in \mathcal{P}_{n,d}\}$. The decomposition (12) induces a decomposition

$$\mathcal{S}_{n,d} = \bigoplus_{d-\ell \in 2\mathbb{N}} V_{n,\ell} \quad \text{with } V_{n,\ell} = \mathcal{H}_{n,\ell}|_{S^n}. \tag{13}$$

Writing $P = \sum_{\ell} P_{\ell}$ with each $P_{\ell} \in \|x\|^{d-\ell} \mathcal{H}_{n,\ell}$ as in (12), when taking restrictions to the unit sphere we have $p = \sum_{\ell} p_{\ell}$ with each p_{ℓ} the restriction to S^n of a polynomial of degree ℓ : in other words, the restriction to the unit sphere “does not see” the $\|x\|^{d-\ell}$ factor, which is constant on the unit sphere.

Here follows the definition of some relevant norms that we will use in this paper. The 1st three of them are induced by an orthogonally invariant scalar product: by property (i) above the decomposition (13) is orthogonal for all of them.

4.2.1 The Bombieri–Weyl norm

Let $P = \sum_{\alpha_0+\dots+\alpha_n=d} a_{\alpha_0,\dots,\alpha_n} \sqrt{\frac{d!}{\alpha_0! \dots \alpha_n!}} x_0^{\alpha_0} \dots x_n^{\alpha_n}$ be homogeneous polynomial of degree d . The Bombieri–Weyl norm of P is defined by

$$\|P\|_{\text{BW}}^2 := \sum_{\alpha_0+\dots+\alpha_n=d} (a_{\alpha_0,\dots,\alpha_n})^2.$$

Comparing with (1), we see that Kostlan polynomials are given by a multivariate standard Gaussian distribution with respect to the Bombieri–Weyl product. This is the

“close connection” we have mentioned earlier in the paper. The Bombieri–Weyl distance is $d_{\text{BW}}(P, Q) = \|P - Q\|$.

4.2.2 *The L^2 -norm*

This norm is the L^2 -norm of p defined by

$$\|P\|_{L^2}^2 := \int_{S^n} p(x)^2 \, dx,$$

where “ dx ” denotes integration with respect to the standard volume form of the sphere.

4.2.3 *The Sobolev q -norm*

Let $P = \sum_{\ell} P_{\ell}$ be the decomposition into the harmonic polynomials basis; see (12). Then the Sobolev q -norm is defined by

$$\|P\|_{H^q}^2 := \|P_0\|_{L^2}^2 + \sum_{d-\ell \in 2\mathbb{N}} \ell^{2q} \|P_{\ell}\|_{L^2}^2.$$

Note that $\|P_0\|_{L^2}^2 = 0$ when d is odd. Moreover, $\|\cdot\|_{H^0} = \|\cdot\|_{L^2}$.

4.2.4 *C^r -norm*

The C^r norm is defined for all C^r functions. The C^r norm for polynomials is then just the restriction to the space of polynomials. We give the general definition.

We fix an orthogonal invariant norm ν on the vector space $\bigoplus_{j=0}^r \mathbb{R}^{mN_j}$. Moreover, we let $\pi : J^r(\mathbb{R}^{n+1}, \mathbb{R}^m) \rightarrow \bigoplus_{j=0}^r \mathbb{R}^{mN_j}$ be the projection that removes the base point. Then, we define the norm of a jet to be $\nu(z) := \nu(\pi(z))$; that is, the norm of a jet is the norm of the point in the fiber. For a given $\eta \in J_x^r(S^n, \mathbb{R}^m)$ we define

$$\hat{\nu}(\eta) = \inf_{z \in J_x^r(\mathbb{R}^{n+1}, \mathbb{R}^m): \rho(z)=\eta} \nu(z), \tag{14}$$

where ρ is the restriction map from (6), and for a function $f \in C^r(S^n, \mathbb{R}^m)$ we then set

$$\|f\|_{C^r}^{\nu} := \sup_{x \in S^n} \hat{\nu}(j^r f(x)). \tag{15}$$

Remark 13. The definition of the C^r norm includes the choice of a norm ν . Yet, the topology it induces is independent of this choice. This is because ν is a norm on a finite-dimensional real vector space, and all norms on finite-dimensional real vector spaces

are equivalent; we call this topology the C^r topology. Here is another way for obtaining it. Consider the natural map $j^r : C^r(S^n, \mathbb{R}^m) \rightarrow C^0(S^n, J^r(S^n, \mathbb{R}^m))$. Because the sphere S^n is compact, the strong and the weak topology on $C^0(S^n, J^r(S^n, \mathbb{R}^m))$ coincide and by [15, Chapter 2, Theorem 4.3] the image of this map is closed in the strong topology. In particular we can immediately induce a topology on the space $C^r(S^n, \mathbb{R}^m)$, which is the C^r topology, [15, page 62, before Theorem 4.4].

4.3 An inequality between the C^r norm and the Sobolev norm

In the last part of this section we want to prove an inequality between the C^r norm and the Sobolev norm. This inequality will be useful for the proofs in the next section. We note that by endowing $\mathcal{P}_{n,d}$ with the product norm, in order to compare norms on $\mathcal{P}_{n,d}$, we can sufficiently reduce to comparison of the corresponding norms for a single polynomial $P \in \mathcal{P}_{n,d}$ rather than a vector of polynomials.

We first recall the following result from [37]:

Theorem 14. Let $\alpha = (\alpha_0, \dots, \alpha_n)$ be a list of nonnegative integers and $\partial^\alpha = \partial_{x_0}^{\alpha_0} \dots \partial_{x_n}^{\alpha_n}$ be the associated differential operator. There are constants $\beta_{(\alpha,n)}$ that only depend on α and n such that

$$\int_{S^n} |\partial^\alpha P(x)|^2 dx \leq \beta_{(\alpha,n)} \ell^{2|\alpha|} \int_{S^n} |P(x)|^2 dx \text{ for every } P \in \mathcal{H}_{n,\ell},$$

where $|\alpha| = \alpha_0 + \dots + \alpha_n$.

The following proposition connects the C^r -norm with the Sobolev q -norm.

