WILKIE'S CONJECTURE FOR PFAFFIAN STRUCTURES

GAL BINYAMINI, DMITRY NOVIKOV, AND BENNY ZAK

Abstract. We prove an effective form of Wilkie's conjecture in the structure generated by restricted sub-Pfaffian functions: the number of rational points of height H lying in the transcendental part of such a set grows no faster than some power of $\log H$. Our bounds depend only on the Pfaffian complexity of the sets involved. As a corollary we deduce Wilkie's original conjecture for $\mathbb{R}_{\exp}$ in full generality.

1. Introduction

1.1. **Sharply o-minimal structures.** Our general results in this paper are stated in the context of *sharply o-minimal structures* admitting *sharp derivatives*. For a complete definition of these notions see §2. Briefly, a sharply o-minimal structure (abbreviated #o-minimal) is a pair $(\mathcal{S}, \Omega)$ where $\mathcal{S}$ is an o-minimal expansion of the real field and we identify $\mathcal{S}$ with the collection of all definable sets in $\mathcal{S}$; and $\Omega = \{\Omega_{\mathcal{F},D} \subset \mathcal{S}\}_{\mathcal{F},D\in\mathbb{N}}$ is filtration on $\mathcal{S}$ depending on the *format* $\mathcal{F}$ and *degree* D. The filtration is increasing with respect to each parameter, and $\mathcal{S} = \cup_{\mathcal{F},D}\Omega_{\mathcal{F},D}$. If $X \in \Omega_{\mathcal{F},D}$ we say that X has format $\mathcal{F}$ and degree D (but note that these are not uniquely defined, as Ω is a filtration rather than a decomposition). One can roughly think of $\mathcal{F}$ as encoding the "logical complexity", e.g. the length of a formula used to define X; and of D as the sum of the algebraic degrees of all polynomials appearing in such a representation. A system of axioms analogous to the standard axioms of o-minimality control how the format and degree grows under the standard logical operations, and give upper bounds on the number of intervals in a definable subset of $\mathbb{R}$ in terms of $\mathcal{F}$ and D.

The structure (S, Ω) is said to have sharp derivatives if the following holds. Let $f : \mathbb{R}^n \to \mathbb{R}$ be definable with format $\mathcal{F}$ and degree D, meaning that its graph is in $\Omega_{\mathcal{F},D}$, and let $\alpha \in \mathbb{N}^n$. Then the derivative $f^{(\alpha)}$ (wherever it is defined) has format depending only on $\mathcal{F}$, and degree depending polynomially on D and $|\alpha|$. The structure $\mathbb{R}_{\mathrm{rPfaff}}$ of restricted sub-Pfaffian sets is an example of a #o-minimal structure with sharp derivatives, see §3. The unfamiliar reader may keep this example in mind in place of the general setting.

2010 *Mathematics Subject Classification.* Primary 03C64,11U09; Secondary 11G50, 14P10.

Key words and phrases. Yomdin-Gromov lemma, point-counting, Pila-Wilkie theorem, o-minimality.

This research was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1167/17) and by funding received from the MINERVA Stiftung with the funds from the BMBF of the Federal Republic of Germany. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 802107).

1.2. **Asymptotic notation and effectivity.** In this paper each appearance of an expression $Z = O_X(Y)$ should be interpreted as shorthand notation for $Z \leqslant C_X \cdot Y$ where $X \to C_X$ is a universally fixed, positive valued real function. Similarly we write $Z = \mathrm{poly}_X(Y)$ as shorthand for $Z \leqslant P_X(Y)$ where $X \to P_X$ is a universally fixed mapping and P_X is a polynomial with positive coefficients. However we suppress dependence of the constants on $(\mathcal{S}, \Omega)$, which we consider to be universally fixed throughout the text.

The definition of a sharply o-minimal structure postulates the existence of a certain upper bound $P_{\mathcal{F}}(\cdot)$ for the number of connected components of a definable set as a function of format and degree. If this upper bound is known effectively, then all asymptotic constants in this paper are also effectively computable unless it is explicitly stated otherwise. In particular, this is indeed the case for the restricted sub-Pfaffian structure $\mathbb{R}_{\mathrm{rPfaff}}$.

1.3. **Main results.** If $X \subset \mathbb{R}^n$ then following [19, 18] we denote by X^{alg} the union of all connected, positive-dimensional semialgebraic sets contained in X, and denote $X^{\mathrm{trans}} := X \setminus X^{\mathrm{alg}}$. For $g, H \in \mathbb{N}$ we denote

$$X(g,H) := \{x \in X : [\mathbb{Q}(x) : \mathbb{Q}] \leqslant g,\ H(x) \leqslant H\}, \qquad X(\mathbb{Q},H) := X(1,H) \quad (1)$$

where $H(\cdot)$ denotes the multiplicative Weil height on $\bar{\mathbb{Q}}$, extended to $\bar{\mathbb{Q}}^n$ as the maximum of the heights of the coordinates. If unfamiliar see §1.2 for the asymptotic notation used below.

Theorem 1. *Let $X \in \Omega_{\mathcal{F},D}$. Then*

$$\#X^{\mathrm{trans}}(g,H) \leqslant \mathrm{poly}_{\mathcal{F}}(D, g, \log H). \quad (2)$$

This establishes, in the restricted Pfaffian setting, a conjecture by Pila [21, Conjecture 1.5]. As a direct corollary we obtain Wilkie's conjecture. Note that since $\mathbb{R}_{\exp}$ is not currently known to be #o-minimal, we allow the asymptotic constant in (3) to depend on the definable set X, not only its format as in Theorem 1.

Corollary 1 (Wilkie's conjecture)**.** *Let X be definable in $\mathbb{R}_{\exp}$. Then*

$$\#X^{\mathrm{trans}}(g,H) \leqslant \mathrm{poly}_X(g, \log H). \quad (3)$$

We stress that unlike in Theorem 1, the asymptotic constants here are not *effectively computable.*

Proof. We adapt the exhaustion argument introduced in [15] in the quantifier free setting and in [1] for general definable sets. By Wilkie's theorem of the complement [29] we have $X = \pi_n(Y)$ where $Y \subset \mathbb{R}^N$ is quantifier-free in $\mathbb{R}_{\exp}$, and $\pi_n : \mathbb{R}^N \to \mathbb{R}^n$ is the projection to the first n coordinates. Let $g, H \in \mathbb{N}$ and choose $M \gg 1$ such that $X(g,H) = X_M(g,H)$ where

$$X_M := \pi_n(Y_M), \quad Y_M := Y \cap [-M, M]^N. \quad (4)$$

Now Y_M is restricted semi-Pfaffian, as it is defined by Pfaffian functions (exponential polynomials) restricted to $[-M,M]^N$. Crucially, $Y_M \in \Omega_{\mathcal{F},D}$ where $\mathcal{F}, D$ depend on Y but not on M. Then the same is true for X_M, and we conclude

$$\#X^{\mathrm{trans}}(g,H) \leqslant \#X_M^{\mathrm{trans}}(g,H) = \mathrm{poly}_X(g, \log H) \quad (5)$$

by Theorem 1. $\square$

In fact, by a result of van den Dries and Miller [26] the structure $\mathbb{R}_{\mathrm{rPfaff,exp}}$ is also model-complete, and Corollary 1 thus holds in this larger structure with the same proof. Moreover if $X := \{X_\lambda\}_{\lambda\in\Lambda}$ is a definable family of sets in this structure then

$$\#X_\lambda^{\mathrm{trans}}(g,H) \leqslant \mathrm{poly}_X(g,\log H), \qquad \forall\lambda\in\Lambda \tag{6}$$

where the asymptotic constants are independent of λ, by essentially the same argument. We thank Alex Wilkie and Raf Cluckers for bringing these generalizations to our attention.

We also have a "blocks" generalization of Theorem 1. Recall from [22] that a definable set $B \subset \mathbb{R}^n$ is called a *basic block* if it is connected and regular (i.e., C^1-smooth), and contained in a connected regular semialgebraic set of the same dimension (which we call a semialgebraic closure of B, though this is not uniquely defined). We denote by Ω^{alg} a filtration making $(\mathbb{R}_{\mathrm{alg}},\Omega^{\mathrm{alg}})$ into a $\#$o-minimal structure (one can take, e.g., the filtration from [8] for the empty Pfaffian chain).

Theorem 2. *Let $X \subset \mathbb{R}^n$ with $X \in \Omega_{\mathcal{F},D}$. Then there exists a collection $\{B_\eta \subset X\}$ of basic blocks with semialgebraic closures S_η such that $X(g,H) \subset \cup_\eta B_\eta$ and*

$$\#\{B_\eta\} = \mathrm{poly}_{\mathcal{F}}(D,g,\log H), \qquad \forall\eta : S_\eta \in \Omega^{\mathrm{alg}}_{O_n(1),\mathrm{poly}_n(g,\log H)}. \tag{7}$$

Theorem 2 clearly implies Theorem 1, since the positive-dimensional basic blocks B_η are subsets of X^{alg} by definition.

1.4. **A C^r-parametrization lemma.** For a C^r-smooth function $f : U \to \mathbb{R}$ on a domain $U \subset \mathbb{R}^m$ we denote

$$\|f\| := \sup_{x\in U} |f(x)|, \qquad\qquad \|f\|_r := \max_{|\boldsymbol{\alpha}|\leqslant r} \|D^{\boldsymbol{\alpha}} f\|. \tag{8}$$

For $F : U \to \mathbb{R}^n$ we set $\|F\| = \max_i \|F_i\|$ and similarly for $\|F\|_r$. For a set $A \subset \mathbb{R}^n$ we write $U_\varepsilon(A)$ for the ε-neighborhood of A with the ℓ_∞-metric. For $A, B \subset \mathbb{R}^n$, we write $A \subseteqq_\varepsilon B$ to mean that $A \subset B$ and $B \subset U_\varepsilon(A)$. We say that A is an *ε-cover* of B.

The main novelty of our approach is the following version of Yomdin's algebraic lemma. Let $I := (0,1)$.

Lemma 2. *Let $r \in \mathbb{N}$ and $\varepsilon > 0$. Let $X \subset [0,1]^n$ be of dimension μ with $X \in \Omega_{\mathcal{F},D}$. Then there exists a collection $\{\phi_\eta : I^\mu \to X\}$ of C^r-smooth maps such that $\|\phi_\eta\|_r \leqslant 1$ and $\cup_\eta \operatorname{Im}\phi_\eta \subseteqq_\varepsilon X$, and*

$$\#\{\phi_\eta\} \leqslant \mathrm{poly}_{\mathcal{F}}(D,r,|\log\varepsilon|), \qquad \forall\eta : \phi_\eta \in \Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}. \tag{9}$$

This formulation is similar in spirit to Yomdin's original formulation [32, 31]. Gromov [13] later refined Yomdin's work by showing that one can formally take $\varepsilon = 0$, avoiding ε-covers and covering the set X completely. However, as we will see, Lemma 2 is sufficient for the applications in Diophantine geometry (as it was for Yomdin's original application in dynamics). The weaker formulation with ε-covers enables us (as was already the case in Yomdin's original work) to restrict to *affine* reparametrizations at some crucial moments, where Gromov's approach involves nonlinear terms. The linearity of the reparametrizing maps turns out to allow for crucial technical simplifications related to achieving polynomial growth of $\#\{\phi_\eta\}$ as a function of r (specifically in §5.4)

Remark 3. *In* [33] *Yomdin considers the case where the parameterizing maps are algebraic, or more generally p-valent with uniform p. He proves that in this case if one takes $\varepsilon = 0$ then the size of $\#\{\phi_\eta\}$ cannot be taken independent of r. The same conclusion is established without the valency assumption in* [5, Section 1.2.1]*. However, with $\varepsilon > 0$ it is still plausible to expect a version of Lemma 2 to hold with no dependence on r. In fact, one may hope to find parameterizing maps with $\|\phi_\eta\|_\infty < 1$. This was recently proved in the algebraic category, albeit without polynomial dependence on D, by Cluckers, Friedland and Yomdin* [10]*. Improving this to depend polynomially on D, and possibly to the context of other #o-minimal structures, would yield an interesting sharpening of Lemma 2.*

1.5. **Background.**

1.5.1. *The Pila-Wilkie theorem.* The origin of the area of point-counting in tame geometry can be traced to the work of Bombieri and Pila [9, 17]. In these papers it was shown that if $\Gamma \subset \mathbb{R}^2$ is a compact analytic curve containing no semialgebraic curves then for every $\varepsilon > 0$ one has $\#\Gamma(\mathbb{Q}, H) = O_{\Gamma,\varepsilon}(H^\varepsilon)$. After some work by Pila on subanalytic surfaces [19, 20], this result was generalized into its canonical form by Pila and Wilkie [18], who proved that the bound $\#X^{\text{trans}}(\mathbb{Q}, H) = O_{X,\varepsilon}(H^\varepsilon)$ holds for *any* X definable in an o-minimal structure. This result has had a profound impact on arithmetic geometry, and we refer the reader to [25] for a survey. Note that the asymptotic constants in both of these theorems are not effective.

