

CANONICAL FORESTS IN DIRECTED FAMILIES

JOSEPH FLENNER AND VINCENT GUINGONA

ABSTRACT. Two uniqueness results on representations of sets constructible in a directed family of sets are given. In the unpackable case, swiss cheese decompositions are unique. In the packable case, they are not unique but admit a quasi-ordering under which the minimal decomposition is unique. Both cases lead to a one-dimensional elimination of imaginaries in VC-minimal and quasi-VC-minimal theories.

1. INTRODUCTION

In this paper, we study canonical forms for sets constructible from a directed family of sets, in the sense of Adler [1]. Every such set is realizable as a disjoint union of swiss cheeses, balls with (finitely many) holes removed. Even after imposing some nontriviality conditions on this decomposition, however, the uniqueness of such a presentation can fail. The dividing line is given by the notion of packability. Without packability, we show in Section 2 that the swiss cheese decomposition is unique. Even in the presence of packability, however, it is possible to canonically choose an ‘optimal’ decomposition. This is described in Section 3.

Directed families arise in logic as the building blocks of the VC-minimal theories. Introduced by Adler in 2008, these theories have garnered interest for being well-situated in the realm of model-theoretic tameness. Classical examples such as strongly minimal and o-minimal theories are VC-minimal. On the other hand, VC-minimal theories can also be seen as a natural ‘simplest case’ among the dependent theories. Some fundamental model-theoretic machinery has already been developed, for example Cotter and Starchenko’s recent analysis of forking in VC-minimal theories [2].

A prototypical example is given by algebraically closed valued fields, from which much of the language of directed families is derived. Our primary goal is to give suitable generalizations of Holly’s study of definable sets in algebraically closed valued fields [6, 7], including the elimination of imaginaries in one dimension which is detailed in Section 4. We also point out how these results can be adapted to the somewhat weaker quasi-VC-minimal setting.

Throughout the paper, we work in a directed family of sets as defined below. It should be acknowledged that this notion is not in accordance with some other uses of the term ‘directed’, such as in category theory.

Date: November 10, 2011.

2010 Mathematics Subject Classification. Primary: 06A07, 03C45.

Key words and phrases. Directed family of sets, swiss cheese decomposition, VC-minimality, elimination of imaginaries.

Both authors were partially supported by the NSF.

1.1. **Directed families.** For any set \mathfrak{U} and $\mathcal{B} \subseteq \mathcal{P}(\mathfrak{U}) \setminus \{\emptyset\}$, \mathcal{B} is *directed* if, for all $B_0, B_1 \in \mathcal{B}$, one of the following holds:

- (i) $B_0 \subseteq B_1$
- (ii) $B_1 \subseteq B_0$
- (iii) $B_0 \cap B_1 = \emptyset$.

\mathfrak{U} is the *universe* of \mathcal{B} . The members of \mathcal{B} are called *balls*, and a *constructible set* is a (finite) boolean combination of balls.

If \mathcal{B} is directed, then (\mathcal{B}, \subseteq) is easily seen to be a forest, that is, a union of trees whose roots are the maximal sets in \mathcal{B} . Moreover, if \mathcal{B} is directed, then so is $\mathcal{B} \cup \{\mathfrak{U}\}$. Thus we may assume that $\mathfrak{U} \in \mathcal{B}$, as will generally be necessary.

2. SWISS CHEESE AND UNPACKABILITY

In this section, we study the representation of constructible sets as boolean combinations of balls and the relation of these representations to unpackability. A directed family is called *unpackable* if no ball is a finite union of proper sub-balls.

Fix a directed family $\mathcal{B} \subseteq \mathcal{P}(\mathfrak{U}) \setminus \{\emptyset\}$ with $\mathfrak{U} \in \mathcal{B}$. A *swiss cheese* is a subset of \mathfrak{U} of the form $S = A \setminus (B_1 \cup \dots \cup B_n)$, where A is a ball and $B_1, \dots, B_n \subsetneq A$ are proper sub-balls of A . In this expression, n may be 0 but it is usually assumed that the expression is nonredundant in the sense that for no $i \neq j$ is $B_i \subseteq B_j$. In this case, A is called a *wheel* of S and each B_i is a *hole*. Note that this notion of wheels and holes is not immediately intrinsic to the set S , but depends on its presentation as a swiss cheese. In fact, we will show in Theorem 2.2 that as long as the holes are pairwise disjoint, the unique determination of wheels and holes is equivalent to unpackability.

If $X \subseteq \mathfrak{U}$ is constructible, a *swiss cheese decomposition* of X is a finite collection of swiss cheeses S_1, \dots, S_n such that

- (i) $S_i \cap S_j = \emptyset$ for $i \neq j$,
- (ii) for no i, j is a wheel of S_i equal to a hole of S_j , and
- (iii) $X = S_1 \cup \dots \cup S_n$.

We again allow $n = 0$, so that \emptyset always has the empty swiss cheese decomposition.

Lemma 2.1. *Every constructible subset $X \subseteq \mathfrak{U}$ has a swiss cheese decomposition.*

Proof. To begin, write X as a boolean combination of balls in the form

$$X = \bigcup_{i=1}^m \bigcap_{j=1}^{n_i} B_{i,j}^{e(i,j)}$$

(where $e(i, j) \in \{0, 1\}$, $B^1 = B$, and $B^0 = \mathfrak{U} \setminus B$ is the complement of B). Considering one of the disjuncts $\bigcap_{j=1}^{n_i} B_{i,j}^{e(i,j)}$, let C_1, \dots, C_r be those balls that appear positively (i.e. $B_{i,j}$ for $e(i, j) = 1$) and D_1, \dots, D_s those that appear negatively. Due to the assumption that the universe \mathfrak{U} is a ball, we can always assume that $r \geq 1$.

By the intersection property of balls, either $C_1 \cap \dots \cap C_r = \emptyset$ or $C_1 \cap \dots \cap C_r = C_i$ for some i . In the former case, the disjunct is empty and can be removed altogether. In the latter case, we write $C = C_i$ and have

$$\bigcap_{j=1}^{n_i} B_{i,j}^{e(i,j)} = C \setminus (D_1 \cup \dots \cup D_s).$$

If this is nonempty, it is a swiss cheese.