Proposition 15. Let ν be the norm defining the C^r norm in equation (15). There exists a constant $c = c(r, n, \nu) > 0$ depending on r, n and ν such that, if $q \geq r + \frac{n-1}{2}$, we have $\|P\|_{C^r}^\nu \leq c \sqrt{d} \|p\|_{H^q}$.

Proof. Let $p = P|_{S^n}$ and $p = \sum_{d-\ell \in 2\mathbb{N}} p_\ell$ be the decomposition of p in the harmonic basis from (13). By definition (15) of the C^r -norm, we have $\|p_\ell\|_{C^r} \leq \sup_{x \in S^n} \nu(j^r P_\ell(x))$. Moreover, there exists a constant $c_1 = c_1(r, n, \nu)$ such that $\nu(j^r P_\ell(x)) \leq c_1 \sup_{x \in S^n} (\sum_{|\alpha| \leq r} |\partial^\alpha P_\ell(x)|)$, where $\partial^\alpha = \partial_{x_0}^{\alpha_0} \dots \partial_{x_n}^{\alpha_n}$ and $|\alpha| = \alpha_0 + \dots + \alpha_n$. This is because the right-hand side of the equation is a multiple of the L^1 -norm on the fibers of the jet

space, and all norms on a finite-dimensional vector space are equivalent. Summarizing, we have

$$\|p_\ell\|_{C^r}^v \leq c_1 \sup_{x \in S^n} \left(\sum_{|\alpha| \leq r} |\partial^\alpha P_\ell(x)| \right). \tag{16}$$

Recall from (13) the definition of $V_{n,\ell} = \mathcal{H}_{n,\ell}|_{S^n}$. For every $\ell = 0, \dots, d$ the space $V_{n,\ell}$ with the L^2 -scalar product is a reproducing kernel Hilbert space, that is, there exists $Z_\ell : S^n \times S^n \rightarrow \mathbb{R}$ such that for every $h_\ell \in V_{n,\ell}$:

$$\partial^\alpha h_\ell(x) = \int_{S^n} h_\ell(\theta) \partial^\alpha Z_\ell(x, \theta) d\theta. \tag{17}$$

The function Z_ℓ (the ‘‘zonal harmonic’’) is defined as follows: letting $\{Y_{\ell,j}\}_{j \in J_\ell}$ be an L^2 -orthonormal basis for $V_{n,\ell}$ we set $Z_\ell(\theta_1, \theta_2) = \sum_{j \in J_\ell} Y_{\ell,j}(\theta_1) Y_{\ell,j}(\theta_2)$ (written in this way (17) is easily verified, see [1, Chapter 5] and [1, Proposition 5.27] for more properties of the Zonal harmonic). Then from this it follows that

$$\|Z_\ell(\theta_1, \cdot)\|_{L^2}^2 = \langle Z_\ell(\theta_1, \cdot), Z_\ell(\theta_1, \cdot) \rangle_{L^2} = Z_\ell(\theta_1, \theta_1) = \frac{\dim(V_{n,\ell})}{\text{vol}(S^n)} = O(\ell^{n-1}), \tag{18}$$

where the last identity follows from (The constant ‘‘ $\frac{1}{\text{vol}(S^n)}$ ’’ appears because in [1] the normalized $L^2(S^n)$ space is used, that is, the convention $\text{vol}(S^n) = 1$ is adopted.) [1, Proposition 5.27 (d)] and [1, Proposition 5.8]. Writing $p_\ell := P_\ell|_{S^n}$ we obtain the following:

$$\begin{aligned} 2|\partial^\alpha P_\ell|_{S^n}(x)| &= \left| \int_{S^n} p_\ell(\theta) \partial^\alpha Z_\ell(x, \theta) d\theta \right|, && \text{by (17)} \\ &\leq \|p_\ell\|_{L^2} \left(\int_{S^n} |\partial^\alpha Z_\ell(x, \cdot)|^2 \right)^{1/2} && \text{by the Cauchy–Schwarz inequality} \\ &\leq c_2(\alpha, n) \ell^{|\alpha| + \frac{n-1}{2}} \|P_\ell\|_{L^2}, && \text{by Theorem 14 and (18),} \end{aligned} \tag{19}$$

where $c_2(\alpha, n)$ is a constant that depends on α and n . From the above inequalities it follows that

$$\begin{aligned} 2\|P\|_{C^r}^v &\leq \sum_{d-\ell \in 2\mathbb{N}} \|p_\ell\|_{C^r}, && \text{by the triangle inequality} \\ &\leq \sum_{d-\ell \in 2\mathbb{N}} c_1 \sup_{x \in S^n} \left(\sum_{|\alpha| \leq r} |\partial^\alpha P_\ell(x)| \right), && \text{by (16)} \\ &\leq c_3 \sum_{d-\ell \in 2\mathbb{N}} \ell^{r + \frac{n-1}{2}} \|p_\ell\|_{L^2}, && \text{by (19),} \end{aligned} \tag{20}$$

where c_3 is a constant that depends on r, n, ν (the dependence on α has been moved into the dependence on r). Then, we use the Cauchy–Schwarz inequality for $v_1 = (\ell^{r+\frac{n-1}{2}} \|p_\ell\|_{L^2})_{d-\ell \in 2\mathbb{N}}$ and $v_2 = (1, \dots, 1)$ so that

$$\begin{aligned} \sum_{d-\ell \in 2\mathbb{N}} \ell^{r+\frac{n-1}{2}} \|p_\ell\|_{L^2} &= \langle v_1, v_2 \rangle \\ &\leq \|v_1\| \|v_2\| \\ &= \left(\sum_{d-\ell \in 2\mathbb{N}} \ell^{2r+n-1} \|p_\ell\|_{L^2}^2 \right)^{1/2} \left(\sum_{d-\ell \in 2\mathbb{N}} 1 \right)^{1/2} \\ &\leq \sqrt{d} \left(\sum_{d-\ell \in 2\mathbb{N}} \ell^{2r+n-1} \|p_\ell\|_{L^2}^2 \right)^{1/2}. \end{aligned}$$

Plugging this into (20) we obtain $\|P\|_{C^r}^\nu \leq c_3 \sqrt{d} \|P\|_{H^q}$ for $q \geq r + \frac{n-1}{2}$. This finishes the proof. ■

5 Proof of the Quantitative Stability Theorem

In order to prove Theorem 7 we need to recall Thom’s isotopy lemma. We give a variant that uses our notation. Recall from (15) the definition of the C^r norm $\|\cdot\|_{C^r}^\nu$. We denote the associated distance function $\text{dist}_{C^{r+1}}^\nu(\cdot, \cdot)$. Furthermore, recall from Definition 5 that $j^r f \pitchfork W$ means that $j^r f$ is transversal to W .