For Pfaffian surfaces, Jones and Thomas [15] established an effective form of the Pila-Wilkie theorem. In the general restricted sub-Pfaffian setting, a recent paper by the first author with Jones, Schmidt and Thomas [1] establishes an effective form of the Pila-Wilkie theorem: if $X \in \Omega_{\mathcal{F},D}$ then in fact $\#X^{\text{trans}}(\mathbb{Q}, H) = \text{poly}_{\mathcal{F},\varepsilon}(D)$ with effective constants. Many of the technical methods for using #o-minimality in our context are inspired by this prior work. Indeed the paper [8] which inspired the notion of #o-minimality grew out of an attempt to provide a suitable foundation for the results in [1].

1.5.2. *The Wilkie conjecture.* Examples by Pila [19, Example 7.5] show that in $\mathbb{R}_{\text{an}}$ the Pila-Wilkie asymptotic is essentially optimal. However, such examples involve “hand-crafted” functions and no “natural” example exhibiting this behavior is known. Wilkie made his conjecture (now Corollary 1) in the original paper [18] as a concrete formulation of this phenomenon. The case of Pfaffian curves was proved by Pila [21], and our approach in the one-dimensional case is indeed somewhat similar to Pila’s approach. Some further examples of surfaces were treated in [23]. Two alternative more sophisticated parametrization techniques were employed in [11] to establish Wilkie’s conjecture for general families of $\mathbb{R}_{\text{an}}$-definable Pfaffian surfaces. However, general definable surfaces in $\mathbb{R}_{\exp}$ already seem difficult to treat with the existing approaches.

The key obstacle to progress on Wilkie’s conjecture has been to establish a C^r-parametrization lemma with polynomial bounds for the number of charts, as a function of the complexity of the set and the smoothness order r. A part of the difficulty was that no fully appropriate notion of complexity was available in $\mathbb{R}_{\exp}$ or $\mathbb{R}_{\text{rPfaff}}$, and the introduction of sharp o-minimality was meant to remedy this situation. However the difficulty of efficient C^r-parametrization goes beyond this issue. In fact, even in the semialgebraic case where suitable notions of format and degree have long been available, this problem was only recently resolved in [5] using complex analytic methods. The problem remains open beyond the semialgebraic case, and the complex analytic methods seem unlikely to directly carry over to the unrestricted exponential case. We remark that for general subanalytic sets (and some more general types of sets), while there is no notion of complexity, one can still ask for polynomial dependence on r. This is achieved in [11, 27, 28].

For general definable sets in an o-minimal structure, the only previously known case of the Wilkie conjecture is [4] by the first two authors. This paper established Wilkie's conjecture for the structure $\mathbb{R}^{\mathrm{RE}}$ generated by the exponential and sine functions *restricted* to compact domains. The proofs were based on an approach avoiding smooth parametrizations altogether, replacing it by complex-geometric ideas. This is only applicable for *holomorphic-Pfaffian* functions, i.e. holomorphic functions whose graph, viewed as a real set, is Pfaffian in the real sense. By comparison, our approach here applies to arbitrary restricted sub-Pfaffian functions without requiring that the complex-analytic continuation is again Pfaffian. The complex-geometric ideas also seem much more difficult to carry out in the presence of unrestricted exponentials. We also point out that a non-archimedean analog of Wilkie's conjecture for Pfaffian varieties over the field $\mathbb{C}(\!(t)\!)$ has been established in [3], and the ideas developed in the present paper may also be of relevance in this and other non-archimedean contexts.

Remark 4. *The proof of Corollary 1 would not be applicable with the methods of* [4] *because the complex analytic nature of these methods would require us to consider* $e^{z_1},\dots,e^{z_M}$ *restricted to large* complex polydiscs $D(0,M)^N$ *rather than large cubes* $[-M,M]^N$ *as we do here. However, while the Pfaffian complexity of* e^x *on* $[-M,M]$ *is bounded independently of* M*, the Pfaffian complexity of* e^z *on* $D(0,M)$ *is roughly* M *as evidenced by the fact that the Pfaffian equation* $e^z=1$ *admits roughly* M/π *solutions in* $D(0,M)$*.*

1.5.3. *Unrestricted exponentials in arithmetic applications.* Unrestricted exponentials are used in many of the most spectacular applications of the Pila-Wilkie theorem, where they arise in uniformizing maps of arithmetic quotients around cusps. Extending the more advanced counting techniques to this case is therefore potentially very useful. In particular, a recent paper by the first author [2] establishes a polylogarithmic counting result in the spirit of Wilkie's conjecture for sets defined using algebraic foliations (not necessarily Pfaffian) over number fields. This result has played an important role in recent progress on the André-Oort conjecture for general Shimura varieties. It was used by the first author, Schmidt and Yafaev [7] to establish Galois lower bounds for special points conditional on certain height bounds. These height bounds were subsequently proved by Pila, Shankar and Tsimerman (with an appendix by Esnault and Groechenig) in [24] thus finishing the proof of André-Oort in general.

The approach of [2] is based on the complex geometric ideas of [4] and suffers from the same limitation to restricted analytic situations, and this leads to technical complications in [7] and in further potential applications of this result in arithmetic geometry. It seems plausible that the new approach developed in the present paper could also lead to progress on unrestricted exponentials in this non-Pfaffian situation, and we have formulated our results in the more general #o-minimal context with this in mind.

1.6. **Acknowledgments.** It is our pleasure to thank Yosi Yomdin for insightful discussions on the algebraic lemma, and Alex Wilkie for alerting us of the potential relevance of his notes [30]. In these notes Wilkie introduces a method for obtaining C^r-parametrizations in the one-dimensional case with a single reparametrization, rather than by the more traditional induction on r. While we did not eventually use this directly in our text, our approach is a kind of discrete version of this idea (so that we can use linear reparametrizations similar to Yomdin's approach) and certainly inspired by it. Pila [21] has used a similar approach earlier for Pfaffian curves, and his idea also inspires our approach. We note further that the interpolation method that we use to control $X(g,H)$ efficiently as a function of g was also introduced in Wilkie's important notes [30].

2. Sharply o-minimal structures

2.1. **#o-minimal structures.** In this section we introduce the notion of a *sharply o-minimal structure* (abbreviated *#o-minimal*). To start, a *format-degree* filtration (abbreviated *FD*-filtration) on a structure $\mathcal{S}$ is a collection $\Omega = \{\Omega_{\mathcal{F},D}\}_{\mathcal{F},D\in\mathbb{N}}$ such that each $\Omega_{\mathcal{F},D}$ is a collection of definable sets (possibly of different ambient dimensions), with $\Omega_{\mathcal{F},D} \subset \Omega_{\mathcal{F}+1,D} \cap \Omega_{\mathcal{F},D+1}$ and $\cup_{\mathcal{F},D}\Omega_{\mathcal{F},D}$ is the collection of all definable sets in $\mathcal{S}$. We call the sets in $\Omega_{\mathcal{F},D}$ sets of *format* $\mathcal{F}$ and *degree* D. We will assume $\Omega_{\mathcal{F},D}$ only contains subsets of $\mathbb{R}^n$ for $n \leqslant \mathcal{F}$.

A #o-minimal structure is a pair $\Sigma := (\mathcal{S},\Omega)$ consisting of an o-minimal expansion of the real field $\mathcal{S}$ and an FD-filtration Ω; and for each $\mathcal{F}\in\mathbb{N}$ a polynomial $P_{\mathcal{F}}(\cdot)$ such that the following holds:

(1) If $A \in \Omega_{\mathcal{F},D}$ with $A\subset\mathbb{R}^n$ then $A^c, \pi_{n-1}(A), A\times\mathbb{R}$ and $\mathbb{R}\times A$ lie in $\Omega_{\mathcal{F}+1,D}$.
(2) If $A_1,\dots,A_k\subset\mathbb{R}^n$ with $A_j\in\Omega_{\mathcal{F},D_j}$ then $\cup_i A_i\in\Omega_{\mathcal{F},D}$ and $\cap_i A_i\in\Omega_{\mathcal{F}+1,D}$ where $D=\sum_j D_j$.
(3) If $P\in\mathbb{R}[x_1,\dots,x_n]$ then $\{P=0\}\in\Omega_{n,\deg P}$.
(4) If $A\in\Omega_{\mathcal{F},D}$ with $A\subset\mathbb{R}$ then it has at most $P_{\mathcal{F}}(D)$ connected components,

Axioms 1-2 bear a close analogy to the standard axioms of a first-order structure, keeping track of the formats and degrees of sets defined using the logical operations. Axiom 3 ensures compatibility with the standard notion of degree in the (semi-)algebraic case. Finally Axiom 4 replaces the mere finiteness postulated in standard o-minimality by polynomial bounds in degrees.

2.2. **Sharp cell decomposition.** The following notion is crucial for working with #o-minimal structures. Below $\mathbb{N}[D]$ (resp. $\mathbb{N}[D,k]$) denotes the set of polynomials in D (resp. in D,k) with coefficients in $\mathbb{N}$.

Definition 5. *We say that* $(\mathcal{S},\Omega)$ *has* sharp cell decomposition *if for every* $\mathcal{F}\in\mathbb{N}$ *there are*

$$a_{\mathcal{F}}\in\mathbb{N}, \qquad b_{\mathcal{F}}\in\mathbb{N}[D,k], \qquad c_{\mathcal{F}}\in\mathbb{N}[D] \tag{10}$$

such that the following holds. For every $k,D\in\mathbb{N}$ *and every* $X_1,\dots,X_k\in\Omega_{\mathcal{F},D}$ *subsets of* $\mathbb{R}^n$*, there exists a cylindrical decomposition* $\{C_\eta\}$ *of* $\mathbb{R}^n$ *compatible with* $X_1,\dots,X_k$ *such that*

$$\#\{C_\eta\}\leqslant b_{\mathcal{F}}(D,k), \qquad \forall\eta: C_\eta\in\Omega_{a_{\mathcal{F}},c_{\mathcal{F}}(D)}. \tag{11}$$

We use the following notation for cells from [5]. For $C \subset \mathbb{R}^{n-1}$ and $a, b : C \to \mathbb{R}$ we set

$$\begin{aligned} C \odot \{a(z)\} &:= \{(z,w) : z \in C,\ w = a(z)\}, \\ C \odot (a(z), b(z)) &:= \{(z,w) : z \in C,\ a(z) < w < b(z)\}. \end{aligned} \tag{12}$$

We will also allow $a(z) \equiv -\infty$ and $b(z) \equiv \infty$ in the second case above.

