



Citation for published version:

Binyamini, G & Vorobjov, N 2022, 'Effective cylindrical cell decompositions for restricted sub-Pfaffian sets', *International Mathematics Research Notices*, vol. 2022, no. 5, pp. 3493-3510.
<https://doi.org/10.1093/imrn/rnaa285>

DOI:

[10.1093/imrn/rnaa285](https://doi.org/10.1093/imrn/rnaa285)

Publication date:

2022

Document Version

Peer reviewed version

[Link to publication](https://doi.org/10.1093/imrn/rnaa285)

This is a pre-copyedited, author-produced version of an article accepted for publication in IMRN following peer review. The version of record Gal Binyamini, Nicolai Vorobjov, Effective Cylindrical Cell Decompositions for Restricted Sub-Pfaffian Sets, *International Mathematics Research Notices*, , rnaa285, is available online at <https://doi.org/10.1093/imrn/rnaa285>

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

EFFECTIVE CYLINDRICAL CELL DECOMPOSITIONS FOR RESTRICTED SUB-PFAFFIAN SETS

GAL BINYAMINI AND NICOLAI VOROBJOV

ABSTRACT. The o-minimal structure generated by the restricted Pfaffian functions, known as restricted *sub-Pfaffian* sets, admits a natural measure of complexity in terms of a *format* \mathcal{F} , recording information like the number of variables and quantifiers involved in the definition of the set, and a *degree* D recording the degrees of the equations involved. Khovanskii and later Gabrielov and Vorobjov have established many effective estimates for the geometric complexity of sub-Pfaffian sets in terms of these parameters. It is often important in applications that these estimates are polynomial in D .

Despite much research done in this area, it is still not known whether cell decomposition, the foundational operation of o-minimal geometry, preserves polynomial dependence on D . We slightly modify the usual notions of format and degree and prove that with these revised notions this does in fact hold. As one consequence we also obtain the first polynomial (in D) upper bounds for the sum of Betti numbers of sets defined using quantified formulas in the restricted sub-Pfaffian structure.

1. STATEMENT OF THE MAIN RESULTS

1.1. Setup. Let $I := [0, 1] \subset \mathbb{R}$ and for $k, n \in \mathbb{N}$, $k \geq n$, denote by $\pi_n^k : I^k \rightarrow I^n$ the projection map. We sometimes omit k if its meaning is clear from the context.

Pfaffian functions, introduced by Khovanskii in [14, 15], are analytic functions satisfying triangular systems of Pfaffian (first order partial differential) equations with polynomial coefficients. We refer the reader to [7] for precise definition of Pfaffian functions, based on *Pfaffian chain*, and examples of Pfaffian functions in an open domain $G \subset \mathbb{R}^k$, which we assume here for simplicity to be given by a product of intervals.

Definition 1 (semi-Pfaffian set). *Let G be an open set in \mathbb{R}^k and $I^k \subset G$. A set $X \subset I^k$ is called (restricted) semi-Pfaffian if it consists of points in I^k satisfying a Boolean combination of atomic equations and inequalities of the kind $f = 0$ or $f > 0$, where f is a Pfaffian functions defined in G . The format of X is the number of variables k and the degree of X is the sum of degrees of all the Pfaffian functions appearing in the atomic formulas (i.e. the degrees of the polynomials defining these Pfaffian functions).*

1991 *Mathematics Subject Classification.* 14P15,03C10,03C64.

Key words and phrases. Pfaffian functions, cell decomposition.

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).

Note that the degree of X bounds from above the number of all atomic equations and inequalities. We assume that a Pfaffian chain has been fixed once and for all, and all semi-Pfaffian sets under consideration are defined from this single Pfaffian chain.

Definition 2 (sub-Pfaffian set). *A set $Y \subset I^n$ is called (restricted) sub-Pfaffian¹ if $Y = \pi_n^k(X)$ for a semi-Pfaffian set $X \subset I^k$.*

In the special case of a *semi-algebraic set* X , the Tarski-Seidenberg theorem states that the set $Y = \pi_n^k(X)$ is also semi-algebraic, i.e., is a set of points satisfying a Boolean combination of polynomial equations and inequalities. By contrast, a sub-Pfaffian set may not be semi-Pfaffian.

It is customary in the literature to define the format and degree of a sub-Pfaffian set as in Definition 2 to be the format and degree of the set X . Below we introduce a variant of these notions which turns out to behave better with respect to cell decompositions. To avoid confusion we refer to these modified notions as **-format* and **-degree*. See Remark 10 for a comparison between the starred and unstarred versions of format and degree.

Definition 3 (**-format and *-degree of projections*). *If $\{X_\alpha\}$, $X_\alpha \subset I^{k_\alpha}$ is a finite collection of semi-Pfaffian sets and X_α° is a connected component of X_α , we define the *-format of the sub-Pfaffian set $Y := \bigcup_\alpha \pi_n^{k_\alpha}(X_\alpha^\circ) \subset I^n$, with this particular representation as a union of projections, to be the maximum among the formats of sets X_α , and the *-degree of Y to be the sum of the degrees of sets X_α .*

Strictly speaking, the **-format* \mathcal{F} and **-degree* D is associated to a particular presentation of a sub-Pfaffian set in the form prescribed by Definition 3, with different presentations possibly giving rise to different pairs (\mathcal{F}, D) . By a slight abuse of notation we will say that a sub-Pfaffian set has format \mathcal{F} and degree D if there exists *some* presentation having this data.

We remark that since the restricted sub-Pfaffian sets form an o-minimal structure, the connected components of semi-Pfaffian (or even sub-Pfaffian) sets, as well as their projections, are again sub-Pfaffian, so the sets Y for which **-format* and **-degree* are introduced in Definition 3 are indeed sub-Pfaffian.

