On Alternation, VC-dimension and k-fold Union of Sets

Amit Roy⋆ and Jayalal Sarma⋆

Abstract. Alternation of Boolean functions is a measure of non-monotonocity of the function. In this paper, we asymptotically characterize the VC-dimension of family of Boolean functions parameterized by the maximum alternation of the Boolean functions in the family. Enroute to our main result, we show exact bounds for VC-dimension of functions which has alternation 1, which strictly contains monotone functions and hence generalizes the bounds in [4]. As an application, we show tightness of VC-dimension bounds for k-fold union, by explicitly constructing a family $\mathcal{F}$ of subsets of $\{0,1\}^n$ such that k-fold union of the family, $\mathcal{F}^k = \left\{\bigcup_{i=1}^k F_i \mid F_i \in \mathcal{F}\right\}$ must have VC-dimension at least $\Omega(dk)$ and that this bound holds even when the union is over disjoint sets from $\mathcal{F}$. This provides a non-geometric set system achieving this bound.

1 Introduction

Vapnik-Chervonenkis Dimension (VC-dimension) is a combinatorial measure of a set system of subsets of a universe, developed by Vapnik and Chervonenkis in the 1960s and has found deep applications the area of statistical learning theory, discrete and computational geometry.

Let U be universe and $\mathcal{F}$ be a set of subsets of U. The family $\mathcal{F}$ is said to *shatter* a subset $S \subseteq U$ if for all subset $S' \subseteq S$, there is a $F \in \mathcal{F}$ such that $S \cap F = S'$. Notice that it is easy to shatter small sets, especially the empty set, and U is not shattered unless $\mathcal{F} = \mathcal{P}(U)$. The VC-dimension of $\mathcal{F}$ is the largest d such that there is a set S of size d shattered by the family.

In the case when $U = \{0,1\}^n$, each $F \in \mathcal{F}$ can be interpreted as the positive inputs of a Boolean function. That is, as the set $f^{-1}(1) = \{x \in \{0,1\}^n \mid f(x) = 1\}$ of some Boolean function $f : \{0,1\}^n \to \{0,1\}$. Hence, any family $\mathcal{F}$ can be equivalently interpreted as a family of Boolean functions. Under this interpretation, the VC-dimension of a class is directly related to the complexity of learning the Boolean functions in the class in the Probably Approximately Correct (PAC) model [14]: where it is known to yield matching upper [3] and lower [5] bounds for the number of samples required in order to learn the functions from the family.

Motivated by this, there has been several works exploring tight VC dimensions bounds various families of Boolean functions with n variables. Tight bounds are known for VC-dimension of various subfamilies of Boolean functions such as Boolean terms (conjunction of literals) - $O(n)$ [10], k-Decision Lists - n^k [11], Symmetric Functions - n [5].

⋆ Indian Institute of Technology Madras, Chennai, India.

An important class of functions for which a characterization is known is the class of monotone functions [4] where the VC-dimension was established to be exactly $\binom{n}{n/2}$. A subclass of this family - namely the monotone terms were also studied in [10] to establish a tight linear bound for VC-dimension. A natural question is how to generalize these bounds to non-monotone functions as well.

For $x, y \in \{0,1\}^n$ we say $x \preceq y$ if $\forall i \in [n], x_i \leq y_i$ where x_i represents the i^{th} bit of x. Recall that a function f is said to be monotone, if $\forall x, y \in \{0,1\}^n, x \preceq y$ then $f(x) \leq f(y)$. Consider a maximal chain of distinct inputs $x_0, x_1, x_2 \ldots, x_n \in \{0,1\}^n$ satisfying $x_0 \preceq x_1 \preceq x_2 \preceq \cdots \preceq x_n$. The alternation of f (denoted by $\mathsf{alt}(f)$) is defined as $\max\{\mathsf{alt}(f, \mathcal{C}) \mid \mathcal{C}$ is a maximal chain in $\mathcal{B}_n\}$ where $\mathsf{alt}(f, \mathcal{C})$ is $|\{i \mid f(x_{i-1}) \neq f(x_i), x_i \in \mathcal{C}, i \in [n]\}|$. Indeed, for a monotone f, $\mathsf{alt}(f) = 1$ and for any Boolean function f, $alt(f) \leq n$. Thus, it forms a measure of how much non-monotone the Boolean function is.