In an upcoming paper we prove, based on ideas from [8], that for every #o-minimal structure $(\mathcal{S}, \Omega)$ there is another FD-filtration Ω^* with $\Omega_{\mathcal{F},D} \subset \Omega^*_{\mathcal{F},D}$ for every $\mathcal{F}, D$ such that $(\mathcal{S}, \Omega^*)$ is #o-minimal with sharp cell decomposition. This implies that Theorem 2 and its consequences actually apply without explicitly assuming that $(\mathcal{S}, \Omega)$ has sharp cell decomposition. However to keep matters clear and avoid dependence on our upcoming text we keep this as an extra condition. Our main example $\mathbb{R}_{\text{rPfaff}}$ does, in any case, admit sharp cell decomposition as explained in §3.

2.3. Some consequences of #o-minimality and sharp cell decomposition.

The axioms of #o-minimality imply that whenever $X_1, \dots, X_k \in \Omega_{\mathcal{F},D}$ and ψ is a first-order formula of depth ℓ with basic predicates $\mathbf{x} \in X_j$ then the set X defined by ϕ satisfies

$$X \in \Omega_{O_{\mathcal{F},\ell}(1), \text{poly}_{\mathcal{F},\ell}(D,k)}, \tag{13}$$

see [8, Section 1.3] for a more precise treatment. Together with sharp cell decomposition, this can be used to effectivize many of the classical constructions of o-minimality in a rather routine fashion. We record a few instances used in our text to familiarize the reader with this technique.

Proposition 6 (Connected components)**.** *Let $X \in \Omega_{\mathcal{F},D}$. Then each connected component of X is in $\Omega_{O_{\mathcal{F}}(1), \text{poly}_{\mathcal{F}}(D)}$, and their number is $\text{poly}_{\mathcal{F}}(D)$.*

Proof. Perform a cell decomposition. Each connected component is a union of cells. ☐

We call a subset $S \subset \mathbb{R}^n$ *regular* if it is an embedded C^1-smooth submanifold of $\mathbb{R}^n$.

Proposition 7 (Stratification)**.** *Let $X \in \Omega_{\mathcal{F},D}$. Then X is a disjoint union $\cup_\eta S_\eta$ where each S_η is connected and regular, and*

$$\#\{S_\eta\} = \text{poly}_{\mathcal{F}}(D), \qquad \forall \eta : S_\eta \in \Omega_{O_{\mathcal{F}}(1), \text{poly}_{\mathcal{F}}(D)}. \tag{14}$$

Proof. Let $\mu := \dim X$. Let $S \subset X$ be the μ-regular part of X, i.e. the set of points $p \in X$ such that for some linear projection $L : \mathbb{R}^n \to \mathbb{R}^\mu$, the map $L|_X$ is locally invertible at p, and the inverse $L' : (\mathbb{R}^\mu, L(p)) \to X \subset \mathbb{R}^n$ is locally C^1 with Jacobian of rank μ. This can be written out explicitly as a first-order formula in ε-δ-type language, so the axioms of #o-minimality give $S \in \Omega_{O_{\mathcal{F}}(1), \text{poly}_{\mathcal{F}}(D)}$. Each connected component of S is a top-dimensional stratum, and the remaining set $X \setminus S$ can be handled by induction on μ. ☐

Below $\operatorname{gr} F$ denotes the graph of a map F.

Proposition 8 (Definable choice)**.** *Let $X \subset \Lambda \times \mathbb{R}^n$ with $X \in \Omega_{\mathcal{F},D}$, and suppose $X_\lambda \neq \emptyset$ for every $\lambda \in \Lambda$. Then there is a map $F : \Lambda \to \mathbb{R}^n$ with $\operatorname{gr} F \in \Omega_{O_{\mathcal{F}}(1), \text{poly}_{\mathcal{F}}(D)}$ such that $\operatorname{gr} F \subset X$.*

Proof. Note that by our assumptions Λ is the projection of X to the first set of variables, and is hence definable. Perform a cylindrical decomposition of $\Lambda \times \mathbb{R}^n$ compatible with X. In particular we obtain a cylindrical decomposition $\{C_\eta\}$ of Λ, and over each C_η a cylindrical decomposition of $C_\eta \times \mathbb{R}^n$ by cells projecting to C_η. It will be enough to handle each C_η separately and then take the unions of the corresponding graphs. Moreover, we may as well consider just one of the cells over C_η that is contained in X for the purpose of defining the choice function. So assume without loss of generality that X is a cell.

Write $X = C \odot (a(z), b(z))$ where $C \in \Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D)}$ is a cell in $\Lambda \times \mathbb{R}^{n-1}$ and $a(z), b(z) : C \to \mathbb{R}$. The cases $a(z) = -\infty$, $b(z) = \infty$ and $C \odot \{a(z)\}$ are treated similarly. We have $\mathrm{gr}\, a(z), \mathrm{gr}\, b(z) \in \Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D)}$ since they can be defined using first-order formulas as the infimum and supremum of the fiber C_z. Then we find a choice function $\hat{F}(\lambda)$ on C by induction on n, and

$$F(\lambda) := \left(\hat{F}(\lambda), \frac{a(\lambda, \hat{F}(\lambda)) + b(\lambda, \hat{F}(\lambda))}{2}\right) \tag{15}$$

is a choice function for X. □

2.4. **Sharp derivatives.** If $f : X \to Y$ is a definable function we will write $f \in \Omega_{\mathcal{F},D}$ as shorthand for $\mathrm{gr}\, f \in \Omega_{\mathcal{F},D}$.

Definition 9. *We say that* $(\mathcal{S}, \Omega)$ *has* sharp derivatives *if for every* $\mathcal{F} \in \mathbb{N}$ *there are*

$$a_{\mathcal{F}} \in \mathbb{N}, \qquad\qquad b_{\mathcal{F}} \in \mathbb{N}[D, k] \tag{16}$$

such that the following holds. Given a definable $f : \mathbb{R}^n \to \mathbb{R}$ *with* $f \in \Omega_{\mathcal{F},D}$*, we have for every* $\alpha \in \mathbb{Z}^n_{\geqslant 0}$

$$f^{(\alpha)} \in \Omega_{a_{\mathcal{F}}, b_{\mathcal{F}}(D,|\alpha|)}. \tag{17}$$

Here $f^{(\alpha)}$ denotes the α-partial derivative of f (in multiindex notation) with domain of definition taken to be equal to the interior of the locus where f is continuously differentiable to order $|\alpha|$.

Remark 10. *In every #o-minimal structure we have* $f^{(\alpha)} \in \Omega_{a_{\mathcal{F},|\alpha|}, b_{\mathcal{F},|\alpha|}(D)}$ *with* $b_{\mathcal{F},|\alpha|} \in \mathbb{N}[D]$*. Sharp derivatives means that as we take derivatives of high order, the format remains fixed and the degree depends polynomially on the order. We do not know whether this holds for general #o-minimal structures.*

3. The restricted sub-Pfaffian structure $\mathbb{R}_{\mathrm{rPfaff}}$

In this section we let Ω denote the #o-minimal filtration on $\mathbb{R}_{\mathrm{rPfaff}}$ introduced in [8]. The main result of loc. cit. is that $(\mathbb{R}_{\mathrm{rPfaff}}, \Omega)$ is a #o-minimal structure admitting sharp cell decomposition.

Remark 11. *A small technical issue is that in* [8] *we considered only subsets of* $[0,1]^n$*, whereas in the o-minimal setting it is of course customary to work in* $\mathbb{R}^n$*. It is a routine matter to translate the results of* [8] *to this alternative context. Say* Ω' *denotes the FD-filtration introduced in* [8]*. Fix an algebraic diffeomorphism* $\phi : \mathbb{R} \to (0,1)$*, and by abuse of notation also write* $\phi : \mathbb{R}^n \to (0,1)^n$ *for* $\phi^{\times n}$*. Then one can define*

$$\Omega_{\mathcal{F},D} := \{\phi^{-1}(X) : X \in \Omega'_{\mathcal{F},D}\}, \tag{18}$$

and deduce #o-minimality and sharp cell decomposition for $(\mathbb{R}_{\text{rPfaff}}, \Omega)$ *from the results of* [8] *for* Ω'.

*Another small issue is that the *-format and *-degree introduced in* [8] *does not exactly satisfy the axioms of #o-minimality as defined in §2.1, though this is a matter of a simple re-indexing. To obtain a #o-minimal structure one can consider* $\Omega_{\mathcal{F},D}$ *to be given by the collection of restricted sub-Pfaffian sets defined by first-order formulas of *-format* $\mathcal{F}$ *and *-degree* D*, as defined in* [8, Definition 7].

In the remainder of the section we will prove the following.

Theorem 3. *The structure* $(\mathbb{R}_{\text{rPfaff}}, \Omega)$ *has sharp derivatives.*

Let $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}$ with $f \in \Omega_{\mathcal{F},D}$ and $\Gamma := \operatorname{gr} f$. By the definition from [8],

$$\Gamma = \cup_i \pi_{n+1} X_i^\circ \tag{19}$$

where i) each X_i° is a connected component of a semi-Pfaffian $X_i \subset \mathbb{R}^N$ of degree $\operatorname{poly}_{\mathcal{F}}(D)$ for some $N = N(\mathcal{F})$, and ii) the number of X_i is $\operatorname{poly}_{\mathcal{F}}(D)$. Moreover according to [8, Lemma 18] we may assume that the projection $\pi_{n+1}|_{X_i}$ has finite fibers.

Fix one $X = X_i$ with $\pi_{n+1}(X_i^\circ)$ of full dimension in Γ. The general case easily reduces to this at the end. By [12] we may further assume that X is effectively smooth, i.e. is defined by a semi-Pfaffian system

$$\{F_1 = \ldots = F_{N-n} = 0\} \cap \{G_1 > 0, \ldots, G_M > 0\} \tag{20}$$

where F_i, G_j Pfaffian functions and the differential $\mathrm{d}F_1 \wedge \cdots \wedge \mathrm{d}F_{N-n}$ non-vanishing on X. The degrees of the Pfaffian functions F_i, G_j in (20) are $\operatorname{poly}_{\mathcal{F}}(D)$. Removing a smaller-dimensional part, we may assume that the projection $\pi_n|_X$ is everywhere submersive.