Remark 4. *It may be useful below for the reader to consider the notions of *-format and *-degree as sub-Pfaffian analogs of the notions of dimension and degree in the theory of algebraic or semi-algebraic geometry. Our main objective is to obtain, as in the semialgebraic case, bounds that depend polynomially on the degree for a fixed format.*

In what follows we write that a is $\text{const}(b)$ (resp., a is $\text{poly}_b(c)$) as shorthand notation meaning that $a \leq \gamma(b)$ (resp., $a \leq (c+1)^{\gamma(b)}$) where $\gamma(\cdot)$ is some *universally fixed* function (which may be different for each occurrence of this notation in the text). Here a and c denote natural numbers, and b can involve one or several arguments of any type. We also allow several arguments in $\text{poly}_b(c_1, \dots, c_n)$ which we interpret as $\text{poly}_b(c_1 + \dots + c_n)$. For each occurrence of an asymptotic notation in this text, the implied functions can be effectively and explicitly computed from the data defining the Pfaffian chain (degrees of the differential equations involved), and we omit these computations in the interest of brevity.

¹Below we will consider only *restricted* sub-Pfaffian sets, and refer to them simply as sub-Pfaffian.

1.2. Cell decompositions. We recall the standard definitions of a cylindrical cell and a cylindrical cell decomposition. Later in the paper we consider only cylindrical decompositions and omit the prefix ‘‘cylindrical’’ for brevity.

Definition 5 (Cylindrical cell). *A cylindrical cell is defined by induction as follows.*

- (1) *Cylindrical 0-cell in \mathbb{R}^n is an isolated point.*
- (2) *Cylindrical 1-cell in \mathbb{R} is an open interval $(a, b) \subset \mathbb{R}$.*
- (3) *For $n \geq 2$ and $0 \leq \ell < n$ a cylindrical $(\ell + 1)$ -cell in \mathbb{R}^n is either a graph of a continuous bounded function $f : C \rightarrow \mathbb{R}$, where C is a cylindrical ℓ -cell in \mathbb{R}^{n-1} , or else a set of the form*

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : (x_1, \dots, x_{n-1}) \in C \\ \text{and } f(x_1, \dots, x_{n-1}) < x_n < g(x_1, \dots, x_{n-1})\}, \quad (1)$$

where C is a cylindrical ℓ -cell in \mathbb{R}^{n-1} , and $f, g : C \rightarrow \mathbb{R}$ are continuous bounded functions such that $f(x_1, \dots, x_{n-1}) < g(x_1, \dots, x_{n-1})$ for all points $(x_1, \dots, x_{n-1}) \in C$.

This definition implies that any ℓ -cell is homeomorphic to an open ℓ -dimensional ball.

Definition 6 (Cylindrical cell decomposition). *A cylindrical cell decomposition \mathcal{D} of a subset $A \subset \mathbb{R}^n$ is defined by induction as follows.*

- (1) *If $n = 1$, then \mathcal{D} is a finite family of pair-wise disjoint cylindrical cells (i.e., isolated points and intervals) whose union is A .*
- (2) *If $n \geq 2$, then \mathcal{D} is a finite family of pair-wise disjoint cylindrical cells in \mathbb{R}^n whose union is A and there is a cylindrical cell decomposition of $\pi_{n-1}^n(A)$ such that $\pi_{n-1}^n(C)$ is its cell for each $C \in \mathcal{D}$.*

Let $B \subset A$. Then \mathcal{D} is compatible with B if for any $C \in \mathcal{D}$ we have either $C \subset B$ or $C \cap B = \emptyset$ (equivalently, some subset $\mathcal{D}' \subset \mathcal{D}$ is a cylindrical cell decomposition of B).

Our main result is as follows.

Theorem 1. *Let $\{Y_\alpha\}$, $Y_\alpha \subset I^n$ be a collection of N sub-Pfaffian sets of $*$ -format \mathcal{F} and $*$ -degree D . Then there exists a sub-Pfaffian cell decomposition of I^n compatible with each Y_α such that the number of cells is $\text{poly}_{\mathcal{F}}(N, D)$, their $*$ -format is $\text{const}(\mathcal{F})$, and their $*$ -degree is $\text{poly}_{\mathcal{F}}(D)$.*

The proof of Theorem 1 is given in §4.1.

1.3. Sub-pfaffian sets defined by quantified formulas. In this paper we will consider formulas in the *restricted sub-Pfaffian language*, which we define to be the first-order language consisting of logical connectives, quantifiers, atomic predicates of the form $(\mathbf{x} \in Y)$ where Y is a restricted sub-Pfaffian set, and no function symbols. All formulas referred to below are assumed to be in this language. We introduce the notion of $*$ -format and $*$ -degree for quantified formulas in this language as follows.

Definition 7 ($*$ -format and $*$ -degree of quantified formulas). *Let ϕ be a formula, not necessarily in a prenex form. We define the $*$ -degree $D(\phi)$ to be the sum of $D(Y)$ for all Y appearing in the atomic predicates of ϕ . We inductively define the $*$ -format $\mathcal{F}(\phi)$ as follows.*

- For $\phi = (\mathbf{x} \in Y)$ we define $\mathcal{F}(\phi)$ to be the $*$ -format of Y .
- For $\phi = \bigvee_{j=1}^k \phi_j$ we define $\mathcal{F}(\phi) = \max_j \mathcal{F}(\phi_j)$.
- For $\phi = \bigwedge_{j=1}^k \phi_j$ we define $\mathcal{F}(\phi) = 1 + \max_j \mathcal{F}(\phi_j)$.
- For $\phi = \exists \mathbf{x} \phi'$ we define $\mathcal{F}(\phi) = \mathcal{F}(\phi')$.
- For $\phi = \neg \phi', \forall \mathbf{x} \phi'$ we define $\mathcal{F}(\phi) = 1 + \mathcal{F}(\phi')$.

In particular, the format $\mathcal{F}(\phi)$ is bounded from above by the maximum over the formats of the atomic predicates of ϕ , plus the depth of the parse-tree for ϕ .

Definition 7 is motivated by the following theorem, which shows that the $*$ -format and $*$ -degree of a set defined by a sub-Pfaffian formula ϕ can be bounded in terms of $\mathcal{F}(\phi)$ and $D(\phi)$.