We have thus shown that X is a union of swiss cheeses. Next suppose two of these swiss cheeses have nonempty intersection, $S_1 \cap S_2 \neq \emptyset$. Writing

$$S_i = A_i \setminus (B_{i,1} \cup \dots \cup B_{i,r_i})$$

with $B_{i,j} \cap B_{i,k} = \emptyset$ for $j \neq k$, $A_1 \cap A_2 \neq \emptyset$ implies $A_1 \subseteq A_2$ or $A_2 \subseteq A_1$. Say the latter. Now let C_1, \dots, C_s be a list of all those balls $C \in \{B_{i,j}\}$ such that one of

- $C = B_{1,j} \subseteq A_1 \setminus A_2$
- $C = B_{2,j}$ and $C \subseteq B_{1,k}$ for some $j \leq r_2$, $k \leq r_1$ —i.e., C is a hole of S_2 contained in a hole of S_1
- $C = B_{1,j}$ and $C \subseteq B_{2,k}$ for some $j \leq r_1$, $k \leq r_2$ —i.e., C is a hole of S_1 contained in a hole of S_2

holds. Then it is straightforward to verify that $S_1 \cup S_2 = A_1 \setminus (C_1 \cup \dots \cup C_s)$. By repeating this process as necessary, we may now assume that we have written X as a *disjoint* union of swiss cheeses: $X = S_1 \cup \dots \cup S_n$, $S_i \cap S_j = \emptyset$ for $i \neq j$.

Finally, suppose that the wheel of one of these is a hole of another,

$$\begin{aligned} S_i &= A_i \setminus (B_1 \cup \dots \cup B_r) \\ S_j &= A_j \setminus (A_i \cup C_1 \cup \dots \cup C_s). \end{aligned}$$

It may of course be assumed that $A_i \cap C_k = \emptyset$ for every $k \leq s$ (recalling that nonredundancy is assumed in the definition of holes). But in this case, clearly

$$S_i \cup S_j = A_j \setminus (B_1 \cup \dots \cup B_r \cup C_1 \cup \dots \cup C_s).$$

Again, this process of elimination can be repeated until arriving at a swiss cheese decomposition of X . \square

Swiss cheese decompositions can be canonically derived from any finite set of balls. Consider such a set \mathcal{S} . Then \mathcal{S} is partially ordered by \subseteq , and as noted in the introduction this ordering is a (finite) forest. Define the levels $\text{Lev}_n(\mathcal{S})$ inductively for $n \geq 0$ by:

$$\text{Lev}_n(\mathcal{S}) = \left\{ B \in \mathcal{S} \mid B \text{ is } \subseteq\text{-maximal in } \mathcal{S} \setminus \bigcup_{i < n} \text{Lev}_i(\mathcal{S}) \right\}.$$

For convenience, define also $\lambda(B, \mathcal{S}) = n$ if $B \in \text{Lev}_n(\mathcal{S})$ and, for $B \in \mathcal{S}$,

$$\text{Sub}(B, \mathcal{S}) = \{ C \in \text{Lev}_{\lambda(B, \mathcal{S})+1}(\mathcal{S}) \mid C \subseteq B \}.$$

Then from \mathcal{S} we construct the swiss cheese decomposition

$$\text{Ch}(\mathcal{S}) = \bigcup_{B \in \text{Lev}_{2n}(\mathcal{S})} B \setminus \left(\bigcup_{C \in \text{Sub}(B, \mathcal{S})} C \right).$$

In other words, the balls on even levels are the wheels and the holes are the wheels' immediate predecessors in (\mathcal{S}, \subseteq) .

However, while constructible sets always admit a swiss cheese decomposition, and such a decomposition can always be realized as $\text{Ch}(\mathcal{S})$ for suitable choice of \mathcal{S} , it may occur that $\text{Ch}(\mathcal{S}) = \text{Ch}(\mathcal{T})$ even though $\mathcal{S} \neq \mathcal{T}$. The notion of unpackability is central in obtaining uniqueness of the decomposition. In fact, it is equivalent. In the next theorem, (4) was first proved by Holly [6] in the case of algebraically closed valued fields. The observation that Holly's theorem holds in this more general unpackable setting is due to Dolich (unpublished).

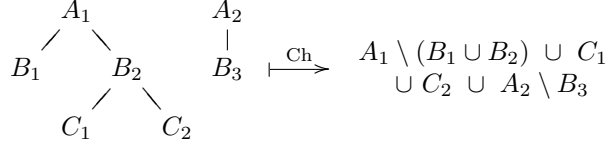


FIGURE 1. The swiss cheese operator

Theorem 2.2. *For a directed family \mathcal{B} containing its universe \mathfrak{A} , the following are equivalent:*

(1) \mathcal{B} is unpackable.

(2) If A, B_1, \dots, B_n are balls such that $A \subseteq \bigcup_{i=1}^n B_i$, then $A \subseteq B_i$ for some i .

(3) If

$$S = A_1 \setminus (B_{1,1} \cup \dots \cup B_{1,m}) = A_2 \setminus (B_{2,1} \cup \dots \cup B_{2,n})$$

is a swiss cheese and $B_{i,j} \cap B_{i,k} = \emptyset$ for $i \in \{1, 2\}$, $j \neq k$, then $A_1 = A_2$ and

$$\{B_{1,1}, \dots, B_{1,m}\} = \{B_{2,1}, \dots, B_{2,n}\}.$$

(4) Every constructible set admits a unique swiss cheese decomposition.

Proof. We first show the equivalence of (1-3), and then show the equivalence of these three conditions with (4).

1 \Rightarrow 2: Suppose $A \subseteq \bigcup_{i=1}^n B_i$. We may assume that $A \cap B_i \neq \emptyset$ for each i . But then either $B_i \subseteq A$ or $A \subseteq B_i$. If $B_i \subsetneq A$ for every i , then A would be a finite union of proper sub-balls, contradicting unpackability.