Denote the coordinates on $\mathbb{R}^N$ by (x, y) where

$$x = (x_1, \ldots, x_n), \qquad y = (y_1, \ldots, y_{N-n}). \tag{21}$$

By the implicit function theorem and our setup above, around every point in X one can express y as a smooth function of x, and

$$F(x, y) = 0 \implies \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial x} = 0 \implies \frac{\partial y}{\partial x} = -\left(\frac{\partial F}{\partial y}\right)^{-1} \frac{\partial F}{\partial x}, \tag{22}$$

where $F = (F_1, \ldots, F_{N-n})$. Note that $\frac{\partial F}{\partial y}$ is invertible everywhere on X by our setup. Note $f(x) = y_1(x)$ on $\pi_n(X^\circ)$. Since F is a vector of Pfaffian functions, all the derivatives in the right hand side are again Pfaffian, and using $A^{-1} = \det^{-1}(A) \operatorname{adj}(A)$ we can write each $\frac{\partial y_i}{\partial x_j}$ in the form P_{ij}/Q where P_{ij} is a Pfaffian function and $Q = \det \frac{\partial F}{\partial y}$. Using this, one can rewrite $f^{(\alpha)}(x) = y_1^{(\alpha)}(x)$ as a ratio of Pfaffian functions $P_\alpha / Q^{2|\alpha|}$ with $\deg P_\alpha = \operatorname{poly}_{\mathcal{F}}(D) \cdot |\alpha|$, the asymptotic constants depending on the Pfaffian chain used to define f (which are part of the format $\mathcal{F}$). This ratio is not formally Pfaffian, but adding a variable z and an equation $Q^{2|\alpha|} z = P_\alpha$ to the equations of X gives a set $Z \subset \mathbb{R}^{N+1}$ with a connected component Z° lying over X°, such that the projection of Z to (x, z) is the graph of $f^{(\alpha)}$ over $\pi_n(X^\circ)$.

Recall that the union of $\operatorname{poly}_{\mathcal{F}}(D)$ sets X° as above define a dense subset of Γ. We have thus seen how to define a dense subset of $\operatorname{gr} f^{(\alpha)}$ with format $O_{\mathcal{F}}(1)$ and degree $\mathrm{poly}_{\mathcal{F}}(D,|\alpha|)$. By #o-minimality, the closure of this dense subset, Γ_α, has similarly bounded format and degree.

Finally, the open set $D_\alpha \subset U$ equal to the interior of the locus where Γ_α is the graph of a continuously differentiable function is also in $\Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,|\alpha|)}$ by #o-minimality. Setting

$$\Gamma'_\alpha = \Gamma_\alpha \cap \bigcap_{|\beta|<|\alpha|} D_\beta \tag{23}$$

defines the graph of $y^{(\alpha)}$ with the correct domain of definition, and the format and degree bounds follow by #o-minimality.

4. Norms on C^r-functions

For

$$P = \sum_{|\alpha|\leqslant r} a_\alpha t^\alpha \in \mathbb{R}[t_1,\dots,t_m] \tag{24}$$

we denote by $\mathcal{M}P = \sum_{|\alpha|\leqslant r} |a_\alpha| t^\alpha$ the *majorant*. We set $\|P\|_{\mathcal{M}} := \mathcal{M}P(1,\dots,1)$.

For a C^r-smooth function $f : U \to \mathbb{R}$ on a domain $U \subset \mathbb{R}^m$ and $x_0 \in U$ we denote by

$$j^r_{x_0} f = \sum_{|\alpha|\leqslant r} \frac{f^{(\alpha)}(x_0)}{\alpha!} t^\alpha \tag{25}$$

the r-jet of f at x_0. Recall from (8) that $\|f\|$ denotes the supremum norm. We define two further norms on f as follows,

$$\|f\|_r := \max_{|\alpha|\leqslant r} \|D^{\boldsymbol{\alpha}} f\|, \qquad \|f\|_{T,r} := \sup_{x\in U} \|j^r_x f\|_{\mathcal{M}}. \tag{26}$$

As in §1.4 we extend this to $F : U \to \mathbb{R}^n$ by coordinate-wise maximum.

For our purposes these two norms are essentially equivalent. Indeed, on the one hand we have

$$\|f\|_{T,r} \leqslant e^m \|f\|_r. \tag{27}$$

On the other hand the following lemma is immediate.

Lemma 12. *Suppose $f : I^n \to \mathbb{R}$ with $\|f\|_{T,r} \leqslant 1$. Let $\phi : I^n \to I^n$ be a diagonal affine map with* $\operatorname{Im}\phi$ *a cube of side-length $1/r$. Then $\|f\circ\phi\|_r \leqslant 1$.*

As a consequence of Lemma 12, given functions of unit (T,r)-norm on I^n we can always rescale to obtain bounded r-norms using r^n charts.

We usually state our results with $\|f\|_r$, but in some cases $\|f\|_{T,r}$ is more technically convenient, mainly because of the following submultiplicativity and subcompositionality properties.

Lemma 13. *The following estimates for products and compositions hold:*

1. *Let $f, g : U \to \mathbb{R}$ be C^r-smooth. Then $\|fg\|_{T,r} \leqslant \|f\|_{T,r}\cdot\|g\|_{T,r}$.*
2. *Let $F : U \to \mathbb{R}^n$ and $g : V \to \mathbb{R}$ be C^r-smooth with* $\operatorname{Im} F \subset V$. *Suppose $\|F_i\|_{T,r} \leqslant 1$ for $i = 1,\dots,n$. Then $\|g\circ F\|_{T,r} \leqslant \|g\|_{T,r}$.*

Proof. Part (1) follows from

$$[\mathcal{M} j^r_x(fg)](1,\dots,1) \leqslant [\mathcal{M} j^r_x f](1,\dots,1)\cdot[\mathcal{M} j^r_x g](1,\dots,1) \tag{28}$$

which holds since $j^r_x(fg)$ is just $j^r_x(f) j^r_x(g)$ truncated to degree r. Part (2) follows in a similar fashion, this time noting that $j^r_x(g\circ F)$ is just $j^r_{F(x)}(g)(j^r_x F_1,\dots,j^r_x F_n)$ truncated to degree r. ☐

We record another simple lemma comparing $\|f\|_{T,r+1}$ and $\|f'\|_{T,r}$.

Lemma 14. *Let $f : U \to \mathbb{R}$ with $U \subset \mathbb{R}$ be C^{r+1}-smooth. Then*

$$\|f'\|_{T,r} \leqslant (r+1)\|f\|_{T,r+1}. \tag{29}$$

Proof. By direct computation,

$$\|f'\|_{T,r} = \sup_{x\in U} \sum_{k=0}^{r} \frac{\left|f^{(k+1)}(x)\right|}{k!} \leqslant \sup_{x\in U} \sum_{k=0}^{r+1} k\frac{\left|f^{(k)}(x)\right|}{k!} \leqslant (r+1)\|f\|_{T,r+1}. \tag{30}$$

□

5. Proof of the algebraic lemma

We will prove the algebraic lemma (Lemma 2) in the following equivalent form which is more suitable for an inductive argument. Recall that below we always work in an #o-minimal structure admitting sharp cell decomposition and sharp derivatives $(\mathcal{S}, \Omega)$. We think of maps $F : \Lambda \times I^n \to I^k$ as definable families of maps $\{F_\lambda : I^n \to I^k\}_{\lambda\in\Lambda}$, where $F_\lambda := F(\lambda, \cdot)$.

Lemma 15. *Let $r \in \mathbb{N}$ and $\varepsilon > 0$. Let $F : \Lambda \times I^n \to I^k$ with $F_j \in \Omega_{\mathcal{F},D}$ for $j = 1,\dots,k$. Then there exists a collection $\{\phi_\eta : \Lambda \times I^n \to I^n\}$ such that for every $\lambda \in \Lambda$ we have i) $\|F_\lambda \circ \phi_{\eta,\lambda}\|_r \leqslant 1$, ii) $\cup_\eta \operatorname{Im}(F_\lambda \circ \phi_{\eta,\lambda}) \subseteqq_\varepsilon \operatorname{Im} F_\lambda$, and iii)*

$$\#\{\phi_\eta\} \leqslant \operatorname{poly}_{\mathcal{F}}(D, r, k, |\log\varepsilon|), \qquad \forall\eta : \phi_\eta \in \Omega_{O_{\mathcal{F}}(1),\operatorname{poly}_{\mathcal{F}}(D,r)}. \tag{31}$$

Lemma 15 with a given n implies the following family version of Lemma 2 with the same n.

Lemma 16. *Let $r \in \mathbb{N}$ and $\varepsilon > 0$. Let $X \subset \Lambda \times I^n$ with $\mu := \max_\lambda \dim X_\lambda$ and $X \in \Omega_{\mathcal{F},D}$. Assume X has no empty fibers. Then there exists a collection $\{\phi_\eta : \Lambda\times I^\mu \to X\}$ such that for every $\lambda \in \Lambda$ we have i) $\|\phi_{\eta,\lambda}\|_r \leqslant 1$, ii) $\cup_\eta \operatorname{Im}\phi_{\eta,\lambda} \subseteqq_\varepsilon X_\lambda$, and iii)*

$$\#\{\phi_\eta\} \leqslant \operatorname{poly}_{\mathcal{F}}(D, r, |\log\varepsilon|), \qquad \forall\eta : \phi_\eta \in \Omega_{O_{\mathcal{F}}(1),\operatorname{poly}_{\mathcal{F}}(D,r)}. \tag{32}$$

Proof. First perform a cell decomposition of $\Lambda \times I^n$ compatible with X, to cover X_λ by $\operatorname{poly}_{\mathcal{F}}(D)$ images $\operatorname{Im} F_{\theta,\lambda}$ for $F_\theta : \Lambda \times I^\mu \to I^n$. The non-empty fibers are required to guarantee we can always do this. Then apply Lemma 15 to each of these maps. The collection of all resulting $F_\theta \circ \phi_\eta$ establishes the conclusion the lemma. □

Remark 17. *Suppose $\Lambda = \Lambda_1 \cup \cdots \cup \Lambda_N$ is a definable subdivision of Λ with $N = \operatorname{poly}_{\mathcal{F}}(D, r, k, |\log\varepsilon|)$ and $\Lambda_i \in \Omega_{O_{\mathcal{F}}(1),\operatorname{poly}_{\mathcal{F}}(D,r)}$. Suppose we prove Lemma 15 for F restricted to each Λ_i separately, say giving collections*

$$\{\phi_{i,j} : \Lambda_i \times I^n \to I^n\} \qquad i = 1,\dots,N, \quad j = 1,\dots,M \tag{33}$$

allowing repetitions to make these collections have the same size M. Fix some arbitrary $c \in I^n$ and define

$$\tilde\phi_{i,j} : \Lambda \times I^n \to I^n, \qquad \tilde\phi_{i,j}(\lambda, x) = \begin{cases} \phi_{i,j}(\lambda, x) & \text{if } \lambda \in \Lambda_i \\ c & \text{otherwise.} \end{cases} \tag{34}$$

Then the collection $\{\tilde\phi_{i,j}\}_{i,j}$ proves Lemma 15 for Λ (the degree and format bounds follow from #o-minimality). A similar remark applies for Lemma 16. In the proof below we will often use this subdivision argument without explicit mention.

To make the notation more suggestive, we sometimes denote the coordinates on $\mathbb{R}^n$ by $(x, y_1, \dots, y_{n-1})$. The proof of Lemma 15 will occupy the remainder of this section. We proceed by induction on n, treating the base case $n = 1$ in the following subsection. We record a simple lemma that is useful in many stages of our argument.

Lemma 18. *Let $F : X \to Y$ be 1-Lipschitz, and suppose $\{\phi_\eta : U_\eta \to X\}$ satisfies $\cup_\eta \operatorname{Im}\phi_\eta \subseteqq_\varepsilon X$. Then $\cup_\eta \operatorname{Im}(F\circ\phi_\eta) \subseteqq_\varepsilon \operatorname{Im} F$.*

In particular this implies that when every F_λ is 1-Lipschitz we can replace the condition $\cup_\eta \operatorname{Im}(F_\lambda \circ \phi_{\eta,\lambda}) \subseteqq_\varepsilon \operatorname{Im} F_\lambda$ in Lemma 15 by $\cup_\eta \operatorname{Im}\phi_{\eta,\lambda} \subseteqq_\varepsilon I^n$. We will often use this remark after performing a pullback to satisfy the 1-Lipschitz condition.