Theorem 2. *Let ϕ be a sub-Pfaffian formula as above, with $*$ -format \mathcal{F} and $*$ -degree D . Then the set defined by ϕ has $*$ -format $\text{const}(\mathcal{F})$ and $*$ -degree $\text{poly}_{\mathcal{F}}(D)$.*

The proof of Theorem 2 is given in §4.2.

1.4. Upper bound on homologies. Theorem 1 implies in particular an upper bound $\text{poly}_{\mathcal{F}}(D)$ for the number of connected components of a sub-Pfaffian set of $*$ -format \mathcal{F} and $*$ -degree D . Below we generalize this to an upper bound for the homology of a sub-Pfaffian set in terms of the $*$ -format and $*$ -degree.

Theorem 3. *The sum of the Betti numbers of a sub-Pfaffian set of $*$ -format \mathcal{F} and $*$ -degree D is bounded by $\text{poly}_{\mathcal{F}}(D)$.*

The proof of Theorem 3 is given in §4.3. Combining Theorem 2 with Theorem 3 we obtain an upper bound for the homologies of (restricted) sub-Pfaffian sets defined by quantified formulas, which is polynomial in the degrees of the Pfaffian functions involved (for formulas of a fixed format). Such bounds have been obtained by Khovanskii [15] for Pfaffian varieties; by Zell [21] for semi-Pfaffian sets; and by Gabrielov, Vorobjov and Zell [10, 9] for sets defined using quantifiers under certain topological restrictions. See §1.6 for some discussion of these results. However in spite of the significant work around the complexity of sub-Pfaffian sets, a polynomial estimate for general definable sets as provided by Theorem 3 does not seem to have been previously known.

1.5. Applications and motivation. Our goal in this paper is to provide a framework that can be used to effectivize most of the results of o-minimal geometry in the restricted sub-Pfaffian structure, with polynomial dependence on the degree. Indeed, many proofs in the theory of o-minimality can be carried out verbatim using Theorems 1 and 2 to obtain such polynomial bounds. As an example we have the following.

Corollary 8. *Let $A \subset I^n$ be a restricted sub-Pfaffian set of $*$ -format \mathcal{F} and $*$ -degree D . Then the topological closure and the smooth part of A have $*$ -format $\text{const}(\mathcal{F})$ and $*$ -degree $\text{poly}_{\mathcal{F}}(D)$.*

Let $f : A \rightarrow B$ be a restricted sub-Pfaffian map of $$ -format \mathcal{F} and $*$ -degree D . Then for every $r \in \mathbb{N}$ the C^r -smooth locus of f has $*$ -format $\text{const}(\mathcal{F}, r)$ and $*$ -degree $\text{poly}_{\mathcal{F}, r}(D)$.*

Proof. All of the statements follow by defining the relevant sets using first-order formulas in the restricted sub-Pfaffian language and applying Theorem 2.

For instance, to define the smooth part of a set A having dimension $m \in \mathbb{N}$ at each point, we can define it as a locus of points $x \in A$ such that a neighbourhood of x in A is the graph of a C^1 -smooth map. More precisely, the smooth part of A is the set of all points $x \in A$ each having a neighbourhood U in A such that there exists a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that the restriction $T|_U$ is one-to-one, $T(U)$ contains a neighbourhood of $T(x)$ in \mathbb{R}^m , and the inverse to $T|_U$ is a C^1 -smooth map with the Jacobian non-vanishing at every point in $T(U)$. It is straightforward to reduce all of this to first-order formulas (an “ ϵ - δ definition”), and we leave the details to the reader. \square

One of our main motivations for developing the theory in this paper is in relation to the Pila-Wilkie counting theorem [17] and its applications in diophantine geometry (see the survey [19]).

Several applications of the counting theorem involve the geometry of elliptic curves and abelian varieties — the most famous example perhaps being the proof of the Manin-Mumford conjecture by Pila-Zannier [18]. In these applications one considers sets defined using elliptic and abelian functions. Since these functions are restricted sub-Pfaffian by an observation of Macintyre [16], the definable sets are restricted sub-Pfaffian. One can therefore hope to effectivize the counting theorem in this context, and subsequently obtain effective results for diophantine problems. Toward this end Jones and Thomas [13] have established a version of the counting theorem for certain surfaces definable in the restricted sub-Pfaffian structure, and Jones and Schmidt have explored several applications of this in diophantine geometry [11, 12].

In an upcoming paper by the first author with Jones, Schmidt and Thomas, we use the framework developed in the present paper to extend the result of [13] to arbitrary restricted sub-Pfaffian sets of arbitrary dimension, and improve the dependence of the effective constants to make them polynomial in the degrees. This can be viewed as a case of effectivizing an o-minimal proof in the sense described above, albeit for a much more technically involved statement. We expect that this result will greatly extend the scope of potential applications in diophantine geometry, as well as give rise to more reasonable (indeed, polynomial in degrees) estimates in these applications.

1.6. Comparison with previous results. Gabrielov and Vorobjov have established many results on effectivity of operations in the restricted sub-Pfaffian category. For a survey we refer the reader to [7]. In these works, the notion of format and degree for sub-Pfaffian sets is similar to ours but more straightforward: one considers simply projections of semi-Pfaffian sets, rather than projections of their connected components. In particular, in [6] Gabrielov and Vorobjov prove a cell decomposition result quite similar to our Theorem 1, as follows (where we omit the explicit, doubly-exponential dependence on the format).

Fact 9. *Let $\{Y_\alpha\}$, $Y_\alpha \subset I^n$ be a collection of N sub-Pfaffian sets of format \mathcal{F} and degree D . Then there exists a linear transformation $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and a cylindrical cell decomposition of \mathbb{R}^n , compatible with each $L(Y_\alpha)$, such that each cell in the decomposition is sub-Pfaffian and has format $\text{poly}_{\mathcal{F}}(N, D)$, and the number of cells and their degree is $\text{poly}_{\mathcal{F}}(N, D)$.*

There are two main differences compared to Theorem 1: first, the application of the linear transformation L ; and second, more crucially, the dependence of the

format of the cells on the degree D . These are crucial limitation, which make it impossible to apply this result to obtain analogs of Theorems 2 and 3: first, because in the recursive proofs it is essential that the cells preserve the order of coordinates; and second, more crucially, because after the first recursive application the format becomes dependent on D , and the complexity of any further operations performed on such cells is no longer polynomial in D .