2 \Rightarrow 3: Since $A_1 \subseteq A_2 \cup B_{1,1} \cup \dots \cup B_{1,m}$, (2) gives $A_1 \subseteq A_2$ or $A_1 \subseteq B_{1,i}$, some i . But the latter cannot occur as the holes in a swiss cheese are presumed to be proper sub-balls of the wheel. Therefore $A_1 \subseteq A_2$, and by symmetry, $A_1 = A_2$. It follows that

$$B_{1,1} \cup \dots \cup B_{1,m} = B_{2,1} \cup \dots \cup B_{2,n}$$

and hence $B_{1,i} \subseteq B_{2,j}$, some $j \leq n$. Similarly, $B_{2,j} \subseteq B_{1,k}$. But here we must have $k = i$, since $B_{1,i} \cap B_{1,k} \neq \emptyset$ for $i \neq k$. Thus $B_{1,i} = B_{2,j}$ and

$$\{B_{1,1}, \dots, B_{1,m}\} \subseteq \{B_{2,1}, \dots, B_{2,n}\}.$$

Again, (3) follows by symmetry.

3 \Rightarrow 1: Suppose a ball A were the disjoint union of proper sub-balls B_1, \dots, B_n . Necessarily $n > 1$. Then the swiss cheese $B_1 = A \setminus (B_2 \cup \dots \cup B_n)$ contradicts (3).

4 \Rightarrow 1: This is similarly clear. If, for example, the ball A were the disjoint union of proper sub-balls B_1, \dots, B_n , then $X = A$ could be decomposed either as simply A , or as the union of the collection B_1, \dots, B_n .

1 \Rightarrow 4: In light of Lemma 2.1, it remains only to prove uniqueness. Suppose we have two decompositions, $X = \bigcup_{i=1}^r S_i = \bigcup_{i=1}^s T_i$ with

$$S_i = A_i \setminus (A_{i,1} \cup \dots \cup A_{i,m_i})$$

$$T_i = B_i \setminus (B_{i,1} \cup \dots \cup B_{i,n_i}).$$

We work by induction on r . If $r = 0$, then $X = \emptyset$. Since no B_i can be the union of its proper sub-balls $B_{i,1}, \dots, B_{i,n_i}$, no T_i can be empty. Thus, $s = 0$ as well.

For $r > 0$, note that $S_1 \subseteq X$ implies

$$A_1 \subseteq \left(\bigcup_{i=1}^s B_i \right) \cup \left(\bigcup_{j=1}^{m_1} A_{1,j} \right).$$

By (2), it follows that A_1 is a subset of one of these balls. But since $S_1 \neq \emptyset$, we must have $A_1 \subseteq B_i$ for some $i \leq s$. By the same reasoning, $B_i \subseteq A_j$ for some $j \leq r$. This can be repeated until one of the balls appears twice, giving equality. So, renumbering for convenience, let us say that $A_1 = B_1$.

Now we claim that $S_1 = T_1$. To this end, note first that if S_1 has no holes, then $S_1 = A_1 = B_1$, and T_1 can have no holes either. To see this, suppose we have a hole $B_{1,1}$ of T_1 . Since $B_{1,1} \subseteq A_1 \subseteq X$, $B_{1,1} \subseteq \bigcup_{i \neq 1} T_i$. By (2), $B_{1,1} \subseteq B_i$ for some $i \neq 1$. Since $B_{1,1} \neq B_i$ by definition of swiss cheese decomposition, $B_{1,1} \subsetneq B_i$. If $B_i \subseteq B_1$, since $T_1 \cap T_i = \emptyset$, B_i would be covered by the holes of T_1 and T_i , contradicting unpackability. If $B_1 \subseteq B_i$, then $T_1 \cap T_i = \emptyset$ implies that T_1 is contained in the holes of T_i . In this case, unpackability gives that B_1 is contained in a hole of T_i , contradicting $B_{1,1} \subseteq T_i$. These contradictions imply that T_1 has no holes and $S_1 = T_1$.

Otherwise, suppose S_1 has at least one hole $A_{1,1}$. Every element $x \in A_{1,1}$ is either

- not in X , in which case $x \in A_1 = B_1$ implies $x \in B_{1,i}$ for some i ; or,
- in S_j for some $j \neq 1$. In this case, since $S_1 \cap S_j = \emptyset$ and $A_j \neq A_{1,1}$, as before we must have $A_j \subsetneq A_{1,1}$.

Altogether,

$$A_{1,1} \subseteq \left(\bigcup_{i=1}^{n_1} B_{1,i} \right) \cup \left(\bigcup_{A_j \subsetneq A_{1,1}} A_j \right)$$

from which (2) implies $A_{1,1} \subseteq B_{1,i}$ for some i . The same argument applies to the other holes of S_1 and T_1 , with the result that

$$A_{1,1} \cup \dots \cup A_{1,m_1} = B_{1,1} \cup \dots \cup B_{1,n_1}$$

and $S_1 = T_1$.

Finally, the induction hypothesis applied to

$$S_2 \cup \dots \cup S_r = T_2 \cup \dots \cup T_s$$

finishes the proof. \square

Combining (3) and (4) of the above theorem, it follows that in an unpackable directed family both the swiss cheeses in a decomposition of a constructible set X and the wheels and holes of these swiss cheeses are uniquely determined by X . We thus obtain:

Corollary 2.3. *If \mathcal{B} is an unpackable directed family with $\mathfrak{U} \in \mathcal{B}$, and $X \subseteq \mathfrak{U}$ is constructible, then $X = \text{Ch}(\mathcal{S})$ for a unique finite set of balls \mathcal{S} .*

3. FORESTS AND PACKABLE FAMILIES

For this section, fix a set \mathfrak{U} and $\mathcal{B} \subseteq \mathcal{P}(\mathfrak{U}) \setminus \{\emptyset\}$ directed, not necessarily unpackable. We again require $\mathfrak{U} \in \mathcal{B}$. Consider Ch as defined before. For any constructible $X \subseteq \mathfrak{U}$, by Lemma 2.1 there exists some finite $\mathcal{S} \subseteq \mathcal{B}$ so that $X = \text{Ch}(\mathcal{S})$. In this case, we will say that \mathcal{S} represents X .

