Bounds on the Fourier Coefficients of the Weighted Sum Function

Igor E. Shparlinski
Department of Computing, Macquarie University
Sydney, NSW 2109, Australia
igor@ics.mq.edu.au

Abstract

We estimate Fourier coefficients of a Boolean function which has recently been introduced in the study of read-once branching programs. Our bound implies that this function has an asymptotically "flat" Fourier spectrum and thus implies several lower bounds of its various complexity measures.

1 Introduction

1.1 Motivation

P. Savický and S. Žák [22], in their study of read-once branching programs, have recently introduced a Boolean function f defined in terms of certain weighted sums in the residue ring modulo a prime. It has also been used by M. Sauerhoff [20, 21] for several more complexity theory applications. In particular, in [21] a certain modification of the same function has been used to prove that quantum read-once branching programs are exponentially more powerful than classical read-once branching programs. Here, motivated by the important role the function f has played in several recent works, we continue to study f and concentrate on estimating its *Fourier coefficients*.

It is well know that there are many close links between Fourier coefficients and various complexity characteristics of any Boolean function, see [2, 3, 4, 5, 6, 10, 11, 12, 14, 16, 18, 19] and references therein. Although we do not present all such implications, we give lower bounds on several complexity characteristics of f.

1.2 Notation

We now fix a sufficiently large integer n and let p be the smallest prime with $p \geq n$.

We also use $\mathcal{B}_r$ to denote the r-dimensional binary cube, that is, $\mathcal{B}_r = \{0,1\}^r$.

Given an n-dimensional binary vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{B}_n$ we define $s(\mathbf{x})$ by the conditions

$$s(\mathbf{x}) \equiv \sum_{k=1}^{n} k x_k \pmod p, \qquad 1 \le s(\mathbf{x}) \le p.$$

Following [22], we consider the Boolean function

$$f(\mathbf{x}) = \begin{cases} x_{s(\mathbf{x})}, & \text{if } 1 \le s(\mathbf{x}) \le n; \\ x_1, & \text{otherwise.} \end{cases} \tag{1}$$

We use some methods of analytic number theory to estimate *Fourier coefficients*

$$\widehat{f}(\mathbf{u}) = \frac{1}{2^n} \sum_{\mathbf{x} \in \mathcal{B}_n} (-1)^{f(\mathbf{x}) + \mathbf{u}\cdot\mathbf{x}},$$

where $\mathbf{u} = (u_1, \ldots, u_n) \in \mathcal{B}_n$, and

$$\mathbf{u}\cdot\mathbf{x} = u_1 x_1 + \ldots + u_n x_n$$

is the inner product.

1.3 Results

We show that all such coefficients are of the size $2^{(-1/2+o(1))n}$ (where the term $o(1)$ depends on our knowledge about the gaps between consecutive primes).

Certainly, the *Parseval identity*

$$\sum_{\mathbf{u} \in \mathcal{B}_n} \widehat{g}(\mathbf{u})^2 = 1, \tag{2}$$

implies that

$$\max_{\mathbf{u} \in \mathcal{B}_n} |\widehat{g}(\mathbf{u})| \ge 2^{-n/2}$$

for Fourier coefficients $\widehat{g}(\mathbf{u})$ of any n-variate Boolean function g. Thus the function f has an asymptotically optimal Fourier spectrum.

We also give present some immediate applications of our bound and derive an asymptotic formula on the *average sensitivity* of f which in turn leads to lower bounds on its circuit complexity and polynomial degree. We also give a lower bounds on the size of a *decision tree* which computes f.

2 Estimating Fourier Coefficients

2.1 Preparations

We start with a bound on the gap between n and p, which follows from [1].

Lemma 1. *We have, $p = n + O(n^{0.525})$.*

We now put $\mathbf{e}(z) = \exp(2\pi\iota z/p)$ where $\iota = \sqrt{-1}$.

Lemma 2. *We have,*

$$\max_{\lambda=1,\dots,p-1} \left| \sum_{j=1}^{n} \mathbf{e}(\lambda j) \right| = O(n^{0.525}).$$

Proof. The result follows immediately from the identity

$$\sum_{\lambda=1}^{p} \mathbf{e}(\lambda z) = \begin{cases} 0, & \text{if } z \not\equiv 0 \pmod{p}, \\ p, & \text{if } z \equiv 0 \pmod{p}, \end{cases} \tag{3}$$

since

$$\left| \sum_{j=1}^{n} \mathbf{e}(\lambda j) \right| = \left| \sum_{j=1}^{p} \mathbf{e}(\lambda j) - \sum_{j=n+1}^{p} \mathbf{e}(\lambda j) \right| = \left| \sum_{j=n+1}^{p} \mathbf{e}(\lambda j) \right| \le p - n = O(n^{0.525})$$

by Lemma 1. □

The following inequality is given in the proof of [13, Theorem 18.2].

