On the Probabilistic Degrees of Symmetric Boolean Functions

Srikanth Srinivasan
Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India
srikanth@math.iitb.ac.in

Utkarsh Tripathi
Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India
utkarshtripathi.math@gmail.com

S. Venkitesh
Department of Mathematics, Indian Institute of Technology Bombay, Mumbai, India
venkitesh.mail@gmail.com

Abstract

The probabilistic degree of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is defined to be the smallest d such that there is a random polynomial $\mathbf{P}$ of degree at most d that agrees with f at each point with high probability. Introduced by Razborov (1987), upper and lower bounds on probabilistic degrees of Boolean functions – specifically symmetric Boolean functions – have been used to prove explicit lower bounds, design pseudorandom generators, and devise algorithms for combinatorial problems.

In this paper, we characterize the probabilistic degrees of all symmetric Boolean functions up to polylogarithmic factors over all fields of fixed characteristic (positive or zero).

2012 ACM Subject Classification Theory of computation → Computational complexity and cryptography

Keywords and phrases Symmetric Boolean function, probabilistic degree

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2019.28

Related Version A full version of the paper is available on [21], `https://eccc.weizmann.ac.il/report/2019/138/download`.

Funding *Srikanth Srinivasan*: Supported by MATRICS grant MTR/2017/000958 awarded by SERB, Government of India.
Utkarsh Tripathi: Supported by the Ph.D. Scholarship of NBHM, DAE, Government of India.
S. Venkitesh: Supported by the Senior Research Fellowship of HRDG, CSIR, Government of India.

1 Introduction

Studying the combinatorial and computational properties of Boolean functions by representing them using multivariate polynomials (over some field $\mathbb{F}$) is an oft-used technique in Theoretical Computer Science. Such investigations into the complexity of Boolean functions have led to many important advances in the area (see, e.g. [2, 14, 23] for a large list of such results).

An "obvious" way of representing a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is via a multilinear polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$ such that $P(a) = f(a)$ for all $a \in \{0,1\}^n$. While such a representation has the advantage of being unique, understanding the computational complexity of f sometimes requires us to understand polynomial representations where we allow some notion of error in the representation. Here again, many kinds of representations have been studied, but we concentrate here on the notion of *Probabilistic degree* of a Boolean function, introduced by Razborov [16]. It is defined as follows.

39th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2019).
Editors: Arkadev Chattopadhyay and Paul Gastin; Article No. 28; pp. 28:1–28:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ **Definition 1** (Probabilistic polynomial and Probabilistic degree). *Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and an $\varepsilon > 0$, an* ε-error probabilistic polynomial *for f is a random polynomial $\mathbf{P}$ (with some distribution having finite support) over $\mathbb{F}[x_1, \ldots, x_n]$ such that for each $a \in \{0,1\}^n$,*

$$\Pr_{\mathbf{P}}[\mathbf{P}(a) \neq f(a)] \leq \varepsilon.$$

We say that the degree of $\mathbf{P}$, denoted $\deg(\mathbf{P})$*, is at most d if the probability distribution defining $\mathbf{P}$ is supported on polynomials of degree at most d. Finally, we define the* ε-error probabilistic degree *of f, denoted* $\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f)$*, to be the least d such that f has an ε-error probabilistic polynomial of degree at most d.*

When the field $\mathbb{F}$ is clear from context, we use $\mathrm{pdeg}_\varepsilon(f)$ *instead of* $\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f)$.

Intuitively, if we think of multivariate polynomials as algorithms and degree as a notion of efficiency, then a low-degree probabilistic polynomial for a Boolean function f is an efficient *randomized* algorithm for f.

The study of the probabilistic degree itself is by now a classical topic, and has had important repercussions for other problems. We list three such examples below, referring the reader to the papers for definitions and exact statements of the results.

- Razborov [16] showed strong upper bounds on the probabilistic degree of the OR function over fields of (fixed) positive characteristic. Along with lower bounds on the probabilistic degree of some symmetric Boolean functions,[1] this led to the first lower bounds for the Boolean circuit class $\mathrm{AC}^0[p]$, for prime p [16, 17, 19].
- Tarui [22] and Beigel, Reingold and Spielman [3] showed upper bounds on the probabilistic degree of the OR function over any characteristic (and in particular over the reals). This leads to probabilistic degree upper bounds for the circuit class AC^0, which was used by Braverman [5] to resolve a long-standing open problem of Linial and Nisan [11] regarding pseudorandom generators for AC^0.
- Alman and Williams [1] showed that for constant error, the probabilistic degree of any symmetric Boolean function is at most $O(\sqrt{n})$, and used this to obtain the first subquadratic algorithm for an offline version of the Nearest Neighbour problem in the Hamming metric.

In all the above results, it was important to understand the probabilistic degree of a certain class of symmetric Boolean functions. However, the problem of *characterizing* the probabilistic degree of symmetric Boolean functions in general does not seem to have been considered. This is somewhat surprising, since this problem has been considered in a variety of other computational models, such as AC^0 circuits of polynomial size [8, 6], $\mathrm{AC}^0[p]$ circuits of quasipolynomial size [12], Approximate degree[2] [15] and Perceptrons[3] of quasipolynomial size [24].

[1] Recall that a *symmetric* Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a function such that $f(x)$ depends only on the Hamming weight of x. Examples include the threshold functions, Parity (counting modulo 2), etc.

[2] A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to have approximate degree at most d if there is a degree d polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ such that at each $a \in \{0,1\}^n$, $|f(a) - P(a)| \leq 1/4$.

[3] Perceptrons are depth-2 circuits with a Majority gate as the output gate with AND and OR gates feeding into it.

Our result. In this paper, we give an almost-complete understanding of the probabilistic degrees of all symmetric Boolean functions over all fields of fixed positive characteristic and characteristic 0. For each Boolean function f on n variables, our upper bounds and lower bounds on $\mathrm{pdeg}(f)$ are separated only by polylogarithmic factors in n.

We now introduce some notation and give a formal statement of our result. We shall use the notation $[a, b]$ to denote an interval in $\mathbb{R}$ as well as an interval in $\mathbb{Z}$; the distinction will be clear from the context. Throughout, fix some field $\mathbb{F}$ of characteristic p which is either a fixed positive constant or 0. Let n be a growing integer parameter which will always be the number of input variables. We use $s\mathcal{B}_n$ to denote the set of all symmetric Boolean functions on n variables. Note that each symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$ is uniquely specified by a string $\mathrm{Spec}\, f : [0, n] \to \{0,1\}$, which we call the *Spectrum* of f, in the sense that for any $a \in \{0,1\}^n$, we have

$$f(a) = \mathrm{Spec}\, f(|a|).$$

Given a $f \in s\mathcal{B}_n$, we define the *period of* f, denoted $\mathrm{per}(f)$, to be the smallest positive integer b such that $\mathrm{Spec}\, f(i) = \mathrm{Spec}\, f(i + b)$ for all $i \in [0, n - b]$. We say f is k*-bounded* if $\mathrm{Spec}\, f$ is constant on the interval $[k, n - k]$; let $B(f)$ denote the smallest k such that f is k-bounded.

Standard decomposition of a symmetric Boolean function [12]. Fix any $f \in s\mathcal{B}_n$. Among all symmetric Boolean functions $f' \in s\mathcal{B}_n$ such that $\mathrm{Spec}\, f'(i) = \mathrm{Spec}\, f(i)$ for all $i \in [\lceil n/3\rceil, \lfloor 2n/3\rfloor]$, we choose a function g such that $\mathrm{per}(g)$ is as small as possible. We call g the *periodic part* of f. Define $h \in s\mathcal{B}_n$ by $h = f \oplus g$. We call h the *bounded part* of f.

We will refer to the pair (g, h) as a *standard decomposition* of the function f. Note that we have $f = g \oplus h$.