We are also unable to sharpen Fact 9 to eliminate the dependence of the format on D , and this appears to be a fundamental difficulty². Our main observation in the present paper is that with the revised notions of $*$ -format and $*$ -degree it is possible, at a crucial point in the cell decomposition algorithm, to produce cells with format independent of D . However other aspects of the algorithm become more delicate with these notions. The main reason is that the known approaches to effective cell decomposition involve reductions using topological closure and frontier, and in our setting one must take care to avoid different components becoming glued along their common boundary when performing these operations. For this reason the strategy of cell decomposition employed in the present paper differs significantly from that of [6].

Remark 10 (Comparing the two notions of format). *Since the number of connected components of a semi-Pfaffian set X_α of format \mathcal{F} and degree D is bounded by $\text{poly}_{\mathcal{F}}(D)$ according to Fact 17, it is clear that a sub-Pfaffian set of format \mathcal{F} and degree D has $*$ -format \mathcal{F} and $*$ -degree $\text{poly}_{\mathcal{F}}(D)$.*

To obtain bounds in the converse direction, it is necessary to compute the format and degree (as a sub-Pfaffian set) of a connected component X_α° , for a given semi-Pfaffian set X_α of format \mathcal{F} and degree D . Currently we only know how to do this by computing a cell-decomposition using Fact 9, giving $\text{poly}_{\mathcal{F}}(D)$ cells of format $\text{poly}_{\mathcal{F}}(D)$. Accordingly a sub-Pfaffian set of $$ -format \mathcal{F} and $*$ -degree D has format and degree $\text{poly}_{\mathcal{F}}(D)$. The dependence of the format on D is what makes the notions of (unstarred) format and degree much less efficient. We do not venture to conjecture whether this dependence can be removed in general, but in any case this appears to be a highly non-trivial problem.*

We also note that analogs of Theorem 3 for quantified formulas have been pursued, using an entirely different topological approach, in the work of Gabrielov, Vorobjov and Zell [10]. This paper establishes similar bounds (and also with an explicit dependence on the format) for sets defined by quantified formulas in a prenex form:

$$\{x \in [0, 1]^n : Q_1 y_1 Q_2 y_2 \cdots Q_\nu y_\nu ((x, y) \in X)\} \quad (2)$$

where X is a semi-Pfaffian set that is either open or closed, and $Q_1, \dots, Q_\nu \in \{\exists, \forall\}$. For formulas involving only existential quantifiers, Gabrielov and Vorobjov [9] have managed to remove the topological condition on X by an approximation method.

Remark 11. *In an unpublished manuscript Clutha [1] extended the method of [9] to an arbitrary number of quantifiers, combining with the ideas of [10], and claimed in particular a result implying Theorem 3 for arbitrary formulas. However the proof of this result contains a substantial gap and we are presently not able to repair it.*

²The linear transformation problem seems less fundamental, and could probably be avoided using a strategy similar to the one used in the present paper.

2. PRELIMINARIES ON SEMI-PFAFFIAN SETS

We will require the notion of a (weak) stratification of a semi-Pfaffian set.

Definition 12 ([5, Definition 5]). *A (weak) stratification of a semi-Pfaffian set X is a subdivision of X into a disjoint union of smooth, not necessarily connected, semi-Pfaffian subsets X_α called strata. The system of equalities and inequalities for each stratum X_α of codimension k includes a set of k equalities $h_{\alpha,1} = \dots = h_{\alpha,k} = 0$ whose differentials define the tangent space of X_α at every point of X_α .*

We will use the following formulation of the main result of [5].

Fact 13 ([5, Theorem 1]). *Let $X \subset I^n$ be semi-Pfaffian of format \mathcal{F} and degree D . Then there is a semi-Pfaffian stratification $X = \bigcup_\alpha X_\alpha$ of X where the number of strata and their degrees are $\text{poly}_{\mathcal{F}}(D)$ and their format is $\text{const}(\mathcal{F})$.*

Remark 14. *We will also need a parametric version of Fact 13: if $X \subset I^n \times I^m$ is semi-Pfaffian of format \mathcal{F} and degree D , then there exists a collection $\{S_\alpha\}$, $S_\alpha \subset I^n \times I^m$ of semi-Pfaffian sets with their number and degree $\text{poly}_{\mathcal{F}}(D)$ and their format $\text{const}(\mathcal{F})$ such that the following holds. For any $x \in I^n$, the collection $\{S_\alpha \cap \pi_n^{-1}(x)\}$ forms a stratification of the fiber $X \cap \pi_n^{-1}(x)$, with $\dim(S_\alpha \cap \pi_n^{-1}(x))$ independent of $x \in \pi_n(S_\alpha)$ for each α . The proof of this is the same as the proof of Fact 13, treating the coordinates in I^n as parameters and performing the construction in I^m .*

We also require the following result on the complexities of the closure \overline{X} and frontier $\partial X := \overline{X} \setminus X$ of a semi-Pfaffian set X .

Fact 15 ([4]). *Let $X \subset I^n$ be a semi-Pfaffian set of format \mathcal{F} and degree D . Then the closure \overline{X} and the frontier ∂X are semi-Pfaffian of format $\text{const}(\mathcal{F})$ and degree $\text{poly}_{\mathcal{F}}(D)$.*

Remark 16. *We will also need a parametric version of Fact 15: if $X \subset I^m \times I^n$ is semi-Pfaffian of format \mathcal{F} and degree D then the union of the fiberwise-closures*

$$\{(x, y) \in I^{n+m} : y \in \overline{X}_x\}, \quad \text{where } X_x := \{y : (x, y) \in X\} \quad (3)$$

is semi-Pfaffian of format $\text{const}(\mathcal{F})$ and degree $\text{poly}_{\mathcal{F}}(D)$. This follows from the proof of Fact 15, see [7, Remark 5.4]. The same remark holds for the fiberwise frontier (by computing the fiberwise closure, and then subtracting the X).

