The Journal of Symbolic Logic
Volume 87, Number 1, March 2022

MODEL THEORY AND COMBINATORICS OF BANNED SEQUENCES

HUNTER CHASE AND JAMES FREITAG

Abstract. We set up a general context in which one can prove Sauer–Shelah type lemmas. We apply our general results to answer a question of Bhaskar [1] and give a slight improvement to a result of Malliaris and Terry [7]. We also prove a new Sauer–Shelah type lemma in the context of op-rank, a notion of Guingona and Hill [4].

§1. Introduction. A single combinatorial notion called *VC dimension* determines important dividing lines in both machine learning (PAC learnability of a class) and model theory (the independence/nonindependence dichotomy, IP/NIP) [5], and the finiteness of this quantity plays an essential role in the development of various structural results in theories without the independence property and in machine learning. Often at the root of these developments is the Sauer–Shelah Lemma, which for a formula $\phi(x; y)$ without the independence property, gives a polynomial bound on the shatter function associated with ϕ—that is, the number of consistent ϕ-types over finite sets. Without NIP, however, the number of ϕ-types can grow exponentially in the size of the finite parameter set. In a recent paper, Bhaskar [1] noticed that when the formula ϕ is actually stable, that is, ϕ has finite Shelah 2-rank (also called Littlestone dimension or thicket dimension in the context of set systems), one can relax the way in which the ϕ-types are constructed, allowing for trees of parameters (explained below) while still proving polynomial bounds on the resulting collection of consistent ϕ-types. Again, in the absence of stability the number of types formed in this manner can grow exponentially in the height of the tree. Following Bhaskar, we refer to this growth dichotomy theorem as the stable Sauer–Shelah Lemma. In [2], we notice that stability also determines an important dividing line in machine learning; stability determines learnability in various settings of *online learning*. In these settings of learning, various results at their core rely on the polynomial growth of the stable shatter function.

In both settings described above, the growth of the number of types being polynomially bounded or exponential is completely determined by whether a simple combinatorial notion of dimension is finite, and the upper bound (which is tight in general) on the number of such types (in terms of the appropriate notion of dimension) is identical in both cases. In light of this, Bhaskar naturally asks if there is a single combinatorial principle which explains both the Sauer–Shelah Lemma and the stable variant. The main purpose of our paper is to set up a general context in

Received February 9, 2018.
2020 *Mathematics Subject Classification*. 03C45, 03C98, 68Q32.
Key words and phrases. VC-dimension, Littlestone dimension, stability, NIP, online learning, query learning, PAC learning.

0022-4812/22/8701-0001
DOI:10.1017/jsl.2019.35

which one can prove Sauer–Shelah type results into which both of the above contexts fit, answering Bhaskar's question as well as proving new results. Our solution to the problem is quite general and deals with what we call *banned sequence problems*.

Our general setup of banned sequence problems is an interesting combinatorial setting in its own right, and we will roughly describe the simplest context here. Suppose that you consider the collection of all binary sequences of length n, and for each subset of the indices of size k, there is at least one "banned subsequence" of length k. How many binary sequences of length n are there which avoid each of the banned sequences on all subsets of the indices of size k? Subject to some very mild assumptions on how the banned sequences are chosen, we show that there are at most

$$\sum_{i=0}^{k-1} \binom{n}{i}$$

such sequences. This bound is the bound of the Sauer–Shelah Lemma. Without the mild assumptions, we show that this bound can be violated. The generality of our setup covers both the settings mentioned above as well as yielding some new results.

We give a slight improvement of a result of Malliaris and Terry [7] regarding sizes of cliques and independent sets in stable graphs. Essentially, their result uses the finiteness of a certain combinatorial dimension, *tree rank*, in order to establish polynomial bounds strong enough to get a version of the Erdos–Hajnal conjecture, among other results (Malliaris and Terry also develop further structural properties of graphs which we will not touch on in this paper). We examine tree rank in the general context of banned sequence problems, and as a result, give a slight improvement to their bounds.

In the last part of the paper, we turn to the setting of op-ranks. For each $s \in \mathbb{N}$, Guingona and Hill [4] define a rank of partial types, op_s-rank. For instance, when $s = 1$, op_1-rank is equal to the Shelah 2-rank. Working with set systems of finite op_s-rank, we establish a new variant of the Sauer–Shelah Lemma using our banned sequence setup.

We note that not every known variant of the Sauer–Shelah Lemma seems to fit into the context of banned sequence problems; the main results of [3] establish a variant of Sauer–Shelah for n-dependent theories which does not seem to easily fit into our context of banned sequence problems. Is there a general setup which also covers the known Sauer–Shelah style results for n-dependent theories? This seems reasonable to ask because n-dependent theories generalize NIP theories in a way similar to how theories with finite op_s-rank generalize stable theories.

1.1. Organization. In Section 2, we give the necessary preliminary notation for our results. In Section 3, we lay out the basic theory of banned sequence problems along with some applications. In Section 4, we generalize our banned sequence problems. In Section 4.2, we apply generalized banned problems to the op-rank setting.

§2. Preliminaries. Our primary combinatorial tool applies to theorems surrounding VC dimension and Littlestone dimension (also known as Shelah's 2-rank in model theory or thicket dimension in [1]), and we recall those definitions and

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

relevant theorems. The next several definitions can be found in various sources, for example, [9].

Throughout, any indexing starts at 0, and $[n] := \{0, 1, \dots, n-1\}$. By $\binom{[n]}{k}$ we mean the collection of all subsets of $[n]$ of size k.

Recall that a set system $(X, \mathcal{F})$ (often referred to as $\mathcal{F}$ when X is understood) consists of a set X and a collection $\mathcal{F} \subseteq \mathcal{P}(X)$ of subsets of X. For $Y \subseteq X$, the projection of $\mathcal{F}$ onto Y is the set system with base set Y and collection of subsets

$$\mathcal{F}_Y := \{F \cap Y \mid F \in \mathcal{F}\}.$$

VC dimension measures the ability of a set system to pick out subsets of a set of a given size.

Definition 2.1. A set system $(X, \mathcal{F})$ *shatters* a set Y if $\mathcal{F}_Y = \mathcal{P}(Y)$. The VC dimension of $\mathcal{F}$ is the largest $k < \omega$ such that $\mathcal{F}$ shatters some set of size k, or is infinite if $\mathcal{F}$ shatters arbitrarily large sets. The *shatter function*

$$\pi_{\mathcal{F}}(n) := \sup_{Y \subseteq X, |Y|=n} |\mathcal{F}_Y|$$

is given by the supremum of the size of the projection onto subsets of a given size.

If a set system has finite VC dimension, then we obtain a polynomial bound on the shatter function.

Theorem 2.2 (Sauer–Shelah Lemma). *Let $\mathcal{F}$ be a set system of VC dimension k. Then the maximum size of a projection from $\mathcal{F}$ onto a set $A = \{a_0, \dots, a_{n-1}\}$ of size n is $\sum_{i=0}^{k} \binom{n}{i}$. In particular,*

$$\pi_{\mathcal{F}}(n) \leq \sum_{i=0}^{k} \binom{n}{i}.$$

Several proofs of the Sauer–Shelah Lemma can be found in various sources, for example, [9, 8].

Littlestone dimension is a variant of VC dimension; our development follows [1]. (Bhaskar calls Littlestone dimension “thicket dimension”—we prefer to use Littlestone dimension, or use “stable” to describe the general setting.) Given a set from the set system, elements are presented sequentially, with the element presented depending on membership of previous elements.

Definition 2.3. A *binary element tree* of height n with labels from X is a function $T : 2^{<n} \to X$. A *node* is a binary sequence $\sigma \in 2^{<n}$ along with its label, $a_\sigma := T(\sigma)$. A *leaf* is a binary sequence of length n, $\tau : [n] \to \{0, 1\}$. A leaf τ is properly labeled by a set A if for all $m < n$,

$$a_{\tau|_{[m]}} \in A \quad \text{iff} \quad \tau(m) = 1.$$

Definition 2.4. The *Littlestone dimension* of a set system $(X, \mathcal{F})$ is the largest $k < \omega$ such that there is a binary element tree of height k with labels from X such that every leaf can be properly labeled by elements of $\mathcal{F}$, or is infinite if there are such trees of arbitrary height. The *stable shatter function* (what Bhaskar calls the

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

“thicket shatter function”) $\rho_{\mathcal{F}}(n)$ is the maximum number of leaves properly labeled by elements of $\mathcal{F}$ in a binary element tree of height n.