Lemma 16 (Isotopy Lemma). Let $f, g \in C^{r+1}(S^n, \mathbb{R}^m)$ such that

$$j^r f \pitchfork W \quad \text{and} \quad \|f - g\|_{C^{r+1}}^\nu < \text{dist}_{C^{r+1}}^\nu(f, \Sigma_W).$$

Then $g \in \text{SN}(f, W)$.

Proof. The condition stated guarantees that the homotopy $f_t = (1-t)f + tg$ has the property that for every $t \in [0, 1]$ the jet $j^r f_t$ is transversal to W . Thus we have an induced homotopy of maps between smooth manifolds $j^r f_t : S^n \rightarrow J^r(S^n, \mathbb{R}^m)$, which is transversal to all the strata of $W \subseteq J^r(S^n, \mathbb{R}^m)$ for every $t \in [0, 1]$. In particular, this holds for $g = f_1$ and the conclusion follows from the definition of stable neighborhood in Definition 6. ■

We also need the following helpful lemma.

Lemma 17. Let $P = \sum_{\alpha_0+\dots+\alpha_n=d} c_\alpha x^\alpha \in \mathcal{P}_{n,\underline{d}}$ and $p = P|_{S^n}$. Furthermore, let $T : \mathcal{P}_{n,\underline{d}} \rightarrow \mathcal{P}_{n,\underline{d}}$ and $Y : \mathcal{P}_{n,\underline{d}} \rightarrow \mathcal{P}_{n,\underline{d}}$ be the linear maps defined by

$$T(P) = \sum_{|\alpha_0| \leq d-r-2} c_\alpha x^\alpha \quad \text{and} \quad Y(P) = \sum_{|\alpha_0| \geq d-r-1} c_\alpha x^\alpha,$$

and let $\tau(p)$ and $\gamma(p)$ be the restrictions of $T(P)$ and $Y(P)$ to S^m and $e_0 = (1, 0, \dots, 0) \in S^n$. Then, $j^{r+1}p(e_0) = j^{r+1}\gamma(p)(e_0)$.

Proof. Note first that $P = Y(P) + T(P)$. All the derivatives of order up to $r + 1$ of $T(P)$ vanish at e_0 , and consequently the restrictions of the corresponding polynomials to $T_{e_0}S^m$ will also be zero. ■

Now we are ready to prove Theorem 7.

Proof of Theorem 7. For every $x \in S^n$ let

$$Z_W(x) := \{j^{r+1}f(x) \mid j^r f \text{ is not transversal to } W \text{ at } x\} \subset J_x^{r+1}(S^n, \mathbb{R}^m);$$

that is, $Z_W(x) = j^{r+1}(\Sigma_W(x))$. For $d_1, \dots, d_m \geq r + 1$, the map $S^n \times \mathcal{P}_{n,\underline{d}} \rightarrow J^{r+1}(S^n, \mathbb{R}^m)$ given by $(x, P) \mapsto j^{r+1}p(x)$, where $p = P|_{S^n}$, is a submersion (see [7, Section 1.7]) and therefore $Z_W(x)$ can be defined using only polynomial functions:

$$Z_W(x) = j^{r+1}(\Sigma_{W,\underline{d}}(x)). \tag{21}$$

Now, let $\text{dist}^{\hat{v}}$ be the distance function induced by the norm \hat{v} on $J_x^{r+1}(S^n, \mathbb{R}^m)$ as defined in (14). The idea of the proof is to show the two inequalities

$$\inf_{x \in S^n} \text{dist}^{\hat{v}}(j^{r+1}p(x), Z_W(x)) \leq \text{dist}_{C^{r+1}}^v(p, \Sigma_W) \tag{22}$$

and

$$\text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}) \leq c'_1 \inf_{x \in S^n} \text{dist}^{\hat{v}}(j^{r+1}p(x), Z_W(x)), \tag{23}$$

for some $c'_1 > 0$. The result will follow from these two inequalities. In fact, setting $c_1 = (c'_1)^{-1}$ and combining (23) with (22) we therefore will have

$$\|f - p\|_{C^{r+1}}^v < c_1 \text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}) \implies \|f - p\|_{C^{r+1}}^v < \text{dist}_{C^{r+1}}^v(p, \Sigma_W) \implies f \in \text{SN}(p, W),$$

where for the 2nd implication we have used the definition of stable neighborhood (Definition 6) and Lemma 16 above.

Let us prove the two inequalities, starting from (23). For every $x \in S^n$ let $g_x \in O(n+1)$ be an orthogonal transformation mapping x to $e_0 = (1, 0, \dots, 0)$ and define P^x to be the polynomial

$$P^x := (P_1 \circ g_x, \dots, P_m \circ g_x), \text{ where } P = (P_1, \dots, P_m).$$

Then, using $\Sigma_{W,\underline{d}} = \bigcup_{x \in S^n} \Sigma_{W,\underline{d}}(x)$, we can write

$$\text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}) = \inf_{x \in S^n} \text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}(x)) = \inf_{x \in S^n} \text{dist}_{\text{BW}}(P^x, \Sigma_{W,\underline{d}}(e_0));$$

the last inequality due to the orthogonal invariance of the Bombieri–Weyl norm. Writing out the definition of the distance we have

$$\text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}) = \inf_{x \in S^n} \inf_{Q \in \Sigma_{W,\underline{d}}(e_0)} \|P^x - Q\|_{\text{BW}}.$$