Our Results: In this paper, we initiate a study of the VC-dimension of Boolean function families parameterized by the *alternation* and show the following results:

- Exact VC-dimension for family of functions with alternation 1: We show that family of functions with alternation 1 has VC-dimension exactly $\binom{n}{n/2} + 1$. Theorem 1. We also show that this family is shattering extremal as defined by [9] and hence has some potentially useful combinatorial properties.
- Tight bounds for VC-dimension for family of functions with alternation k: We show (Theorem 2 and 3) that the family of functions $\mathcal{F}$ with alternation k has VC-dimension satisfying : $\sum_{i=\frac{n-k}{2}}^{\frac{n+k}{2}} \binom{n}{i} \leq \mathsf{VC}(\mathcal{F}_k) \leq O\left(k \times \binom{n}{n/2}\right)$. For $k \leq \sqrt{n}$ the upper and lower bounds is asymptotically of the same order and hence the bound is tight in general.
- Application to VC-dimension of disjoint union of families: If a family $\mathcal{F}$ is of VC-dimension $\leq d$, how large can the VC-dimension of the k-fold union family, defined as, $\mathcal{F}^k = \{\cup_{i=1}^k A_i \mid A_i \in \forall i, A_i \in \mathcal{F}\}$ be? Blumer *et al* [2] and Haussler and Welzl [7] showed that the VC-dimension is at most $O(dk \log k)$. This bound was shown to be tight by Eisentat and Angluin [6] who shows existence of a geometric family with VC-dimension at most d and the k-fold union has VC-dimension at least $\Omega(dk \log k)$. The family constructed were point sets in a plane.Using our methods, we construct a family of Boolean functions such that k-fold union has VC-dimension at least $\Omega(dk)$ even when the unions in the k-fold union are restricted to k-fold disjoint union.

2 VC-dimension bounds for functions with alternation 1

As a warm up towards later sections, in this section, we describe VC-dimension bounds and extremal properties of a family of functions related to alternation and monotonicity. As mentioned in the introduction, [4] computes the VC-dimension of family of monotone functions (denoted by $\mathcal{M}$) as : $\mathsf{VC}(\mathcal{M}) = \binom{n}{n/2}$

A function $f : \{0,1\}^n \to \{0,1\}$ is said to be k-slice function if $f(x) = 0$ if for every $x \in \{0,1\}^n$, $\sum_{i=1}^n x_i < k$ and 1 if $\sum_{i=1}^n x_i > k$. Define $\mathcal{M}^*$ to be the family of all *slice*-functions. We compute the VC-dimension$(\mathcal{M}^*)$ (see App. A.1).

Proposition 1. $\mathsf{VC}(\mathcal{M}^*) = \binom{n}{n/2}$

Now consider the family of functions where each function has alternation at most 1. $\mathcal{F}_1 = \{f \mid \text{either } f \text{ or } \neg f \text{ is monotone}\}$. We compute the VC-dimension of this family exactly.

Theorem 1. $\mathsf{VC}(\mathcal{F}_1) = \binom{n}{\lfloor n/2 \rfloor} + 1$

Proof. Lower Bound: We show the lower bound by shattering a set $S \subseteq U$ of cardinality $\binom{n}{\lfloor n/2 \rfloor} + 1$. $S = \{x \in \{0,1\}^n \mid \sum_{i=1}^{n} x_i = \lfloor n/2 \rfloor\} \bigcup \{w\}$ where w is any arbitrary point in $\{0,1\}^n$ such that $\sum_{i=1}^{n} x_i < n/2$. Let $S' \subseteq S$. We need to give an $F \in \mathcal{F}$ such that $F \cap S = S'$. We consider the following cases:

- **Case 1:** $S' = \{x \in \{0,1\}^n \mid \sum_{i=1}^{n} x_i = n/2\}$. F such that $F \cap S = S'$ is given by the characteristic function $f = \bigvee_{z \in S'} \bigwedge_{z_i=1} x_i$.
- **Case 2:** $S' = X \cup \{w\}$ where $\forall x \in X, \sum_{i=1}^{n} x_i = n/2$. F such that $F \cap S = S'$ is given by the characteristic function $f = \bigwedge_{z \in S'} \bigvee_{z_i=1} \overline{x_i}$. Observe that negation of this function is monotone.

Upper Bound: Suppose $\mathcal{F}_1$ shatters a set S such that $|S| \geq \binom{n}{n/2} + 2$. We first obtain certain properties that S cannot have through the following lemma.

Lemma 1. *$\mathcal{F}_1$ cannot shatter a set S if it has either of the properties.*

1. **Parallel Chain :** *When there are elements $p, p', q, q' \in S$ such that all of them are distinct and $p \preceq p'$ and $q \preceq q'$.*
2. **Triplet Chain :** *If $p, q, r \in S$ such that $p \preceq q \preceq r$ then it is said to form a triplet chain.*

Proof. For parallel chain, without loss of generality let us suppose that $p \preceq p'$ and $q \preceq q'$. Suppose $S' = \{p', q\}$. Now any monotonically increasing function will obtain the set $\{p', q, q'\}$ and any monotonically decreasing function will obtain the set $\{p, p', q\}$ but never $\{p', q\}$ alone. Hence it cannot be shattered.
For triplet chain, suppose $p, q, r \in S$ such that $p \preceq q \preceq r$. Consider the set $S' = \{p, r\}$. We observe that there does not exist a function $f \in \mathcal{F}_1$ such that it is true on p and r and evaluates to false on q. Hence S cannot be shattered.

Now we claim that set S which is shattered must be a disjoint unions of at most 2 maximal antichains. To see this, suppose $S = S_1 \uplus S_2 \uplus S_3$ such that the sets are maximal antichain. Without loss of generality, consider an element $p \in S_1$. Now there exists a point $q \in S_2$ such that p and q are comparable(otherwise they will be in the same set). Now if this is the case then neither p nor q can be related to any point in S_3 as it will either form a chain length of 3 i.e. $p \prec q \prec r$(triplet chain) or parallel chains because of which the set S cannot be shattered(see Lemma 1). Now if p and q are not comparable then we can have a larger antichain by including either p or q in the set S_3 which contradicts the maximality of antichain set S_3. The same argument can be given for each of the sets. Hence set S can be disjoint union of at most 2 maximal antichains.

Hence we conclude that there must be $S = S_1 \uplus S_2$ where S_1, S_2 are maximal antichains i.e. no elements from S_1 can be put into S_2 and vice versa. Using Lemma 1 again, we have that S_1 and S_2 does not have either parallel chain or triplet chain. But that contradicts the maximality of S_1 and S_2. Consider $p', q' \in S_2$. Then $\exists p \in S_1$ such that $p \preceq p'$ and $p \preceq q'$. But we can obtain a larger antichain by including p', q' into S_1. Thus contradicting the maximality.

We now show that $\mathcal{M}$ and $\mathcal{F}_1$ family exhibit a special property (Proposition 2) which is also known as *s-extremal* or *shattering extremal* family. Following is proved in Appendix A.2.

Proposition 2. *The families $\mathcal{M}$ and $\mathcal{F}_1$ are shattering extremal.*

3 Bounds for VC-dimension in terms of alt(f)

In this section, we derive VC-dimension bounds for families of Boolean functions, parameterized by the maximum alternation of functios in the family. We need the following known theorem.