5.1. **The case $n = 1$ in Lemma 15.** The main difficulty in proving Lemma 15 is to get polynomial growth with respect to r and D *simultaneously*. For a fixed r the classical proof of Yomdin-Gromov gives the same statement, even with a true cover in place of the ε-cover. We record below the C^1-version that we will need in the sequel.

Lemma 19. *Let $F : \Lambda \times I \to I^k$ with $F_i \in \Omega_{\mathcal{F},D}$. Then there exists a collection $\{\phi_\eta : \Lambda \times I \to I\}$ such that for every $\lambda \in \Lambda$ we have i) $\|F_\lambda \circ \phi_{\eta,\lambda}\|_1 \leqslant 1$, ii) $\cup_\eta \phi_{\eta,\lambda}(I) = I$, and iii)*

$$\#\{\phi_\eta\} \leqslant \operatorname{poly}_{\mathcal{F}}(D,k), \qquad \forall \eta : \phi_\eta \in \Omega_{O_{\mathcal{F}}(1),\operatorname{poly}_{\mathcal{F}}(D)}. \tag{35}$$

Proof. Assume without loss of generality that $f(x) = x$ is among the F_i. Denote $(\cdot)' = \frac{\partial}{\partial x}(\cdot)$. Perform a cell decomposition of $\Lambda \times I$ compatible with the sets of zeros of all the functions $|F_i'| - |F_j'|$ for $i, j = 1, \dots, k$ as well as with the sets of points where F_i' is undefined. We have $\operatorname{poly}_{\mathcal{F}}(D,k)$ cells, each in $\Omega_{O_{\mathcal{F}}(1),\operatorname{poly}_{\mathcal{F}}(D)}$.

It will suffice to handle each cell separately. For cells of the form $C \odot \{a(\lambda)\}$ one can cover their image by a constant map, so consider a cell $C \odot (a(\lambda), b(\lambda))$. Since each $|F_i'| - |F_j'|$ is either identically vanishing or identically non-vanishing on the cell, there is one F_i, without loss of generality F_1, such that

$$|F_1'| \geqslant |F_j'| \quad \forall j = 2, \dots, k \tag{36}$$

uniformly over the cell. In particular $|F_1'| \geqslant 1$. Set

$$F_1(\lambda, I) = (A(\lambda), B(\lambda)) \tag{37}$$

and define $\tilde\phi : C \odot (A(\lambda), B(\lambda)) \to I$ by $\tilde\phi(\lambda, s) = F_{1,\lambda}^{-1}(s)$. By (36) we have

$$\begin{aligned} |(F_{j,\lambda} \circ \tilde\phi_\lambda(s))'| &= |F_{j,\lambda}'(\tilde\phi_\lambda(s))\tilde\phi_\lambda(s)'| \\ &= |F_{j,\lambda}'(\tilde\phi_\lambda(s))/F_{1,\lambda}'(\tilde\phi_\lambda(s))| \leqslant 1 \end{aligned} \tag{38}$$

so $\|F_\lambda \circ \tilde\phi_\lambda\|_1 \leqslant 1$. Finally, let $\phi : C \times I \to I$ be the pullback of $\tilde\phi$ by a linear rescaling map $C \times I \to C \odot (A(\lambda), B(\lambda))$. Since $(A(\lambda), B(\lambda)) \subset I$ this only decreases derivatives, so the collection of maps ϕ thus obtained satisfies the conditions of the lemma. □

To prove Lemma 15 in the $n = 1$ case, we first apply Lemma 19 to F. Pulling back F by each of the ϕ_η thus obtained, we may assume without loss of generality that $\|F_\lambda\|_1 \leqslant 1$ for every $\lambda \in \Lambda$. In particular, each F_λ is 1-Lipschitz (with respect to the ℓ_∞-norm).

Perform a cell decomposition of $\Lambda \times I$ compatible with the sets of zeros of all the functions $F_i^{(j)}$ for $i = 1,\dots,k$ and $j = 0,\dots,r+1$, as well as the sets of points where these functions are undefined. We have $\mathrm{poly}_{\mathcal{F}}(D,k,r)$ cells, each in $\Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}$. It will suffice to handle each cell separately. For cells of the form $C \odot \{a(\lambda)\}$ there is nothing to prove, so we consider cells $C \odot (a(\lambda), b(\lambda))$. Pulling back by the affine map $C \odot I \to C \odot (a(\lambda), b(\lambda))$ only decreases derivatives, so without loss of generality it now suffices to prove Lemma 15 assuming each F_λ is 1-Lipschitz and has constant-signed derivatives up to order $r+1$. The result now follows from the following lemma.

Lemma 20. *Let $f : I \to I$ be C^{r+1}-smooth and suppose that $f^{(j)}$ has constant sign for $j = 0,\dots,r+1$. Then for every $M > 1$ and $j = 0,\dots,r$ we have*

$$|f^{(j)}(x)| < M^j \text{ whenever } \operatorname{dist}(x, \partial I) > j/M. \tag{39}$$

Proof. We proceed by induction, the case $j = 0$ being trivial. Suppose the claim is proved for $f^{(j)}$. Assume $f^{(j+1)}$ is weakly-increasing (or weakly-decreasing, which is analogous) and suppose toward contradiction that $f^{(j+1)}(x) \geqslant M^{j+1}$ for some x with $\operatorname{dist}(x, \partial I) > (j+1)/M$ (the case $f^{(j+1)}(x) \leqslant -M^{j+1}$ being analogous). Then $f^{(j+1)} > M^{j+1}$ throughout the interval $[x, x+1/M]$. Thus $f^{(j)}(x+1/M) - f^{(j)}(x) \geqslant M^j$. This contradicts the inductive hypothesis, since both x and $x + 1/M$ have distance at least j/M to ∂I, and $f^{(j)}$ is constant-signed and bounded in absolute value by M^j at both points. □

It follows from Lemma 20 that if $\phi : I \to I$ is an affine translation with the length of $\phi(I)$ smaller than $\operatorname{dist}(\phi(I), \partial I)/r$ then $\|F_\lambda \circ \phi\|_r \leqslant 1$ for every λ. It is an elementary exercise that with $\mathrm{poly}(r, |\log \varepsilon|)$ such maps we can cover $I_\varepsilon := (\varepsilon, 1-\varepsilon)$. Finally, since F_λ is 1-Lipschitz for every $\lambda \in \Lambda$ we have $F_\lambda(I_\varepsilon) \subseteqq_\varepsilon F_\lambda(I)$ so this covering satisfies the conditions of Lemma 15.

We record a corollary of the proof above for later use, where we cover the domain I by linear maps but skip the 1-Lipschitz preparation step, so we only get an ε-cover of the domain I but not necessarily of the image under F.

Corollary 21. *Let $r \in \mathbb{N}$ and $\varepsilon > 0$. Let $F : \Lambda \times I \to I^k$ with $F_j \in \Omega_{\mathcal{F},D}$ for $j = 1,\dots,k$. Then there exists a collection $\{\phi_\eta : \Lambda \times I \to I\}$ such that for every $\lambda \in \Lambda$ we have i) $\|F_\lambda \circ \phi_{\eta,\lambda}\|_r \leqslant 1$, ii) $\cup_\eta \phi_{\eta,\lambda}(I) \subseteqq_\varepsilon I$, iii) $\phi_{\eta,\lambda}$ is affine, and iv)*

$$\#\{\phi_\eta\} \leqslant \mathrm{poly}_{\mathcal{F}}(D, r, k, |\log \varepsilon|), \qquad \forall \eta : \phi_\eta \in \Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}. \tag{40}$$

5.2. **Reduction to bounded y-derivatives.** We now continue inductively with the case of general n, assuming that Lemma 15 and Lemma 16 are already established for $n-1$.

Apply the inductive hypothesis to $F : (\Lambda \times I_x) \times I_y^{n-1} \to I^k$ with $\varepsilon/2$ in place of ε, viewing x as an additional parameter (the subscripts I_x, I_y are introduced only for clarity). Let $\{\phi_\eta\}$ be the resulting collection. It will suffice to establish the conclusion of Lemma 15 with $\varepsilon/2$ in place of ε and each $F \circ (\lambda, x, \phi_\eta)$ in place of F. In other words, without loss of generality it suffices to prove Lemma 15 assuming that $\|F_\lambda(x, \cdot)\|_r \leqslant 1$ for every $\lambda \in \Lambda$ and every $x \in I$. In particular, the derivatives up to order r with respect to the y-variables are defined everywhere. Below we keep working with ε rather than $\varepsilon/2$ to simplify notations, and we make similar reductions later in the proof.

5.3. **Reduction to 1-Lipschitz.** Our goal in this section is to reduce to the case where F_λ is 1-Lipschitz for every $\lambda \in \Lambda$, while still preserving the bound $\|F_\lambda(x,\cdot)\|_r \leqslant 1$ for every $\lambda \in \Lambda$ established in the previous section.

Recall the following lemma from [6]. For the convenience of the reader we also give a proof using Corollary 21 in §5.6.

Lemma 22 ([6, Lemma 16])**.** *Let $f : I_x \times I_y^{n-1} \to I$ be definable and suppose that for each $i = 1, \dots, n-1$ the derivative $f_{y_i}(x,y)$ is bounded by a finite number $c(x)$ for all y where it is defined, for almost every $x \in I_x$. Then $f_x(x,y)$ is similarly bounded by a finite number $\tilde c(x)$ for all y where it is defined, for almost every $x \in I_x$.*

For $i = 1, \dots, k$ we define $S_i \subset \Lambda \times I^n$ by

$$S_i := \big\{(\lambda,x,y) : |\tfrac{\partial F_i}{\partial x}(\lambda,x,y)| \geqslant \frac{1}{2} \sup_{\tilde y \in I^{n-1}} |\tfrac{\partial F_i}{\partial x}(\lambda,x,\tilde y)|\big\} \tag{41}$$

where the supremum is taken over the points where $\frac{\partial F_i}{\partial x}(\lambda,x,\tilde y)$ is defined. By Lemma 22, the supremum is finite, for each $\lambda \in \Lambda$, for almost every x. For x where the supremum is infinite we consider that the condition is vacuous, i.e. every (λ,x,y) is included in S_i in this case. Clearly $S_i \in \Omega_{O_{\mathcal F}(1),\mathrm{poly}_{\mathcal F}(D)}$.

By definable choice we may choose subsets $\Gamma_i \subset S_i$ such that Γ_i contains exactly one (λ,x,y) for every (λ,x). In particular, sharp cylindrical decomposition shows that $\Gamma_i \in \Omega_{O_{\mathcal F}(1),\mathrm{poly}_{\mathcal F}(D)}$. By definition Γ_i is a graph of an $(n-1)$-tuple of functions

$$\gamma_{i,1},\dots,\gamma_{i,n-1} : \Lambda \times I_x \to I \tag{42}$$

and $\gamma_{i,j} \in \Omega_{O_{\mathcal F}(1),\mathrm{poly}_{\mathcal F}(D)}$ as well.