Lemma 3. *For any complex numbers $z, z_1, \dots, z_N$ on the unit circle, $|z| = |z_1| = \dots = |z_N| = 1$, we have*

$$\left| \prod_{k=1}^{N} (z + z_k) \right| \le 2^{N/2} \left(1 + \frac{1}{N} \left| \sum_{k=1}^{N} z_k \right| \right)^{N/2}.$$

2.2 Main Result

Theorem 4. *For the function f given by (1), we have*

$$\max_{\mathbf{u} \in \mathcal{B}_n} |\widehat{f}(\mathbf{u})| = 2^{-n/2 + O(n^{0.525})}.$$

Proof. As we have remarked the lower bound follows immediately from (2), so we now concentrate on deriving the upper bound.

For every $j \in \{1, \dots, p\}$, let $\mathcal{X}_j$ be the set of $\mathbf{x} \in \mathcal{B}_n$ with $s(\mathbf{x}) = j$. We now write

$$\widehat{f}(\mathbf{u}) = \frac{1}{2^n} \sum_{j=1}^{p} F_j(\mathbf{u}) \tag{4}$$

and estimate each of the inner sums

$$F_j(\mathbf{u}) = \sum_{\mathbf{x} \in \mathcal{X}_j} (-1)^{f(\mathbf{x}) + \mathbf{u} \cdot \mathbf{x}}$$

separately.

We start with considering the sum $F_j(\mathbf{u})$ for $j \in \{1, \dots, n\}$. In this case, for every pair $(\alpha, \beta) = \mathcal{B}_2$ we use $\mathcal{X}_{j,\alpha,\beta}$ to denote the set of $\mathbf{x} \in \mathcal{X}_j$ with

$$x_j = \alpha, \qquad \mathbf{u} \cdot \mathbf{x} = \beta.$$

Therefore

$$F_j(\mathbf{u}) = \sum_{\alpha,\beta \in \mathcal{B}_2} \#\mathcal{X}_{j,\alpha,\beta}(-1)^{\alpha+\beta}. \tag{5}$$

From the identity (3) we have

$$\begin{aligned}
\#\mathcal{X}_{j,\alpha,\beta} &= \sum_{\substack{\mathbf{x}\in\mathcal{B}_n \\ x_j=\alpha}} \frac{1}{2p}\left(1+(-1)^{\mathbf{u}\cdot\mathbf{x}-\beta}\right)\sum_{\lambda=1}^{p}\mathbf{e}\left(\lambda\left(s(\mathbf{x})-j\right)\right) \\
&= \frac{1}{2p}\sum_{\lambda=1}^{p}\mathbf{e}(-\lambda j)\sum_{\mu=0}^{1}(-1)^{\mu\beta}\sum_{\substack{\mathbf{x}\in\mathcal{B}_n \\ x_j=\alpha}}(-1)^{\mu\mathbf{u}\cdot\mathbf{x}}\mathbf{e}\left(\lambda\sum_{k=1}^{n}kx_k\right) \\
&= \frac{1}{2p}\sum_{\lambda=1}^{p}\mathbf{e}\left(\lambda j(\alpha-1)\right)\sum_{\mu=0}^{1}(-1)^{\mu(\alpha u_j+\beta)}\prod_{\substack{k=1 \\ k\neq j}}^{n}\left(1+(-1)^{\mu u_k}\mathbf{e}\left(\lambda k\right)\right).
\end{aligned}$$

We say that $\mathbf{u} \in \mathcal{B}_n$ is j-vanishing if $u_k = 0$ for every $k \in \{1, \dots, n\}$ with $k \neq j$. The we see that in the above sum the term corresponding to $\lambda = 0$ is

$$\frac{1}{2p}\sum_{\mu=0}^{1}(-1)^{\mu(\alpha u_j+\beta)}\prod_{\substack{k=1 \\ k\neq j}}^{n}\left(1+(-1)^{\mu u_k}\right) = \sigma_j(\mathbf{u}, \alpha, \beta),$$

where

$$\sigma_j(\mathbf{u}, \alpha, \beta) = \begin{cases} 2^{n-2}p^{-1}, & \text{if } \mathbf{u} \text{ is not } j\text{-vanishing}, \\ 2^{n-2}p^{-1}\left(1+(-1)^{\alpha u_j+\beta}\right), & \text{otherwise.} \end{cases}$$

The contribution from other terms can be estimated as

$$\frac{1}{2p}\sum_{\lambda=1}^{p}\sum_{\mu=0}^{1}\left|\prod_{\substack{k=1 \\ k\neq j}}^{n}\left(1+(-1)^{\mu u_k}\mathbf{e}\left(\lambda k\right)\right)\right| = O\left(2^{n/2+O(n^{0.525})}\right)$$

by Lemma 2 and Lemma 3. Thus we see from (5) that

$$F_j(\mathbf{u}) = \sum_{\alpha,\beta\in\mathcal{B}_2}\sigma_j(\mathbf{u},\alpha,\beta)(-1)^{\alpha+\beta} + O\left(2^{n/2+O(n^{0.525})}\right)$$

and one can easily verify that

$$\sum_{\alpha,\beta\in\mathcal{B}_2}\sigma_j(\mathbf{u},\alpha,\beta)(-1)^{\alpha+\beta} = 0$$