▶ **Observation 2.** *Let $f \in s\mathcal{B}_n$ and let (g, h) be a standard decomposition of f. Then,* $\mathrm{per}(g) \leq \lfloor n/3\rfloor$ *and* $B(h) \leq \lceil n/3\rceil$.

In this paper, we prove the following upper and lower bounds for the probabilistic degrees of symmetric Boolean functions. While the most important setting for understanding the probabilistic degree is the setting of constant error (i.e. $\varepsilon = \Omega(1)$), we state the upper bound results for arbitrary $\varepsilon > 0$ since the inductive construction naturally gives rise to this stronger statement.

▶ **Theorem 3** (Upper bounds on probabilistic degree). *Let $\mathbb{F}$ be a field of constant characteristic p (possibly 0) and $n \in \mathbb{N}$ be a growing parameter. Let $f \in s\mathcal{B}_n$ be arbitrary and let (g, h) be a standard decomposition of f. Then we have the following for any* $\varepsilon > 0$.

1. *If* $\mathrm{per}(g) = 1$, *then* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(g) = 0, .$
   If $\mathrm{per}(g)$ *is a power of p, then* $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(g) \leq \mathrm{per}(g)$, *[12]*
   (Note that $\mathrm{per}(g)$ *cannot be a power of p if $p = 0$.)*
2. $\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(h) = \widetilde{O}(\sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon))$,
3. $$\mathrm{pdeg}^{\mathbb{F}}_{\varepsilon}(f) = \begin{cases} O(\sqrt{n\log(1/\varepsilon)}) & \text{if } \mathrm{per}(g) > 1 \text{ and not a power of } p, \text{ [1]} \\ \widetilde{O}(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g) + & \text{otherwise.} \\ \quad \sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon)\}) & \end{cases}$$

When p is positive, we can replaced the $\widetilde{O}(\cdot)$ with $O(\cdot)$ in all the above bounds.

We obtain almost (up to polylogarithmic factors) matching lower bounds for all symmetric Boolean functions over all fields.

▶ **Theorem 4** (Lower bounds on probabilistic degree). *Let $\mathbb{F}$ be a field of constant characteristic p (possibly 0) and $n \in \mathbb{N}$ be a growing parameter. Let $f \in s\mathcal{B}_n$ be arbitrary and let (g, h) be a standard decomposition of f. Then for any* constant $\varepsilon \leq 1/3$, *we have*

1. $\mathrm{pdeg}_{\varepsilon}^{\mathbb{F}}(g) = \widetilde{\Omega}(\sqrt{n})$ *if* $\mathrm{per}(g) > 1$ *and is not a power of p and* $\widetilde{\Omega}(\min\{\sqrt{n}, \mathrm{per}(g)\})$ *otherwise.*
2. $\mathrm{pdeg}_{\varepsilon}^{\mathbb{F}}(h) = \widetilde{\Omega}(\sqrt{B(h)})$,
3. $\mathrm{pdeg}_{\varepsilon}^{\mathbb{F}}(f) = \begin{cases} \widetilde{\Omega}(\sqrt{n}) & \textit{if } \mathrm{per}(g) > 1 \textit{ and not a power of } p, \\ \widetilde{\Omega}(\min\{\sqrt{n}, \mathrm{per}(g) + \sqrt{B(h)}\}) & \textit{otherwise.} \end{cases}$

where the $\widetilde{\Omega}(\cdot)$ *hides* $\mathrm{poly}(\log n)$ *factors.*

▶ Remark 5. A natural open question following our results is to remove the polylogarithmic factors separating our upper and lower bounds. We remark that in characteristic 0, such gaps exist even for the very simple OR function despite much effort [13, 9, 4]. Over positive characteristic, there is no obvious barrier, but our techniques fall short of proving tight lower bounds for natural families of functions such as the Exact Threshold functions (defined below).

Many proofs are omitted for lack of space. They appear in the full version of the paper.

1.1 Proof Outline

For the outline below, we assume that the field is of fixed positive characteristic p.

Upper bounds. Given a symmetric Boolean function f on n variables with standard decomposition (g, h), it is easy to check that $\mathrm{pdeg}_{\varepsilon}(f) = O(\mathrm{pdeg}_{\varepsilon}(g) + \mathrm{pdeg}_{\varepsilon}(h))$. So it suffices to upper bound the probabilistic degrees of periodic and bounded functions respectively.

For periodic functions g with period a power of p, Lu [12] showed that the *exact* degree of the Boolean functions is at most $\mathrm{per}(g)$. If the period is not a power of p, then we use the upper bound of Alman and Williams [1] that holds for all symmetric Boolean functions (as we show below, this is nearly the best that is possible).

For a t-constant function h (defined in Section 3), we use the observation that any t-constant function is essentially a linear combination of the threshold functions $\mathrm{Thr}_n^0, \ldots, \mathrm{Thr}_n^t$ (see Section 2 for the definition) and so it suffices to construct probabilistic polynomials for Thr_n^i, for $i \in [0, t]$.[4]

Our main technical upper bound is a new probabilistic degree upper bound of $O(\sqrt{t \log(1/\varepsilon)} + \log(1/\varepsilon))$ for any threshold function Thr_n^t. This upper bound interpolates smoothly between a classical upper bound of $O(\log(1/\varepsilon))$ due to Razborov [16] for $t = 1$ and a recent result of Alman and Williams [1] that yields $O(\sqrt{n \log(1/\varepsilon)})$ for $t = \Omega(n)$.

The proof of our upper bound is based on the beautiful inductive construction of Alman and Williams [1] which gives their above-mentioned result. The key difference between our proof and the proof of [1] is that we need to handle separately the case when the error $\varepsilon \leq 2^{-\Omega(t)}$.[5] In [1], this is a trivial case since any function on n Boolean variables has an exact polynomial of degree n which is at most $O(\sqrt{n \log(1/\varepsilon)})$ when $\varepsilon \leq 2^{-\Omega(n)}$. In our setting, the correct bound in this case is $O(\log(1/\varepsilon))$, which is non-obvious. We obtain this bound by a suitable modification of Razborov's technique (for $t = 1$) to handle larger thresholds.

[4] We actually need to construct probabilistic polynomials for all the threshold functions simultaneously. We ignore this point in this high-level outline.

[5] This case comes up naturally in the inductive construction, even if one is ultimately only interested in the case when ε is a constant.

Lower bounds. Here, our proof closely follows a result of Lu [12], who gave a characterization of symmetric Boolean functions that have quasipolynomial-sized $\mathrm{AC}^0[p]$ circuits.[6] To show circuit lower bounds for a symmetric Boolean function h, Lu showed how to convert a circuit C computing h to a circuit C' computing either the Majority or a MOD_q function (where q and p are relatively prime). Since both of these are known to be hard for $\mathrm{AC}^0[p]$ [16, 17], we get the lower bound.

We show how to use Lu's reductions (and variants thereof) but in the setting of probabilistic polynomials. This works because

- We also have strong probabilistic degree lower bounds for the Majority and MOD_q functions (in fact, this is the source of the $\mathrm{AC}^0[p]$ lower bound).
- Lu's constructions of the hard functions from h (and our variants) involve taking ANDs and ORs of a few copies of (restrictions of) h. This also gives a reduction from the hard functions to h in the setting of probabilistic degree, since ANDs and ORs are known to have small probabilistic degree [16].

With these observations in place, the proof reduces to a careful case analysis to get the correct lower bound in each case. Interestingly, while it is not clear whether these ideas give a tight understanding of the $\mathrm{AC}^0[p]$-circuit complexity of symmetric Boolean functions, they do give nearly tight (up to log factors) lower bounds for probabilistic degree.