Let T, Y, τ, γ be as in Lemma 17. This lemma implies that, if $Q \in \Sigma_{W,\underline{d}}(e_0)$ is in the discriminant, any other polynomial of the form $\tilde{Q} = Q + T(F)$, for $F \in \mathcal{P}_{n,\underline{d}}$, is also in the discriminant $\Sigma_{W,\underline{d}}(e_0)$, because all the derivatives of $T(F)$ up to order $r+1$ vanish at e_0 . In particular, if we want to minimize the quantity $\|P^x - Q\|_{\text{BW}}$ for $Q \in \Sigma_{W,\underline{d}}(e_0)$, we can restrict ourself to the polynomials Q such that: $T(Q) = T(P^x)$ and $Y(Q) \in \Sigma_{W,\underline{d}}(e_0)$ (notice that $Y(P)$ and $T(Q)$ are orthogonal for any pairs of polynomials P and Q). Therefore, we get

$$\text{dist}_{\text{BW}}(P, \Sigma_{W,\underline{d}}) = \inf_{x \in S^n} \inf_{Q \in \Sigma_{W,\underline{d}}(e_0)} \left(\sum_{i=1}^m \left\| Y(P_i^x) - Y(Q_i) \right\|_{\text{BW}}^2 \right)^{1/2} \quad (24)$$

For $i = 1, \dots, m$, let

$$P_i^x(y) = \sum_{\alpha_1 + \dots + \alpha_n = d} c_\alpha(P_i^x) y^\alpha \quad \text{and} \quad Q(y) = \sum_{\alpha_1 + \dots + \alpha_n = d} c_\alpha(Q_i) y^\alpha$$

be the expansions of $P_i^x = P_i \circ g_x$ and Q_i (the i -th entry of Q) in the monomial basis. Then, following (24), we have

$$\begin{aligned} \text{dist}_{\text{BW}}(P, \Sigma_{W, \underline{d}}) &= \inf_{x \in S^n} \inf_{Q \in \Sigma_{W, \underline{d}}(e_0)} \left(\sum_{i=1}^m \sum_{|\alpha_0| \geq d-r-1} |c_\alpha(P_i^x) - c_\alpha(Q_i)|^2 \frac{\alpha_0! \cdots \alpha_n!}{d!} \right)^{1/2} \\ &\leq \inf_{x \in S^n} \inf_{Q \in \Sigma_{W, \underline{d}}(e_0)} \left(\sum_{i=1}^m \sum_{|\alpha_1| + \dots + |\alpha_n| \leq r+1} |c_\alpha(P_i^x) - c_\alpha(Q_i)|^2 \right)^{1/2}. \end{aligned}$$

Let $q = Q|_{S^n}$ and $p^x = P^x|_{S^n}$. Observe now that $\left(\sum_{i=1}^m \sum_{|\alpha_1| + \dots + |\alpha_n| \leq r+1} |c_\alpha(P_i^x) - c_\alpha(Q_i)|^2 \right)^{1/2}$ is the Frobenius norm of $\pi(j^{r+1}(p^x - q)(e_0))$, where, as before, π is the projection on $J_{e_0}^{r+1}(S^n, \mathbb{R}^m)$ that removes the base point. Since all norms on finite-dimensional spaces are equivalent, there exists c'_1 such that

$$\left(\sum_{i=1}^m \sum_{|\alpha_1| + \dots + |\alpha_n| \leq r+1} |c_\alpha(P_i^x) - c_\alpha(Q_i)|^2 \right)^{1/2} \leq c'_1 \hat{v}(j^{r+1}(p^x - q)(e_0)).$$

In particular, we have

$$\begin{aligned} \text{dist}_{\text{BW}}(P, \Sigma_{W, \underline{d}}) &\leq c'_1 \inf_{x \in S^n} \inf_{Q \in \Sigma_{W, \underline{d}}(e_0)} \hat{v}_{e_0}(j^{r+1}(p^x - q)(e_0)) \\ &= c'_1 \inf_{x \in S^n} \text{dist}^{\hat{v}}(j^{r+1}p^x(e_0), Z_W(e_0)) \quad \text{by the definition of } Z_W(e_0) \text{ in (21),} \\ &= c'_1 \inf_{x \in S^n} \text{dist}^{\hat{v}}(j^{r+1}p(x), Z_W(x)) \quad \text{by orthogonal invariance.} \end{aligned}$$

This proves (23).

For the proof of (22) we argue as follows. Recalling the definition of $\Sigma_W(x)$ given in (8) we have

$$\begin{aligned} \inf_{x \in S^n} \text{dist}^{\hat{v}}(j^{r+1}p(x), Z_W(x)) &= \inf_{x \in S^n} \inf_{f \in \Sigma_W(x)} \hat{v}(j^{r+1}(p - f)(x)) \\ &\leq \inf_{x \in S^n} \inf_{f \in \Sigma_W(x)} \sup_{y \in S^n} \hat{v}(j^{r+1}(p - f)(y)). \end{aligned}$$

By definition of the C^r norm (15) we have $\sup_{y \in S^n} \hat{v}(j^{r+1}(p-f)(y)) = \|f-p\|_{C^{r+1}}^v$, and therefore

$$\begin{aligned} \text{dist}_{\text{BW}}(P, \Sigma_{W, \underline{d}}) &\leq \inf_{x \in S^n} \inf_{f \in \Sigma_W(x)} \|f-p\|_{C^{r+1}}^v \\ &= \inf_{x \in S^n} \text{dist}_{C^{r+1}}^v(p, \Sigma_W(x)) \\ &= \text{dist}_{C^{r+1}}^v(p, \Sigma_W), \end{aligned}$$

where in the last line we have used the fact that $\Sigma_W = \bigcup_{x \in S^n} \Sigma_W(x)$. This gives (22). ■

6 Low-Degree Approximation

The goal of this section is proving Theorem 4. For this we need to estimate the probability for the following projection of a polynomial stay inside the stable neighborhood.

Let us recall from the introduction our definition of the projection operator on polynomial maps of smaller degree. For an integer $L \in \{0, \dots, d\}$, we set

$$p|_L := \sum_{\ell \leq L, d-\ell \in 2\mathbb{N}} p_\ell.$$

where $p = \sum_{d-\ell \in 2\mathbb{N}} p_\ell$ is the harmonic decomposition of (each component of) $p = P|_{S^n}$. We extend this definition to polynomial maps $p = (p_1, \dots, p_m)$ as done in (3). Next, we define the event of this projection to stay inside the stable neighborhood.