However, since \mathcal{B} is potentially packable, we may have $\mathcal{S}, \mathcal{T} \subseteq \mathcal{B}$ distinct but $\text{Ch}(\mathcal{S}) = \text{Ch}(\mathcal{T})$. Nevertheless, in this section we describe a way to choose a canonical \mathcal{S} representing X .

Define, on the set of all finite forests, a quasi-ordering \trianglelefteq so that $\mathcal{S} \trianglelefteq \mathcal{T}$ iff:

- (i) $|\mathcal{S}| = |\mathcal{T}|$ and for all n , $|\text{Lev}_n(\mathcal{S})| = |\text{Lev}_n(\mathcal{T})|$; or,
- (ii) $|\mathcal{S}| < |\mathcal{T}|$; or,
- (iii) $|\mathcal{S}| = |\mathcal{T}|$ and for some n and all $i < n$, $|\text{Lev}_i(\mathcal{S})| = |\text{Lev}_i(\mathcal{T})|$ but $|\text{Lev}_n(\mathcal{S})| > |\text{Lev}_n(\mathcal{T})|$.

So, roughly speaking, forests are ordered first by cardinality then by top-heaviness.

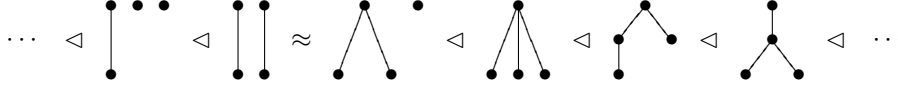


FIGURE 2. Some forests of size 4, quasi-ordered by \trianglelefteq

We use this order to get a uniqueness of decomposition result:

Theorem 3.1. *Let $X \subseteq \mathfrak{U}$ be constructible and $\text{Ch}(\mathcal{S}) = \text{Ch}(\mathcal{T}) = X$. If (\mathcal{S}, \subseteq) and (\mathcal{T}, \subseteq) are both \trianglelefteq -minimal among all representatives of X in \mathcal{B} , then $\mathcal{S} = \mathcal{T}$.*

Proof. Fix $\mathcal{S}, \mathcal{T} \subseteq \mathcal{B}$ as in the hypothesis. If $N = |\mathcal{S}| = 0$, then $\mathcal{S} = \mathcal{T} = \emptyset$ and we are done. We proceed now by induction on N .

Assume that $N > 0$. We aim to prove

$$\text{Lev}_0(\mathcal{S}) = \text{Lev}_0(\mathcal{T}).$$

By induction, this suffices. Define, for $B \in \mathcal{S}$, $R_{\mathcal{S}}(B) = B \setminus \bigcup \text{Sub}(B, \mathcal{S})$, and similarly $R_{\mathcal{T}}(C)$ for $C \in \mathcal{T}$. For B and C in even levels, these are the swiss cheeses comprising the decompositions generated by \mathcal{S} and \mathcal{T} . Note that since $R_{\mathcal{S}}(B) = \emptyset$ or $R_{\mathcal{T}}(C) = \emptyset$ clearly violates minimality of N , we can rule this possibility out.

So, consider a ball $B \in \text{Lev}_0(\mathcal{S})$. Since $R_{\mathcal{S}}(B) \subseteq \text{Ch}(\mathcal{S}) = \text{Ch}(\mathcal{T})$, there must exist $C \in \text{Lev}_{2n}(\mathcal{T})$ for some n such that $R_{\mathcal{S}}(B) \cap R_{\mathcal{T}}(C) \neq \emptyset$. It follows that $B \cap C \neq \emptyset$. If $C \notin \text{Lev}_0(\mathcal{T})$, then replace C with the ball containing it in $\text{Lev}_0(\mathcal{T})$. In other words, we have shown that for any $B \in \text{Lev}_0(\mathcal{S})$, there is $C \in \text{Lev}_0(\mathcal{T})$ such that $B \cap C \neq \emptyset$. Since \mathcal{B} is directed, $B \subseteq C$ or $C \subseteq B$. Likewise, for any $C \in \text{Lev}_0(\mathcal{T})$, there is $B \in \text{Lev}_0(\mathcal{S})$ such that $B \subseteq C$ or $C \subseteq B$.

Now, if we did not have $\text{Lev}_0(\mathcal{S}) = \text{Lev}_0(\mathcal{T})$, then by the above observation there would be $B \in \text{Lev}_0(\mathcal{S})$ and $C \in \text{Lev}_0(\mathcal{T})$ such that $B \subsetneq C$ or $C \subsetneq B$. Thus say, for instance, that we have found $B \subsetneq C$. Note that, as $B \in \text{Lev}_0(\mathcal{S})$, there can be no $B' \in \mathcal{S}$ for which $C \subseteq B'$.

Define

$$\mathcal{S}' = \{B' \in \text{Lev}_0(\mathcal{S}) \mid B' \subseteq C, \text{ but } B' \not\subseteq C' \text{ for any } C' \in \text{Sub}(C, \mathcal{T})\}$$

$$\mathcal{T}' = \{C' \in \text{Sub}(C, \mathcal{T}) \mid C' \cap B' = \emptyset \text{ for all } B' \in \mathcal{S}'\}.$$

Claim 1: $C \setminus \bigcup \mathcal{T}' = \bigcup \mathcal{S}'$.

First, if $x \in \bigcup \mathcal{S}'$, then $x \in C$. If $C' \in \mathcal{T}'$, then $C' \cap (\bigcup \mathcal{S}') = \emptyset$ gives $x \notin C'$. In other words, $x \in C \setminus \bigcup \mathcal{T}'$.