We will also need a bound on the sum of Betti numbers of semi-Pfaffian sets (in fact we will only require the zeroth Betti number, i.e. the number of connected components). This type of bound was proved by Khovanskii [15] for Pfaffian sets, and extended by Zell to the semi-Pfaffian class. We state only the part we need, omitting the more precise dependence on the parameters which is achieved in [21].

Fact 17 ([21, Main result]). *The sum of the Betti numbers of a semi-Pfaffian set of format \mathcal{F} and degree D does not exceed $\text{poly}_{\mathcal{F}}(D)$.*

3. CELL DECOMPOSITION OF SUB-PFAFFIAN SETS

In this section we prove a result on cell decomposition of sub-Pfaffian sets, that is the main ingredient of the proof of Theorem 1. As a shorthand, if Δ is a collection of sets we write $\pi_k(\Delta)$ for the collection of the projections of the elements of Δ . Furthermore, $\bigcup \Delta$ will stand for the union of sets in Δ . Thus, $\bigcup \pi_k(\Delta)$ is the union

of projections of sets in Δ . We will freely use the results on effective bounds for stratifications and frontiers from §2 without explicit reference.

Lemma 18 (Effective fiber cutting). *Let $X \subset I^{n+m}$ be a semi-Pfaffian set of format \mathcal{F} and degree D . Then there exists a semi-Pfaffian set \widehat{X} of format $\text{const}(\mathcal{F})$ and degree $\text{poly}_{\mathcal{F}}(D)$ such that $\pi_n(X) = \pi_n(\widehat{X})$ and $\pi_n|_{\widehat{X}}$ has zero-dimensional fibers.*

Proof. We proceed by induction on the maximal fiber dimension k of $\pi_n|_X$. If $k = 0$ we are done, and otherwise we will construct a set X' with $\pi_n(X) = \pi_n(X')$ and maximal fiber dimension smaller than k .

Let $\{S_\alpha\}$, $S_\alpha \subset I^{n+m}$ denote a fiberwise stratification of X , as in Remark 14. For any S_α with fiber dimension smaller than k , we put S_α into X' . For any S_α with fiber dimension k we do the following.

- (1) We put the fiberwise frontier of S_α into X' using Remark 16.
- (2) For any $j = 1, \dots, m$ let $C_{\alpha,j}$ be the set of fiberwise critical points of the coordinate function x_{n+j} on the fibers of $\pi_n|_{S_\alpha}$. Then we stratify the fibers of $\pi_n|_{C_{\alpha,j}}$ again by Remark 14 and put all strata of fiberwise dimension smaller than k into X' .

Clearly $\pi_n(X') \subset \pi_n(X)$. To prove the converse, let $x \in \pi_n(X)$. If the fiber of X over x has strata of dimension less than k , or strata of dimension k with non-empty frontier, we are done. Otherwise this fiber consists of smooth compact k -dimensional manifolds. Then one of the coordinate functions x_{n+j} must be non-constant on this fiber, and the corresponding set $C_{\alpha,j}$ is non-empty and locally closed in the fiber. Then the fiberwise stratification of $C_{\alpha,j}$ must contain strata of dimension less than k that we put into X' . This proves the claim. \square

The following proposition is quite similar to Theorem 1: the difference is that here we essentially assume that we are given sub-Pfaffian sets with bounded format and degree (rather than their *-analogs), but produce cell decompositions with bounded *-format and *-degree. In §4.1 we show how to deduce the general case of Theorem 1 from this statement.

Proposition 19. *Let $\{X_\alpha\}$, $X_\alpha \subset I^\ell$, where $\ell \geq n$, be a collection of N semi-Pfaffian sets of format \mathcal{F} and degree D . Then there exists a sub-Pfaffian cell decomposition of I^n compatible with each $\pi_n(X_\alpha)$ such that the number of cells is $\text{poly}_{\mathcal{F}}(N, D)$, their *-format is $\text{const}(\mathcal{F})$ and their *-degree is $\text{poly}_{\mathcal{F}}(D)$.*

Proof. We will work by lexicographic induction on (n, k) where

$$k := \max_{\alpha} \dim \pi_{n-1}(X_\alpha).$$

By Lemma 18 we may assume that $\pi_n|_{X_\alpha}$ has zero-dimensional fibers. Refining each X_α into its stratification, we may further assume without loss of generality that X_α is smooth, that $\pi_n|_{X_\alpha}$ has rank $\dim \pi_n(X_\alpha)$, and that $\pi_{n-1}|_{X_\alpha}$ has constant rank. Let $\Pi := \{X_\alpha\}$. Using Fact 15 we may also assume that Π is closed under taking frontiers.

Let Π_k (resp. Π_{k+1} and $\Pi_{<k}$) denote the collection of sets X_α with $\dim X_\alpha = k$ (resp. $\dim X_\alpha = k + 1$ and $\dim X_\alpha < k$).

For each $X_\alpha \in \Pi_{k+1}$ we also apply Lemma 18 to $\pi_{n-1}|_{X_\alpha}$ and add the resulting sets to Π (after stratifying as above). With this modification we may assume that

$$\pi_{n-1}(\Pi_{k+1}) \subset \pi_{n-1}(\Pi_k \cup \Pi_{<k}). \quad (4)$$

We construct a set Σ of closed semi-Pfaffian sets of bounded *-format and *-degree, satisfying $\dim \bigcup \pi_{n-1}(\Sigma) < k$. First, we put the closures of the sets in $\Pi_{<k}$ into Σ .