Definition 18 (Stability event). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type and $c_1 > 0$ be the constant given by the quantitative stability Theorem 7. For an integer $L \in \{0, \dots, d\}$, we denote by E_L the event $E_L = \{P \in \mathcal{P}_{n, \underline{d}} \mid \|p-p|_L\|_{C^{r+1}} < c_1 \text{dist}_{\text{BW}}(P, \Sigma_W)\}$, where, as before, $p = P|_{S^n}$.

Notice that the quantitative stability Theorem 7 implies that $E_L \subseteq A_L$ (the low-degree approximation event defined in Definition 3) and in particular a lower bound on the probability of E_L serves also as a lower bound for the probability of A_L .

Lemma 19. With the above notations, we have $E_L \subseteq A_L$.

Proof. Let $P \in E_L$ and $p = P|_{S^n}$. Theorem 7 implies $p|_L \in \text{SN}(p, W)$. Then $j^r p$ and $j^r p|_L$ are connected by a homotopy transversal to W . The conclusion follows from [40, Théorème 2.D.2]. ■

Note that type- W singular locus of polynomials in the event A_L is ambient isotopic to the type- W singular locus of a polynomial of degree L . In other words, polynomials in E_L and in A_L “look like” polynomials of lower degree.

For obtaining a bound on the probability of E_L we first need to prove the next proposition.

Proposition 20. Let $W \subseteq J^r(S^n, \mathbb{R}^m)$. There exist constants $a_1, a_2, a_3 > 0$, which depend on W , such that the following holds. For every $s \geq a_1 d^{a_2}$ we have for systems of Kostlan polynomials $P = (P_1, \dots, P_m) \in \mathcal{P}_{n, \underline{d}}$, with $\underline{d} = (d_1, \dots, d_n)$,

$$\mathbb{P}\{\|P\|_{\text{BW}} \leq s \operatorname{dist}_{\text{BW}}(P, \Sigma_W)\} \geq 1 - \frac{a_3 d^{a_2}}{s},$$

where $d = \max d_i$.

Proof. The proof of this proposition is the same as the proof of [6, Proposition 4], of which the current statement is just a simple generalization. Let Q be the polynomial given by Proposition 11. Then Σ_W is contained in $Z(Q)$, the zero set of Q , and we can apply [2, Theorem 21.1] as follows.

Let $N = \dim(\mathcal{P}_{n, \underline{d}}) = \sum_{k=1}^m \binom{d_k+n}{n} \leq d^{nm}$ and $D = \deg(Q)$. Denoting by d_{sin} the sine distance in the sphere, [2, Theorem 21.1] tells that there exists a constant $C_3 > 0$ such that for all $s \geq 2DN$ we have

$$\frac{\operatorname{vol}\left(\left\{p \in S^{N-1} \text{ such that } \frac{1}{d_{\text{sin}}(p, \bar{\Sigma})} \geq s\right\}\right)}{\operatorname{vol}(S^{N-1})} \leq C_3 D N s^{-1}.$$

Taking the cone over the set $\{p \in S^{N-1} \text{ such that } \frac{1}{d_{\text{sin}}(p, \bar{\Sigma})} \geq s\}$, we can rewrite the previous inequality in terms of the Kostlan distribution, obtaining that for all $s \geq 2DN$:

$$\mathbb{P}\{\|P\|_{\text{BW}} \geq s \cdot \operatorname{dist}_{\text{BW}}(P, \Sigma_W)\} \leq C_3 D N s^{-1}. \quad (25)$$

By Proposition 11, we have $D \leq u d^u$. This implies that for some constants $a_1, a_2, a_3 > 0$ we have

$$2DN \leq a_1 d^{a_2} \quad \text{and} \quad C_3 D N \leq a_3 d^{a_2}.$$

In particular (25) finally implies that for all $s \geq a_1 d^{a_2}$:

$$\mathbb{P}\left\{\|p\|_{\text{BW}} \leq s \cdot \text{dist}_{\text{BW}}(P, \Sigma_W)\right\} \geq 1 - a_3 d^{a_2} s^{-1}.$$

This finishes the proof. ■

The following theorem estimates the probability that the stability event holds.

Theorem 21 (Probability estimation). There exist constants $c_2, c_3, c_4 > 0$ (depending on W) such that for every $L \in \{0, \dots, d\}$ with $d - L \in 2\mathbb{N}$ we have

$$\mathbb{P}(E_L) \geq 1 - c_2 d^{c_3} L^{c_4} e^{-\frac{L^2}{3d}}.$$

Proof. By Proposition 15 there exists a constant $c = c(\ell, n, \nu)$ depending on ℓ, n, ν such that

$$\|p - p|_L\|_{C^{\ell+1}}^\nu \leq cd^{1/2} \|p - p|_L\|_{H^q} = (*)$$

Moreover, observe that the proof of [6, Proposition 2] works also for a polynomial list and gives the existence of a constant $c_2 = c_2(n) > 0$ (which depends on n) such that for all $t, q \geq 0$ and for every $L \in \{0, \dots, d\}$ we have that

$$(*) \leq cd^{1/2} t \|p\|_{\text{BW}} \quad \text{holds with probability } \mathbb{P}_1 \geq 1 - c_2 \frac{d^{-\frac{3n}{2}+1} L^{2q+n-2} e^{-\frac{L^2}{d}}}{t^2}.$$

At the same time, by Proposition 20, for every $s \geq a_1 d^{a_2}$:

$$\|p\|_{\text{BW}} \leq s \cdot \text{dist}_{\text{BW}}(P, \Sigma_W) \quad \text{holds with probability } \mathbb{P}_2 \geq 1 - a_3 \frac{d^{a_2}}{s}.$$

Now, choosing s, t as in the proof of [6, Theorem 5], there exist constants $c_1, c'_2, c'_3, c'_4 > 0$ such that for every $\sigma \geq 0$ we have $\|p - p|_L\|_{C^{r+1}} < c_1 \text{dist}(P, \Sigma_W)$ with probability

$$\mathbb{P}(E_L) \geq 1 - (1 - \mathbb{P}_1) - (1 - \mathbb{P}_2) \geq 1 - \left(c'_2 d^{c'_3} L^{c'_4} e^{-\frac{L^2}{d}} \sigma^2 + \frac{1}{\sigma}\right). \quad (26)$$