2 Preliminaries

Some Boolean functions. Fix some positive $n \in \mathbb{N}$. The *Majority* function Maj_n on n Boolean variables accepts exactly the inputs of Hamming weight at least $n/2$. For $t \in [0, n]$, the *Threshold* function Thr_n^t accepts exactly the inputs of Hamming weight at least t; and similarly, the *Exact Threshold* function EThr_n^t accepts exactly the inputs of Hamming weight exactly t. Finally, for $b \in [2, n]$ and $i \in [0, b-1]$, the function $\mathrm{MOD}_n^{b,i}$ accepts exactly those inputs a such that $|a| \equiv i \pmod{b}$. In the special case that $i = 0$, we also use MOD_n^b.

▶ **Fact 6.** *We have the following simple facts about probabilistic degrees. Let $\mathbb{F}$ be any field.*

1. *(Error reduction [9]) For any $\delta < \varepsilon \le 1/3$ and any Boolean function f, if $\mathbf{P}$ is an ε-error probabilistic polynomial for f, then $\mathbf{Q} = M(\mathbf{P}_1, \ldots, \mathbf{P}_\ell)$ is a δ-error probabilistic polynomial for f where M is the exact multilinear polynomial for Maj_ℓ and $\mathbf{P}_1, \ldots, \mathbf{P}_\ell$ are independent copies of $\mathbf{P}$. In particular, we have $\mathrm{pdeg}_\delta^{\mathbb{F}}(f) \le \mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f) \cdot O(\log(1/\delta)/\log(1/\varepsilon))$.*
2. *(Composition) For any Boolean function f on k variables and any Boolean functions $g_1, \ldots, g_k$ on a common set of m variables, let h denote the natural composed function $f(g_1, \ldots, g_k)$ on m variables. Then, for any $\varepsilon, \delta > 0$, we have $\mathrm{pdeg}_{\varepsilon+k\delta}^{\mathbb{F}}(h) \le \mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f) \cdot \max_{i\in[k]} \mathrm{pdeg}_\delta^{\mathbb{F}}(g_i)$.*
3. *(Sum) Assume that $f, g_1, \ldots, g_k$ are all Boolean functions on a common set of m variables such that $f = \sum_{i\in[k]} g_i$. Then, for any $\delta > 0$, we have $\mathrm{pdeg}_{k\delta}^{\mathbb{F}}(f) \le \max_{i\in[k]} \mathrm{pdeg}_\delta^{\mathbb{F}}(g_i)$.*

2.1 Some previous results on probabilistic degree

The following upper bounds on probabilistic degrees of OR and AND functions were proved by Razborov [16] in the case of positive characteristic and Tarui [22] and Beigel, Reingold and Spielman [3] in the general case.

[6] Recall that an $\mathrm{AC}^0[p]$ circuit is a constant-depth circuit made up of gates that can compute the Boolean functions AND, OR, NOT and MOD_p (defined below).

▶ **Lemma 7** (Razborov's upper bound on probabilistic degrees of OR and AND)**.** *Let $\mathbb{F}$ be a field of characteristic p. For $p > 0$, we have*

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{OR}_n) = \mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{AND}_n) = O(p \log(1/\varepsilon)). \tag{1}$$

For any p, we have

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{OR}_n) = \mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{AND}_n) = O(\log n \cdot \log(1/\varepsilon)). \tag{2}$$

We now recall two probabilistic degree lower bounds due to Smolensky [18, 20], building on the work of Razborov [16].

▶ **Lemma 8** (Smolensky's lower bound for close-to-Majority functions)**.** *For any field $\mathbb{F}$, any $\varepsilon \in (1/2^n, 1/5)$, and any Boolean function g on n variables that agrees with Maj_n on a $1-\varepsilon$ fraction of its inputs, we have*

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(g) = \Omega(\sqrt{n \log(1/\varepsilon)}).$$

▶ **Lemma 9** (Smolensky's lower bound for MOD functions)**.** *For $2 \leq b \leq n/2$, any $\mathbb{F}$ such that char$(\mathbb{F}) = p$ is coprime to b, any $\varepsilon \in (1/2^n, 1/(3b))$, there exists an $i \in [0, b-1]$ such that*

$$\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{MOD}_n^{b,i}) = \Omega(\sqrt{n \log(1/b\varepsilon)}).$$

▶ Remark 10. From the above lemma, it also easily follows that if $b \leq n/4$, then for *every* $i \in [0, b-1]$, we have $\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(\mathrm{MOD}_n^{b,i}) = \Omega(\sqrt{n \log(1/b\varepsilon)})$. This is the usual form in which Smolensky's lower bound is stated. The above form is slightly more useful to us because it holds for b up to $n/2$.

We will also need the following result of Alman and Williams [1].

▶ **Lemma 11.** *Let $\mathbb{F}$ be any field. For any $n \geq 1, \varepsilon > 0$ and $f \in s\mathcal{B}_n$, $\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f) = O(\sqrt{n \log(1/\varepsilon)})$.*

2.2 A string lemma

Given a function $w : I \to \{0,1\}$ where $I \subseteq \mathbb{N}$ is an interval, we think of w as a string from the set $\{0,1\}^{|I|}$ in the natural way. For an interval $J \subseteq I$, we denote by $w|_J$ the substring of w obtained by restriction to J.

The following simple lemma can be found, e.g. as a consequence of [10, Chapter I, Section 2, Theorem 1].

▶ **Lemma 12.** *Let $w \in \{0,1\}^+$ be any non-empty string and $u, v \in \{0,1\}^+$ such that $w = uv = vu$. Then there exists a string $z \in \{0,1\}^+$ such that w is a power of z (i.e. $w = z^k$ for some $k \geq 2$).*

▶ **Corollary 13.** *Let $g \in s\mathcal{B}_n$ be arbitrary with $\mathrm{per}(g) = b > 1$. Then for all $i, j \in [0, n-b+1]$ such that $i \not\equiv j \pmod b$, we have $\mathrm{Spec}\, g|_{[i,i+b-1]} \neq \mathrm{Spec}\, g|_{[j,j+b-1]}$.*

Proof. Suppose $\mathrm{Spec}\, g|_{[i,i+b-1]} = \mathrm{Spec}\, g|_{[j,j+b-1]}$ for some $i \not\equiv j \pmod b$. Assume without loss of generality that $i < j < i + b$. Let $u = \mathrm{Spec}\, g|_{[i,j-1]}, v = \mathrm{Spec}\, g|_{[j,i+b-1]}, w = \mathrm{Spec}\, g|_{[i+b,j+b-1]}$. Then $u = w$ and the assumption $uv = vw$ implies $uv = vu$. By Lemma 12, there exists a string z such that $uv = z^k$ for $k \geq 2$ and therefore $\mathrm{per}(g) < b$. This contradicts our assumption on b. ◀

▶ **Lemma 14.** *Let $n \in \mathbb{N}$ be a growing parameter and let $f \in s\mathcal{B}_n$ with periodic part g. For any $1 \leq b \leq \lfloor n/3 \rfloor$, either $\mathrm{per}(g) \leq b$ or for all distinct $i,j \in [\lceil n/3 \rceil - b, \lceil n/3 \rceil]$, $\mathrm{Spec}\, f|_{[i,i+(\lceil n/3 \rceil + b)]} \neq \mathrm{Spec}\, f|_{[j,j+(\lceil n/3 \rceil + b)]}$.*

Proof. W.l.o.g. say $i < j$. Assume $\mathrm{per}(g) > b$ (otherwise, we are done trivially). Then, for any $b' \leq b$, it follows that there is an $k \in [\lceil n/3 \rceil, \lfloor 2n/3 \rfloor - b']$ such that $\mathrm{Spec}\, f(k) \neq \mathrm{Spec}\, f(k+b')$. In particular, we see that $\mathrm{Spec}\, f|_{[i,i+(\lceil n/3 \rceil + b)]} \neq \mathrm{Spec}\, f|_{[i+b',i+b'+(\lceil n/3 \rceil + b)]}$. Fixing $b' = j - i$ yields the result. ◀

3 Upper bounds

In this section, we will first prove upper bounds on the probabilistic degree of a smaller class of symmetric Boolean functions, called *t-constant functions*, and then use it to prove Theorem 3.