Conversely, suppose $x \in C \setminus \bigcup \mathcal{T}'$. If $x \notin X$, then since $x \in C$ we must have $x \in C'$ for some $C' \in \text{Sub}(C, \mathcal{T})$. But $C' \notin \mathcal{T}'$, so $C' \cap B' \neq \emptyset$ for some $B' \in \mathcal{S}'$. Now $C' \subseteq B'$, whence $x \in \bigcup \mathcal{S}'$. If on the other hand $x \in X$, then $x \in B'$ for

some $B' \in \text{Lev}_0(\mathcal{S})$. Suppose $B' \notin \mathcal{S}'$. Since $B' \subsetneq C$ by choice of C , it follows that $B' \subseteq C'$ for some $C' \in \text{Sub}(C, \mathcal{T})$. Since $x \in C'$ but not $\bigcup \mathcal{T}'$, $C' \cap B'' \neq \emptyset$ for some $B'' \in \mathcal{S}'$. Now $B' \subsetneq B''$ contradicts $B' \in \text{Lev}_0(\mathcal{S})$, and the claim is proven.

Next, let

$$\begin{aligned}\mathcal{S}^* &= (\mathcal{S} \setminus \mathcal{S}') \cup \{C\} \cup \mathcal{T}' \\ \mathcal{T}^* &= (\mathcal{T} \setminus (\{C\} \cup \mathcal{T}')) \cup \mathcal{S}',\end{aligned}$$

noting that $\mathcal{S}' \subseteq \text{Lev}_0(\mathcal{T}^*)$, $C \in \text{Lev}_0(\mathcal{S}^*)$, and $\mathcal{T}' \subseteq \text{Lev}_1(\mathcal{S}^*)$.

Claim 2: $\text{Ch}(\mathcal{S}^*) = X$.

For $\text{Ch}(\mathcal{S}^*) = X$, suppose first that $x \in X$. Then x resides in a maximal chain

$$(1) \quad x \in B_{2n} \subseteq B_{2n-1} \subseteq \dots \subseteq B_0$$

for some $B_i \in \text{Lev}_i(\mathcal{S})$, $x \notin \bigcup \text{Sub}(B_{2n}, \mathcal{S})$. There are several cases to consider:

- If $B_0 \notin \mathcal{S}'$, then either
 - $B_0 \cap C = \emptyset$ in which case $B_i \in \text{Lev}_i(\mathcal{S}^*)$ as well and $x \in \text{Ch}(\mathcal{S}^*)$; or
 - $B_0 \subseteq C$ and, since $B_0 \notin \mathcal{S}'$, $B_0 \subseteq C'$ for some $C' \in \text{Sub}(C, \mathcal{T})$. This C' must be in \mathcal{T}' since $\bigcup \mathcal{S}' = C \setminus \bigcup \mathcal{T}'$. So in this case we have $B_i \in \text{Lev}_{i+2}(\mathcal{S}^*)$ in (1) and $x \in \text{Ch}(\mathcal{S}^*)$.
- If $B_0 \in \mathcal{S}'$, then $B_0 \subseteq C$ but $B_0 \cap C' = \emptyset$ for all $C' \in \mathcal{T}'$. It follows that

$$x \in B_{2n} \subseteq \dots \subseteq B_1 \subseteq C$$

with $B_i \in \text{Lev}_i(\mathcal{S}^*)$, and $x \in \text{Ch}(\mathcal{S}^*)$.

This shows $X \subseteq \text{Ch}(\mathcal{S}^*)$.

For the converse, suppose again (1) but this time with $B_i \in \text{Lev}_i(\mathcal{S}^*)$, $x \notin \bigcup \text{Sub}(B_{2n}, \mathcal{S}^*)$. If $B_0 \in \mathcal{S} \setminus \mathcal{S}'$, then by choice of C , $B_0 \cap C = \emptyset$ and $x \in \text{Ch}(\mathcal{S}) = X$. The other possibility is that $B_0 = C$. Again there are several cases:

- If $B_1 \in \mathcal{T}'$, then $B_2 \in \text{Lev}_0(\mathcal{S})$ and $x \in \text{Ch}(\mathcal{S})$.
- If $B_1 \in \mathcal{S} \setminus \mathcal{T}'$, then $B_1 \not\subseteq C'$ for any $C' \in \mathcal{T}'$ but $B_1 \notin \mathcal{S}'$ implies $B_1 \notin \text{Lev}_0(\mathcal{S})$. Since $B_1 \in \text{Lev}_1(\mathcal{S}^*)$, it follows that $B_1 \in \text{Lev}_1(\mathcal{S})$ as well. So there is $B \in \text{Lev}_0(\mathcal{S})$, $B_1 \subseteq B$ and

$$x \in B_{2n} \subseteq \dots \subseteq B_1 \subseteq B$$

with $B_i \in \text{Lev}_i(\mathcal{S})$, giving $x \in \text{Ch}(\mathcal{S}) = X$.

We thus have shown $\text{Ch}(\mathcal{S}^*) = X$.

Claim 3: $\text{Ch}(\mathcal{T}^*) = X$.

Suppose, similarly, $x \in X$ with a maximal chain

$$(2) \quad x \in C_{2n} \subseteq \dots \subseteq C_0,$$

$C_i \in \text{Lev}_i(\mathcal{T})$. So,

- If $C_0 = C$, then either
 - $C_1 \in \mathcal{T}'$, in which case $C_1 \cap B' = \emptyset$ for every $B' \in \mathcal{S}'$, and $C_i \in \text{Lev}_{i-2}(\mathcal{T}^*)$ for any $i \geq 2$. This shows $x \in \text{Ch}(\mathcal{T}^*)$.
 - $C_1 \notin \mathcal{T}'$, so that $x \in C \setminus \bigcup \mathcal{T}' = \bigcup \mathcal{S}'$. So $x \in B'$ for some $B' \in \mathcal{S}'$, and the definition of \mathcal{S}' gives $C_1 \subsetneq B'$. Thus the chain

$$x \in C_{2n} \subseteq \dots \subseteq C_1 \subseteq B'$$

with $C_i \in \text{Lev}_i(\mathcal{T}^*)$ gives $x \in \text{Ch}(\mathcal{T}^*)$.

- If $C_0 \neq C$, then in (2), $C_i \in \text{Lev}_i(\mathcal{T}^*)$ as well, so again $x \in \text{Ch}(\mathcal{T}^*)$.