THEOREM 2.5 (Stable (Thicket) Sauer–Shelah Lemma [1]). *Let $\mathcal{F}$ be a set system of Littlestone dimension k. Then the maximum number of properly labeled leaves in a binary element tree of height n is $\sum_{i=0}^{k} \binom{n}{i}$. In particular,*

$$\rho_{\mathcal{F}}(n) \leq \sum_{i=0}^{k} \binom{n}{i}.$$

VC dimension and the (VC) shatter function can be viewed in the context of binary element trees where every node of the same height is labeled with the same element, that is, $a_\sigma = a_{\sigma'}$ whenever $|\sigma| = |\sigma'|$.

There are dual notions of both VC dimension and Littlestone dimension, and their corresponding shatter functions, where the roles of elements and sets are reversed.

DEFINITION 2.6. Given a set system $(X, \mathcal{F})$, the dual set system $(X, \mathcal{F})^*$, or just $\mathcal{F}^*$, is the set system with base set $\mathcal{F}$ where the subsets are given by

$$\{F \mid F \in \mathcal{F}, x \in F\}$$

for each $x \in X$. The dual VC (resp., Littlestone) dimension of $\mathcal{F}$ is the VC (resp., Littlestone) dimension of $\mathcal{F}^*$.

Dual Littlestone dimension can be calculated by examining *binary decision trees*, where nodes are labeled by sets in the set system, and leaves are labeled by elements. Dual VC dimension can be calculated similarly.

In model theory, given a model $\mathcal{M}$, the VC (resp., Littlestone) dimension of a partitioned formula $\phi(x; y)$ is the VC (resp., Littlestone) dimension of the set system

$$(M^{|x|}, \{\phi(M^{|x|}, b) \mid b \in M^{|y|}\}).$$

These combinatorial notions encode model-theoretic dividing lines. A formula is NIP iff it has finite VC dimension, and is stable iff it has finite Littlestone dimension.

§3. The combinatorics of banned sequences. The binary element tree structure used to define Littlestone dimension allows us to identify a leaf of the tree with the binary sequence corresponding to the path through the tree to that leaf. Then counting properly labeled leaves amounts to counting the corresponding binary sequences. We establish a framework for counting binary sequences under certain conditions reflecting the tree structure, from which we will obtain a unified proof of the Sauer–Shelah Lemmas.

3.1. Banned binary sequences and Sauer–Shelah Lemmas. Our framework for counting binary sequences will reflect the height of the tree as well as the dimension (either Littlestone or VC) of the set system. We find it easier to count banned sequences. Having Littlestone dimension $k - 1$ says that in a tree of height k, there are some leaves which cannot be properly labeled, and those leaves correspond to sequences that we ban.

DEFINITION 3.1. A *k- fold banned binary sequence problem* (*BBSP*) *of length n*, for $0 \leq k \leq n$ is a function

$$f : \binom{[n]}{k} \times 2^{n-k} \to \mathcal{P}(2^k) \setminus \{2^k\}.$$

Informally, for each k-subset of $[n]$ and each binary sequence of length $n - k$, the binary sequences of length k not selected by f are banned, and we ban at least one such sequence. Sometimes we will refer to the sequences omitted by the function f as *banned subsequences*.

REMARK 3.2. It will be convenient to view binary sequences as functions, where the domain is the appropriate set of indices. Given $S \in \binom{[n]}{k}$, let $\bar{S} := [n] \setminus S$. When we consider $f(S, Y)$ for some fixed S, we view $Y \in 2^{n-k}$ as a function $Y : \bar{S} \to \{0, 1\}$, and elements of $f(S, Y)$ as functions $Z : S \to \{0, 1\}$, identifying 2^{n-k} with $2^{\bar{S}}$ and 2^k with 2^S.

Given $X : [n] \to \{0, 1\}$ and $S \subseteq [n]$, let X_S denote the restriction $X|_S$ of X to S, that is, the subsequence obtained by restricting to the indices in S.

We shall denote the union of two binary sequences Y and Z with disjoint domains as $Y \sqcup Z$. For example, if Y has domain $\{0, 2\}$, with $Y(0) = Y(2) = 0$, and Z has domain $\{1\}$ with $Z(1) = 1$, then $Y \sqcup Z$ is the binary sequence 010. When we wish to extend a sequence by appending some $j \in \{0, 1\}$, we will merely write $Y \sqcup j$, with the index of j usually understood from the context.

For a fixed $S \in \binom{[n]}{k}$, we denote the elements of S by $\{s_0, \dots, s_{k-1}\}$, where $s_0 < s_1 < \cdots < s_{k-1}$.

DEFINITION 3.3. A *solution* to a k-fold banned binary sequence problem f of length n is a binary sequence $X \in 2^n$ such that for any $S \in \binom{[n]}{k}$,

$$X_S \in f(S, X_{\bar{S}}).$$

A sequence which is not a solution is *banned* .

Intuitively, a solution to a banned binary sequence problem is a sequence which avoids every banned subsequence. In applications to binary element trees, properly labeled leaves will correspond to solutions of a certain banned binary sequence problem.

Without further assumptions, the number of solutions of a BBSP can grow exponentially in n for a fixed k.

PROPOSITION 3.4. *A k-fold BBSP f of length n has at most* $(2^k - 1)2^{n-k}$ *solutions.*

PROOF. Fix $S \in \binom{n}{k}$. For $Y : \bar{S} \to \{0, 1\}$ and $Z : S \to \{0, 1\}$, $Y \sqcup Z$ can only be a solution if $Z \in f(S, Y)$, and for each of 2^{n-k} many such Y's, there are at most $2^k - 1$ many Z's. ⊣

We observe that to obtain this bound, and so have only 2^{n-k} banned sequences, we must be able to find a collection $\mathcal{B}$ of 2^{n-k} sequences $X : [n] \to \{0, 1\}$ such that for all $S \in \binom{[n]}{k}$ and all $Y : \bar{S} \to \{0, 1\}$, there is some $X \in \mathcal{B}$ such that $Y \subseteq X$. Then we can set $f(S, Y) := \{X_S\}$, and then every $X \in 2^n \setminus \mathcal{B}$ is a solution. In general this

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

is not possible. It is possible for $k = n$, where we simply pick a sequence of length n to ban, $k = n - 1$, where $\mathcal{B}$ can consist of, say, the two constant sequences, $k = 1$, given below, and $k = 0$, which is trivial. But this condition already cannot be met for $k = 2$ and $n = 4$. In this case, one can verify that the minimum size of $\mathcal{B}$ to satisfy the above condition is 5, and so a 2-fold BBSP of length 4 can have at most 11 solutions.

EXAMPLE 3.5. Let f be the 1-fold BBSP of length n given by

$$f(\{s\}, Y) = \begin{cases} 1 & Y \text{ has an even number of 1s,} \\ 0 & Y \text{ otherwise.} \end{cases}$$

Then f has 2^{n-1} solutions, given by those binary sequences which have an even number of 1s.

We therefore need stronger hypotheses in order to bound the number of solutions by the Sauer–Shelah bound.

DEFINITION 3.6. A k-fold banned binary sequence problem f of length n is *not hereditary* if there is $S \in \binom{[n]}{k}$ and a function $g : 2^S \to 2^{\bar{S}}$ such that

- for all $Z : S \to \{0, 1\}$, we have $Z \in f(S, g(Z))$, and
- for all $Z_\alpha \neq Z_\beta$, the first index at which $g(Z_\alpha) \sqcup Z_\alpha$ and $g(Z_\beta) \sqcup Z_\beta$ differ is in S.

Otherwise, say f is *hereditary*.

One can think of the second condition as stating that g is continuous in the sense that for any $t \in \bar{S}$, $g(Z_\alpha)(t) = g(Z_\beta)(t)$ whenever $(Z_\alpha)_{S\cap[t]} = (Z_\beta)_{S\cap[t]}$, that is, $g(Z)(t)$ depends only on $Z_{S\cap[t]}$.