3.1 Upper bound on probabilistic degree of t-constant functions

▶ **Definition 15** (t-constant function)**.** *For any positive $n \in \mathbb{N}$ and $t \in [0,n]$, a Boolean function $f \in s\mathcal{B}_n$ is said to be* t-constant *if $f|_{\{x:|x|\geq t\}}$ is a constant, that is, $\mathrm{Spec}\, f|_{[t,n]}$ is a constant.*

The following observation is immediate.

▶ **Observation 16.** *A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is t-constant if and only if $f = \sum_{j=0}^{t} a_j \mathrm{Thr}_n^j$, for some $a_0, \ldots, a_t \in \{-1,0,1\}$. In other words, f is t-constant if and only if there exists a linear polynomial $g(Y_0, \ldots, Y_t) = a_0 Y_0 + \cdots + a_t Y_t \in \mathbb{F}[Y_0, \ldots, Y_t]$ with $a_j \in \{-1,0,1\}$, $j \in [0,t]$ such that $f = g(\mathrm{Thr}_n^0, \ldots, \mathrm{Thr}_n^t)$.*

We will prove an upper bound on the probabilistic degree of t-constant Boolean functions. For this, we first generalize the notion of probabilistic polynomial and probabilistic degree to a *tuple* of Boolean functions. This generalization was implicit in [1].

▶ **Definition 17** (Probabilistic poly-tuple and probabilistic degree)**.** *Let $f = (f_1, \ldots, f_m) : \{0,1\}^n \to \{0,1\}^m$ be an m-tuple of Boolean functions and $\varepsilon \in (0,1)$. An ε-error probabilistic poly-tuple for f is a random m-tuple of polynomials $\mathbf{P}$ (with some distribution having finite support) from $\mathbb{F}[X_1, \ldots, X_n]^m$ such that*

$$\Pr_{P \sim \mathbf{P}}[P(x) \neq f(x)] \leq \varepsilon, \quad \textit{for all } x \in \{0,1\}^n.$$

We say that the degree of $\mathbf{P}$ is at most d if $\mathbf{P}$ is supported on m-tuples of polynomials $P = (P_1, \ldots, P_m)$ where each P_i has degree at most d. Finally we define the ε-error probabilistic degree of f, denoted by $\mathrm{pdeg}_\varepsilon^{\mathbb{F}}(f)$, to be the least d such that f has an ε-error probabilistic poly-tuple of degree at most d.

We make a definition for convenience.

▶ **Definition 18** (Threshold tuple)**.** *For positive $n \in \mathbb{N}$, $t \in [0,n]$, an (n,t)-threshold tuple is any tuple of Boolean functions $(\mathrm{Thr}_n^{t_1}, \ldots, \mathrm{Thr}_n^{t_m})$, with $t_1, \ldots, t_m \in [0,t]$ and $\max\{t_1, \ldots, t_m\} \leq t$.*

The main theorem of this subsection is the following.

▶ **Theorem 19.** *For any positive $n \in \mathbb{N}$, $t \in [0,n]$, if T is an (n,t)-threshold tuple and $\varepsilon \in (0,1/3)$, then*

$$\mathrm{pdeg}_\varepsilon(T) = \begin{cases} \widetilde{O}(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathit{char}(\mathbb{F}) = 0, \\ O(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathit{char}(\mathbb{F}) = p > 0. \end{cases}.$$

As a corollary to the above theorem, we get an upper bound for the probabilistic degree of t-constant functions.

▶ **Corollary 20.** *For any t-constant Boolean function $f : \{0,1\}^n \to \{0,1\}$ and $\varepsilon \in (0,1/3)$,*

$$\mathrm{pdeg}_\varepsilon(f) = \begin{cases} \widetilde{O}(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathit{char}(\mathbb{F}) = 0, \\ O(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathit{char}(\mathbb{F}) = p > 0. \end{cases}.$$

Proof. By Observation 16, there exists $g(Y_0,\ldots,Y_t) = a_0Y_0 + \cdots + a_tY_t \in \mathbb{F}[Y_0,\ldots,Y_t]$ with $a_j \in \{-1,0,1\}$, $j \in [0,t]$ such that $f = g(\mathrm{Thr}_n^0,\ldots,\mathrm{Thr}_n^t)$. We note that $\deg g = 1$. So by Theorem 19, we get

$$\mathrm{pdeg}_\varepsilon(f) = \deg g \cdot \mathrm{pdeg}_\varepsilon(\mathrm{Thr}_n^0,\ldots,\mathrm{Thr}_n^t) = \begin{cases} \widetilde{O}(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathrm{char}(\mathbb{F}) = 0, \\ O(\sqrt{t\log(1/\varepsilon)} + \log(1/\varepsilon)), & \mathrm{char}(\mathbb{F}) = p > 0. \end{cases}.$$

◀

Before we prove Theorem 19, we will gather a few results that we require. The following lemma is a particular case of Bernstein's inequality (Theorem 1.4, [7]).

▶ **Lemma 21.** *Let $X_1,\ldots,X_m$ be independent and identically distributed Bernoulli random variables with mean p. Let $X = \sum_{i=1}^m X_i$. Then for any $\theta > 0$,*

$$\Pr\left[|X - mp| > \theta\right] \le 2\exp\left(-\frac{\theta^2}{2mp(1-p) + 2\theta/3}\right).$$

We will also need the following polynomial construction.

▶ **Theorem 22** (Lemma 3.1, [1])**.** *For any symmetric Boolean function $f : \{0,1\}^n \to \{0,1\}$ and integer interval $[a,b] \subseteq [0,n]$, there exists a symmetric multilinear polynomial $EX_{[a,b]}f \in \mathbb{Z}[X_1,\ldots,X_n]$ such that $\deg(EX_{[a,b]}f) \le b - a$ and $\mathrm{Spec}\,(EX_{[a,b]}f)|_{[a,b]} = \mathrm{Spec}\,f|_{[a,b]}$.*

We will now prove Theorem 19.

Proof of Theorem 19. For any $a = (a_1,\ldots,a_k), b = (b_1,\ldots,b_k) \in \mathbb{F}^k$, fix the notation $a * b = (a_1b_1,\ldots,a_kb_k)$. Throughout, the notation $\mathbf{1}$ will denote the constant-1 vector of appropriate length.

For positive characteristic p, we prove that for any positive $n \in \mathbb{N}$, $t \in [0,n]$ and $\varepsilon \in (0,1/100)$, any (n,t)-threshold tuple T has an ε-error probabilistic poly-tuple $\mathbf{T}$ of degree at most $A_p\sqrt{t\log(1/\varepsilon)} + B_p\log(1/\varepsilon)$, for constants $A_p = B_p = 4800000p$ (we make no effort to minimize the constants). For $p = 0$, we prove a similar result with a degree bound of $A_0\log n \cdot \sqrt{t\log(1/\varepsilon)} + B_0\log n \cdot \log(1/\varepsilon)$, for $A_0 = B_0 = 5000000$. This will prove the theorem for $\varepsilon < 1/100$. To prove the theorem for all $\varepsilon \le 1/3$, we use error reduction (Fact 6) and reduce the error to $1/100$ and then apply the result for small error.

The proof is by induction on the parameters n, t and ε. At any stage of the induction, given an (n,t)-threshold tuple with error parameter ε, we construct the required probabilistic poly-tuple by using the probabilistic poly-tuples (guaranteed by inductive hypothesis) for suitable threshold poly-tuples with $n/10$ inputs and error parameter $\varepsilon/4$. Thus the base cases of the induction are as follows.