Next, we wish to ensure that $G := \bigcup \pi_{n-1}(\Pi) \setminus \bigcup \pi_{n-1}(\Sigma)$ is smooth, and that for every $X_\alpha \in \Pi_{k+1}$ the intersection $\pi_n(X_\alpha) \cap (G \times I)$ is open in $G \times I$. By (4) and since we already put $\Pi_{<k}$ in Σ we have $G = \bigcup \pi_{n-1}(\Pi_k) \setminus \bigcup \pi_{n-1}(\Sigma)$. Since $\pi_{n-1}|_{X_\alpha}$ has constant rank k for every $X_\alpha \in \Pi_k$, the set $\bigcup \pi_{n-1}(\Pi_k)$ can be thought of as an immersed smooth manifold with possible self-intersections. It may be non-smooth only in points $p \in \bigcup \pi_{n-1}(\Pi_k)$ where:

- (1) there exist two points $p_\alpha \in X_\alpha \in \Pi_k$ and $p_\beta \in X_\beta \in \Pi_k$ such that $p = \pi_{n-1}(p_\alpha) = \pi_{n-1}(p_\beta)$ but the projections of the germs (X_α, p_α) and (X_β, p_β) under π_{n-1} (which are both smooth manifolds) are different;
- (2) the point p is also in $\bigcup \pi_{n-1}(\{\partial X_\alpha \mid X_\alpha \in \Pi_k\})$.

We show how to put the points corresponding to these two cases in Σ , and it then follows that G is smooth. The second case is rather simple: for any $X_\alpha \in \Pi_{<k} \cup \Pi_k$ we add the frontier ∂X_α to Σ . We also note for later purposes that, since the sets X_α are bounded, this implies

$$\partial \pi_n(X_\alpha) \subset \bigcup \pi_n(\Sigma). \quad (5)$$

To handle the first case, consider for every pair $X_\alpha, X_\beta \subset I^\ell$ in Π_k the set

$$\{(x, y) \in X_\alpha \times X_\beta : x_1 = y_1, \dots, x_{n-1} = y_{n-1}\} \subset I^\ell \times I^\ell \quad (6)$$

where x denotes the coordinates on the first I^ℓ and y on the second. In what follows we will be routinely adjusting the dimension ℓ so that the set I^ℓ contains the new semi-Pfaffian sets we create, in this case the set in (6). We stratify this set and put (the closure of) every stratum of dimension less than k into Σ . It is clear that these strata cover the points described in the first case. Thus we established that G is smooth.

For $X_\alpha \in \Pi_{k+1}$ we know that $\pi_n|_{X_\alpha}$ has constant rank $k+1$ and $\pi_{n-1}|_{X_\alpha}$ has constant rank k . Therefore, the projection $X_\alpha \cap \pi_{n-1}^{-1}(G) \rightarrow G \times I$ is a submersion, and indeed $\pi_n(X_\alpha) \cap (G \times I)$ is open in $G \times I$.

For later purposes we put one more set into Σ , controlling the possible interactions between different elements of Π_k . For any $X_\alpha, X_\beta \in \Pi_k$ we consider the set

$$Z_{\alpha, \beta} := (\{(x, y) \in X_\alpha \times X_\beta : x_1 = y_1, \dots, x_n = y_n\}) \subset I^\ell \times I^\ell. \quad (7)$$

We stratify $Z_{\alpha, \beta}$ and add the closures of the strata of dimension at most $k-1$, and the frontier of the strata of dimension k , to Σ .

Now apply induction to obtain a cell decomposition of I^{n-1} compatible with $\pi_{n-1}(\Pi)$ and with $\pi_{n-1}(\Sigma)$. If a cell C is disjoint from $\bigcup \pi_{n-1}(\Pi)$ then $C \times I$ is a cell compatible with every $\pi_n(X_\alpha)$.

Let C be a cell contained in $G = \bigcup \pi_{n-1}(\Pi) \setminus \bigcup \pi_{n-1}(\Sigma)$. We will show how to construct cells over C that are compatible with every $\pi_n(X_\alpha)$. First consider $X_\alpha \in \Pi_{k+1}$. Recall that $\pi_n(X_\alpha) \cap (G \times I)$ is open in $G \times I$. It follows that its boundary in $G \times I$ agrees with its frontier. If a cell in $C \times I$ is compatible with the boundary of $\pi_n(X_\alpha)$ (or, equivalently in this case, with the frontier) then it is also compatible with $\pi_n(X_\alpha)$ by elementary topology. Since we assume that Π is closed under taking frontiers, it will be enough to construct cells over C compatible with $\pi_n(\Pi_k)$.

We claim that for any $X_\alpha \in \Pi_k$, the set $\pi_n(X_\alpha)$ is a union of graph cells over C . Consider an arbitrary point $c_0 \in C$. Then the fiber of $\pi_n(X_\alpha)$ over c_0 consists of finitely many points. By construction, $\pi_n(X_\alpha)$ maps submersively to C , so as we move c_0 these points locally move continuously. Moreover the points must remain in $\pi_n(X_\alpha)$ as we continue globally to C , for otherwise one of these points would meet $\partial\pi_n(X_\alpha)$ contradicting (5) since $C \subset G$ is disjoint from $\pi_{n-1}(\Sigma)$. Note that for the same reasons X_α itself is also a union of graphs of continuous maps over C . If $s : C \rightarrow I$ is such a section of $\pi_n(X_\alpha)$ we denote by \hat{s} the corresponding extension to a section of X_α .

If $s_\alpha, s_\beta : C \rightarrow I$ is a pair of such sections obtained from X_α, X_β then one of

$$s_\alpha < s_\beta \qquad s_\alpha = s_\beta \qquad s_\alpha > s_\beta \qquad (8)$$

holds uniformly over C . Indeed, otherwise there should be a point $c_0 \in C$ where $s_\alpha(c_0) = s_\beta(c_0)$ and a neighborhood where this does not hold identically. This implies that $(c_0, c_\alpha(c_0))$ belongs to a stratum of dimension at most $(k-1)$, or to the frontier of a k -dimensional stratum of $Z_{\alpha,\beta}$, and hence $c_0 \notin G$ contradicting $C \subset G$.

By the above, the set of sections obtained from any of the $X_\alpha \in \Pi_k$ is

$$\{s_1 < \dots < s_q\},$$

where we may suppose for simplicity that $s_1 = 0$ and $s_q = 1$. Also note since $\pi_n(X_\alpha)$ is a union of disjoint graphs of continuous functions, their number is bounded by the number of connected components of X_α . By Fact 17 we therefore get at most $\text{poly}_{\mathcal{F}}(D)$ sections from each X_α , and in total $q = \text{poly}_F(N, D)$.