Lemma 2 (Characterization of Alternation [2]). *Let $f : \{0,1\}^n \to \{0,1\}$. Then there exists $k = \mathsf{alt}(f)$ monotone functions $g_1, \ldots, g_k$ each from $\{0,1\}^n$ to $\{0,1\}$ such that $f(x) = \oplus_{i=1}^k g_i$ if $f(0^n) = 0$ and $f(x) = \neg \oplus_{i=1}^k g_i$ if $f(0^n) = 1$.*

We use this theorem to establish the following upper bound:

Theorem 2. *Let $k > 1$. If $\mathcal{F}_k$ is the family of Boolean functions f such that $\mathsf{alt}(f) \leq k$. Then, $\mathsf{VC}(\mathcal{F}_k) \leq O\left(k\binom{n}{n/2}\right)$*

Proof. We apply Lemma 2 to conclude that $\mathcal{F}_k$ can be equivalently written as $\mathcal{F}_k = \left\{(\neg \oplus \text{ or }) \oplus_{i=1}^k f_i \mid f_i \in \mathcal{M}\right\}$ where $\mathcal{M}$ is the family of Monotone Boolean functions. We look at a family $\mathcal{G} = \left\{f \oplus g \mid f = \oplus_{i=1}^k f_i, f_i \in \mathcal{M}, g = \mathit{const}\right\}$ where $g(x) = 1$ if $f(0^n) = 1$ and $g(x) = 0$ if $f(0^n) = 0$. Observe that $\mathcal{F}_k \subseteq \mathcal{G}$ and hence $\mathsf{VC}(\mathcal{F}_k) \leq \mathsf{VC}(\mathcal{G})$. We turn to the following lemma to show an upper bound in general for such constructed families. Given k classes of n-bit Boolean functions $\mathcal{F}_1, \mathcal{F}_2 \ldots, \mathcal{F}_k$, and a fixed Boolean function $f : \{0,1\}^k \to \{0,1\}$. We define
$\mathcal{F}(\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_k) = \{f(f_1(.), ..., f_k(.)) \mid f_i \in \mathcal{F}_i, i \in [k]\}$. We have,

Lemma 3 ([3,7,12]). *Let $d = \max_{i\in[k]}(\mathsf{VC}(\mathcal{F}_i))$. $\mathsf{VC}(\mathcal{F}(\mathcal{F}_1, \ldots, \mathcal{F}_k)) \leq O(dk \log k)$*

Applying this lemma we obtain for the above family $\mathsf{VC}(\mathcal{F}_k) \leq O\left(k\binom{n}{n/2}\log k\right)$. We show below how to improve the bound. The idea is simple counting : we have $|\mathcal{G}| \leq |\mathcal{M}|^{k+1}$. We know that $\mathsf{VC}(\mathcal{G}) \leq \log(|\mathcal{G}|)$, which gives us $\mathsf{VC}(\mathcal{G}) \leq (k + 1)\log(\mathcal{M})$. This bound is the *Dedekind's number* and we use the following bound due to *Kleitman et al.* (refer [8]) : $\log(\mathcal{M}) \leq \binom{n}{n/2}\left(1 + O(\frac{\log n}{n})\right)$. This gives, $\mathsf{VC}(\mathcal{F}_k) \leq (k+1)\binom{n}{n/2}\left(1 + O\left(\frac{\log n}{n}\right)\right)$ Hence, we have $\mathsf{VC}(\mathcal{F}_k) \leq O(k\binom{n}{n/2})$.

Now we turn to the lower bound. Using the fact that for any $k \geq 1$, the family $\mathcal{F}_k$ also includes the set of monotone functions $\mathcal{M}$, the VC-dimension$(\mathcal{F}_k) \geq$ VC-dimension$(\mathcal{M})$. Hence VC-dimension$(\mathcal{F}_k) \geq \binom{n}{n/2}$. We can improve this:

Theorem 3. *Let $k > 1$. If $\mathcal{F}_k$ is the family of Boolean functions f such that $\mathsf{alt}(f) \leq k$. Then, $\mathsf{VC}(\mathcal{F}_k) \geq \sum_{i=n/2-k/2}^{n/2+k/2} \binom{n}{i}$*