We will usually suppress the function g, and instead use indices to indicate the mapping—given $Z_\alpha : S \to \{0, 1\}$, we let $Y_\alpha := g(Z_\alpha)$. Then being not hereditary amounts to finding $S \in \binom{[n]}{k}$ such that for all $Z_\alpha : S \to \{0, 1\}$, we can associate a $Y_\alpha : \bar{S} \to \{0, 1\}$ such that $Z_\alpha \in f(S, Y_\alpha)$, and for any $Z_\alpha \neq Z_\beta$, the first index at which $Y_\alpha \sqcup Z_\alpha$ and $Y_\beta \wedge Z_\beta$ differ is in S.

For our purposes, being hereditary is the desirable property; hereditary BBSPs allow us to obtain the Sauer–Shelah bound on the number of solutions. We can also study binary element trees using hereditary BBSPs, and thus derive the corresponding Sauer–Shelah Lemmas. We choose to call these BBSPs hereditary because proving the Sauer–Shelah bound on the number of solutions uses derivative BBSPs in the inductive step, and being hereditary is preserved in these derivative problems.

THEOREM 3.7. *Any hereditary k-fold banned binary sequence problem of length n has at most $\sum_{i=0}^{k-1} \binom{n}{i}$ solutions.*

The proof is by induction. We will make use of the recursive property of binomial coefficients,

$$\binom{n}{i} = \binom{n-1}{i-1} + \binom{n-1}{i},$$

from which it follows that

$$\sum_{i=0}^{k-1}\binom{n}{i} = \sum_{i=0}^{k-2}\binom{n-1}{i} + \sum_{i=0}^{k-1}\binom{n-1}{i}.$$

In particular, we use the inductive strategy suggested by these equalities. The base cases will be $k = 0$ and $k = n$. In the inductive step, given a hereditary k-fold banned binary sequence problem f of length n, we seek two derivative problems: a $(k-1)$-fold banned binary sequence problem of length $n-1$, and a k-fold banned binary sequence problem of length $n-1$, both of which are hereditary. We will use banned sequences of these derivative problems to construct banned sequences of the original problem.

Definition 3.8. Let f be a k-fold banned binary sequence problem of length n, for $1 \leq k \leq n-1$.

- Let $\hat{f}$ be the $(k-1)$-fold banned binary sequence problem of length $n-1$ given as follows: for all $T \in \binom{[n-1]}{k-1}$, all $Y \in 2^{n-k}$, and all $Z \in 2^{k-1}$, let
$$Z \notin \hat{f}(T, Y) \qquad \text{iff} \qquad \exists j \in \{0, 1\}\ Z \sqcup j \notin f(T \sqcup \{n-1\}, Y).$$
- Let f' be the k-fold banned binary sequence problem of length $n-1$ given as follows: for all $S \in \binom{[n-1]}{k}$, all $Y \in 2^{n-k-1}$, and all $Z \in 2^k$, let
$$Z \notin f'(S, Y) \qquad \text{iff} \qquad \forall j \in \{0, 1\}\ Z \notin f(S, Y \sqcup j).$$

That is, the banned subsequences in $\hat{f}(T, Y)$ are those subsequences which can be extended by a particular j to a banned subsequence in $f(T \cup \{n-1\}, Y)$. In particular, any banned sequence of $\hat{f}$ has some extension which is a banned sequence of f.

The banned subsequences in $f'(S, Y)$ are those subsequences which are banned subsequences in $f(S, Y \sqcup j)$ for any extension of Y by j. In particular, any extension of any banned sequence of f' is a banned sequence of f.

Lemma 3.9. *Suppose f is a hereditary k-fold banned binary sequence problem of length n, for $1 \leq k \leq n-1$. Then both $\hat{f}$ and f' are also hereditary.*

Proof. Suppose for contradiction that $\hat{f}$ is not hereditary. Then there exists $T \in \binom{[n-1]}{k-1}$ such that for each $Z_\alpha : T \to \{0, 1\}$, there is $Y_\alpha : \bar{T} \to \{0, 1\}$ such that $Z_\alpha \in \hat{f}(T, Y_\alpha)$, and for any $Z_\alpha \neq Z_\beta$, the first index at which $Y_\alpha \sqcup Z_\alpha$ and $Y_\beta \sqcup Z_\beta$ differ belongs to T. Note that for some Z_α and some $j \in \{0, 1\}$, we have that

$$Z_\alpha \sqcup j \notin f(T \cup \{n-1\}, Y_\alpha),$$

or else associating each $Z_\alpha \sqcup j$ with Y_α would witness that f itself is not hereditary. Then, by definition of $\hat{f}$, $Z_\alpha \notin \hat{f}(T, Y_\alpha)$, a contradiction. So $\hat{f}$ is hereditary.

Suppose for contradiction that f' is not hereditary. Then there exists $S \in \binom{[n-1]}{k}$ such that for all $Z_\alpha : S \to \{0, 1\}$, there is $Y_\alpha : \bar{S} \to \{0, 1\}$ such that $Z_\alpha \in f'(S, Y_\alpha)$, and for any $Z_\alpha \neq Z_\beta$, the first index at which $Y_\alpha \sqcup Z_\alpha$ and $Y_\beta \sqcup Z_\beta$ differ belongs

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

to S. By definition of f', for each Z_α, there is $j_\alpha \in \{0, 1\}$ such that

$$Z_\alpha \in f(S, Y_\alpha \sqcup j_\alpha).$$

Let Y'_α be $Y_\alpha \sqcup j_\alpha$. Then associating Z_α with Y'_α witnesses that f is not hereditary, a contradiction. So f' is hereditary. ⊣

PROOF OF THEOREM 3.7. We prove the result by induction on n and k. Let f be a hereditary k-fold banned binary sequence problem of length n. Let $B(f)$ denote the number of sequences banned by f. It suffices to prove that

$$B(f) \geq 2^n - \sum_{i=0}^{k-1} \binom{n}{i}.$$

The base cases are $k = n$ and $k = 0$. When $k = n$, we have $2^n - \sum_{i=0}^{k-1} \binom{n}{i} = 1$, and any BBSP has at least one banned sequence. When $k = 0$, for all $Y \in 2^n$, we have $f(\emptyset, Y) = \emptyset$. Then for all $X \in 2^n$, we have $X_\emptyset = \emptyset \notin f(\emptyset, X_{[n]})$. So all $X \in 2^n$ are banned, and f has no solutions.

Otherwise, we proceed by induction. We show

$$B(f) \geq B(\hat{f}) + B(f').$$

For each sequence $\hat{X}$ that is banned by $\hat{f}$, there is at least one extension X which is banned by f, and we pick one such extension. For each sequence X' banned by f', at most one extension of X' was already obtained by extending a sequence $\hat{X}$ banned by $\hat{f}$. So there is at least one extension X of X' which is banned by f (by definition of f') but was not obtained by extending banned sequences for $\hat{f}$. Therefore these banned sequences of f constructed from f' and $\hat{f}$ have no common members, and so we have

$$B(f) \geq B(\hat{f}) + B(f'),$$

as desired. By induction, we have that

$$\begin{aligned} B(f) &\geq \left(2^{n-1} - \sum_{i=0}^{k-2} \binom{n-1}{i}\right) + \left(2^{n-1} - \sum_{i=0}^{k-1} \binom{n-1}{i}\right) \\ &\geq 2^n - \sum_{i=0}^{k-1} \binom{n}{i}. \end{aligned}$$

Thus f has at most $\sum_{i=0}^{k-1} \binom{n}{i}$ solutions. ⊣

It shall be useful to identify a stronger banned binary sequence problem, namely those in which $f(S, Y)$ depends only on S.

DEFINITION 3.10. A banned binary sequence problem f is *independent* if $f(S, Y) = f(S, Y')$ for any $Y, Y' : \bar{S} \to 0, 1$. When f is independent, we write $f(S)$.

COROLLARY 3.11. *Any independent k-fold banned binary sequence problem f of length n has at most $\sum_{i=0}^{k-1} \binom{n}{i}$ solutions.*

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

PROOF. We check that f is hereditary. If not, then there is $S \in \binom{[n]}{k}$ such that for all $Z_\alpha : S \to \{0,1\}$, there is $Y_\alpha : \bar{S} \to \{0,1\}$ with $Z_\alpha \in f(S, Y_\alpha) = f(S)$. But then $f(S) = 2^k$, a contradiction. The result follows from Theorem 3.7. ⊣

Banned binary sequence problems provide a common framework to prove Sauer–Shelah type bounds.