We apply Lemma 19 to the tuple including the functions γ_i and x, as well as $F \circ (\lambda, x, \gamma_i)$ for every $i = 1, \dots, k$. For every ϕ_η thus obtained let

$$\Phi_\eta(\lambda,t,y) = (\lambda, \phi_\eta(\lambda,t), y). \tag{43}$$

Denote $F_\eta := F \circ \Phi_\eta$. It will suffice to establish the conclusion of Lemma 15 for each F_η in place of F. Moreover we obviously still have $\|F_{\eta,\lambda}(t,\cdot)\|_r \leqslant 1$ for every $\lambda \in \Lambda$ and every $t \in I$. Denote $\Gamma_{i,\eta} := \Phi_\eta^{-1}(\Gamma_i)$. Then $\Gamma_{i,\eta}$ is the graph of an $(n-1)$-tuple of functions $\gamma_{i,j,\eta} := \gamma_{i,j} \circ (\lambda, \phi_\eta)$ with $\|\gamma_{i,j,\eta,\lambda}\|_1 \leqslant 1$ for every $\lambda \in \Lambda$. Note that for every $(\lambda,t,y) \in \Gamma_{i,\eta}$ we have

$$\begin{aligned}|\tfrac{\partial F_{i,\eta}}{\partial t}(\lambda,t,y)| = |\tfrac{\partial F_i}{\partial x} \circ \Phi_\eta(\lambda,t,y) \cdot \tfrac{\partial \phi_\eta}{\partial t}(\lambda,t)| \geqslant \\ \frac{1}{2} \sup_{\tilde y \in I^{n-1}} |\tfrac{\partial F_i}{\partial x}(\lambda, \phi_\lambda(t), \tilde y) \cdot \tfrac{\partial \phi_\eta}{\partial t}(\lambda,t)| = \\ \frac{1}{2} \sup_{\tilde y \in I^{n-1}} |\tfrac{\partial F_{i,\eta}}{\partial t}(\lambda,t,\tilde y)|.\end{aligned} \tag{44}$$

In other words, $\Gamma_{i,\eta}$ satisfies the same definition in the (λ,t,y) coordinates as Γ_i in the (λ,x,y) coordinates. Clearly the sets and functions defined above have format $O_{\mathcal F}(1)$ and degree $\mathrm{poly}_{\mathcal F}(D)$.

We claim that $|\frac{\partial F_{i,\eta}}{\partial t}| = O_n(1)$ whenever it is defined. According to (44) it is enough to check this on the curves $\Gamma_{i,\eta}$. We compute the derivative of $F_{i,\eta}$ along this curve,

$$1 \geqslant |\tfrac{\partial}{\partial t} F_{i,\eta}(\lambda, t, \gamma_{i,1,\eta,\lambda}, \dots, \gamma_{i,n-1,\eta,\lambda})| = \\ |\tfrac{\partial F_{i,\eta}}{\partial t} + \sum_{j=1}^{n-1} \tfrac{\partial F_{i,\eta}}{\partial y_j} \tfrac{\partial \gamma_{i,j,\eta,\lambda}}{\partial t}| \geqslant |\tfrac{\partial F_{i,\eta}}{\partial t}| - n - 1, \quad (45)$$

and rearranging we see that $|\frac{\partial F_{i,\eta}}{\partial t}| = O_n(1)$ as claimed.

According to the following lemma, each $F_\eta(\lambda, \cdot)$ is $O_n(1)$-Lipschitz for each $\lambda \in \Lambda$ after a subdivision of I_x into intervals. The proof is postponed to §5.6.

Lemma 23. *Let $f : I_x \times I_y^{n-1} \to I$ be definable with $\|f(x, \cdot)\|_1 \leqslant 1$ everywhere and $|f_x| \leqslant 1$ whenever f_x is defined. Then the locus of discontinuity of f is contained in finitely many hyperplanes $x = x_1, \dots, x_N$, and f is $O_n(1)$-Lipschitz on each $(x_i, x_{i+1}) \times I_y^{n-1}$ (where we take $x_1 = 0$ and $x_N = 1$).*

Perform a cell decomposition of $\Lambda \times I_x$ compatible with the projections of the loci of discontinuity of $F_{i,\eta}$ for each $i = 1, \dots, k$ to $\Lambda \times I_x$, giving cells of the form either $C = C_\lambda \odot \{a(\lambda)\}$ or $C = C_\lambda \odot (a(\lambda), b(\lambda))$ where in the latter case $F_{\eta,\lambda}$ is $O_n(1)$-Lipschitz in $(a(\lambda), b(\lambda)) \times I^{n-1}$ for every $\lambda \in \Lambda$. In the former case we can handle $F_\eta|_{C \times I^{n-1}}$ by the inductive hypothesis. In the latter case, rescaling $(a(\lambda), b(\lambda))$ back to $(0, 1)$ only improves the Lipschitz constant in $F_{\eta,\lambda}|_{C \times I^{n-1}}$. Applying a further linear subdivision in the x, y coordinates we may further reduce the Lipschitz constant to 1 to simplify our notations.

Returning to the original notation, we conclude that it will suffice to establish the conclusion of Lemma 15 assuming that $\|F_\lambda(x, \cdot)\|_r \leqslant 1$ for every $\lambda \in \Lambda$ and every $x \in I$, and F_λ is 1-Lipschitz for every λ.

5.4. **Controlling higher derivatives.** We start off similarly to §5.3. For $i = 1, \dots, k$ and $\alpha \in \mathbb{Z}^n_{\geqslant 0}$ with $|\alpha| \leqslant r$ we define $S_{i,\alpha} \subset \Lambda \times I^n$ by

$$S_{i,\alpha} := \big\{(\lambda, x, y) : |F_i^{(\alpha)}(\lambda, x, y)| \geqslant \frac{1}{2} \sup_{\tilde{y} \in I^{n-1}} |F_i^{(\alpha)}(\lambda, x, \tilde{y})|\big\} \quad (46)$$

where the supremum is restricted to those points where $F_i^{(\alpha)}(\lambda, x, \tilde{y})$ is defined. Repeatedly applying Lemma 22 and using the fact that for $\alpha_1 = 0$ all the $F^{(\alpha)}$ are defined and bounded everywhere, we again conclude that the supremum is finite, for each $\lambda \in \Lambda$, for almost every x. For x where the supremum is infinite we consider that the condition is vacuous, i.e. every (λ, x, y) is included in $S_{i,\alpha}$ in this case. By the sharp derivatives condition one sees that $S_{i,\alpha} \in \Omega_{O_{\mathcal{F}}(1), \mathrm{poly}_{\mathcal{F}}(D,r)}$.

By definable choice we may choose subsets $\Gamma_{i,\alpha} \subset S_{i,\alpha}$ such that $\Gamma_{i,\alpha}$ contains exactly one (λ, x, y) for every (λ, x). In particular, sharp cylindrical decomposition shows that $\Gamma_{i,\alpha} \in \Omega_{O_{\mathcal{F}}(1), \mathrm{poly}_{\mathcal{F}}(D,r)}$. By definition $\Gamma_{i,\alpha}$ is a graph of an $(n-1)$-tuple of functions $\gamma_{i,\alpha}$ given by

$$\gamma_{i,\alpha,1}, \dots, \gamma_{i,\alpha,n-1} : \Lambda \times I_x \to I \quad (47)$$

and $\gamma_{i,\alpha,j} \in \Omega_{O_{\mathcal{F}}(1), \mathrm{poly}_{\mathcal{F}}(D,r)}$ as well.

We apply Corollary 21 to the tuple including the functions $\gamma_{i,\alpha}$ as well as $F^{(\beta)} \circ (\lambda, x, \gamma_{i,\alpha})$ for every $i = 1, \dots, k$, every $\alpha \in \mathbb{Z}^n_{\geqslant 0}$ satisfying $|\alpha| \leqslant r$, and every $\beta \in \mathbb{Z}^n_{\geqslant 0}$ satisfying $|\beta| \leqslant r$ and $\beta_1 = 0$. Note that $F^{(\beta)}$ is indeed bounded by 1 according to our previous steps, so the corollary applies. For every ϕ_η thus obtained let

$$\Phi_\eta(\lambda, t, y) = (\lambda, \phi_\eta(\lambda, t), y). \tag{48}$$

Denote $F_\eta := F \circ \Phi_\eta$. It will suffice to establish the conclusion of Lemma 15 for each F_η in place of F: indeed, $\cup_\eta \phi_{\eta,\lambda}(I) \subseteqq_\varepsilon I$ for every $\lambda \in \Lambda$, and since F_λ is 1-Lipschitz for every λ it follows that $\cup_\eta F_\lambda \circ \Phi_{\eta,\lambda}(I^n) \subseteqq_\varepsilon F_\lambda(I^n)$. Moreover we obviously still have $\|F_{\eta,\lambda}(t,\cdot)\|_r \leqslant 1$ for every $\lambda \in \Lambda$ and every $t \in I$.

Denote $\Gamma_{i,\alpha,\eta} := \Phi_\eta^{-1}(\Gamma_{i,\alpha})$. Then $\Gamma_{i,\alpha,\eta}$ is the graph of an $(n-1)$-tuple $\gamma_{i,\alpha,\eta} := \gamma_{i,\alpha} \circ (\lambda, \phi_\eta)$ of functions $\gamma_{i,\alpha,j,\eta} := \gamma_{i,\alpha,j} \circ (\lambda, \phi_\eta)$ with $\|\gamma_{i,\alpha,\eta,\lambda}\|_r \leqslant 1$ for every $\lambda \in \Lambda$. Denote $\frac{\partial}{\partial t}(\cdot) = (\cdot)'$ and recall that

$$\phi_\eta'' = 0 \tag{49}$$

since these maps are affine in t. Then for every $(\lambda, t, y) \in \Gamma_{i,\alpha,\eta}$ we have

$$\begin{aligned} |F_{i,\eta}^{(\alpha)}(\lambda, t, y)| = |F_i^{(\alpha)} \circ \Phi_\eta(\lambda, t, y) \cdot (\phi_\eta'(\lambda, t))^{\alpha_1}| \geqslant \\ \frac{1}{2} \sup_{\tilde{y} \in I^{n-1}} |F_i^{(\alpha)}(\lambda, \phi_\lambda(t), \tilde{y}) \cdot (\phi_\eta'(\lambda, t))^{\alpha_1}| = \\ \frac{1}{2} \sup_{\tilde{y} \in I^{n-1}} |F_{i,\eta}^{(\alpha)}(\lambda, t, \tilde{y})|. \end{aligned} \tag{50}$$

In other words, $\Gamma_{i,\alpha,\eta}$ satisfies the same definition in the (λ, t, y) coordinates as $\Gamma_{i,\alpha}$ in the (λ, x, y) coordinates. Note that (49) was crucial at this point: otherwise we would get numerous additional terms involving mixed derivatives of F_i and ϕ_η. As before, the sets and functions defined above have format $O_{\mathcal{F}}(1)$, and their number and degree are bounded by $\operatorname{poly}_{\mathcal{F}}(D, r)$. Recall the norm $\|f\|_{T,r}$ introduced in (26).

Lemma 24. *Let $\lambda \in \Lambda$. Let Γ be one of $(\Gamma_{i,\alpha,\eta})_\lambda$ and $\gamma : I \to \Gamma$ be the tuple $\gamma_{i,\alpha,\eta,\lambda}$. Let $G = F_{l,\eta,\lambda}$ for some $l = 1, \ldots, k$. Then for each $\beta \in \mathbb{Z}^n_{\geqslant 0}$ with $|\beta| \leqslant r$ we have*

$$\|G^{(\beta)} \circ (t, \gamma)\|_{T, r-\beta_1} \leqslant E(\beta_1), \qquad E(\beta_1) := e \cdot (r + (n-1)e)^{\beta_1}. \tag{51}$$

Proof. We will work by induction on β_1. For $\beta_1 = 0$ the estimate (51) follows from the fact that $F^{(\beta)} \circ (\lambda, x, \gamma_{i,\alpha})$ was included in the tuple of functions to which Corollary 21 was applied. Indeed, $F^{(\beta)} \circ \Phi_\eta = (F \circ \Phi_\eta)^{(\beta)}$ and it follows that

$$\begin{aligned} G^{(\beta)} \circ (t, \gamma) &= (F_l \circ \Phi_\eta)^{(\beta)} \circ (\lambda, t, \gamma_{i,\alpha,\eta,\lambda}) \\ &= F_l^{(\beta)} \circ \Phi_\eta \circ (\lambda, t, \gamma_{i,\alpha,\eta,\lambda}) \\ &= F_l^{(\beta)} \circ (\lambda, \phi_\eta(\lambda, t), \gamma_{i,\alpha} \circ \phi_\eta(\lambda, t)) \\ &= F_l^{(\beta)} \circ (\lambda, x, \gamma_{i,\alpha}) \circ \phi_\eta(\lambda, t). \end{aligned} \tag{52}$$

so the (T, r)-norm of the left hand side is bounded according to Corollary 21 and (27).