Conversely, suppose (2) but now with $C_i \in \text{Lev}_i(\mathcal{T}^*)$, $x \notin \bigcup \text{Sub}(C_{2n}, \mathcal{T}^*)$. If $C_0 \in \mathcal{S}'$, then $x \in C \setminus \bigcup \mathcal{T}'$. Since $C_1 \subseteq C_0$, $C_1 \cap C' = \emptyset$ for all $C' \in \mathcal{T}'$. Thus we have in \mathcal{T} the (maximal) chain

$$x \in C_{2n} \subseteq \dots \subseteq C_1 \subseteq C$$

with $C_i \in \text{Lev}_i(\mathcal{T})$.

On the other hand, suppose $C_0 \in \mathcal{T} \setminus \mathcal{S}'$. Then since $C_0 \in \text{Lev}_0(\mathcal{T}^*)$, $C_0 \cap B' = \emptyset$ for all $B' \in \mathcal{S}'$. It follows that either $C_0 \cap C = \emptyset$, or $C_0 \subseteq C'$ for some $C' \in \mathcal{T}'$. In the first case, the chain C_i is the same in \mathcal{T} and $x \in \text{Ch}(\mathcal{T})$. In the second case, since $C_0 \notin \mathcal{T}'$, the chain becomes

$$x \in C_{2n} \subseteq \dots \subseteq C_0 \subseteq C' \subseteq C$$

with $C_i \in \text{Lev}_{i+2}(\mathcal{T})$. This shows again that $x \in \text{Ch}(\mathcal{T})$, and the claim is proven.

Finally, depending on the relative sizes of $|\mathcal{S}'|$ and $|\mathcal{T}'|$, we derive a contradiction to \leq -minimality of \mathcal{S} and \mathcal{T} as follows:

- If $|\mathcal{S}'| < |\mathcal{T}'| + 1$, then \mathcal{T}^* represents X with strictly fewer balls than \mathcal{T} .
- If $|\mathcal{S}'| > |\mathcal{T}'| + 1$, then \mathcal{S}^* represents X with strictly fewer balls than \mathcal{S} .
- If $|\mathcal{S}'| = |\mathcal{T}'| + 1 = 1$, say $\mathcal{S}' = \{B\}$, then $C \setminus \bigcup \mathcal{T}' = C = B$ contradicts our choice of C .
- If $|\mathcal{S}'| = |\mathcal{T}'| + 1 \geq 2$, then \mathcal{S}^* represents X with N balls, but with $|\text{Lev}_0(\mathcal{S}^*)| > |\text{Lev}_0(\mathcal{S})|$.

The contradiction gives $\text{Lev}_0(\mathcal{S}) = \text{Lev}_0(\mathcal{T})$, and the result follows. \square

4. VC-MINIMALITY

While the previous sections relied purely on the combinatorial properties of directed families, the questions originate in logic with the notion of VC-minimality.

A complete theory T is *VC-minimal* if there is a family of formulas Ψ such that the set

$$\left\{ \psi(x, \bar{a}) \mid \psi(x, \bar{y}) \in \Psi, \bar{a} \in \mathfrak{M}^{|\bar{y}|} \right\}$$

of instances of Ψ form a directed family for every $\mathfrak{M} \models T$; and such that every definable subset of \mathfrak{M} is a boolean combination of instances of Ψ . Here we do not distinguish between $\psi(x, \bar{a})$ and the set it defines in a model \mathfrak{M} of T . Note moreover that the length of the tuples \bar{y} may vary with $\psi(x, \bar{y})$, but x is exclusively a single variable. Ψ is called a *generating family*. A formula $\delta(x; \bar{z})$ in T is *directed* if for any $\mathfrak{M} \models T$, the set of instances of δ in \mathfrak{M} forms a directed family.

The terminology around directed families carries over naturally to VC-minimal theories, with the definable subsets of \mathfrak{M} being the constructible sets. As with directed families, for our purposes it will be most convenient to stick to the convention that $x = x \in \Psi$ but $x \neq x \notin \Psi$, i.e. the whole universe is always a ball, the empty set is never a ball.

Likewise, Theorems 2.2 and 3.1 can be applied immediately to the family of balls generated by any model of a VC-minimal theory.

4.1. Uniform definability of levels. In this subsection, we observe that the levels of a canonical decomposition as in 3.1 are uniformly definable. This fact will be useful in type counting arguments in VC-minimal theories (see [4]).

Fix a formula $\varphi(x; \bar{y})$ in a VC-minimal theory. By compactness, there exists a single directed $\delta(x; \bar{z})$ and $N < \omega$ so that all instances of φ are a boolean combination of at most N instances of $\delta(x; \bar{z})$. (More precisely, compactness gives us finitely many $\psi \in \Psi$, then we can use coding tricks to combine them into a single directed δ .) As we will only be working with instances of φ , we disregard Ψ and work instead in the directed family of instances of δ .

There are only finitely many forests of size at most N ; call this set \mathcal{F}_N . For each $F \in \mathcal{F}_N$, let $\psi_F(\bar{y})$ denote the formula which says that there exists \bar{z}_f for $f \in F$ so that $\varphi(x; \bar{y})$ is equivalent to $\text{Ch}(\{\delta(x; \bar{z}_f) \mid f \in F\})$ and $\delta(x; \bar{z}_f)$ are ordered so that $f \mapsto \delta(x; \bar{z}_f)$ is an isomorphism of forests.

Finally, for any $n < N$, let $\gamma_n(x; \bar{y})$ denote the formula that says, for the \leq -least $F \in \mathcal{F}_N$ such that $\psi_F(\bar{y})$ holds, there exists witnesses \bar{z}_f for $f \in F$ as above such that $\delta(x; \bar{z}_f)$ holds for some $f \in \text{Lev}_n(F)$. Thus, for any \bar{b} , $\gamma_n(x; \bar{b})$ holds if and only if x appears in the n th level of a \leq -minimal decomposition $\{\delta(x; \bar{c}_f) \mid f \in F\}$ of $\varphi(x; \bar{b})$. However, by Theorem 3.1, the set $\{\delta(x; \bar{c}_f) \mid f \in F\}$ is unique up to T -equivalence. Therefore, $\gamma_n(x; \bar{b})$ holds if and only if x is in the n th level of the \leq -minimal decomposition.