PROOF OF THEOREM 2.2. We obtain a $k+1$-fold independent BBSP f of length n as follows. Given $S = \{a_{s_0}, \dots a_{s_k}\} \in \binom{A}{k+1}$, let $f(S)$ be the set of binary sequences Z of length $k+1$ such that there is some $F \in \mathcal{F}$ such that $a_{s_i} \in F$ iff $Z(i) = 1$. We have that $f(S) \neq 2^{k+1}$ since the VC dimension of $\mathcal{F}$ is k, and f is clearly independent. Then a subset B of A is in the projection from $\mathcal{F}$ onto A iff the characteristic sequence of B (i.e., the sequence where the jth entry is 1 iff $a_j \in B$) is a solution to f. The result follows from Corollary 3.11. ⊣

PROOF OF THEOREM 2.5. Let T be a binary element tree of height n, with nodes a_σ for $\sigma \in 2^{<n}$. We obtain a $k+1$-fold hereditary BBSP of length n, f, as follows. Given $S = \{s_0, \dots, s_k\} \in \binom{[n]}{k+1}$ where $s_0 < s_1 < \cdots < s_k$ and $Y : \bar{S} \to \{0,1\}$, we obtain a binary element tree of height $k+1$ by taking all paths $\tau \in 2^n$ through T such that $Y \subseteq \tau$. Any two such paths first differ at some node a_σ where $|\sigma| \in S$, so removing all other nodes gives us the binary element tree $T_{S,Y}$ of height $k+1$. Since $\mathcal{F}$ has Littlestone dimension k, not all leaves of $T_{S,Y}$ can be properly labeled, so let $f(S, Y)$ be the set of all sequences whose corresponding leaves in $T_{S,Y}$ can be properly labeled. Then a leaf in T can only be properly labeled if the corresponding sequence is a solution to f.

We now show that f as constructed above is hereditary. Fix $S = \{s_0, \dots, s_k\}$, and suppose for contradiction that this choice of S witnesses that f is *not* hereditary. Then, for each $Z_\alpha : S \to \{0,1\}$, there is $Y_\alpha : \bar{S} \to \{0,1\}$ such that $Z_\alpha \in f(S, Y_\alpha)$. We obtain a complete binary tree of height $k+1$ specified by each path $Y_\alpha \sqcup Z_\alpha$ constructed in this manner, restricted to S. In particular, any two paths constructed in this manner first differ at some index in S, as the first index at which $Y_\alpha \sqcup Z_\alpha$ and $Y_\beta \sqcup Z_\beta$ differ is in S. Since each Z_α is not banned, we have a complete binary tree of height $k+1$ in which every leaf can be properly labeled, a contradiction.

The result then follows from Theorem 3.7. ⊣

3.2. An application to type trees. Banned binary sequence problems can be applied to other problems with a tree structure. We use this to improve a result of Malliaris and Terry [7].

DEFINITION 3.12. Given a graph $G = (V, E)$ on n vertices and $A \subseteq 2^{<n}$, closed under initial segments, we say that a labeling $V = \{a_\eta \mid \eta \in A\}$ is a *type tree* if for each $\eta \in A$:

(1) If $\eta \sqcup 0 \in A$, then $a_{\eta \sqcup 0}$ is nonadjacent to a_η. If $\eta \sqcup 1 \in A$, then $a_{\eta \sqcup 1}$ is adjacent to a_η.
(2) If $\eta \subsetneq \eta' \subsetneq \eta''$, then a_η is adjacent to $a_{\eta'}$ if and only if a_η is adjacent to $a_{\eta''}$.

A type tree has height h if $A \subseteq 2^{<h}$ but $A \not\subseteq 2^{<h-1}$.

More generally, given a model $\mathcal{M}$, a finite set $B \subseteq M$, a finite collection Δ of partitioned formulas closed under cycling of the variables, and $A \subseteq \omega^{<\omega}$ closed

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

under initial segments, a type tree is a labeling $B = \{b_\eta \mid \eta \in A\}$ such that, for any $\eta, \eta' \in A$, b_η and $b_{\eta'}$ have the same Δ-type over their common predecessors $\{b_\zeta \mid \zeta \subsetneq \eta, \beta \subsetneq \eta'\}$ iff $\eta \subseteq \eta'$ or $\eta' \subseteq \eta$. Type trees are used in more generality in [6], but we restrict our attention to type trees of graphs.

DEFINITION 3.13. The *tree rank* of a graph $G = (V, E)$ is the largest integer t such that there is a subset $V' \subset V$ and some indexing $V' = \{a_\eta \mid \eta \in 2^{<t}\}$ which is a type tree for the induced graph on V', that is, the type tree of V' is a full binary tree of height t.

The main interest in type trees for graphs lies in the fact that if we have a branch of length h for a graph (V, E) with tree rank t, there is a clique or independent set of size at least $\max\{\frac{h}{2}, t\}$ [7, Lemma 4.4]. More generally, branches through a type tree can be used to extract indiscernible sequences [6, Theorem 3.5]. In both cases, stability establishes the length of long branches through the type tree. For graphs, this is by way of tree rank—observe that the edge relation having Littlestone dimension k implies that the tree rank is at most $k + 1$. We use banned binary sequence problems to improve the bounds from [7, Theorem 4.6]. The improvement is modest, but it demonstrates how banned binary sequence problems accommodate the combinatorics of type trees, at least in the case of the graph edge relation.

THEOREM 3.14. *Let $G = (V, E)$ be a graph with $|V| = n$ and tree rank $t \geq 2$. Suppose $A \subseteq 2^{<n}$ and $V = \{a_\eta \mid \eta \in A\}$ is a type tree with height h, where $h \geq 2t$. Then*

$$h \geq \big(n \cdot (t-2)!\big)^{\frac{1}{t}} + 1.$$

The assumptions on t and h are not restrictive if our aim is to obtain cliques or independent sets. If $t = 1$, then there is no branching, and we obtain a clique or independent set of size $\frac{n}{2}$. If $h < 2t$, then the largest clique or independent set guaranteed by [7, Lemma 4.4] is just the tree rank t.

PROOF. We will associate a hereditary t-fold banned binary sequence problem of length $h - 1$ with the type tree. Fix any subset $S = \{s_0, \dots, s_{t-1}\}$ in $\binom{[h-1]}{t}$ and any $Y : \bar{S} \to \{0, 1\}$. Let $f(S, Y)$ consist of all $Z : S \to \{0, 1\}$ such that $(Y \sqcup Z)_{[s_{t-1}+1]}$ is an element of $2^{<h}$ which is in the index set A of the type tree.

Suppose for contradiction that $f(S, Y) = 2^t$. For each $\eta \in 2^{<t+1}$, we identify η with a partial function $Z_\eta : S \rightharpoonup \{0, 1\}$, where $\eta(i) = Z_\eta(s_i)$. For each $i < t$ and each $\eta : [i] \to \{0, 1\}$ in $2^{<t+1} \setminus 2^t$, let $b_\eta = a_{(Y \sqcup Z_\eta)_{[s_i]}}$. For each $\eta : [t] \to \{0, 1\}$ in 2^t, let $b_\eta = a_{(Y \sqcup Z_\eta)_{[s_{t-1}+1]}}$. Note that each b_η is well-defined—in particular, for $\eta \in 2^t$, if $b_\eta = a_{(Y \sqcup Z_\eta)_{[s_{t-1}+1]}}$ was not an element of the type tree, then we would have $Z_\eta \notin f(S, Y)$. The rest of the elements are well-defined since the index set of a type tree is closed under initial segments. Then the b_η define a full binary type tree of height $t + 1$, contradicting our assumption that the tree rank of G is t. So f is a t-fold BBSP of length $h - 1$.