A cell decomposition of $C \times I$, compatible with each X_α , is given by the cells

$$\begin{aligned} &\{x_n = s_1(x_1, \dots, x_{n-1})\}, \\ &\{s_1(x_1, \dots, x_{n-1}) < x_n < s_2(x_1, \dots, x_{n-1})\}, \\ &\dots, \{x_n = s_q(x_1, \dots, x_{n-1})\}. \end{aligned} \quad (9)$$

We show that each of these cells is a sub-Pfaffian set with appropriately bounded *-format and *-degree. Specifically we show that the graph of each section s_j over C is a sub-Pfaffian set with the appropriate *-bounds, and it is then a simple exercise to construct each of the cells above (where for the interval cells one repeats the construction for s_j, s_{j+1} and works in the direct product).

Recall that $C = \pi_n^\ell(Z^\circ)$ is the projection of the connected component Z° of some semi-Pfaffian set $Z \subset I^\ell$ (for an appropriate ℓ). Suppose that s_j is a section corresponding to X_α . Then the graph of s_j is a connected component of the set $\pi_n(X_\alpha) \cap (C \times I)$. In fact, since the section s_j extends to a section of X_α , the graph of s_j the projection $\pi_n(W_j^\circ)$ of a connected component W_j° of the set

$$W_j := \{(x, y) \in Z \times X_\alpha : x_1 = y_1, \dots, x_{n-1} = y_{n-1}\} \quad (10)$$

given by Z° in the x -coordinates and by the graph of $y = \hat{s}_j(x)$. Here we ordered the coordinates in such a way that $\pi_n(x, y) = (y_1, \dots, y_n)$. This proves that the graph of s_j is indeed sub-Pfaffian with appropriately bounded *-format and *-degree. Since each cell is constructed using one or two of these graphs, their *-degree is indeed bounded by $\text{poly}_{\mathcal{F}}(D)$ (noting in particular that there is no dependence on N).

Finally, we must construct cells covering $\bigcup \pi_{n-1}(\Sigma) \times I$ and compatible with every $\pi_n(X_\alpha)$. Let

$$Y_\alpha := \left(\bigcup \pi_{n-1}(\Sigma) \times I \right) \cap \pi_n(X_\alpha) = \pi_n(\widehat{Y}_\alpha), \quad (11)$$

where it is easy to choose \widehat{Y}_α to be a semi-Pfaffian set of bounded complexity. It will suffice to construct a cell decomposition of I^n compatible with $\bigcup \pi_n(\Sigma) \times I$ and $\{Y_\alpha\}$ and take from it only the cells covering $\bigcup \pi_n(\Sigma) \times I$. This can now be achieved by induction on k , noting that for any Z in this collection $\pi_{n-1}(Z) \subset \pi_{n-1}(\Sigma)$ has dimension strictly less than k . \square

4. PROOFS OF THE MAIN RESULTS

4.1. Proof of Theorem 1. We will require a couple of elementary lemmas, whose proofs are left an exercises for the reader.

Lemma 20. *Suppose that a cell C is compatible with a set X . Then it is compatible with each connected component of X .*

Lemma 21. *Suppose that a cell decomposition of I^k is compatible with a set A . Then the induced decomposition on I^n is compatible with $\pi_n^k(A)$.*

We are now ready to finish the proof. Let $X_{\alpha,\beta} \subset I^k$ denote semi-Pfaffian sets and $X_{\alpha,\beta}^\circ$ connected components such that $Y_\alpha = \bigcup_\beta \pi_n^k(X_{\alpha,\beta}^\circ)$. Use Proposition 19 to find a sub-Pfaffian cell decomposition of I^k compatible with $\{X_{\alpha,\beta}\}$ with suitably bounded *-format and *-degree. Then by the preceding lemmas this cell decomposition is compatible with $X_{\alpha,\beta}^\circ$, and the induced decomposition on I^n is therefore compatible with Y_α .

Remark 22. *Note that in the proof of Theorem 1, even though we are interested in a cell decomposition of I^n , we apply Proposition 19 to obtain a cell decomposition of the full space I^k . This allows us to ensure the compatibility with the projections of each connected component $X_{\alpha,\beta}^\circ$ separately. Applying Proposition 19 to the projections $\pi_n(X_{\alpha,\beta})$ would not suffice. It is therefore crucial at this point that Proposition 19 produces a cell decomposition in the original coordinates, without applying a linear transformation (or at least preserving the projection π_n^k). For this reason, the cell decomposition algorithm of Gabrielov-Vorobjov [6], which applies such a linear transformation to the coordinates, would not suffice for our purposes.*

4.2. Proof of Theorem 2. The proof is a routine recursive argument on the structure of ϕ . The statements for existential quantifiers and disjunctions hold by definition. Up to re-writing

$$\forall \mathbf{x} \phi \equiv \neg \exists \mathbf{x} (\neg \phi), \quad (12)$$

$$\bigwedge_{i=1}^k \phi_i \equiv \neg \bigvee_{i=1}^k (\neg \phi_i), \quad (13)$$

everything else follows from the statement for negations $\phi = \neg \phi'$. This is proved by constructing a cell decomposition of I^n compatible with the set X defined by ϕ' , and taking the union of all cells disjoint from X in this decomposition.

4.3. Proof of Theorem 3. We first establish an effective result on the existence of triangulations in the restricted sub-Pfaffian class. Recall that a *finite simplicial complex* in \mathbb{R}^n is a finite collection $K = \{\bar{\sigma}_1, \dots, \bar{\sigma}_p\}$ of (closed) simplices $\bar{\sigma}_i \subset \mathbb{R}^n$ such that the intersection of any pairs $\bar{\sigma}_i \cap \bar{\sigma}_j$, if not empty, is a common face of $\bar{\sigma}_i$ and $\bar{\sigma}_j$; and such that any face of any $\bar{\sigma}_i$ also belongs to K . We write $|K|$ for the union of all simplices in K .