Proof. We shatter the set $S = \{x \in \{0,1\}^n \mid n/2 - k/2 \leq \sum_{i=1}^{n} x_i \leq n/2 + k\}$. To obtain any $S' \subseteq S$, we give $f \in \mathcal{F}_k$ as $f(x) = 1$ whenever $x \in S'$ and $f(x) = 0$ otherwise. It remains to show that $\mathsf{alt}(f) \leq k$. Since the number of 1s in $x \in S'$ can only be in the range $[n/2 - k/2, n/2 + k/2]$, any chain in this part will have alternation at most k and in the remaining part 0. Hence we obtain all $S' \subseteq S$.

VC-dimension of Read-once functions: A Boolean function is said to be *read-once* if there is a Formula (a Boolean circuit where every gate has fanout at most 1) computing the function f such that every variable appears only once in the formula. Monotone read-once functions have no negations in the formula computing the function. The following lemma motivates the study of VC-dimension under composition between read-once functions and functions with alternation at most 1. We prove (see sec.A.3).

Lemma 4. *Let $f : \{0,1\}^n \to \{0,1\}$ such that $\mathsf{alt}(f) \leq k$. Then $f = g(f_1, f_2, \ldots, f_k)$ where g is a monotone-Read Once formula and $f_i \in \mathcal{F}_1$ where $\mathcal{F}_1$ is the family of functions with alternation at most 1.*

This gives us a motivation to study VC-dimension of monotone read-once functions which could potentially be applied to improve the bound of alt-k family. Let $\mathcal{R} = \{f : \{0,1\}^n \to \{0,1\} \mid f \text{ is monotone read-once}\}$. We prove the following in the Appendix A.4.

Proposition 3. $n \leq \mathsf{VC}(\mathcal{R}) \leq O(n \log n)$

4 Application to VC-dimension of k-fold Union

In this section we show VC-dimension bound for a non-geometric family which is a k-union. We also remark in the end that even if we restrict our family to have only disjoint union of k functions, we obtain a VC-dimension bound of $\Omega(dk)$.

Lemma 5. *Let $\mathcal{F}_{2k} = \{f : \{0,1\}^n \to \{0,1\} \mid \mathsf{alt}(f) \leq 2k\}$. Then this family is same[1] as $\mathcal{G} = \left\{\bigcup_{i=1}^{k} g_i \mid g_i : \{0,1\}^n \to \{0,1\}, \mathsf{alt}(g_i) \leq 2\right\}$*

Proof. Let $f \in \mathcal{F}_{2k}$. Due to alternation characterization described in Lemma 2 we have, $f = \bigoplus_{i=1}^{2k} f_i = \bigoplus_{i=1}^{k} (\neg\, f_{2i-1} \wedge f_{2i}) \vee (f_{2i-1} \wedge \neg\, f_{2i})$. It can be observed from the construction of [2], that $f_i \to f_{i+1}$. Now using this fact we obtain $f = \bigvee_{i=1}^{k} (\neg\, f_{2i-1} \wedge f_{2i}) = \bigvee_{i=1}^{k} g_i$ such that $\mathsf{alt}(g_i) \leq 2$. In fact something stronger is true - we argue that f is the disjoint union of k sets. See Appendix A.6.

[1] Under the interpretation of the sets in the system as $f^{-1}(1)$ for Boolean functions, $\vee$ and $\cup$ are used interchangably.

Now we need to argue the reverse direction. We have a Boolean function $g : \{0,1\}^n \to \{0,1\}$, $g = \bigvee g_i$ where $g_i : \{0,1\}^n \to \{0,1\}$ and $\mathsf{alt}(g_i) \leq 2$. We need to show that $\mathsf{alt}(g) \leq 2k$. We use the property (see Appendix A.5) that $\mathsf{alt}(g_1 \vee g_2) \leq \mathsf{alt}(g_1) + \mathsf{alt}(g_2)$ iteratively to conclude that $\mathsf{alt}(g) \leq 2k$.