We check that f is hereditary. Suppose for contradiction that f is not hereditary, witnessed by some $S \in \binom{[h-1]}{t}$. So for each $Z_\alpha : S \to \{0, 1\}$, there is $Y_\alpha : \bar{S} \to \{0, 1\}$ such that $Z_\alpha \in f(S, Y_\alpha)$, and for $\alpha \neq \beta$, the first index at which $Y_\alpha \sqcup Z_\alpha$ and

$Y_\beta \sqcup Z_\beta$ differ is in S. Identify each $\eta \in 2^{<t+1}$ with Z_η as above. For each $i < t$ and each $\eta : [i] \to \{0, 1\}$, let $b_\eta = a_{(Y_\alpha \sqcup Z_\alpha)_{[s_i]}}$ for any α such that $Z_\eta \subseteq Z_\alpha$. For each $\eta : [t] \to \{0, 1\}$, let $b_\eta = a_{(Y_\eta \sqcup Z_\eta)_{[s_{t-1}+1]}}$. These b_η are defined since $Z_\eta \notin f(S, Y_\eta)$ by hypothesis. All other b_η, for $\eta : [i] \to \{0, 1\}$, $i < t$, are defined since type trees are closed under initial segments, and well-defined since if $Z_\eta \subseteq Z_\alpha, Z_\beta$, then the first index at which $Y_\alpha \sqcup Z_\alpha$ and $Y_\beta \sqcup Z_\beta$ differ is in S and is at least s_i. Then the b_η form a type tree of height $t + 1$, a contradiction.

Thus a type tree of height h gives a hereditary t-fold banned binary sequence problem of length $h - 1$. Now, by Theorem 3.7, the number of nodes at level h_0, $h_0 = 0, \dots , h - 1$, is at most

$$\sum_{i=0}^{t-1} \binom{h_0}{i}.$$

Thus, the total number of nodes of a type tree of height h and tree rank t is at most

$$\begin{aligned}
\sum_{h_0=0}^{h-1} \sum_{i=0}^{t-1} \binom{h_0}{i} &= 1 + \sum_{h_0=1}^{h-1} \sum_{i=0}^{t-1} \binom{h_0}{i} \\
&= 1 + \sum_{h_0=1}^{h-1} \left(1 + \sum_{i=1}^{t-1} \binom{h_0}{i}\right) \\
&\leq \sum_{h_0=1}^{h-1} \sum_{i=1}^{t-1} \binom{h-1}{i} \qquad (1) \\
&\leq \sum_{h_0=1}^{h-1} \sum_{i=1}^{t-1} \frac{(h-1)^{t-1}}{(t-1)!} \\
&\leq \sum_{h_0=1}^{h-1} \frac{(h-1)^{t-1}}{(t-2)!} \\
&\leq \frac{(h-1)^{t}}{(t-2)!}, \qquad (2)
\end{aligned}$$

where estimates in (1) and (2) follow from hypotheses on t and h. Then

$$\frac{(h-1)^t}{(t-2)!} \geq n,$$

so

$$h \geq \left(n \cdot (t-2)!\right)^{\frac{1}{t}} + 1. \qquad \dashv$$

Under the hypotheses of Theorem 3.14, applying [7, Lemma 4.4] gives us a clique or independent set of size at least

$$\frac{\left(n \cdot (t-2)!\right)^{\frac{1}{t}} + 1}{2}.$$

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

This is an improvement of the lower bound given by Malliaris and Terry [7, Corollary 4.7].

§4. Generalized banned sequence problems and applications. In this section we generalize Theorem 3.7 to the setting of j-ary sequences, and apply the resulting combinatorics to prove Sauer–Shelah type lemmas in the op-rank context [4].

4.1. Banned *j*-ary sequence problems.

DEFINITION 4.1. A *k- fold banned j- ary sequence problem of length n*, for $0 \leq k \leq n$, is a function

$$f : \binom{[n]}{k} \times j^{n-k} \to \mathcal{P}(j^k) \setminus \{j^k\}.$$

A *solution* to g is a j-ary sequence $X \in j^n$ such that for any $S \in \binom{[n]}{k}$,

$$X_S \in f(S, X_{\bar{S}}).$$

As before, for a fixed $S \in \binom{[n]}{k}$, we denote the elements of S by $\{s_0, \dots, s_{k-1}\}$, where $s_0 < s_1 < \cdots < s_{k-1}$. When we consider $f(S, Y)$, we view $Y \in j^{n-k}$ as a function $Y : \bar{S} \to [j] = \{0, 1, \dots, j-1\}$, and elements of $f(S, Y)$ as functions $Z : S \to [j]$, identifying j^{n-k} with $j^{\bar{S}}$ and j^k with j^S.

DEFINITION 4.2. A k-fold banned j-ary sequence problem (j-ary BSP) f of length n is *not hereditary* if there is $S \in \binom{[n]}{k}$ and a function $g : j^S \to j^{\bar{S}}$ such that

- for all $Z : S \to [j]$, we have $Z \in f(S, g(Z))$, and
- for all $Z_\alpha \neq Z_\beta$, the first index at which $g(Z_\alpha) \sqcup Z_\alpha$ and $g(Z_\beta) \sqcup Z_\beta$ differ is in S.

Otherwise, say f is *hereditary* .

As before, we suppress g and use indices to indicate the mapping, letting Y_α denote Z_α.

THEOREM 4.3. *Any hereditary k-fold banned j-ary sequence problem of length n has at most $\sum_{i=0}^{k-1} (j-1)^{n-i} \binom{n}{i}$ solutions.*

The proof is similar to the proof of Theorem 3.7. We use the generalized versions of the derivative problems for the induction.

DEFINITION 4.4. Let f be a k-fold banned j-ary sequence problem of length n, for $1 \leq k \leq n-1$.

- Let $\hat{f}$ be the $(k-1)$-fold banned j-ary sequence problem of length $n-1$ given as follows: for all $T \in \binom{[n-1]}{k-1}$, all $Y \in j^{n-k}$, and all $Z \in j^{k-1}$, let

$$Z \notin \hat{f}(T, Y) \qquad \text{iff} \qquad \exists l \in [j]\ Z \sqcup l \notin f(T \sqcup \{n-1\}, Y).$$

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

- Let f' be the k-fold banned j-ary sequence problem of length $n-1$ given as follows: for all $S \in \binom{[n-1]}{k}$, all $Y \in j^{n-k-1}$, and all $Z \in j^k$, let

$$Z \notin f'(S, Y) \qquad \text{iff} \qquad \forall l \in [j]\ Z \notin f(S, Y \sqcup l).$$

LEMMA 4.5. *Suppose f is a hereditary k-fold banned j-ary sequence problem of length n, for $1 \leq k \leq n-1$. Then both $\hat{f}$ and f' are also hereditary.*

The proof is a straightforward generalization of Lemma 3.9.

PROOF OF THEOREM 4.3. The proof is by induction on n and k. Let f be a hereditary k-fold banned j-ary sequence problem of length n.

Let $B(f)$ denote the number of sequences banned by f. It suffices to prove that

$$B(f) \geq j^n - \sum_{i=0}^{k-1} (j-1)^{n-i} \binom{n}{i}.$$

The base cases are $k = n$ and $k = 0$. When $k = n$, $j^n - \sum_{i=0}^{k-1}(j-1)^{n-i}\binom{n}{i} = 1$, and any j-ary BSP has at least one banned sequence. When $k = 0$, for all $X \in j^n$, we have $X_\emptyset = \emptyset \notin f(\emptyset, X_{[n]}) = \emptyset$. So all $X \in j^n$ are banned.

Otherwise, we proceed by induction. We show

$$B(f) \geq B(\hat{f}) + B(f') \cdot (j-1).$$

For each sequence $\hat{X}$ that is banned by $\hat{f}$, there is at least one extension X which is banned by f, and we pick one such extension. For each sequence X' banned by f', there are at least $j-1$ extensions X of X' which are banned by f but were not obtained by extending banned sequences for $\hat{f}$. Therefore these banned sequences constructed from f' and $\hat{f}$ have no common members, and so we have

$$B(f) \geq B(\hat{f}) + B(f') \cdot (j-1),$$

as desired. By induction, we have that

$$\begin{aligned} B(f) &\geq j^{n-1} - \sum_{i=0}^{k-2} (j-1)^{n-1-i} \binom{n-1}{i} \\ &\quad + (j-1)\left(j^{n-1} - \sum_{i=0}^{k-1} (j-1)^{n-1-i} \binom{n-1}{i}\right) \\ &= j^n - \sum_{i=0}^{k-1} (j-1)^{n-i} \binom{n}{i}. \end{aligned}$$

Thus, f has at most $\sum_{i=0}^{k-1}(j-1)^{n-i}\binom{n}{i}$ solutions. ⊣

4.2. On the op-rank shatter function. The context of banned j-ary sequences allows us to work in the op-rank context of [4], which we reframe in terms of set systems. Whereas VC dimension and Littlestone dimension make use of binary trees, op_s-rank makes use of 2^s-ary trees.