Suppose now that (51) is proved for all β' with $\beta_1' < \beta_1$. Compute

$$(G^{(\beta - e_1)} \circ (t, \gamma))' = G^{(\beta)} \circ (t, \gamma) + \sum_{j=2}^{n} (G^{(\beta - e_1 + e_j)} \circ (t, \gamma)) \cdot \gamma_j', \tag{53}$$

where $e_1,\dots,e_n \in \mathbb{Z}^n_{\geqslant 0}$ and $e_{i,j}=\delta_{i,j}$. We will use this to bound the first summand on the right hand side. By the inductive hypothesis

$$\|G^{(\beta-e_1)}\circ(t,\gamma)\|_{T,r-\beta_1+1} \leqslant E(\beta_1-1). \tag{54}$$

Thus using Lemma 14 the left-hand side of (53) has $(T,r-\beta_1)$-norm bounded by $rE(\beta_1-1)$. Similarly

$$\|G^{(\beta-e_1+e_j)}\circ(t,\gamma))\|_{T,r-\beta_1} \leqslant E(\beta_1-1) \tag{55}$$

by induction and $\|\gamma'_j\|_{T,r-\beta_1}\leqslant e$ since $\|\gamma_j\|_r \leqslant 1$. Rearranging and using Lemma 13 we conclude that

$$\|G^{(\beta)}\circ(t,\gamma)\|_{T,r-\beta_1} \leqslant rE(\beta_1-1)+(n-1)eE(\beta_1-1) \tag{56}$$

as claimed. □

Finally we conclude that $|F^{(\alpha)}_{i,\eta,\lambda}| = O_n(r^{|\alpha|})$ whenever it is defined, for every $\lambda\in\Lambda$, every $i=1,\dots,k$ and every $|\alpha|\leqslant r$. Indeed, according to (50) it is enough to check this on the curve $\Gamma_{i,\alpha,\eta}$, and there it holds by Lemma 24 taking $G=F^{(\alpha)}_{i,\eta,\lambda}$ since the (T,r)-norm bounds the maximum norm. A further linear subdivision into cubes of length $O_n(1/r)$ as in Lemma 12 one can replace the upper bound $O_n(r^{|\alpha|})$ by 1.

Returning again to the original notation, we conclude that it will suffice to establish the conclusion of Lemma 15 assuming that F_λ is 1-Lipschitz and $\|F^{(\alpha)}_\lambda\|\leqslant 1$ wherever it is defined, for every $\lambda\in\Lambda$ and every $|\alpha|\leqslant r$.

5.5. **Final clean up.** By now we have satisfied the conclusions of Lemma 15, except that some derivatives of F may be undefined at some points. Let $V_{i,\alpha}$ denote the locus where $F^{(\alpha)}_i$ is undefined, for $i=1,\dots,k$ and $|\alpha|\leqslant r$. Perform a cell decomposition of $\Lambda\times I^n$ compatible with every $V_{i,\alpha}$ giving $\mathrm{poly}_{\mathcal{F}}(D,r,k)$ cells, each in $\Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}$. This induces a cell decomposition $\{C_\nu\}$ of $\Lambda\times I^{n-1}$.

By Lemma 16 for families in $\Lambda\times I^{n-1}$, which is part of our inductive hypothesis, we get a collection $\{\phi_{\nu,\eta}:\Lambda\times I^{n-1}\to C_\nu\}$ such that

$$\cup_\eta(\phi_{\nu,\eta,\lambda}(I^{n-1}))\subseteq_\varepsilon C_{\nu,\lambda} \tag{57}$$

and $\|\phi_{\nu,\eta,\lambda}\|_r\leqslant 1$. Moreover,

$$\#\{\phi_{\nu,\eta}\}\leqslant \mathrm{poly}_{\mathcal{F}}(D,r,|\log\varepsilon|), \qquad \forall\nu,\eta:\phi_{\nu,\eta}\in\Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}. \tag{58}$$

Since F_λ is 1-Lipschitz for every λ it will be enough to prove the conditions of Lemma 15 for each pullback $F\circ(\phi_{\nu,\eta},x_n)$, and we may further restrict to the case where $C_{\nu,\lambda}$ has full dimension in I^{n-1}. Then $\|F_\lambda\circ(\phi_{\nu,\eta,\lambda},x_n)\|_r$ is bounded for every λ by Lemma 13, and by linear subdivision as in Lemma 12 one can return to unit r-norms (and after further subdivision to 1-Lipschitz).

In other words we may replace F by each $F\circ(\phi_{\nu,\eta},x_n)$ and assume without loss of generality that $V:=\cup_{i,\alpha}V_{i,\alpha}$, i.e. the locus of non-smoothness of F, is already given by graphs of functions $G_1,\dots,G_q\in\Omega_{O_{\mathcal{F}}(1),\mathrm{poly}_{\mathcal{F}}(D,r)}$ with $q=\mathrm{poly}_{\mathcal{F}}(D,k,r)$, where $G_i:\Lambda\times I^{n-1}\to I$. We also may assume $G_1=0$ and $G_q=0$ for simplicity.

Now apply the inductive hypothesis, i.e. Lemma 15 in dimension $n-1$, to the tuple $x_1,\dots,x_{n-1}$ and to $G_1,\dots,G_q$ on $\Lambda\times I^{n-1}$. Making the same reduction as above, we may now assume without loss of generality that $\|G_{1,\lambda}\|_r, \dots, \|G_{q,\lambda}\|_r \leqslant 1$ for every λ. We cover

$$(\Lambda \times I^n) \setminus V = \bigcup_{i=1}^{q} (\Lambda \times I^{n-1}) \odot (G_i, G_{i+1}) \tag{59}$$

by images of affine maps

$$(\lambda, x_1, \dots, x_{n-1}, t) \to (\lambda, x_1, \dots, x_{n-1}, tG_{i+1} + (1-t)G_i) \tag{60}$$

with similarly bounded r-norms, and finally by another pullback step as above reduce to the case $V = \emptyset$, finishing the proof.

5.6. **Proofs of auxiliary lemmata.** In this section we prove the two auxiliary Lemmata 22 and 23.

Proof of Lemma 22. This is essentially [6, Lemma 16], but we sketch the argument to illustrate the connection with the material above, especially Corollary 21. Suppose toward contradiction that $f_x(x,y)$ is unbounded in y for every x in some infinite interval $J \subset I$. Without loss of generality assume $J = I$. Then by definable choice one can choose a function $\gamma : I \to I^{n-1}$ such that each $f_{y_i}(x, \gamma(x))$ is defined and $f_x(x, \gamma(x)) > M$ for every $x \in I$ (or $f_x(x, \gamma(x)) < -M$, which is analogous). Moreover, treating M as an additional parameter, we see that the format and degree of γ can be taken to be bounded independently of M. By Corollary 21 applied to γ and $f(x, \gamma(x))$ with $r = 1$ and say $\varepsilon = 0.1$ we find a subinterval $I' \subset I$ where both $\|\gamma'(x)\|$ and $|f(x, \gamma(x))'|$ are bounded from above by a constant independent of M. One can take the longest of the intervals $\phi_\eta(I)$ for example, which has length bounded from below uniformly in M. This is now a contradiction for $M \gg 1$ because

$$M < f_x(x, \gamma(x)) = f(x, \gamma(x))' - \sum_{i=1}^{n-1} f_{y_i}(x, \gamma(x)) \gamma_i'(x) \tag{61}$$

and the right-hand side is uniformly bounded. □

Proof of Lemma 23. Since f is $O_n(1)$-Lipschitz in the y-direction, it is easy to check that if f is discontinuous at (x_0, y_0) it is also discontinuous at every point of $\{x_0\} \times U_\delta(y_0)$ for some $\delta \ll 1$. Since the locus of discontinuity has empty interior, the first claim follows.

Let $V \subset (x_i, x_{i+1}) \times I_y^{n-1}$ denote the locus where f_x is undefined. Consider two points in $(x_i, x_{i+1}) \times I_y^{n-1}$ and the straight line γ connecting them. If $\gamma \cap V$ is finite then $f|_\gamma$ is piecewise $O_n(1)$-Lipschitz and continuous, so it is $O_n(1)$-Lipschitz. In general, we deform γ into a curve $\gamma + v$ with $v \in \mathbb{R}^n$. For the same reason as above, we can choose $\|v\|$ arbitrarily small so that $(\gamma + v) \cap V$ is finite, and f is $O_n(1)$-Lipschitz on $\gamma + v$. The claim follows by continuity of f. □

6. Point counting

In this section we give the proof of Theorem 2. Since the argument is fairly standard by now we focus mostly on the novel parts. The key proposition is the following. Recall the notation $X(g, H)$ from (1).

Proposition 25. *There are appropriate choices of* $r, d = \operatorname{poly}_m(g, \log H)$ *and* $\log(1/\varepsilon) = \operatorname{poly}_m(g, \log H)$ *such that the following holds. Let* $\phi : I^m \to X$ *such that* $\phi(I^m) \subseteqq_\varepsilon X \subset I^{m+1}$ *and* $\|\phi\|_r \leqslant 1$*. Then there exists a polynomial* $P \in \mathbb{R}[x_1, \ldots, x_{m+1}] \setminus \{0\}$ *of degree* d *such that* $X(g, H) \subset \{P = 0\}$*.*

Proof. There are two approaches to proving such a statement. The first, due to Bombieri-Pila [9], is based on interpolation determinants. The second, due to Wilkie [30] is based on the Siegel lemma. Both of these are normally stated with $\phi(I^m) = X$. We briefly show that for both methods the weaker assumption $\phi(I^m) \subseteqq_\varepsilon X$ is sufficient. After cutting into 2^m pieces we may assume that the domain of ϕ is J^m where $J := (0, 1/2)$ instead of I^m.

We start with the interpolation determinant method, which directly applies to the case $g = 1$ (see Remark 26 for a discussion). Recall that an *interpolation determinant* is

$$\Delta^d(p_1, \ldots, p_\mu) = \det(p_i^\alpha) \quad \text{for} \quad {}_{\alpha \in \mathbb{Z}^{m+1}_{\geqslant 0}, |\alpha| \leqslant d}^{i=1,\ldots,\mu} \tag{62}$$

where $p_i \in \mathbb{R}^{m+1}$ and μ is the dimension of the space of polynomials of degree at most d in $m + 1$ variables. By linear algebra, the existence of a polynomial $P \in \mathbb{R}[x_1, \ldots, x_{m+1}] \setminus \{0\}$ of degree d such that $X(\mathbb{Q}, H) \subset \{P = 0\}$ is equivalent to the vanishing of $\Delta^d(\cdot)$ for every μ-tuple of points in $X(\mathbb{Q}, H)$. The vanishing of this determinant is proved as follows.