Theorem 4. *Let* $\mathcal{F}_2 = \{f : \{0,1\}^n \to \{0,1\} \mid \mathsf{alt}(f) \leq 2\}$. *Consider the family* $\mathcal{F}^{k\cup} = \left\{\bigcup_{i=1}^{k} f_i \mid f_i \in \mathcal{F}_2\right\}$. *For* $k \leq \Theta(\sqrt{n})$, *we have* $\mathsf{VC}(\mathcal{F}^{k\cup}) = \Theta(k\binom{n}{n/2})$.

Proof. From Lemma 5, we have that the family $\mathcal{F}^{k\cup}$ can alternately be represented as parity-composition of a family of monotone Boolean functions. So $\mathcal{F}^{k\cup} = \mathcal{F}_{2k}$. We conclude using Theorem 2 that $\mathsf{VC}(\mathcal{F}^{k\cup}) \leq O(k\binom{n}{n/2})$. From Theorem 3 we have: $\mathsf{VC}(\mathcal{F}^{k\cup}) \geq \sum_{i=n/2-k}^{n/2+k} \binom{n}{i}$. Now we use the following bounds due to [13] $\sum_{i=n/2-k}^{n/2+k} \binom{n}{i} = \binom{n}{n/2} + 2\sum_{i=1}^{k} \binom{n}{n/2+i}$. When i is $o(n^{2/3})$, each summand in the second term is equal to $\binom{n}{n/2} e^{-\frac{2i^2}{n} + O(i^3 n^{-2})}$. Hence for $i \leq c\sqrt{n}$, we obtain $\binom{n}{n/2+i} = c_1\binom{n}{n/2}$, $c_1 > 0$. Hence we obtain the lower bound as $\Omega(k\binom{n}{n/2})$.

References

1. J. Barnett, H.E. Correia, P. Johnson, M. Laughlin, and K. Wilson. Darwin meets graph theory on a strange planet: Counting full n-ary trees with labeled leafs. 2015.
2. E. Blais, C. Canonne, I. Oliveira, R. Servedio, and L. Tan. Learning Circuits with few Negations. In *Conference on Approximation, Randomization, and Combinatorial Optimization (APPROX/RANDOM 2015)*, volume 40, pages 512–527, 2015.
3. A. Blumer, A. Ehrenfeucht, D. Haussler, and Manfred K. Warmuth. Learnability and the vapnik-chervonenkis dimension. *J. ACM*, 36(4):929–965, October 1989.
4. A. D Procaccia and J. Rosenschein. Exact vc-dimension of monotone formulas. *Neural Information Processing -Letters and Reviews*, 10, 08 2006.
5. A. Ehrenfeucht, D. Haussler, M. Kearns, and L. Valiant. A general lower bound on the number of examples needed for learning. *Information and Computation*, 82(3):247 – 261, 1989.
6. David Eisenstat and Dana Angluin. The vc dimension of k-fold union. *Inf. Process. Lett.*, 101(5):181–184, March 2007.
7. David Haussler and Emo Welzl. e-nets and simplex range queries. *Discrete & Computational Geometry*, 2(2):127–151, 1987.
8. D. Kleitman and G. Markowsky. On dedekind's problem: The number of isotone boolean functions. ii. *Trans. of the American Math. Soc.*, 213:373–390, 1975.
9. Tamás Mészáros and Lajos Rónyai. Shattering-extremal set systems of small vc-dimension. *International Scholarly Research Notices*, 2013, 2013.
10. Thomas Natschläger and Michael Schmitt. Exact vc-dimension of boolean monomials. *Information Processing Letters*, 59(1):19 – 20, 1996.
11. R. L. Rivest. Learning decision lists. *Machine Learning*, 2(3):229–246, Nov 1987.
12. E.D. Sontag. Vc dimension of neural networks. In C.M. Bishop, editor, *Neural Networks and Machine Learning*, pages 69–95. Springer, Berlin, 1998.
13. J.H. Spencer and L. Florescu. *Asymptopia*. Student mathematical library. American Mathematical Society.
14. L. G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, November 1984.