Let us denote now $\alpha(d, L) = c'_2 d^{c'_3} L^{c'_4} e^{-\frac{L^2}{d}}$, so that we can rewrite (26) as

$$\mathbb{P}(E_L) \geq 1 - \left(\alpha(d, L) \sigma^2 + \frac{1}{\sigma}\right),$$

for all $\sigma > 0$. In particular

$$\mathbb{P}(E_L) \geq 1 - \inf_{\sigma > 0} \left(\alpha(d, L)\sigma^2 + \frac{1}{\sigma} \right).$$

The above infimum is actually a minimum that is reached at

$$\sigma_0 = (2\alpha(d, L))^{-\frac{1}{3}}$$

and equals to $3 \left(\frac{\alpha(d, L)}{4} \right)^{1/3}$. Plugging the value of the infimum into (26) gives the claim. ■

The above estimate is very general, and one has to consider that we would like to have L as small as possible and at the same time σ as large as possible. We approach this issue as follows.

First, we choose L to be of the order $O(\sqrt{d \log d})$ and we prove that with fast growing probability $p|_L$ stably approximates p ; in particular, Theorem 7 implies that for a given W , the W -type singular locus of a random Kostlan polynomial of degree d , with high probability as d grows rapidly, is ambient isotopic to the W -type singular locus of a polynomial of degree $O(\sqrt{d \log d})$. In the 2nd step, we deal with exponential rarefactions of maximal configurations. More precisely, when choosing L to be a root or a fraction of d , we can tune σ so that the probability of stably approximate goes to 1 exponentially fast.

This distinction is the motivation for having three different regimes in Theorem 4. Now we give a proof for this Theorem.

Proof of Theorem 4. Since $E_L \subseteq A_L$, and by Theorem 21 we have

$$\mathbb{P}(A_L) \geq \mathbb{P}(E_L) \geq 1 - c_2 d^{c_3} L^{c_4} e^{-\frac{L^2}{3d}}.$$

In the rest of the proof we plug in different values for L and evaluate the right hand side of this inequality. ■

Proof of Theorem 4.1. Let $L = b\sqrt{d \log d}$. We see that with this choice

$$\begin{aligned} \mathbb{P}(A_L) &\geq 1 - c_2 d^{c_3} (b\sqrt{d \log d})^{c_4} e^{-\frac{b^2 \log d}{3}} \\ &= 1 - c_2 b^{c_4} d^{c_3 + c_4/2 - b^2/3} (\log d)^{c_4/2} \\ &\geq 1 - c_2 b^{c_4} d^{c_3 + c_4 - b^2/3}. \end{aligned} \tag{27}$$

Thus, if $b > b_0 := \sqrt{3(c_3 + c_4)}$, then we have

$$\mathbb{P}(A_L) \geq 1 - a_1 d^{-a_2},$$

where $a_1 = c_2 b^{c_4} > 0$ and $a_2 = -(c_3 + c_4 - b^2/3) > 0$. This proves the 1st part of the theorem. ■

Proof of Theorem 4.2 Let $L = d^b$ for some $b \in (1/2, 1)$. With this choice we have

$$\mathbb{P}(A_L) \geq 1 - c_2 d^{c_3 + bc_4} e^{-\frac{d^{2b-1}}{3}} = 1 - c_2 \frac{d^{c_3 + bc_4}}{e^{\frac{d^{2b-1}}{6}}} e^{-\frac{d^{2b-1}}{6}}.$$

Since $2b - 1 > 0$, we have that $\frac{d^{c_3 + bc_4}}{e^{\frac{d^{2b-1}}{6}}}$ is bounded and so

$$\mathbb{P}(A_L) \geq 1 - a_1 e^{-d^{a_2}}$$

for some $a_1, a_2 > 0$ with $a_2 < 1$, because $2b - 1 < 1$. This proves the 2nd part of the theorem. ■

Proof of Theorem 4.3. Let $L = bd$ with $b \in (0, 1)$. With this choice we have

$$\mathbb{P}(A_L) \geq 1 - c_2 b^{c_4} d^{c_3 + c_4} e^{-\frac{b^2 d}{3}} = c_2 b^{c_4} \frac{d^{c_3 + c_4}}{e^{\frac{b^2 d}{6}}} e^{-\frac{b^2 d}{6}}.$$

We have that $\frac{d^{c_3 + c_4}}{e^{\frac{b^2 d}{6}}}$ is bounded and so

$$\mathbb{P}(A_L) \geq 1 - a_1 e^{-a_2 d}$$

for some $a_1, a_2 > 0$. This proves the 3rd part of the theorem. ■

Proof of Theorem 4.4. Let $a > 0$ and set $L = b\sqrt{d \log d}$. As in (27) we get

$$\mathbb{P}(A_L) \geq 1 - c_2 b^{c_4} d^{c_3+c_4-b^2/3}.$$

If we pick b and d large enough, such that $c_2 b^{c_4} d^{c_3+c_4-a-b^2/3} \leq 1$, then

$$\mathbb{P}(A_L) \geq 1 - d^{-a}.$$

This proves the fourth part of the theorem. The proof of points (5) and (6) proceeds along the same lines. \blacksquare

7 Maximal Configurations Are Rare Events

In this section we provide a couple of examples of tail probabilities, proving exponential rarefaction of maximal configurations. For a semialgebraic set Y , we denote by $b(Y)$ the sum of its Betti numbers (this sum is finite by semialgebraicity).