Theorem 4. *Let $Y \subset [0, 1]^n$ be a closed sub-Pfaffian set, and $X_1, \dots, X_k \subset Y$ be sub-Pfaffian subsets, and suppose all these sets have $*$ -format bounded by \mathcal{F} and $*$ -degree bounded by D .*

Then there exists a finite simplicial complex K with vertices in \mathbb{Q}^n and a definable homeomorphism $\Phi : |K| \rightarrow Y$ such that each X_i is a union of images by Φ of open simplices of K . Moreover K contains $\text{poly}_{\mathcal{F}}(D)$ simplices, and Φ has $$ -format $\text{const}(\mathcal{F})$ and $*$ -degree $\text{poly}_{\mathcal{F}}(D)$.*

A proof of the triangulation theorem in the o-minimal setting can be found in [20], where it follows a similar proof for the semialgebraic class in [3]. For convenience we refer the reader to the alternative presentation given in [2, Theorem 4.4], which also establishes Theorem 4 without the effective estimates for arbitrary o-minimal structures. Deriving the effective estimates from this proof in the restricted sub-Pfaffian context is a routine exercise in the application of Theorems 1 and 2: one only needs to verify that in the proof of [2, Theorem 4.4], the triangulating map Φ is indeed defined by a first-order formula ϕ with $\mathcal{F}(\phi) = \text{const}(\mathcal{F})$ and $D(\phi) = \text{poly}_{\mathcal{F}}(D)$. As this verification is entirely straightforward from the presentation of loc. cit., we leave the details as an exercise for the reader.

To deduce Theorem 3 we first apply Theorem 4 with $Y = [0, 1]^n$ and X the given sub-Pfaffian set. We obtain a homeomorphism $\Phi : |K| \rightarrow [0, 1]^n$. In particular, the sum of the Betti numbers of X is equal to that of $\Phi^{-1}(X)$, which is a union of at most $\text{poly}_{\mathcal{F}}(D)$ simplices. This set being semialgebraic, the bound on the sum of Betti numbers now follows, e.g., from [8, Theorem 1].

REFERENCES

- [1] M. Clutha. Bounding betti numbers of sets definable in o-minimal structures over the reals, 2012.
- [2] M. Coste. *An Introduction to O-minimal Geometry*. 1999.
- [3] H. Delfs and M. Knebusch. Homology of algebraic varietie over real closed fields. *J. reine angew. Math.*, 335:122–163, 1982.
- [4] A. Gabrielov. Frontier and closure of a semi-Pfaffian set. *Discrete Comput. Geom.*, 19(4):605–617, 1998.
- [5] A. Gabrielov and N. Vorobjov. Complexity of stratifications of semi-Pfaffian sets. *Discrete Comput. Geom.*, 14(1):71–91, 1995.
- [6] A. Gabrielov and N. Vorobjov. Complexity of cylindrical decompositions of sub-Pfaffian sets. *J. Pure Appl. Algebra*, 164(1-2):179–197, 2001.
- [7] A. Gabrielov and N. Vorobjov. Complexity of computations with Pfaffian and Noetherian functions. In *Normal forms, bifurcations and finiteness problems in differential equations*, volume 137 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 211–250. Kluwer Acad. Publ., Dordrecht, 2004.
- [8] A. Gabrielov and N. Vorobjov. Betti numbers of semialgebraic sets defined by quantifier-free formulae. *Discrete Comput. Geom.*, 33:395–401, 2005.
- [9] A. Gabrielov and N. Vorobjov. Approximation of definable sets by compact families, and upper bounds on homotopy and homology. *J. Lond. Math. Soc. (2)*, 80(1):35–54, 2009.
- [10] A. Gabrielov, N. Vorobjov, and T. Zell. Betti numbers of semialgebraic and sub-Pfaffian sets. *J. London Math. Soc. (2)*, 69(1):27–43, 2004.

- [11] G. Jones and H. Schmidt. A Manin-Mumford theorem for the maximal compact subgroup of a universal vectorial extension of a product of elliptic curves. *Int. Math. Res. Not.*, To appear.
- [12] G. Jones and H. Schmidt. Pfaffian definitions of Weierstrass elliptic functions. *Math. Ann.*, To appear.
- [13] G. Jones and M. E. M. Thomas. Effective Pila–Wilkie bounds for unrestricted Pfaffian surfaces. *Math. Ann.*, To appear.
- [14] A. Khovanskii. On a class of systems of transcendental equations. *Soviet Math. Dokl.*, 22:762–765, 1980.
- [15] A. Khovanskii. *Fewnomials*. Amer. Math. Soc., Providence, RI, 1991.
- [16] Angus Macintyre. Some observations about the real and imaginary parts of complex Pfaffian functions. In *Model theory with applications to algebra and analysis. Vol. 1*, volume 349 of *London Math. Soc. Lecture Note Ser.*, pages 215–223. Cambridge Univ. Press, Cambridge, 2008.
- [17] J. Pila and A. J. Wilkie. The rational points of a definable set. *Duke Math. J.*, 133(3):591–616, 2006.
- [18] Jonathan Pila and Umberto Zannier. Rational points in periodic analytic sets and the Manin-Mumford conjecture. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 19(2):149–162, 2008.
- [19] Thomas Scanlon. Counting special points: logic, Diophantine geometry, and transcendence theory. *Bull. Amer. Math. Soc. (N.S.)*, 49(1):51–71, 2012.
- [20] L. van den Dries. *Tame Topology and O-minimal Structures*. Cambridge University Press, Cambridge, 1998.
- [21] T. Zell. Betti numbers of semi-Pfaffian sets. *J. Pure Appl. Algebra*, 139(1-3):323–338, 1999. Effective methods in algebraic geometry (Saint-Malo, 1998).

WEIZMANN INSTITUTE OF SCIENCE, REHOVOT, ISRAEL
E-mail address: gal.binyamini@weizmann.ac.il

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF BATH, BATH, BA2 7AY, UK
E-mail address: masnrv@bath.ac.uk