A Appendix

A.1 Proof of Proposition 1

Proof. To show the lower bound, we demonstrate how to shatter the following set : $S = \{x \in \{0,1\}^n \mid \sum_{i=1}^{n} x_i = n/2\}$ where x_i represents i^{th} bit of x and $|S| = \binom{n}{n/2}$. $\forall S' \subseteq S$, F such that $F \cap S = S'$ is given by the characteristic function $f = \bigvee_{z \in S'} \bigwedge_{z_i=1} x_i$. Observe that f is $\frac{n}{2}$-*slice* function. To show the upper bound, observe that every slice function is a monotone function and hence slice family is a subset of monotone family. Hence $\mathsf{VC}(\mathcal{M}^*) \leq \mathsf{VC}(\mathcal{M}) \leq \binom{n}{n/2}$.

A.2 Proof of Proposition 2

Proof. For a family $\mathcal{F}$, $Sh(\mathcal{F})$ denotes the set of shattered sets. Then,

Proposition 4. $|Sh(\mathcal{M})| = |\mathcal{M}|$.

Proof. For monotone family, we can only shatter sets which forms an antichain and there is one to one correspondence between an antichain and a monotone Boolean function(see [8]). Hence all the shattered sets must be an antichain.
Let us suppose $f : \{0,1\}^n \to \{0,1\}$ be a monotone Boolean function and $A \subseteq \{0,1\}^n$ be the antichain corresponding to f. Here is a way to obtain A. First obtain minimal monotone-DNF formula for f. Each term in the formula will correspond to an element in A. For each positive literal appearing in the term, corresponding bit will be set to 1 and rest as 0. Observe that, in this construction if the A is not an antichain then the monotone-DNF formula is not minimal. To see the construction, for a term $x_1 \wedge x_2$ the $w \in A$ will be $110\ldots0$. Lets argue that this is a *one-one* relation. Suppose $f_i \neq f_j$ then their minimal monotone-DNF formula will differ in, at least 1 term and hence their corresponding antichain A_i, A_j will also differ. Using a similar reverse construction, we can have a monotone-DNF for a monotone Boolean function for each antichain.

Proposition 5. *Let* $\mathcal{F}_1 = \{f : \{0,1\}^n \to \{0,1\} \mid \mathsf{alt}(f) \leq 1\}$. *Then* $|\mathcal{F}_1| = |Sh(\mathcal{F}_1)|$.

Proof. Observe that $|\mathcal{F}_1| = 2|\mathcal{M}|$, where $|\mathcal{M}|$ is size of monotone family. Also observe that $\mathcal{F}_1$ shatters all the antichains of the Boolean hypercube. We also note that this family shatters another kind of set shown in Theorem 1 i.e. an antichain union with one extra element which has neighbours in the antichain ($S \cup \{a\}$ where S is an antichain). Hence we obtain some more shattered sets by combining each antichain with one comparable element. Thus shattering $|\mathcal{M}|$ more sets. These are the only sets which can be shattered by this family. Hence this family is *s-extremal*.

A.3 Proof of Lemma 4

Proof. Let $f : \{0,1\}^n \to \{0,1\}$ such that $\mathsf{alt}(f) \leq k$. Due to the alternation characterization in Lemma 2, we have $f = \bigoplus_{i=1}^{k} f_i = \bigoplus_{i=1}^{k/2} (\neg\, f_{2i-1} \wedge f_{2i}) \vee (f_{2i-1} \wedge \neg\, f_{2i})$. It can be observed from the construction in [2] that $f_i \to f_{i+1}$. Now using this fact we obtain $f = \bigvee_{i=1}^{k/2} (\neg\, f_{2i-1} \wedge f_{2i})$ if k is even and $f = f_k \wedge \bigwedge_{i=1}^{\frac{k-1}{2}} (f_{2i-1} \vee \neg f_{2i})$ if k is odd. Observe that in the formula each f_i appears only once and no negation in the formula is involved. Since f_i's are monotone, hence $f_i, \neg f_i \in \mathcal{F}_1$.