We summarize in the following theorem:

Theorem 4.1. *If $\varphi(x; \bar{y})$ is any formula in a VC-minimal theory T , there exists a directed $\delta(x; \bar{z})$, $N_\varphi < \omega$ and $\gamma_{\varphi, n}(x; \bar{y})$ for all $n < N_\varphi$ such that:*

- (i) *For all \bar{b} , $\gamma_{\varphi, n}(x; \bar{b})$ is T -equivalent to a disjoint union of (at most N_φ) instances of δ .*
- (ii) *$\varphi(x; \bar{y})$ is T -equivalent to*

$$\bigvee_{n < N_\varphi} (\gamma_{\varphi, 2n}(x; \bar{y}) \wedge \neg \gamma_{\varphi, 2n+1}(x; \bar{y})).$$

4.2. VC-minimality and imaginaries. In this subsection, we outline an application which extends the analogy to Holly's work with algebraically closed valued fields [7]. The main observation is that the canonical representation of definable sets from 3.1 leads to a one-dimensional elimination of imaginaries.

A theory is said to *eliminate imaginaries* if, for every model \mathfrak{M} , $n \in \mathbb{N}$, and definable set $X \subseteq \mathfrak{M}^n$, there is a formula $\varphi(\bar{x}; \bar{y})$ and tuple $\bar{a} \in \mathfrak{M}^{|\bar{y}|}$ such that for all $\bar{b} \in \mathfrak{M}^{|\bar{y}|}$, $\varphi(\bar{x}; \bar{b})$ defines X iff $\bar{b} = \bar{a}$. In this case, \bar{a} is called a *code* (or *canonical parameter*) of X . The existence of codes allows one, in a sense, to treat definable sets as elements of the model. See [8] for further discussion.

It should be noted that it is always possible to expand a model \mathfrak{M} to a (usually multi-sorted) structure \mathfrak{M}^{eq} which eliminates imaginaries by explicitly adding to the language a code for every definable set. This suffices for many applications, but in other situations one may find a better understanding of the definable sets in a structure by finding a way to expand the language to achieve elimination of imaginaries in a more efficient, or natural, way.

The notion of codes also specializes naturally to definable sets of a certain dimension. To this end, say a theory has *n -prototypes* if there is a family $\Phi = \{\varphi(\bar{x}; \bar{y})\}$ with $|\bar{x}| = n$ such that for every model \mathfrak{M} and every definable set $X \subseteq \mathfrak{M}^n$, there is exactly one $\varphi \in \Phi$ and $\bar{a} \in \mathfrak{M}^{|\bar{y}|}$ such that $\varphi(\bar{x}; \bar{a})$ defines X . Holly proves in [7] that a theory eliminates imaginaries iff it has n -prototypes for every $n \geq 1$. It is also clear from the proof that a theory has 1-prototypes iff every definable subset of a model (in one variable) has a code.

Returning to VC-minimal theories, the definable sets in more than one variable are not yet well understood. The favorite example of algebraically closed valued fields indicates that the situation can be quite complex (see for instance [5]). However, on the question of 1-prototypes, the work of the preceding sections does the trick. We present this in two forms.

Suppose T is VC-minimal, with generating family Ψ . Expand the language of T to add to any model $\mathfrak{M} \models T$ a new sort consisting of the finite sets of balls. Add also, for each $\langle \psi_1(x; \bar{y}_1), \dots, \psi_n(x; \bar{y}_n) \rangle \in \Psi^n$, a new function symbol from the main sort to the new sort,

$$f_{\langle \psi_1, \dots, \psi_n \rangle}(\bar{y}_1, \dots, \bar{y}_n) : \langle \bar{a}_1, \dots, \bar{a}_n \rangle \mapsto \{B_1, \dots, B_n\}$$

taking the tuple of parameters $\langle \bar{a}_1, \dots, \bar{a}_n \rangle$ to the set of balls defined by $\psi_i(x; \bar{a}_i) = B_i$. Let T^* be the theory of a model of T expanded in this way.

Theorem 4.2. *If T is VC-minimal, then T^* has 1-prototypes.*

Proof. We show only that every definable set $X \subseteq \mathfrak{M}$ has a code. Assume $X \neq \mathfrak{M}$; that \mathfrak{M} has a code is obvious. By Theorem 3.1, there is a finite set $\mathcal{S} = \{B_1, \dots, B_n\}$ of balls so that $\text{Ch}(\mathcal{S}) = X$ and if $\text{Ch}(\mathcal{T}) = X$ and $(\mathcal{T}, \subseteq) \cong (\mathcal{S}, \subseteq)$ (as forests), then $\mathcal{S} = \mathcal{T}$.

Now, for $i \leq n$ let $B_i = \psi_i(x; \bar{a}_i)$, $\psi \in \Psi$. Let $\varphi(x; \mathcal{Y})$ be the formula stating that

$$\forall \langle \bar{y}_1, \dots, \bar{y}_n \rangle \in f_{\langle \psi_1, \dots, \psi_n \rangle}^{-1}(\mathcal{Y}) \left(\begin{array}{c} \left(\{\psi_i(x; \bar{y}_i)\}_{i \leq n}, \subseteq \right) \cong (\mathcal{S}, \subseteq) \rightarrow \\ x \in \text{Ch} \left(\{\psi_i(x; \bar{y}_i)\}_{i \leq n} \right) \end{array} \right).$$

Since $X \neq \mathfrak{M}$, if $\varphi(x; \mathcal{T})$ defines X then $(\mathcal{T}, \subseteq) \cong (\mathcal{S}, \subseteq)$, and hence $\mathcal{T} = \mathcal{S}$. So \mathcal{S} is in fact a code for X . \square

This language is complicated somewhat by the need to allow for all finite sets of balls. This is necessary, as it is not generally possible, for example, to distinguish $\langle B_1, B_2 \rangle$ from $\langle B_2, B_1 \rangle$ in terms of the definable set represented by this pair of balls. This phenomenon is commonplace enough to earn its own terminology.