First, for any such tuple p one estimates the height of $\Delta^d(p)$, concluding that either $\Delta^d(p) = 0$ or $|\Delta^d(p)| \geqslant H^{-d\mu}$. On the other hand, using Taylor expansions of order r for ϕ (crucially using the assumption $\|\phi\|_r \leqslant 1$) one shows that for an appropriate choice of r, d as above we have $|\Delta^d(p)| < \frac{1}{2}H^{-d\mu}$. These contradicting estimates force $\Delta^d(p) = 0$ on $\phi(J^m)$ and finish the proof.

In our context, we would like to extend this to the case $p_1, \ldots, p_\mu \in U_\varepsilon(\phi(J^m))$. Let $q_1, \ldots, q_\mu \in \phi(J^m)$ with $\operatorname{dist}(p_i, q_i) \leqslant \varepsilon$. Then $\Delta^d(q) < \frac{1}{2}H^{-d\mu}$ as above, and if we show $|\Delta^d(q) - \Delta^d(p)| < \frac{1}{2}H^{-d\mu}$ we can finish as above. One easily estimates $\|\partial\Delta^d(p)/\partial p\| \leqslant d\mu!$ in $I^{\mu\times(m+1)}$, so choosing $\varepsilon \sim \frac{1}{2}H^{-d\mu}/[d(\mu + 2)!]$ suffices.

We now consider Wilkie's approach. Using Siegel's lemma one constructs a polynomial $P \in \mathbb{Z}[x_1, \ldots, x_{m+1}]$ of degree d and coefficients bounded by some N such that $P(\phi_1, \ldots, \phi_{m+1})$ has many small Taylor coefficients. Liouville's inequality gives for any $x \in X(g, H)$ that either $P(x) = 0$ or

$$|P(x)| \geqslant (d^{m+1} N H^{d(m+1)})^{-g}. \tag{63}$$

Denote the right-hand side by R. Wilkie shows that for an appropriate choice of r, d as above (and also $\log N = \operatorname{poly}_m(g, \log H)$) we have $|P(x)| < R/2$ whenever $x \in \phi(J^m)$, thus forcing P to vanish on x.

In fact Wilkie [30] considered only sets of the form $X(\mathbb{Q}, H)$ with X a definable curve. We remark that in the special case where X is the graph of the Riemann zeta function, a similar idea was used earlier by Masser [16]. A similar method was later used for $X(g, H)$ and X of arbitrary dimension in [14, Proposition 16] and in [2, Proposition 28]. The polynomial dependence $d = \operatorname{poly}_m(g, \log H)$ is explicitly demonstrated in this latter reference, though it follows easily also from the computations of the earlier work by Wilkie and Habegger.

If we now take $x \in U_\varepsilon(\phi(I^m))$ instead, and $x_0 \in \phi(I^m)$ with $\operatorname{dist}(x, x_0) \leqslant \varepsilon$, then as above it will suffice to show that $|P(x_0) - P(x)| < R/2$. The bounds on d, N easily imply $\log \|\partial P/\partial x\| < \operatorname{poly}_m(g, \log H)$ in $[0, 1]^n$, so choosing an appropriate $\log(1/\varepsilon) = \operatorname{poly}_m(g, \log H)$ suffices to finish the proof. □

Remark 26. *The first proof sketched above applies directly to the $g = 1$ case. In [22] Pila reduces the general case of $X(g, H)$ to the $g = 1$ case $Y(1, H)$ for an auxiliary set Y using his blocks formalism. The coordinates of Y essentially encode the coefficients for the minimal polynomials for the coordinates of X. Unfortunately this means that the dimension, and in particular the format, of Y depends on g, and we are thus unable to establish bounds polynomial in g with this method.*

We proceed to the proof of Theorem 2, which is very similar to [22] and [1] modulo the sharper Proposition 25. We proceed by induction on dimension $m := \dim X$, the zero-dimensional case being trivial by #o-minimality. Suppose the claim is proved for X of dimension smaller than m.

Up to inverting and negating some coordinates (which does not affect height) one can cut X into pieces contained in $[0, 1]^n$, so assume this without loss of generality. Apply Lemma 2 to X to obtain a collection $\{\phi_\eta\}$ of size $\text{poly}_{\mathcal{F}}(D, g, \log H)$ and $\cup_\eta \phi_\eta(I^m) \subseteq_\varepsilon X$ where $m = \dim X$. It will be enough to consider each $\phi_\eta(I^m)$ separately, so fix one $\phi = \phi_\eta$ and suppose $\phi(I^m) \subseteq_\varepsilon X$. Using Proposition 25 we can find for each $T \subset \{1, \ldots, n\}$ of size $m + 1$ a polynomial P_T of degree $d = \text{poly}_n(g, \log H)$ in the coordinates $(x_i : i \in T)$ vanishing on $X(g, H)$. The zero loci of these polynomials cut out an algebraic variety V of dimension at most m in $\mathbb{R}^n$. Stratify V (e.g. using #o-minimality of $\mathbb{R}^{\text{alg}}$) and let $\{S_i\}$ denote the top dimensional strata and S' the union of the rest. Note $S_i, S' \in \Omega^{\text{alg}}_{O_n(1), \text{poly}_n(d)}$ by #o-minimality. The points in $S' \cap X$ are handled by induction on m. Now stratify $X \cap S_i$ and denote by $\{B_{ij}\}$ the top dimensional strata and by B' the union of the rest (over all S_i). Note $\#\{B_{ij}\}$ is again $\text{poly}_{\mathcal{F}}(D, g, \log H)$ by #o-minimality. Finally $B' \cap X$ is similarly handled by induction on m, while B_{ij} are by definition basic blocks with semialgebraic closures S_i, which finishes the proof.

References

[1] G. Binyamini, G. Jones, H. Schmidt, and M. Thomas. An effective Pila-Wilkie theorem for sets definable using Pfaffian functions, with some diophantine applications . *arXiv:2109.08788 preprint*, 2023.

[2] Gal Binyamini. Point counting for foliations over number fields. *Forum Math. Pi*, 10:e6, 2022.

[3] Gal Binyamini, Raf Cluckers, and Dmitry Novikov. Point counting and Wilkie's conjecture for non-Archimedean Pfaffian and Noetherian functions. *Duke Mathematical Journal*, 171(9):1823 – 1842, 2022.

[4] Gal Binyamini and Dmitry Novikov. Wilkie's conjecture for restricted elementary functions. *Ann. of Math. (2)*, 186(1):237–275, 2017.

[5] Gal Binyamini and Dmitry Novikov. Complex cellular structures. *Ann. of Math. (2)*, 190(1):145–248, 2019.

[6] Gal Binyamini and Dmitry Novikov. The Yomdin-Gromov algebraic lemma revisited. *Arnold Math. J.*, 7(3):419–430, 2021.

[7] Gal Binyamini, Harry Schmidt, and Andrei Yafaev. Lower bounds for Galois orbits of special points on Shimura varieties: a point-counting approach. *Math. Ann.*, 385:961–973, 2023.

[8] Gal Binyamini and Nicolai Vorobjov. Effective cylindrical cell decompositions for restricted sub-Pfaffian sets. *Int. Math. Res. Not. IMRN*, 11 2020.

[9] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[10] Raf Cluckers, Omer Friedland, and Yosef Yomdin. Doubling coverings via resolution of singularities and preparation. *Commun. Contemp. Math.*, 23(2):Paper No. 2050018, 23, 2021.

[11] Raf Cluckers, Jonathan Pila, and Alex Wilkie. Uniform parameterization of subanalytic sets and diophantine applications. *Ann. Sci. Éc. Norm. Supér. (4)*, 53(1):1–42, 2020.

[12] A. Gabrièlov and N. Vorobjov. Complexity of stratifications of semi-Pfaffian sets. *Discrete Comput. Geom.*, 14(1):71–91, 1995.

[13] M. Gromov. Entropy, homology and semialgebraic geometry. *Astérisque*, 1985/86(145-146):5, 225–240, 1987. Séminaire Bourbaki, Vol. 1985/86.
[14] P. Habegger. Diophantine approximations on definable sets. *Selecta Math. (N.S.)*, 24(2):1633–1675, 2018.
[15] Gareth O. Jones and Margaret E. M. Thomas. Effective Pila-Wilkie bounds for unrestricted Pfaffian surfaces. *Math. Ann.*, 381(1-2):729–767, 2021.
[16] D. Masser. Rational values of the Riemann zeta function. *J. Number Theory*, 131(11):2037–2046, 2011.
[17] J. Pila. Geometric postulation of a smooth function and the number of rational points. *Duke Math. J.*, 63(2):449–463, 1991.
[18] J. Pila and A. J. Wilkie. The rational points of a definable set. *Duke Math. J.*, 133(3):591–616, 2006.
[19] Jonathan Pila. Integer points on the dilation of a subanalytic surface. *Q. J. Math.*, 55(2):207–223, 2004.
[20] Jonathan Pila. Rational points on a subanalytic surface. *Ann. Inst. Fourier (Grenoble)*, 55(5):1501–1516, 2005.
[21] Jonathan Pila. The density of rational points on a Pfaff curve. *Ann. Fac. Sci. Toulouse Math. (6)*, 16(3):635–645, 2007.
[22] Jonathan Pila. On the algebraic points of a definable set. *Selecta Math. (N.S.)*, 15(1):151–170, 2009.
[23] Jonathan Pila. Counting rational points on a certain exponential-algebraic surface. *Ann. Inst. Fourier (Grenoble)*, 60(2):489–514, 2010.
[24] Jonathan Pila, Ananth Shankar, and Jacob Tsimerman. Canonical heights on shimura varieties and the André-Oort conjecture (with an appendix by H. Esnault and M. Groechenig). *arXiv:2109.08788 preprint*, 2021.
[25] Thomas Scanlon. A proof of the André-Oort conjecture via mathematical logic [after Pila, Wilkie and Zannier]. *Astérisque*, 2010/2011(348):Exp. No. 1037, ix, 299–315, 2012. Séminaire Bourbaki: Vol. 2010/2011. Exposés 1027–1042.
[26] Lou van den Dries and Chris Miller. On the real exponential field with restricted analytic functions. *Israel J. Math.*, 85(1-3):19–56, 1994.
[27] Siegfried Van Hille. On a family of mild functions. *Int. J. Number Theory*, 17(6):1379–1390, 2021.
[28] Siegfried Van Hille. Smooth parameterizations of power-subanalytic sets and compositions of Gevrey functions. *Proc. Edinb. Math. Soc. (2)*, 64(3):532–554, 2021.
[29] A. J. Wilkie. Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *J. Amer. Math. Soc.*, 9(4):1051–1094, 1996.
[30] A. J. Wilkie. Rational points on definable sets. In *O-minimality and diophantine geometry*, volume 421 of *London Math. Soc. Lecture Note Ser.*, pages 41–65. Cambridge Univ. Press, Cambridge, 2015.
[31] Y. Yomdin. C^k-resolution of semialgebraic mappings. Addendum to: "Volume growth and entropy". *Israel J. Math.*, 57(3):301–317, 1987.
[32] Y. Yomdin. Volume growth and entropy. *Israel J. Math.*, 57(3):285–300, 1987.
[33] Y. Yomdin. Analytic reparametrization of semi-algebraic sets. *J. Complexity*, 24(1):54–76, 2008.

Weizmann Institute of Science, Rehovot, Israel
Email address: gal.binyamini@weizmann.ac.il

Weizmann Institute of Science, Rehovot, Israel
Email address: dmitry.novikov@weizmann.ac.il

Weizmann Institute of Science, Rehovot, Israel
Email address: binyamin.zackkutuzov@weizmann.ac.il