Base Case: Suppose $n \leq 10$. Let $T = (T_1, \ldots, T_m)$ be an (n,t)-threshold tuple. Let $Q_1, \ldots, Q_m$ be the unique multilinear polynomial representations of $T_1, \ldots, T_m$ respectively. Then $Q = (Q_1, \ldots, Q_m)$ is an ε-error probabilistic poly-tuple for T, for all $\varepsilon \in (0, 1/100)$, with $\deg Q \leq n = 10$.

Base Case: Suppose $\varepsilon \leq 2^{-t/160000}$. Let $T = (T_1, \ldots, T_m) = (\mathrm{Thr}_n^{t_1}, \ldots, \mathrm{Thr}_n^{t_m})$ be any (n,t)-threshold tuple and let $r = 160000 \log(1/\varepsilon)$.

Suppose $n \leq r$. Let $Q_1, \ldots, Q_m$ be the unique multilinear representations of $T_1, \ldots, T_m$ respectively. Then $Q = (Q_1, \ldots, Q_m)$ is an ε-error probabilistic polynomial with $\deg Q \leq n \leq r = \lceil \log(1/\varepsilon) \rceil$. Now suppose $n > r$. Let $P_1 = (\mathrm{EX}_{[0,r]} T_1, \ldots, \mathrm{EX}_{[0,r]} T_m)$. Then $\deg P_1 \leq r$. Choose a uniformly random hash function $\mathbf{H} : [n] \to [r]$ and let $\mathbf{S}_j = \mathbf{H}^{-1}(j)$, $j \in [r]$.

First let us suppose that $\mathrm{char}(\mathbb{F}) = p > 0$. Choose $\alpha_i \sim \mathbb{F}_p$, $i \in [n]$ independently and uniformly at random and define $\mathbf{L}_j(x) = \sum_{i \in \mathbf{S}_j} \alpha_i x_i$, for $x \in \{0,1\}^n$, $j \in [r]$. For $i \in [m]$, let $\mathbf{P}_2^{(i)} = Q_r^{(i)}(\mathbf{L}_1^{p-1}, \ldots, \mathbf{L}_r^{p-1})$, where $Q_r^{(i)}$ is the unique multilinear polynomial representation of $\mathrm{Thr}_r^{t_i}$. Let $\mathbf{P}_2 = (\mathbf{P}_2^{(1)}, \ldots, \mathbf{P}_2^{(m)})$. Define $\mathbf{P} = \mathbf{1} - (\mathbf{1} - P_1) * (\mathbf{1} - \mathbf{P}_2)$, that is, $\mathbf{P} = (\mathbf{P}^{(1)}, \ldots, \mathbf{P}^{(m)})$, where $\mathbf{P}^{(i)} = \mathrm{OR}_2(P_1^{(i)}, \mathbf{P}_2^{(i)})$, for all $i \in [m]$.

Note that since $\varepsilon \leq 2^{-t/160000}$, we have $r = 4800000 \log(1/\varepsilon) \geq t$. Thus $t_i \leq t \leq r$, for all $i \in [m]$. Now fix any $a \in \{0,1\}^n$. Let $Z_a = \{i \in [m] : \mathrm{Thr}_n^{t_i}(a) = 0\}$ and $N_a = \{i \in [m] : \mathrm{Thr}_n^{t_i}(a) = 1\}$. So we have $|a| < t_i \leq t \leq r$ and hence $\mathrm{EX}_{[0,r]} T_i(a) = 0$, for all $i \in Z_a$. Also $|(\mathbf{L}_1^{p-1}(a), \ldots, \mathbf{L}_r^{p-1}(a))| \leq |a| < t_i$ w.p.1, and so $\mathbf{P}_2^{(i)}(a) = Q_r^{(i)}((\mathbf{L}_1^{p-1}(a), \ldots, \mathbf{L}_r^{p-1}(a))) = 0$ w.p.1, for all $i \in Z_a$ simultaneously. Thus $\mathbf{P}^{(i)}(a) = 0$ w.p.1, for all $i \in Z_a$ simultaneously.

Further we have $|a| \geq t_i$, for all $i \in N_a$. We will now show that $\mathbf{P}^{(i)}(a) = 1$ w.p. at least $1 - \varepsilon$, for all $i \in N_a$ simultaneously. If $|a| \leq r$, then again $P_1^{(i)}(a) = 1$, for all $i \in N_a$ and so $\mathbf{P}^{(i)}(a) = 1$ w.p.1. Now suppose $|a| \geq r$. Without loss of generality, assume $t_1 \leq \cdots \leq t_m = t$. Then we have $\mathbf{P}_2^{(1)}(a) \geq \cdots \geq \mathbf{P}_2^{(m)}(a)$ w.p.1, under the order $1 > 0$. So it is enough to show that $\mathbf{P}^{(m)}(a) = 1$ w.p. at least $1 - \varepsilon$.

Define $I(\mathbf{H}) = \{j \in [r] : \mathrm{supp}(a) \cap \mathbf{S}_j \neq \emptyset\}$. We get

$$\begin{aligned}
\Pr\left[\mathbf{P}_2^{(m)}(a) = 0\right] &= \Pr\left[\mathbf{P}_2^{(m)}(a) = 0 \mid |I(\mathbf{H})| < r/10\right] \cdot \Pr\left[|I(\mathbf{H})| < r/10\right] \\
&\quad + \Pr\left[\mathbf{P}_2^{(m)}(a) = 0 \mid |I(\mathbf{H})| \geq r/10\right] \cdot \Pr\left[|I(\mathbf{H})| \geq r/10\right] \\
&\leq \Pr\left[|I(\mathbf{H})| < r/10\right] + \max_{\mathbf{H} : |I(\mathbf{H})| \geq r/10} \Pr\left[\mathbf{P}_2^{(m)}(a) = 0 \mid \mathbf{H}\right].
\end{aligned}$$

By Union Bound, we get

$$\Pr\left[|I(\mathbf{H})| < r/10\right] \leq \sum_{I \subseteq [r],\, |I| = r/10} \Pr\left[I(\mathbf{H}) \subset I\right] \leq \binom{r}{r/10} \frac{1}{10^r} \leq \frac{1}{4^r} \leq \frac{1}{4} \cdot \frac{1}{2^r} \leq \frac{\varepsilon}{4}.$$

Now fix any $\mathbf{H}$ such that $|I(\mathbf{H})| \geq r/10$, and let $\ell = |I(\mathbf{H})|$. Note that $\mathbf{P}_2^{(m)}(a)$ is 0 if and only if at most $t - 1$ many $\mathbf{L}_j(a)$ are non-zero. We consider only $j \in I(\mathbf{H})$. For each $j \in I(\mathbf{H})$, the probability that $\mathbf{L}_j(a)$ is non-zero is $1 - 1/p \geq 1/2$. Thus, the expected number of $\mathbf{L}_j(a)$ $(j \in I(\mathbf{H}))$ that are non-zero is at least $\ell/2 \geq r/20$. Thus, by Lemma 21,

$$\Pr\left[\mathbf{P}_2^{(m)}(a) = 0 \mid \mathbf{H}\right] = \Pr\left[|I(\mathbf{H}) \cap \{j : \mathbf{L}_j(a) = 1\}| \leq t-1 \mid \mathbf{H}\right] \leq 2\exp\left(-\frac{r}{240}\right) < \frac{\varepsilon}{2}.$$

where for the inequality we have used the fact that $t \leq r/40$. Thus $\Pr\left[\mathbf{P}_2^{(m)}(a) = 0\right] \leq \varepsilon$, proving the base case when $\mathrm{char}(\mathbb{F}) = p > 0$.