Proposition 22 (Exponential rarefaction of maximal configurations). Let $W \subseteq J^r(S^n, \mathbb{R}^m)$ be a singularity type. For every $C > 0$ there exist $a_1, a_2 > 0$ such that

$$\mathbb{P}\left(\left\{p \in \mathcal{P}_{n,d} \mid b(j^r p^{-1}(W)) \geq Cd^n\right\}\right) \leq a_1 e^{-a_2 d}.$$

Proof. Let $W \subseteq J^r(S^n, \mathbb{R}^m)$. We recall that if $q : S^n \rightarrow \mathbb{R}^m$ is a polynomial map with degree bounded by L such that $j^r q \pitchfork W$, the following estimate is proved in [27] for the sum of the Betti numbers of $j^r q^{-1}(W)$:

$$b(j^r q^{-1}(W)) \leq C_1 L^n.$$

In particular we see that if for $p \in \mathcal{P}_{n,d}$ we have $j^r p \pitchfork W$ (which happens with probability 1 by [27, Theorem 1, point (4)]) and $b(j^r p^{-1}(W)) \geq Cd^n$, then if moreover $p \in A_L$ (the stability event from Definition 3) we must have $L \geq \left(\frac{C}{C_1}\right)^{1/n} d$. Set now

$$L = \min \left\{ d, \left\lfloor \left(\frac{C}{C_1}\right)^{1/n} d \right\rfloor + 1 \right\}.$$

Then for every $p \in A_L$, we have

$$b(j^r p^{-1}(W)) = b(j^r p|_L^{-1}(W)) < Cd^n.$$

In particular, we must have

$$\{b(j^r p^{-1}(W) \geq Cd^n)\} \subseteq (A_L)^c \quad (28)$$

(the complement of the stability event A_L). Point (3) of Theorem 4 implies that there exists constants $a_1, a_2 > 0$ such that $\mathbb{P}(A_L) \geq 1 - a_1 e^{-a_2 d}$, which combined with (28) gives

$$\mathbb{P}(\{b(j^r p^{-1}(W) \geq Cd^n)\}) \leq a_1 e^{-a_2 d}.$$

This proves the proposition. ■

We can apply the previous result to the examples discussed in the Introduction.

1. In the case $W = S^n \times \{0\} \subset J^0(S^n, \mathbb{R}^m)$, Harnack's bound implies that

$$b(j^0 p^{-1}(W)) = b(\{p = 0\}) \leq O(d^n)$$

and the probability that $b(\{p = 0\}) \geq C_3 d^n$ is exponentially small. This example is extensively discussed in [6].

2. In the case $W = \{j^1 f = 0\} \subset J^1(S^n, \mathbb{R})$ then $j^1 f^{-1}(W)$ is the set of critical points of $p : S^n \rightarrow \mathbb{R}$; by [5] for a polynomial P of degree d

$$\#j^1 p^{-1}(W) \leq 2(d-1)^n + \dots + (d-1) + 1$$

(this bound follows from complex algebraic geometry), and this estimate was recently proved to be sharp by Kozhasov [22]. In this context [27, Theorem 14] tells that the expected number of critical points is of the order $\Theta(d^{n/2})$ and Theorem 22 implies that for every $C > 0$ there exists $a_1, a_2 > 0$ such that the probability that a random polynomial of degree d has at least Cd^n many critical points is smaller than $a_1 e^{-a_2 d}$.

3. $W = \{df = 0, d^2 f > 0\} \subset J^2(S^n, \mathbb{R})$, and $j^2 f^{-1}(W)$ is the set of non-degenerate *minima* of $f : S^n \rightarrow \mathbb{R}$; the same argument can be applied here: the expectation of the number of minima is of the order $\Theta(d^{n/2})$ ([27, Theorem 14]). There are polynomials of degree d with $\Theta(d^n)$ many non-degenerate minima, but the measure of the sets of such polynomials becomes exponentially small as d grows.

4. $W = \{\text{Whitney cusps}\} \subset J^3(S^2, \mathbb{R}^2)$, and $j^3 f^{-1}(W)$ is the set of points where $f : S^2 \rightarrow \mathbb{R}^2$ has critical points of type Whitney cusp, that is, in some coordinates near this point the map has the form $(x_1, x_2) \mapsto (x_1, x_2^3 - x_1 x_2)$, see [4] (we need a condition on the 3rd jet in order to make sure that this local form exists). The number of such cusps, for a polynomial map of degree d , is $O(d^2)$. On average there are $\Theta(d)$ many of them and the probability of having $\Theta(d^2)$ Whitney cusps becomes exponentially small as d grows.

Remark 23. By applying points (1) and (2) from Theorem 4 one can obtain other similar tail probabilities. For example, choosing $L = b\sqrt{d \log d}$ as in Theorem 4 we get that

$$b(Z(p)) \leq O((d \log d)^{n/2}) \quad \text{with } \mathbb{P} \rightarrow 1 \text{ as } d \rightarrow \infty.$$

This result follows using the same ideas as before, combined with the deterministic estimate $b(Z(Q)) \leq O(L^n)$ for the sum of the Betti numbers of the zero set of a polynomial Q of degree L in projective space (see [28]).

Acknowledgments

This article was written in parts while the 3rd author was on a visiting fellowship of the Max-Planck Institute for Mathematics in the Sciences in Leipzig. The authors also wish to thank the anonymous referees, whose comments and suggestions improved the structure of the paper.

References

- [1] Axler, S., P. Bourdon, and W. Ramey. *Harmonic Function Theory*, vol. 137 of Graduate Texts in Mathematics, 2nd edn. New York: Springer, 2001.
- [2] Bürgisser, P. and F. Cucker. *Condition: The Geometry of Numerical Algorithms*, vol. 349 of Grundlehren der Mathematischen Wissenschaften. Heidelberg: Springer, 2013.
- [3] Bochnak, J., M. Coste, and M. F. Roy. *Real Algebraic Geometry*, vol. 36 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3). Berlin: Springer, 1998.
- [4] Callahan, J. "Singularities and plane maps." *Amer. Math. Monthly* 81 (1974): 211–40.
- [5] Cartwright, D. and B. Sturmfels. "The number of eigenvalues of a tensor." *Linear Algebra Appl.* 438, no. 2 (2013): 942–52.
- [6] Diatta, D-N. and A. Lerario. "Low degree approximation of random polynomials." (2018): preprint arXiv:1812.10137.
- [7] Eliashberg, Y. and N. Mishachev. *Introduction to the h-Principle*, vol. 48 of Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2002.