Definition 4.6. A 2^s- *ary element tree* T of height n with labels from X is a labeling of each node $v \in (2^s)^{<n}$ by s-tuples $x_v = (x_{v,0}, \dots, x_{v,s-1})$ from X. A *leaf* of T is an element of $(2^s)^n$. A leaf ξ is properly labeled by a set A if, for all $j < n$ and for all $i < s$, $x_{\xi|_{[j]},i} \in A$ iff $\xi(j)(i) = 1$.

While this will be the definition that we use in practice, it is often useful think of such trees as binary trees with certain requirements on uniformity of labels within levels.

Definition 4.7. An *alternative* 2^s- *ary element tree* T of height n with labels from X is a labeling of $2^{<ns}$ by elements of X such that given any two nodes σ and σ' with labels x_σ and $x_{\sigma'}$, if $|\sigma| = |\sigma'| = l$ and $\sigma|_{s[\lfloor \frac{l}{s} \rfloor]} = \sigma'|_{s[\lfloor \frac{l}{s} \rfloor]}$, then $x_\sigma = x_{\sigma'}$. A *leaf* of T is an element of 2^{ns}, that is, a binary sequence of length ns. A leaf τ is properly labeled by a set A if, for all $j < ns$, $x_{\tau|_{[j]}} \in A$ iff $\tau(j) = 1$.

Definition 4.8. The op_s- *rank* of a set system $(X, \mathcal{F})$, written $\mathrm{opR}_s(X, \mathcal{F})$ or $\mathrm{opR}_s(\mathcal{F})$, is the largest $k < \omega$ such that there is a 2^s-ary element tree of height k with labels from X such that every leaf can be properly labeled by elements of $\mathcal{F}$, or is infinite if there are such trees of arbitrary height. As a convention, we set $\mathrm{opR}_s(\mathcal{F}) = -\infty$ if $\mathcal{F} = \emptyset$. The op_s *shatter function* $\psi^s_{\mathcal{F}}(n)$ is the maximum number of leaves properly labeled by elements of $\mathcal{F}$ in a 2^s-ary element tree of height n.

It is easy to verify that the op_s-rank and op_s shatter function do not depend on which definition of 2^s-ary element tree we take.

The op_s context is therefore intermediate between the stable context and VC context. Instead of picking labels node by node (as in the stable context) or uniformly for a single level (as in the VC context), we pick labels s at a time. We observe that Littlestone dimension is just the op_1-rank, and VC dimension is the greatest integer s such that the op_s-rank is at least 1.

Likewise, the op_s shatter function is a natural generalization of both the VC and stable shatter functions—observe that the VC shatter function $\pi_{\mathcal{F}}(n)$ is exactly $\psi^n_{\mathcal{F}}(1)$, and the stable shatter function $\rho_{\mathcal{F}}(n)$ is exactly $\psi^1_{\mathcal{F}}(n)$. Although the op-ranks as developed in [4] were indeed intended as a generalization of Littlestone dimension (there referred to as Shelah's 2-rank) and have natural connections with VC dimension, they more strongly considered the geometric properties as they pertained to model theory, and did not study the combinatorics surrounding the shatter function. We study the shatter function here, in particular examining connections between finite op-ranks and growth rates of op-shatter functions.

As before, the dual op_s-rank and dual op_s shatter function of a set system are the op_s-rank and op_s shatter function of the dual set system.

Corollary 4.9. *Let $\mathcal{F}$ be a set system with* $\mathrm{opR}_s(\mathcal{F}) = k$. *Then*

$$\psi^s_{\mathcal{F}}(n) \leq \sum_{i=0}^{k} (2^s - 1)^{n-i} \binom{n}{i}.$$

The proof follows our proof of Theorem 2.5, using j-ary banned sequence problems.

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

PROOF. Let T be an 2^s-ary element tree of height n. Identifying the 2^s binary sequences of length s with $[2^s]$, we obtain a hereditary $(k+1)$-fold banned 2^s-ary sequence problem f of length n as follows. Given $S = \{s_0, \dots, s_k\} \in \binom{[n]}{k+1}$, where $s_0 < s_1 < \cdots < s_k$ and $Y : \bar{S} \to 2^s$, we obtain a 2^s-ary element tree of height $k+1$ by taking all paths $\xi \in (2^s)^n$ through T such that $Y \subset \tau$. Decisions will only be made at nodes v, where $|v| \in S$, so removing all other nodes gives us a 2^s-ary element tree $T_{S,Y}$ of height $k+1$. Since $\mathrm{opR}_s(\mathcal{F}) = k$, not all leaves of $T_{S,Y}$ can be properly labeled, so let $f(S, Y)$ be the set of all sequences whose corresponding leaves in $T_{S,Y}$ can be properly labeled. Then a leaf in T can only be properly labeled if the corresponding sequence is a solution to f.

It remains to show that f is hereditary. Fix $S = \{s_0, \dots, s_k\}$, and suppose for contradiction that this choice of S witnesses that f is *not* hereditary. Then, for any $Z_\alpha : S \to [2^s]$, there is $Y_\alpha : \bar{S} \to [2^s]$ such that $Z_\alpha \in f(S, Y_\alpha)$. We obtain a complete 2^s-ary tree of height $k+1$ specified by each path $Y_\alpha \sqcup Z_\alpha$ constructed in this manner, restricted to S. Since each Z_α is not banned, we have a 2^s-ary tree of height $k+1$ in which every leaf can be properly labeled, a contradiction.

The result then follows from Theorem 4.3. ⊣

The bound of Corollary 4.9 can be improved by using more information—in particular, when bounding the op_s shatter function, we can consider op_r-ranks for $r \le s$. We can already give a better bound for the case where a set system has op_r-rank 0 for some r.

PROPOSITION 4.10. *Let $\mathcal{F}$ be a set system with $\mathrm{opR}_r(\mathcal{F}) = 0$. Then*

$$\psi^s_{\mathcal{F}}(n) \le \left(\sum_{i=0}^{r-1} \binom{s}{i} \right)^n .$$

PROOF. Call a node *live* if it is the initial segment of a leaf that can be properly labeled. At each node of the tree, we consider s elements. Observing that $\mathrm{opR}_r(\mathcal{F}) = 0$ says precisely that the VC dimension of $\mathcal{F}$ is strictly less than r, Theorem 2.2 tells us that we can find sets which properly label at most $\sum_{i=0}^{r-1} \binom{s}{i}$ of the possible boolean combinations of the s elements. That is, each live node has at most $\sum_{i=0}^{r-1} \binom{s}{i}$ live successors in the next level. Therefore, there are at most $\left(\sum_{i=0}^{r-1} \binom{s}{i}\right)^m$ live nodes at the level of height m (counting from 0). Since leaves in a tree of height n appear at the nth level, the result follows. ⊣

The set system of half-spaces in $\mathbb{R}^r$ achieves the bound of Proposition 4.10 for the *dual* op_s shatter function. (This is the famous cake-cutting problem.)

PROPOSITION 4.11. *Let $\mathcal{F}$ be the dual set system to the set system of $\mathbb{R}^r$ consisting of half-spaces. Then*

$$\psi^s_{\mathcal{F}}(n) = \left(\sum_{i=0}^{r} \binom{s}{i} \right)^n .$$

In particular, $\mathrm{opR}_{r+1}(\mathcal{F}) = 0$.

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

PROOF. It suffices to verify that taking s hyperplanes in general position (i.e., so that any m hyperplanes intersect in a $(r-m)$-dimensional subspace) partitions $\mathbb{R}^r$ into $\sum_{i=0}^{r}\binom{s}{i}$ pieces, each of which contains an open set (in the Euclidean topology). Such a partition corresponds to one level in the 2^s-ary tree. Each piece may then be partitioned further in the same manner for each successive level of the tree.

We proceed by induction. The $s = 1$ case is obvious, for all r. The $r = 1$ case is obvious, for all s.