Now suppose $\mathrm{char}(\mathbb{F}) = 0$. Then for $i \in [m]$ we let $\mathbf{P}_2^{(i)} = Q_r^{(i)}(\mathbf{O}_1, \ldots, \mathbf{O}_r)$, where $Q_r^{(i)}$ is the unique multilinear polynomial representation of $\mathrm{Thr}_r^{t_i}$, and for $j \in [r]$, $\mathbf{O}_j$ is a 1/3-error probabilistic polynomial for $\mathrm{OR}_{\mathbf{S}_j}$, the OR function on variables $(X_k : k \in \mathbf{S}_j)$. Let $\mathbf{P}_2 = (\mathbf{P}_2^{(1)}, \ldots, \mathbf{P}_2^{(m)})$. Define $\mathbf{P} = \mathbf{1} - (\mathbf{1} - P_1) * (\mathbf{1} - \mathbf{P}_2)$, that is, $\mathbf{P} = (\mathbf{P}^{(1)}, \ldots, \mathbf{P}^{(m)})$, where $\mathbf{P}^{(i)} = \mathrm{OR}_2(P_1^{(i)}, \mathbf{P}_2^{(i)})$, for all $i \in [m]$. The rest of the analysis follows similarly, proving the base case when $\mathrm{char}(\mathbb{F}) = 0$.

Inductive Construction. For any positive characteristic p, any $n' < n$, $t' \in [0, n']$ and $\varepsilon' \in (0, 1/100)$, assume the existence of an ε'-error probabilistic poly-tuple for any (n', t')-threshold tuple, with degree at most $A_p\sqrt{t' \log(1/\varepsilon')} + B_p \log(1/\varepsilon')$; similarly, for characteristic zero, assume we have a probabilistic poly-tuple of degree $A_0 \log n \cdot \sqrt{t' \log(1/\varepsilon')} + B_0 \log n \cdot \log(1/\varepsilon')$.

We now consider an (n, t)-threshold tuple $T = (T_1, \ldots, T_m) = (\mathrm{Thr}_n^{t_1}, \ldots, \mathrm{Thr}_n^{t_m})$. Assume that the parameter $\varepsilon > 2^{-t/160000}$ since otherwise can use the construction from the base case. Define

$$T' = (T_1', \ldots, T_m') = \left(\mathrm{Thr}_{n/10}^{t_1/10}, \ldots, \mathrm{Thr}_{n/10}^{t_m/10}\right),$$
$$T_+'' = (T_{1,+}'', \ldots, T_{m,+}'') = \left(\mathrm{Thr}_{n/10}^{t_1/10 + 20\sqrt{t\log(1/\varepsilon)}}, \ldots, \mathrm{Thr}_{n/10}^{t_m/10 + 20\sqrt{t\log(1/\varepsilon)}}\right),$$
$$T_-'' = (T_{1,-}'', \ldots, T_{m,-}'') = \left(\mathrm{Thr}_{n/10}^{t_1/10 - 20\sqrt{t\log(1/\varepsilon)}}, \ldots, \mathrm{Thr}_{n/10}^{t_m/10 - 20\sqrt{t\log(1/\varepsilon)}}\right).$$

By induction hypothesis, let $\mathbf{T}', \mathbf{T}_+'', \mathbf{T}_-''$ be $\varepsilon/4$-error probabilistic poly-tuples for T', T_+'', T_-'' respectively. Let $\mathbf{N}'' = (\mathbf{1} - \mathbf{T}_+'') * \mathbf{T}_-''$. For any $x \in \{0,1\}^n$, choose a random subvector $\hat{\mathbf{x}} \in \{0,1\}^{n/10}$ with each coordinate chosen independently with probability 1/10, with replacement. Define

$$\mathbf{T}(x) = \mathbf{N}''(\hat{\mathbf{x}}) * E(x) + (\mathbf{1} - \mathbf{N}'')(\hat{\mathbf{x}}) * \mathbf{T}'(\hat{\mathbf{x}}),$$

where $E = (E_1, \ldots, E_m)$, with $E_i = \mathrm{EX}_{[t_i - 300\sqrt{t\log(1/\varepsilon)}, t_i + 300\sqrt{t\log(1/\varepsilon)}]}\mathrm{Thr}_n^{t_i}$, $i \in [m]$. We will now prove that $\mathbf{T}$ is an ε-error probabilistic poly-tuple for T.

Correctness of Inductive Construction. We now check that the construction above gives an ε-error probabilistic poly-tuple for T. Fix any $a \in \{0,1\}^n$. Let $\hat{\mathbf{a}} \in \{0,1\}^{n/10}$ be chosen as given in the inductive construction.

Suppose $|a| \leq 2t$. Let $\theta = 10\sqrt{t\log(1/\varepsilon)}$. Applying Lemma 21, we get $\Pr\left[||\hat{\mathbf{a}}| - |a|/10| > \theta\right] < \varepsilon/4$. By induction hypothesis, the probability that $\mathbf{T}'(\hat{\mathbf{a}})$ does not agree with $T'(\hat{\mathbf{a}})$ is at most $\varepsilon/4$, and similarly for $\mathbf{T}_+''$ and $\mathbf{T}_-''$. Let $\mathcal{G}_a$ be the event that none of the above events occur; by a union bound, the event $\mathcal{G}_a$ occurs with probability at least $1 - \varepsilon$. In this case, we show that $\mathbf{T}(a) = T(a)$, which will prove the correctness of the construction in the case that $|a| \leq 2t$.

To see the above, observe the following for each $i \in [m]$.

- $\mathbf{T}_i'(\hat{\mathbf{a}}) = T_i(a)$ if $||a| - t_i| > 10\theta$. This is because $\mathbf{T}_i'(\hat{\mathbf{a}}) = T_i'(\hat{\mathbf{a}})$ by our assumption that the event $\mathcal{G}_a$ has occurred. Further, we also have $T_i'(\hat{\mathbf{a}}) = T_i(a)$ since $|\hat{a} - |a|/10| \leq \theta$ (by occurrence of $\mathcal{G}_a$) and hence $|a| \geq t_i$ if and only if $|\hat{a}| \geq t_i/10$.

- If $||a| - t_i| > 30\theta$, then $\mathbf{N}''_i(\hat{\mathbf{a}}) = 0$. This is because $||\hat{a}| - |a|/10| \leq \theta$ and hence $||\hat{a}| - t_i/10| > 2\theta$. Hence, either $\mathbf{T}''_{i,+}(\hat{a}) = 1$ or $\mathbf{T}''_{i,-}(\hat{a}) = 0$ and therefore, $\mathbf{N}''_i(\hat{\mathbf{a}}) = 0$. Thus, when $||a| - t_i| > 30\theta$, the definition of $\mathbf{T}$ yields $\mathbf{T}_i(a) = \mathbf{T}'_i(\hat{\mathbf{a}}) = T_i(a)$. We are therefore done in this case.
- If $||a| - t_i| < 10\theta$, then $\mathbf{N}''_i(\hat{\mathbf{a}}) = 1$. This is similar to the analogous statement above. Therefore, when $||a| - t_i| < 10\theta$, we have $\mathbf{T}_i(a) = E_i(a) = T_i(a)$ as $|a| \in [t_i - 300\sqrt{t\log(1/\varepsilon)}, t_i + 300\sqrt{t\log(1/\varepsilon)}]$. Hence, we are done in this case also.
- If $10\theta \leq ||a| - t_i| \leq 30\theta$, then $E_i(a) = \mathbf{T}'(\hat{\mathbf{a}}) = T_i(a)$. Since $\mathbf{N}''_i(\hat{\mathbf{a}}) \in \{0, 1\}$ for each $i \in [m]$, we again obtain $\mathbf{T}_i(a) = T_i(a)$.

This shows that for any a such that $|a| \leq 2t$, whenever $\mathcal{G}_a$ does not occur, $\mathbf{T}(a) = T(a)$.