- [8] Fyodorov, Y. V., A. Lerario, and E. Lundberg. "On the number of connected components of random algebraic hypersurfaces." *J. Geom. Phys.* 95 (2015): 1–20.
- [9] Gudkov, D. "Ovals of sixth order curves." *Nine Papers on Hilbert's 16th Problem*, vol. 112. 9–14. American Mathematical Society Translations, 1978.
- [10] Gayet, D. and J.-Y. Welschinger. "Lower estimates for the expected Betti numbers of random real hypersurfaces." *J. Lond. Math. Soc.* 90 (2014): 105–20.
- [11] Gayet, D. and J.-Y. Welschinger. "Expected topology of random real algebraic submanifolds." *J. Inst. Math. Jussieu* 14 (2015): 673–702.
- [12] Gayet, D. and J.-Y. Welschinger. "Betti numbers of random real hypersurfaces and determinants of random symmetric matrices." *J. Eur. Math. Soc.* 18 (2016): 733–72.
- [13] Heintz, J. "Definability and fast quantifier elimination in algebraically closed fields." *Theoret. Comput. Sci.* 24, no. 3 (1983): 239–77.
- [14] Hilbert, D. "Über die reellen züge algebraischer curven." *Math. Annalen* 38 (1891): 115–38.
- [15] Hirsch, M. W. *Differential Topology*, vol. 33 of Graduate Texts in Mathematics. New York: Springer, 1994. Corrected reprint of the 1976 original.
- [16] Kharlamov, V. "Isotopic types of nonsingular surfaces of degree 4 in $\mathbb{R}P^3$." *Funct. Anal. Appl.* 12 (1978): 86–7.
- [17] Kharlamov, V. "Rigid classification up to isotopy of real plane curves of degree 5." *Funct. Anal. Appl.* 15 (1981): 73–4.
- [18] Kharlamov, V. "Classification of nonsingular surfaces of degree 4 in $\mathbb{R}P^3$ with respect to rigid isotopies." *Funct. Anal. Appl.* 18 (1984): 39–45.
- [19] Kaihnsa, N., M. Kummer, D. Plaumann, M. Sayyary Namin, and B. Sturmfels. "Sixty-four curves of degree six." *J. Exp. Math.* 28 (2019): 132–50.
- [20] Klein, F. "A comparative review of recent researches in geometry." *Bull. Amer. Math. Soc.* 2, no. 10 (1893): 215–49.
- [21] Kostlan, E. "On the Distribution of Roots of Random Polynomials." In *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*. 419–31. New York: Springer, 1993.
- [22] Kozhasov, K. "On fully real eigenconfigurations of tensors." *SIAM J. Appl. Algebra Geom.* 2, no. 2 (2018): 339–47.
- [23] Lerario, A. "Random matrices and the average topology of the intersection of two quadrics." *Proc. Amer. Math. Soc.* 143 (2015): 3239–325.
- [24] Lerario, A. and E. Lundberg. "Statistics on Hilbert's 16th problem." *Int. Math. Res. Not. IMRN* 12 (2015): 4293–321.
- [25] Lerario, A. and E. Lundberg. "Gap probabilities and Betti numbers of a random intersection of quadrics." *Discrete Comput. Geom.* 55 (2016): 462–96.
- [26] Lerario, A. and E. Lundberg. "On the geometry of random lemniscates." *Proc. Lond. Math. Soc.* 113 (2016): 649–73.
- [27] Lerario, A. and M. Stecconi. "Differential topology of gaussian random fields: applications to random algebraic geometry." (2019).

- [28] Milnor, J. "On the Betti numbers of real varieties." *Proc. Amer. Math. Soc.* 15 (1964): 275–80.
- [29] Nikulin, V. "Integer symmetric bilinear forms and some of their geometric applications." *USSR-Izv.* 14 (1980): 103–67.
- [30] Nazarov, F. and M. Sodin. "On the number of nodal domains of random spherical harmonics." *Amer. J. Math.* 131 (2009): 1337–57.
- [31] Nazarov, F. and M. Sodin. "Asymptotic laws for the spatial distribution and the number of connected components of zero sets of Gaussian random functions." *Zh. Mat. Fiz. Anal. Geom.* 12 (2016): 205–78.
- [32] Orevkov, S. Y. and V. M. Kharlamov. "Growth order of the number of classes of real plane algebraic curves as the degree grows." *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 266 (Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 5) (2000): 218–33, 339.
- [33] Petrovsky, I. "On the topology of real plane algebraic curves." *Ann. Math.* 39 (1938): 189–209.
- [34] Rohn, K. "Die Maximalzahl und Anordnung der Ovale bei der ebenen Kurve 6. Ordnung und bei der Fläche 4. Ordnung." *Math. Annalen* 73 (1913): 177–229.
- [35] Rokhlin, V. "Complex topological characteristics of real algebraic curves." *Russian Math. Surveys* 33 (1978): 85–98.
- [36] Sarnak, P. "Letter to B. Gross and J. Harris on Ovals of Random Planes Curve." <http://publications.ias.edu/sarnak/section/515>, 2011.
- [37] Seeley, R. T. "Spherical harmonics." *Amer. Math. Monthly* 73 (4, part II) (1966): 115–21.
- [38] Shafarevich, I. R. *Basic Algebraic Geometry*, vol. 1, 3rd ed. Heidelberg: Springer, 2013. Varieties in projective space.
- [39] Sarnak, P. and I. Wigman. "Topologies of nodal sets of random band limited functions." *Contemp. Math.* 664 (2016): 351–65.
- [40] Thom, R. "Ensembles et morphismes stratifiés." *Bull. Amer. Math. Soc.* 75 (1969): 240–84.
- [41] Varchenko, A. N., V. I. Arnold, and S. M. Gusein-Zade. *Singularities of Differentiable Maps*. Monographs in Mathematics. Boston: Birkhäuser, 1985.
- [42] Viro, O. *Gluing of Plane Real Algebraic Curves and Constructions of Curves of Degrees 6 and 7*. Topology (Leningrad, 1982), Lecture Notes in Math. Berlin: Springer, 1984.
- [43] Viro, O. "Progress in the topology of real algebraic varieties over the last six years." *Russian Math. Surveys* 41 (1986): 55–82.
- [44] Viro, O. "From the sixteenth Hilbert problem to tropical geometry." *Japanese J. Math.* 3 (2008): 185–214.