Consider the $s + 1$ and $r + 1$ case. Removing one of the $s + 1$ hyperplanes, we have $\sum_{i=0}^{r+1}\binom{s}{i}$ pieces by induction. Restore the hyperplane that we removed. Viewing that hyperplane as a copy of $\mathbb{R}^r$, it is partitioned into $\sum_{i=0}^{r}\binom{s}{i}$ pieces by the other hyperplanes, by induction. Each such piece corresponds to a piece in $\mathbb{R}^{r+1}$ which is cut into two pieces by the restored hyperplane. We therefore find that the total number of pieces is

$$\sum_{i=0}^{r+1}\binom{s}{i}+\sum_{i=0}^{r}\binom{s}{i}=\sum_{i=0}^{r+1}\binom{s+1}{i}.$$

as desired. ⊣

We can further refine our methods. Fix a base set X. We identify any set system $(X, \mathcal{F})$ on X with $\mathcal{F}$.

PROPOSITION 4.12.

(1) *Let* $\mathcal{F}_1 \subseteq \mathcal{F}_2$. *Then, for any* s, $\operatorname{opR}_s(\mathcal{F}_1) \leq \operatorname{opR}_s(\mathcal{F}_2)$.
(2) *Let* $s_1 < s_2$. *Then* $\operatorname{opR}_{s_1}(\mathcal{F}) \geq \lfloor \frac{s_2}{s_1} \rfloor \operatorname{opR}_{s_2}(\mathcal{F})$.

PROOF. (1) is trivial. For (2), suppose that we have a 2^{s_2}-ary element tree T of height $n_2 := \operatorname{opR}_{s_2}(\mathcal{F})$, with labels $x_v = (x_{v,0}, \dots, x_{v,s_2-1})$ for each $v \in (2^{s_2})^{<n_2}$, in which every leaf can be properly labeled. Then we can obtain a 2^{s_1}-ary element tree T' of height $n_1 := \lfloor \frac{s_2}{s_1} \rfloor n_2$ in which every leaf can be properly labeled. Let $t = \lfloor \frac{s_2}{s_1} \rfloor$. Intuitively, we split each level of the 2^{s_2}-ary tree into t levels of the 2^{s_1}-ary tree, with any label $x_v = (x_{v,0}, \dots, x_{v,s_2-1})$ splitting into t labels

$$(x_{v,0}, \dots, x_{v,s_1-1}), (x_{v,s_1}, x_{v,2s_1-1}), \dots, (x_{v,(t-1)s_1}, \dots, x_{v,ts_1-1}).$$

More formally, suppose $\xi \in (2^{s_1})^i$, for $i < n_1$. Suppose $i = jt + k$, for $0 \leq k < t$. Then label ξ with

$$x_\xi = (x_{v_\xi,ks_1}, \dots x_{v_\xi,(k+1)s_1-1}),$$

where $v_\xi \in (2^{s_2})^j$ is as follows. Let $\sigma_l = \xi(l) \in 2^{s_1}$. Then let $\tau_m \in 2^{s_2}$ be the concatenation of $\sigma_{mt}, \dots, \sigma_{(m+1)t-1}$, appending as many 0s as needed to obtain a sequence of length s_2. Then let

$$v_\xi := (\tau_0, \dots, \tau_{j-1}).$$

Then the labeling of T' by the x_ξ gives a 2^{s_1}-ary tree of height n_1 in which every leaf can be properly labeled (in particular, by one of the labels of the leaves of the 2^{s_2}-ary tree). ⊣

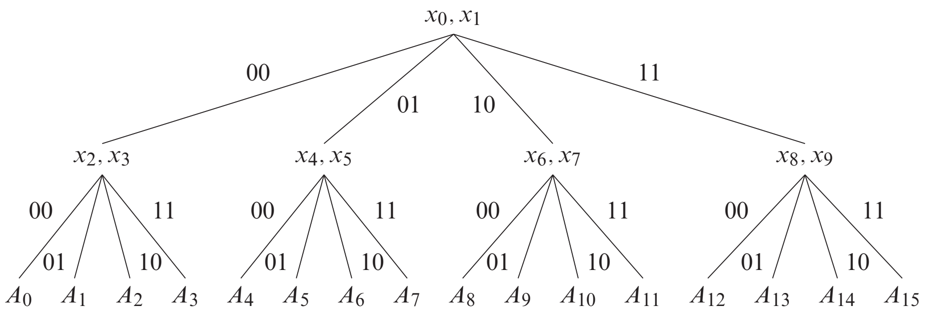

FIGURE 1. A 2^2-ary element tree of height 2. A_9 properly labels its leaf if it contains x_0 and x_7, but does not contain x_1 and x_6, with no requirements on membership of the other elements.

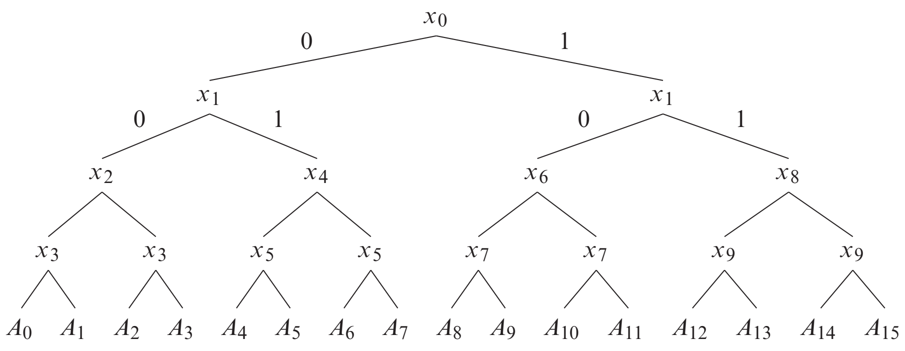

FIGURE 2. An alternative 2^2-ary element tree of height 2. Observe that the labels on the first two levels are uniform. Then, on the fourth level (and, trivially, the third level), labels are uniform across all nodes with the same initial segment of length 2. We identify 1 with the right branch. As before, A_9 properly labels its leaf if it contains x_0 and x_7, but does not contain x_1 and x_6, with no requirements on membership of the other elements.

(2) shows how different finite op_s ranks can interact; in particular, a finite op_s rank establishes upper bounds on $\mathrm{op}_{s'}$ ranks, for $s < s'$. (2) above is somewhat easier to see using the alternative definition—we simply view the tree as a 2^{s_1}-ary tree instead of a 2^{s_2}-ary tree, possibly after removing some levels. Figure 2 is the 2^1-ary tree obtained from Figure 1 by this process.

Given $\mathcal{F}$, $x_0, \dots, x_{s-1} \in X$, and $\sigma : [s] \to 2$, let

$$\mathcal{F}_\sigma := \{Y \in \mathcal{F} \mid \text{for all } i < n,\ x_i \in Y \text{ iff } \sigma(i) = 1\}.$$

Call each $\mathcal{F}_\sigma$ a child of $\mathcal{F}$. Then, in an op_s-tree with root $(x_0, \dots, x_{s-1})$, $\mathcal{F}_\sigma$ consists of all sets in $\mathcal{F}$ which properly label a leaf whose path begins with σ. Observe that if for all $\sigma : [s] \to 2$, $\mathrm{opR}_s(\mathcal{F}_\sigma) \geq a$, then $\mathrm{opR}_s(\mathcal{F}) \geq a + 1$; we can obtain a 2^s-ary tree of height $a+1$ by labeling the root with $(x_0, \dots, x_{s-1})$, and appending 2^s-ary trees of height a witnessing $\mathrm{opR}_s(\mathcal{F}_\sigma) \geq a$ at the appropriate successor nodes.

The following lemma generalizes the observation that, given $\mathcal{F}$ with Littlestone dimension $a < \infty$ and any $x \in X$, at most one of $\{F \in \mathcal{F} \mid x \in F\}$ and $\{F \in \mathcal{F} \mid x \notin F\}$ has Littlestone dimension a; if both had Littlestone dimension a, joining the two binary element trees witnessing this with root x would witness that $\mathcal{F}$ has Littlestone dimension $a+1$.

LEMMA 4.13. *Suppose* $\mathrm{opR}_r(\mathcal{F}) = a < \infty$. *Then, given any* $x_0, \dots, x_{s-1} \in X$, *we have* $\mathrm{opR}_r(X_\sigma) \leq a-1$ *for at least* $2^s - \sum_{i=0}^{r-1} \binom{s}{i}$ *children* $\mathcal{F}_\sigma$. *More generally, we have* $\mathrm{opR}_r(X_\sigma) \leq a-l$ *for at least* $2^s - \sum_{i=0}^{lr-1} \binom{s}{i}$ *children* $\mathcal{F}_\sigma$.