Now suppose $|a| > 2t$. Then by a Chernoff bound (follows from Lemma 21), we get $\Pr[|\hat{\mathbf{a}}| < 1.5t/10] < 2\exp(-t/400) < \varepsilon/2$. Also, by the induction hypothesis, the probability that $\mathbf{T}'(\hat{\mathbf{a}})$ does not agree with $T'(\hat{\mathbf{a}})$ is at most $\varepsilon/4$, and similarly for $\mathbf{T}''_+$ and $\mathbf{T}''_-$. Let $\mathcal{G}_a$ denote the event that none of the above events occur; we have $\Pr[\mathcal{G}] \geq 1 - \varepsilon$. As above, we show that when $\mathcal{G}_a$ occurs, then $\mathbf{T}(a) = T(a)$.

To see this, we proceed as follows.

- Since $|a| \geq 2t$ and $|\hat{\mathbf{a}}| \geq 1.5t/10$, both $T(a)$ and $\mathbf{T}'_i(\hat{a})$ are both the constant-1 vector.
- Further, we note that we have $\mathbf{N}''_i(\hat{\mathbf{a}}) = 0$ for each $i \in [m]$. This is because $||\hat{\mathbf{a}}| - t_i/10| \geq (|\hat{\mathbf{a}}| - t/10) \geq t/20 > 20\sqrt{t\log(1/\varepsilon)}$.

This implies that $\mathbf{T}_i(a) = \mathbf{T}'_i(\hat{\mathbf{a}}) = 1$ for each $i \in [m]$.

Hence, when $\mathcal{G}_a$ does not occur, we have $\mathbf{T}(a) = T(a)$, which proves the correctness of the construction.

Correctness of Degree. The computation that shows that $\deg(\mathbf{T})$ satisfies the inductive claim is omitted here and is in the full version of the paper. ◀

3.2 Upper bounds from Theorem 3

Upper bound for $\mathrm{pdeg}_\varepsilon(g)$. This result is due to Lu [12].

Upper bound for $\mathrm{pdeg}_\varepsilon(h)$. Let $B(h) = k$. Thus we can write $h = h_1 + (1 - \widetilde{h_2})$, for k-constant symmetric Boolean functions h_1, h_2, where $\widetilde{h_2}(x_1, \ldots, x_n) = h_2(1 - x_1, \ldots, 1 - x_n)$. But then by Corollary 20, $\mathrm{pdeg}_\varepsilon(h_1) = \mathrm{pdeg}_\varepsilon(h_2) = O(\sqrt{k\log(1/\varepsilon)} + \log(1/\varepsilon))$ and so $\mathrm{pdeg}_\varepsilon(h) = O(\sqrt{k\log(1/\varepsilon)} + \log(1/\varepsilon))$ over positive characteristic p. For $p = 0$, we obtain the same upper bound up to log-factors.

Upper bound for $\mathrm{pdeg}_\varepsilon(f)$. Let (g, h) be the standard decomposition of f. So $f = g \oplus h = g + h - 2gh$. Further, we already have the Alman-Williams bound of $O(\sqrt{n\log(1/\varepsilon)})$ on $\mathrm{pdeg}_\varepsilon(f)$ (Lemma 11). So we get $\mathrm{pdeg}_\varepsilon(f) = O(\min\{\sqrt{n\log(1/\varepsilon)}, \mathrm{per}(g) + \sqrt{B(h)\log(1/\varepsilon)} + \log(1/\varepsilon)\})$ over positive characteristic and the same bound up to log-factors over characteristic 0. This concludes the proof of Theorem 3.

4 Lower bounds

In this section, we prove the lower bounds from Theorem 4.

Throughout, $\mathbb{F}$ is fixed to be some arbitrary field of characteristic p (possibly 0). We use $\mathrm{pdeg}_\varepsilon(\cdot)$ instead of $\mathrm{pdeg}^{\mathbb{F}}_\varepsilon(\cdot)$ and $\mathrm{pdeg}(\cdot)$ instead of $\mathrm{pdeg}^{\mathbb{F}}_{1/3}(\cdot)$.

4.1 Preliminary lemmas

We need some preliminary lemmas. The proofs are omitted for lack of space.

▶ **Lemma 23.** *Let $n, m \in \mathbb{N}$, $n > m$ and $\varepsilon \leq 1/3$. Let $f \in s\mathcal{B}_n$ and $g \in s\mathcal{B}_m$ be such that* $\mathrm{per}(g)$ *divides* $\mathrm{per}(f)$ *and* $\mathrm{per}(f) \leq (n-m+2)/2$. *Then* $\mathrm{pdeg}_\varepsilon(f) = \Omega(\mathrm{pdeg}_\varepsilon(g)/\log^3(n/\varepsilon))$.

▶ **Lemma 24.** *For any* $1/2^n \leq \varepsilon \leq 1/3$, $\mathrm{pdeg}_\varepsilon(\mathrm{EThr}_n^{\lceil n/2 \rceil}) = \widetilde{\Omega}(\sqrt{n \log(1/\varepsilon)})$.

4.2 Lower bounds from Theorem 4

We recall the lower bound for periodic functions from Theorem 4. In light of Observation 2, this is a slightly more general statement.

▶ **Lemma 25.** *Let $g \in s\mathcal{B}_n$ be any function with* $\mathrm{per}(g) = b \leq n/3$, *then* $\mathrm{pdeg}_\varepsilon(g) = \widetilde{\Omega}(\sqrt{n})$ *if* $\mathrm{per}(g) > 1$ *and is not a power of p and* $\widetilde{\Omega}(\min\{\sqrt{n}, \mathrm{per}(g)\})$ *otherwise.*

Proof. Assume $\mathrm{per}(g) = b > 1$. Consider the two cases below.

b is not a power of p. Let b' be any non-trivial divisor of b which is coprime to p (if $p = 0$, we simply take $b' = b$). For $i \in [0, b'-1]$, define $g_i = \mathrm{MOD}_{\lceil n/3 \rceil}^{b',i}$. The functions g and g_i satisfy the hypotheses of Lemma 23 and therefore for any constant $\varepsilon \leq 1/3$, $\mathrm{pdeg}_\varepsilon(g) = \widetilde{\Omega}(\mathrm{pdeg}_\varepsilon(g_i))$.

Note that as $b \leq n/3$, we have $b' \leq n/6 \leq \frac{1}{2}\lceil n/3 \rceil$. Hence, by Lemma 9, for some $i \in [0, b'-1]$, $\mathrm{pdeg}_{1/n^2}(g_i) = \Omega(\sqrt{n \log(n^2/b)}) = \widetilde{\Omega}(\sqrt{n})$.

Therefore $\mathrm{pdeg}_{1/n^2}(g) = \widetilde{\Omega}(\sqrt{n})$ and hence by Fact 6 item 1 $\mathrm{pdeg}(g) = \widetilde{\Omega}(\sqrt{n})$.

$b = p^k$ for some $k \in \mathbb{N}$. Let $m = \min(b^2/100, \lceil n/3 \rceil)$. Let $g' \in s\mathcal{B}_m$ with $\mathrm{per}(g') = b$ be such that $\mathrm{Spec}\,(g')(i) = 0$ whenever $\lfloor m/2 \rfloor - \lfloor b/2 \rfloor \leq i \leq \lfloor m/2 \rfloor$ and $\mathrm{Spec}\,(g')(i) = 1$ whenever $\lfloor m/2 \rfloor < i \leq \lfloor m/2 \rfloor + b - \lfloor b/2 \rfloor - 1$.

Again, the functions g and g' satisfy the hypotheses of Lemma 23 and therefore for any constant $\varepsilon \leq 1/3$, $\mathrm{pdeg}_\varepsilon(g) = \widetilde{\Omega}(\mathrm{pdeg}_\varepsilon(g'))$.