A.4 Proof of Proposition 3

Proof. To see the lower bound a subclass of this family was studied in [10] known as Monotone monomial whose VC-dimension is n. To see the upper bound we will count all possible monotone read-once formulas on n-variables.

Claim. $|\mathcal{R}| = n!\binom{2n}{n}$

Proof. A formula is rooted binary tree such that all the internal nodes are $\wedge, \vee$ or$\neg$. The number of full binary trees with $n+1$ labeled leaves is $\prod_{i=1}^{n}(2i-1)$ [1]. But in our case we also have to consider the two choices $\{\wedge, \vee\}$ for each internal nodes. Hence we get $|\mathcal{R}| = 2^n . \prod_{i=1}^{n}(2i-1)$ which gives $|\mathcal{R}| = n!\binom{2n}{n}$.

Using the above claim we obtain, $\mathsf{VC}(\mathcal{R}) \leq \log|\mathcal{R}| = O(n \log n)$.

A.5 Alternation under Disjunction

Proof. Claim. Let $g_1 : \{0,1\}^n \to \{0,1\}$ and $g_2 : \{0,1\}^n \to \{0,1\}$ with $\mathsf{alt}(g_1) = k_1$ and $\mathsf{alt}(g_2) = k_2$. Then $\mathsf{alt}(g_1 \vee g_2) \leq k_1 + k_2$.

Proof. Suppose on the Boolean hypercube, $x, y \in \{0,1\}^n, x \preceq y$ and they differ exactly at one bit. Then we will say (x, y) forms an edge. Now we say (x, y) is monochromatic when function $f(x) = f(y)$ and bichromatic when $f(x) \neq f(y)$. Clearly for any function $f : \{0,1\}^n \to \{0,1\}$ with $\mathsf{alt}(f) \leq k$, there is at most k-bichromatic edges on any chain and vice versa. Now we count the number of bichromatic edges on any chain of $g_1 \vee g_2$. Suppose on some chain in g_1, there is a monochromatic edge with $g_1(x) = g_1(y) = 1$ then this edge will stay monochromatic even after OR with g_2. But if the edge (x, y) is monochromatic and $g_1(x) = g_1(y) = 0$ then they can become bichromatic if they are ORed with a bichromatic edge. Hence the number of bichromatic edges on any chain of $g_1 \vee g_2$ can at most be increased by number of bichromatic edges on any chain in g_2. Hence $\mathsf{alt}(g_1 \vee g_2) \leq k_1 + k_2$.

A.6 Disjointness of the Decomposition

We now show how the $\mathcal{F}_{2k}$ family is also a k-disjoint union of alt-2 family.

Claim. Let $f : \{0,1\}^n \to \{0,1\}$ and $\mathsf{alt}(f) \leq 2k$. Then $f = \uplus_{i=1}^{k} g_i$ i.e. disjoint union of k sets.

Proof. From Lemma 4 we have, $f = \bigvee_{i=1}^{k}(\neg f_{2i-1} \wedge f_{2i})$. Let $g_i = (\neg f_{2i-1} \wedge f_{2i})$. We will show that g_1 and g_2 are disjoint which will be applicable for all $i, j \in [k]$. Suppose for $x \in \{0,1\}^n$, $g_1(x) = 1 \Rightarrow f_1(x) = 0, f_2(x) = 1$. To show a contradiction, suppose $g_2(x) = 1 \Rightarrow f_3(x) = 0, f_4(x) = 1$. We also have $f_i \to f_{i+1}$. We obtained $f_2(x) = 1$ but $f_3(x) = 0$ which contradicts $f_i \to f_{i+1}$. Hence g_1 and g_2 are disjoint. The same argument holds for any g_i and g_j, $i \neq j$.