PROOF. We obtain an independent r-fold banned binary sequence problem f of length s as follows. For each $S \in \binom{[s]}{r}$, let $f(S)$ be those functions $\eta : S \to 2$ such that $\mathrm{opR}_r(\mathcal{F}_\eta) \leq a-1$, where

$$\mathcal{F}_\eta := \{Y \in \mathcal{F} \mid \text{for all } i \in S,\ x_i \in Y \text{ iff } \eta(i) = 1\}.$$

Each $f(S)$ is nonempty, or else those $\mathcal{F}_\eta$ witness that that $\mathrm{opR}_r(\mathcal{F}) \geq a+1$, a contradiction. Then $\sigma : [s] \to 2$ is banned by f if there is some $S \in \binom{[s]}{r}$ such that $\mathrm{opR}_r(\mathcal{F}_{\sigma_S}) \leq a-1$, whence $\mathrm{opR}_r(\mathcal{F}_\sigma) \leq a-1$. So sequences banned by f have the corresponding child drop in op_r-rank, of which there are at least $2^s - \sum_{i=0}^{r-1} \binom{s}{i}$ many.

For the more general case, we instead obtain an independent lr-fold banned binary sequence problem. For each $S \in \binom{[s]}{lr}$, let $f(S)$ be those $\eta : S \to 2$ such that $\mathrm{opR}_r(\mathcal{F}_\eta) \leq a-l$. Each $f(S)$ is nonempty, or else those $\mathcal{F}_\eta$ witness that $\mathrm{opR}_r(\mathcal{F}) \geq a+1$. Then sequences banned by f have the corresponding child drop in op_r-rank by at least l, of which there are at least $2^s - \sum_{i=0}^{lr-1} \binom{s}{i}$ many. ⊣

The boundary between finite and infinite op-ranks serves as an important parameter in obtaining better bounds. It is also of model-theoretic interest, coinciding with other known properties.

DEFINITION 4.14. The op- *dimension* of a set system $\mathcal{F}$ is

$$\sup\{r \mid \mathrm{opR}_r(\mathcal{F}) = \infty\}.$$

Expressed in model-theoretic terms, the op-dimension of a (type-)definable set X in some model is the supremum of the op-dimension of set systems on X generated finite sets of formulas. In this context, op-dimension coincides with o-minimal dimension in o-minimal theories and dp-rank in distal theories [4].

We use Lemma 4.13 to obtain better bounds on the op_s shatter function by using op-dimension.

DEFINITION 4.15. Let $\psi^s_{r,b}(n)$ be the greatest possible number of properly labeled leaves in a 2^s-ary tree of height n by any set system $\mathcal{F}$ with $\mathrm{opR}_r(\mathcal{F}) \leq b < \omega$.

THEOREM 4.16. *Let* $a_0 := \sum_{i=0}^{r-1} \binom{s}{i}$ *and* $a_1 = 2^s - a_0$. *Then*

$$\psi^s_{r,b}(n) \leq \sum_{i=0}^{b} \binom{n}{i} a_0^{n-i} a_1^i.$$

https://doi.org/10.1017/jsl.2019.35 Published online by Cambridge University Press

PROOF. The case $n = 0$ is trivial for all b. We proceed by induction on b. The case $b = 0$ is Proposition 4.10.

For the case $b + 1$, we observe that, by monotonicity of $\psi^s_{r,b}(n)$ in b, we maximize the possible number of properly labeled leaves by having as many children as possible not decrease in op_r-rank. We now proceed by induction on n. By Lemma 4.13, we can have at most a_0 such children, and the remaining a_1 children must drop in op_r-rank by at least 1. We therefore obtain the recurrence

$$\begin{aligned}
\psi^s_{r,b+1}(n) &\le a_0\psi^s_{r,b+1}(n-1) + a_1\psi^s_{r,b}(n-1)\\
&\le a_0\sum_{i=0}^{b+1}\binom{n-1}{i}a_0^{n-i-1}a_1^i + a_1\sum_{i=0}^{b}\binom{n-1}{i}a_0^{n-i-1}a_1^i \quad \text{by induction}\\
&\le \sum_{i=0}^{b+1}\binom{n-1}{i}a_0^{n-i}a_1^i + \sum_{i=0}^{b}\binom{n-1}{i}a_0^{n-i-1}a_1^{i+1}\\
&\le \binom{n-1}{0}a_0^n + \sum_{i=1}^{b+1}\binom{n-1}{i}a_0^{n-i}a_1^i + \sum_{i=1}^{b+1}\binom{n-1}{i-1}a_0^{n-i}a_1^i\\
&\le \binom{n}{0}a_0^n + \sum_{i=1}^{b+1}\binom{n}{i}a_0^{n-i}a_1^i\\
&\le \sum_{i=0}^{b+1}\binom{n}{i}a_0^{n-i}a_1^i
\end{aligned}$$

as desired. ⊣

In particular, for a set system with op-dimension d, we take $r = d + 1$. Then the op shatter function is bounded by an exponential function with the base a_0 determined by d. Furthermore, coefficients for lower order terms can be improved when $r \le \frac{s}{2}$, as then the more general case of Lemma 4.13 dictates that some children must drop in op_r-rank by more than 1. This creates a more complicated recurrence, but the result remains exponential in a_0.

Finally, we observe that we can recover both the VC and stable Sauer–Shelah bounds from Theorem 4.16. If $\mathcal{F}$ has VC dimension r, then $\mathrm{opR}_{r+1}(\mathcal{F}) = 0$. Then

$$\pi_{\mathcal{F}}(s) = \psi^s_{\mathcal{F}}(1) \le \psi^s_{r+1,0}(1) \le \sum_{i=0}^{r}\binom{s}{i}.$$

Similarly, if $\mathcal{F}$ has Littlestone dimension b, this says that $\mathrm{opR}_1(\mathcal{F}) = b$. Then

$$\rho_{\mathcal{F}}(n) = \psi^1_{\mathcal{F}}(n) \le \psi^1_{1,b}(n) \le \sum_{i=0}^{b}\binom{n}{i}.$$

Acknowledgments. The authors would like to thank Dave Marker, Vincent Guingona, Caroline Terry, Maryanthe Malliaris, Gabriel Conant, and Cameron Hill for useful conversations and suggestions during the preparation of this article. The authors would also like to especially thank Alex Kruckman and Siddharth Bhaskar for a number of useful comments as well as pointing out some mistakes in early versions of several proofs. The authors would also like to thank the anonymous referee for several suggestions that significantly improved the exposition. James Freitag was supported by NSF grant no. 1700095.

REFERENCES

[1] S. Bhaskar, *Thicket density*, arXiv preprint, 2017. arXiv:1702.03956.

[2] H. Chase and J. Freitag, *Model theory and machine learning*. ***Bulletin of Symbolic Logic***, vol. 25 (2019), pp. 319–332.

[3] A. Chernikov, D. Palacin, and K. Takeuchi, *On n-dependence*, arXiv preprint, 2014. arXiv:1411.0120.

[4] V. Guingona and C. D. Hill, *On a common generalization of Shelah's 2-rank, dp-rank, and o-minimal dimension*. ***Annals of Pure and Applied Logic***, vol. 166 (2015), no. 4, pp. 502–525,

[5] M. C. Laskowski, *Vapnik-Chervonenkis classes of definable sets*. ***Journal of the London Mathematical Society***, vol. 2 (1992), no. 2, pp. 377–384,

[6] M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*. ***Transactions of the American Mathematical Society***, vol. 366 (2014), pp. 1551–1585.

[7] M. Malliaris and C. Terry, *On unavoidable induced subgraphs in large prime graphs*. ***Journal of Graph Theory***, vol. 88 (2017), 2, pp. 255–270.

[8] H. Q. Ngo, *Three proofs of Sauer–Shelah Lemma*, Course notes, 2010. Available at https://www.cse.buffalo.edu/hungngo/classes/2010/711/lectures/sauer.pdf.

[9] P. Simon, ***A Guide to NIP Theories***, Cambridge University Press, Cambridge, MA, 2015.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS CHICAGO
CHICAGO, IL, USA
E-mail: freitagj@gmail.com

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND, COLLEGE PARK
COLLEGE PARK, MD, USA
E-mail: hschase@umd.edu