Note that g' agrees with the Maj_m function on all inputs $x \in \{0,1\}^m$ such that $||x| - (m/2)|$ is at most $b/2$. By a Chernoff bound (follows from Lemma 21),

$$\Pr_{x \in \{0,1\}^m}[||x| - m/2| > b/2] \leq 2e^{-\frac{b^2}{2m}} = 2e^{-50} < 1/5.$$

Therefore g' agrees with Maj_m on more than $4/5$ fraction of inputs and hence by Lemma 8, $\mathrm{pdeg}(g') = \Omega(\sqrt{m})$. Therefore, $\mathrm{pdeg}(g) = \widetilde{\Omega}(\sqrt{m}) = \min(\widetilde{\Omega}(b), \widetilde{\Omega}(\sqrt{n}))$. ◀

We now recall the lower bound for bounded symmetric Boolean functions from Theorem 4.

▶ **Lemma 26.** *Let $h \in s\mathcal{B}_n$ be such that $B(h) \leq \lceil n/3 \rceil$. Then,* $\mathrm{pdeg}_\varepsilon(h) = \widetilde{\Omega}(\sqrt{B(h)})$,

Proof. Let $B(h) = b$. Then, we know that $\mathrm{Spec}\,h(i) = 0$ for $i \in [b, n-b]$ and further either $\mathrm{Spec}\,h(b-1)$ or $\mathrm{Spec}\,h(n-b+1)$ is 1. We assume w.l.o.g. that $\mathrm{Spec}\,h(n-b+1) = 1$ (the other case is similar).

Fix some integer $b' = b - O(1)$ so that $2b + 2\lfloor b'/2 \rfloor \leq n$. Define $h' \in s\mathcal{B}_{b'}$ as

$$h'(x) = \bigvee_{i \in [0, \lfloor b'/2 \rfloor]} h(x 1^{n-b-2\lfloor b'/2 \rfloor + i} 0^{b-b'+2\lfloor b'/2 \rfloor - i}), \quad \text{for all } x \in \{0,1\}^{b'}.$$

We claim that $h' = \mathrm{Maj}_{b'}$. To show this, we proceed as follows.

Let $|x| \le b'/2$ and therefore $|x| \le \lfloor b'/2 \rfloor$. Then for all $i \in [0, \lfloor b'/2 \rfloor]$, using our choice of b', we have

$$b \le n - b - 2\lfloor b'/2 \rfloor \le |x1^{n-b-2\lfloor b'/2 \rfloor + i}0^{b-b'+2\lfloor b'/2 \rfloor - i}| = |x| + (n - b - 2\lfloor b'/2 \rfloor + i) \le n - b$$

and therefore none of the terms in the OR above evaluate to 1. Thus $h'(x) = 0$.

Let $|x| > b'/2$ and therefore $|x| \ge \lfloor b'/2 \rfloor + 1$. Let $|x| = \lfloor b'/2 \rfloor + j$ for some $j \in [1, \lceil b'/2 \rceil]$. Let $i = \lfloor b'/2 \rfloor - j + 1$. Then $|x1^{n-b-2\lfloor b'/2 \rfloor + i}0^{b-b'+2\lfloor b'/2 \rfloor - i}| = n - b + 1$. Therefore the OR evaluates to 1 and $h'(x) = 1$.

From the above we see that $h' = \mathrm{Maj}_b$. Now,

$$\begin{aligned} \mathrm{pdeg}(h') &\le \mathrm{pdeg}_{2/n}(h') \\ &\le \mathrm{pdeg}_{1/n}(\mathrm{OR}) \cdot \mathrm{pdeg}_{1/n^2}(h) \\ &\le O(\log^2 n) \cdot O(\log n) \cdot \mathrm{pdeg}(h) \\ &\le \widetilde{O}(\mathrm{pdeg}(h)). \end{aligned}$$

The second inequality follows from Fact 6 item 2 and the third inequality follows from Lemma 7 and Fact 6 item 1.

Since $\mathrm{pdeg}(\mathrm{Maj}_{b'}) = \Omega(\sqrt{b'}) = \Omega(\sqrt{b})$, it follows that $\mathrm{pdeg}(h) = \widetilde{\Omega}(\sqrt{b})$. ◀

Using the above, a short case analysis yields the lower bound on $\mathrm{pdeg}(f)$ for general $f \in s\mathcal{B}_n$. The proof is omitted.

References

1 Josh Alman and Ryan Williams. Probabilistic polynomials and hamming nearest neighbors. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 136–150. IEEE, 2015.

2 Richard Beigel. The polynomial method in circuit complexity. *[1993] Proceedings of the Eigth Annual Structure in Complexity Theory Conference*, pages 82–95, 1993.

3 Richard Beigel, Nick Reingold, and Daniel A. Spielman. The Perceptron Strikes Back. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*, pages 286–291, 1991. doi:10.1109/SCT.1991.160270.

4 Siddharth Bhandari, Prahladh Harsha, Tulasimohan Molli, and Srikanth Srinivasan. On the Probabilistic Degree of OR over the Reals. In Sumit Ganguly and Paritosh Pandya, editors, *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*, volume 122 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:12, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.FSTTCS.2018.5.

5 Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *J. ACM*, 57(5), 2010. doi:10.1145/1754399.1754401.

6 Bettina Brustmann and Ingo Wegener. The Complexity of Symmetric Functions in Bounded-Depth Circuits. *Inf. Process. Lett.*, 25(4):217–219, 1987. doi:10.1016/0020-0190(87)90163-3.

7 Devdatt P Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

8 Ronald Fagin, Maria M. Klawe, Nicholas Pippenger, and Larry J. Stockmeyer. Bounded-Depth, Polynomial-Size Circuits for Symmetric Functions. *Theor. Comput. Sci.*, 36:239–250, 1985. doi:10.1016/0304-3975(85)90045-3.

9 Prahladh Harsha and Srikanth Srinivasan. On Polynomial Approximations to AC^0. In Klaus Jansen, Claire Mathieu, José D. P. Rolim, and Chris Umans, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, volume 60 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:14, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.APPROX-RANDOM.2016.32.

10 David Lawrence Johnson, David Leroy Johnson, and SS Johnson. *Topics in the theory of group presentations*, volume 42. Cambridge University Press, 1980.

11 Nathan Linial and Noam Nisan. Approximate inclusion-exclusion. *Combinatorica*, 10(4):349–365, 1990.

12 Chi-Jen Lu. An exact characterization of symmetric functions in $\mathrm{qAC}^0[2]$. *Theoretical Computer Science*, 261(2):297–303, 2001.

13 Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for Polynomials of Independent Random Variables. *Theory of Computing*, 12(1):1–17, 2016. doi:10.4086/toc.2016.v012a011.

14 Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, New York, NY, USA, 2014.

15 Ramamohan Paturi. On the degree of polynomials that approximate symmetric Boolean functions (preliminary version). In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 468–474. ACM, 1992.

16 Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematicheskie Zametki*, 41(4):598–607, 1987. (English translation in *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987). doi:10.1007/BF01137685.

17 Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 77–82. ACM, 1987.

18 Roman Smolensky. Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

19 Roman Smolensky. On representations by low-degree polynomials. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 130–138. IEEE, 1993.

20 Roman Smolensky. On Representations by Low-Degree Polynomials. In *FOCS*, pages 130–138, 1993. doi:10.1109/SFCS.1993.366874.

21 Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. On the Probabilistic Degrees of Symmetric Boolean functions. *Electronic Colloquium on Computational Complexity (ECCC)*, 138:1–24, 2019. URL: https://eccc.weizmann.ac.il/report/2019/138.

22 Jun Tarui. Probabilistic Polynomials, AC^0 Functions, and the Polynomial-Time Hierarchy. *Theoretical Computer Science*, 113(1):167–183, 1993.

23 Richard Ryan Williams. The polynomial method in circuit complexity applied to algorithm design (invited talk). In *34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS 2014)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2014.

24 Zhi-Li Zhang, David A Mix Barrington, and Jun Tarui. Computing symmetric functions with AND/OR circuits and a single MAJORITY gate. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 535–544. Springer, 1993.