A theory *weakly eliminates imaginaries* if for every model \mathfrak{M} , $n \in \mathbb{N}$, and definable set $X \subseteq \mathfrak{M}^n$, there is a formula $\varphi(\bar{x}; \bar{y})$ and nonempty finite set $A \subseteq \mathfrak{M}^{|\bar{y}|}$ such that for all $\bar{b} \in \mathfrak{M}^{|\bar{y}|}$, $\varphi(\bar{x}; \bar{b})$ defines X iff $\bar{b} \in A$. Analogously, a theory has *weak 1-prototypes* if there is a family Φ such that for every model \mathfrak{M} and every definable set $X \subseteq \mathfrak{M}$, there is exactly one $\varphi(x; \bar{y}) \in \Phi$ and finitely many $\bar{a} \in \mathfrak{M}^{|\bar{y}|}$ such that $\varphi(x; \bar{a})$ defines X .

Now from a VC-minimal T construct an expanded theory T° by adding a new sort consisting of the balls, as well as for each $\psi \in \Psi$ a new function symbol

$$f_\psi : \bar{a} \mapsto B = \psi(x, \bar{a}).$$

As the codes from Theorem 4.2 depended only on the set of balls $\{B_1, \dots, B_n\}$, we may replace them with an ordered tuple of balls $\langle B_1, \dots, B_n \rangle$ at the expense of allowing as many as $n!$ potential codes rather than only one. We thus obtain

Corollary 4.3. *If T is VC-minimal, then T° has weak 1-prototypes.*

4.3. Quasi-VC-minimality. A theory T is called *quasi-VC-minimal* if there exists a family Ψ such that the set of all instances of formulas from Ψ is directed and every parameter-definable subset of a model of T is a boolean combination of instances of Ψ and \emptyset -definable sets. An example of a quasi-VC-minimal but not VC-minimal theory is Presburger arithmetic, $\text{Th}(\mathbb{Z}; +, <)$. See [3] for details.

We outline how the above results can be adapted to apply to quasi-VC-minimal theories. As the main differences in the proofs are notational annoyances, these are omitted.

Given $\mathfrak{M} \models T$ and \emptyset -definable $Q \subseteq \mathfrak{M}$, the restriction of the balls to Q , $\{B \cap Q \mid B \text{ a ball}\}$ is again a directed family. Thus a boolean combination of balls intersected with Q admits a canonical, \leq -minimal swiss cheese decomposition as in Theorem 3.1. Here, the balls themselves are not uniquely defined, but only their intersection with Q . For a finite set $\mathcal{S} = \{B_i\}_i$ of balls, write $\mathcal{S} \cap Q = \{B_i \cap Q\}_i$.

Now, given a formula $\varphi(x; \bar{y})$, there are formulas $\theta_1(x), \dots, \theta_k(x)$ over \emptyset such that every instance of φ is a boolean combination of balls and $\theta_1, \dots, \theta_k$. For $e : \{1, \dots, k\} \rightarrow \{0, 1\}$, write

$$\theta^e(x) = \bigwedge_{i=1}^k \theta_i(x)^{e(i)}.$$

Then as in 2.1 it is proved that every instance of φ can be written as

$$\bigvee_e (\theta^e(x) \wedge \sigma_e(x, \bar{a}_e))$$

where σ_e defines a swiss cheese decomposition. Again, the balls in this swiss cheese decomposition are not uniquely determined, but by 3.1 they can be chosen so that their intersections with θ^e are:

Theorem 4.4. *If $X \subseteq \mathfrak{M}$ is definable, then there are pairwise disjoint \emptyset -definable $Q_1, \dots, Q_k \subseteq \mathfrak{M}$ partitioning \mathfrak{M} and, for each $i \leq k$ a finite set of balls \mathcal{S}_i such that*

- (i) $X \cap Q_i = \text{Ch}(\mathcal{S}_i \cap Q_i)$, and
- (ii) for any \mathcal{T} , if also $X \cap Q_i = \text{Ch}(\mathcal{T} \cap Q_i)$ and $(\mathcal{T} \cap Q_i, \subseteq) \cong (\mathcal{S}_i \cap Q_i, \subseteq)$, then $\mathcal{T} \cap Q_i = \mathcal{S}_i \cap Q_i$.

Finally, for quasi-VC-minimal T , let T^\sharp be the theory obtained by adding

- a new sort consisting of the intersections of balls with \emptyset -definable sets, and
- for each $\psi(x; \bar{y}) \in \Psi$ and each formula θ over \emptyset , a new function symbol

$$f_{[\theta, \psi]}(\bar{y}) : \bar{a} \mapsto \theta(x) \wedge \varphi(x; \bar{a}).$$

We then obtain as in 4.3:

Corollary 4.5. T^\sharp has weak 1-prototypes.

REFERENCES

- [1] H. Adler, *Theories controlled by formulas of Vapnik-Chervonenkis codimension 1*, Preprint (2008).
- [2] S. Cotter and S. Starchenko, *Forking in VC-minimal theories*, Preprint (2011).
- [3] J. Flenner and V. Guingona, *Two theories which are not VC-minimal*, In preparation.
- [4] V. Guingona, *Low VC-density in VC-minimal theories*, In preparation.
- [5] D. Haskell, E. Hrushovski, and D. Macpherson, *Definable sets in algebraically closed valued fields: elimination of imaginaries*, J. Reine Angew. Math. **597** (2006), 175–236.

- [6] J. E. Holly, *Canonical forms for definable subsets of algebraically closed and real closed valued fields*, J. Symbolic Logic **60** (1995), no. 3, 843–860.
- [7] ———, *Prototypes for definable subsets of algebraically closed valued fields*, J. Symbolic Logic **62** (1997), no. 4, 1093–1141.
- [8] B. Poizat, *A course in model theory*, Universitext, Springer-Verlag, New York, 2000. An introduction to contemporary mathematical logic; Translated from the French by Moses Klein and revised by the author.

UNIVERSITY OF NOTRE DAME, DEPARTMENT OF MATHEMATICS, 255 HURLEY HALL, NOTRE DAME, IN 46556, U.S.A.

E-mail address: `jflenner@nd.edu`

E-mail address: `guingona.1@nd.edu`