whether $\mathbf{u}$ is j-vanishing or not. Hence

$$|F_j(\mathbf{u})| \le 2^{n/2+O(n^{0.525})}. \qquad (6)$$

It remains to estimate $F_j(\mathbf{u})$ for $j \in \{n+1, \ldots, p\}$. In this case, for every pair $(\alpha, \beta) = \mathcal{B}_2$ we use $\mathcal{Y}_{j,\alpha,\beta}$ to denote the set of $\mathbf{x} \in \mathcal{X}_j$ with

$$x_1 = \alpha, \qquad \mathbf{u} \cdot \mathbf{x} = \beta.$$

Exactly the same arguments as before lead to the bound

$$F_j(\mathbf{u}) = \sum_{\alpha,\beta \in \mathcal{B}_2} \sigma_1(\mathbf{u}, \alpha, \beta)(-1)^{\alpha+\beta} + O\left(2^{n/2+O(n^{0.525})}\right).$$

Therefore (6) still holds. Substituting (6) in (4) we finish the proof. □

3 Applications

3.1 Average Sensitivity, Circuit Complexity and Polynomial Representations

We recall that the *average sensitivity* $\sigma_{av}(g)$ of an n-variate Boolean function g is defined as

$$\sigma_{av}(g) = 2^{-n} \sum_{\mathbf{x} \in \mathcal{B}_n} \sum_{i=1}^{n} \left| g(\mathbf{x}) - g(\mathbf{x}^{(i)}) \right|.$$

where $\mathbf{x}^{(i)}$ is the vector obtained from $\mathbf{x}$ by flipping its ith coordinate.

Theorem 5. *For the function f given by* (1)*, we have*

$$\sigma_{av}(f) = (1 + o(1))n$$

Proof. It is shown in [12] that

$$\sigma_{av}(f) = \sum_{\mathbf{u} \in \mathcal{B}_n} \mathrm{wt}\,(\mathbf{u}) |\widehat{f}(\mathbf{u})|^2$$

where $\mathrm{wt}\,(\mathbf{u})$ is the Hamming weight of $\mathbf{u}$.

Therefore, for any $w \le n$, from the Parseval identity (2), we obtain

$$\begin{aligned}
\sigma_{av}(f) &\ge \sum_{\substack{\mathrm{wt}\,(\mathbf{u}) \in \mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u}) < w}} \mathrm{wt}\,(\mathbf{u}) \left|\widehat{f}(\mathbf{u})\right|^2 + w \sum_{\substack{\mathrm{wt}\,(\mathbf{u}) \in \mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u}) \ge w}} \left|\widehat{f}(\mathbf{u})\right|^2 \\
&= \sum_{\substack{\mathrm{wt}\,(\mathbf{u}) \in \mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u}) < w}} \mathrm{wt}\,(\mathbf{u}) \left|\widehat{f}(\mathbf{u})\right|^2 + w \left(1 - \sum_{\substack{\mathrm{wt}\,(\mathbf{u}) \in \mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u}) < w}} \left|\widehat{f}(\mathbf{u})\right|^2\right) \\
&\ge w - (w-1) \sum_{\substack{\mathrm{wt}\,(\mathbf{u}) \in \mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u}) < w}} \left|\widehat{f}(\mathbf{u})\right|^2.
\end{aligned}$$

Using the bound of Theorem 4 we see that

$$\sum_{\substack{\mathrm{wt}\,(\mathbf{u})\in\mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u})<w}} \left|\widehat{f}(\mathbf{u})\right|^2 \le 2^{-n+O(n^{0.525})} \sum_{\substack{\mathrm{wt}\,(\mathbf{u})\in\mathcal{B}_n \\ \mathrm{wt}\,(\mathbf{u})<w}} 1 = 2^{-n+O(n^{0.525})} \sum_{j=0}^{w-1} \binom{n}{j}.$$

We recall that for any $w \le n/2$ we have the bound

$$\sum_{j=0}^{w-1} \binom{n}{j} \le 2^{nH(w/n)+o(n)},$$

where

$$H(\gamma) = -\gamma \log \gamma - (1-\gamma)\log(1-\gamma), \qquad 0 < \gamma < 1,$$

and $\log z$ denotes the binary logarithm, see [15, Section 10.11]. Hence, for $w \le n/2$,

$$\sigma_{av}(f) \ge w - (w-1)2^{n(H(w/n)-1)+\delta(n)}$$

for some function $\delta(n) \to 0$ as $n \to \infty$. One easily verifies that, as $\eta \to 0$,

$$H(1/2-\eta) = 1 - a\eta^2 + O(\eta^3)$$

where $a = 2\log e = 2.885\ldots$. Taking $w = n/2 - \delta(n)^{1/2}$ gives the desired result. $\square$

By the Boppana result [4] if an unbounded fan-in Boolean circuit of depth d and size S computes a Boolean function g, then $d \log\log S \ge \log \sigma_{av}(g)$. Thus we see from Theorem 5 that if an unbounded fan-in Boolean circuit of depth d and size S computes the function f given by (1), then

$$d \log\log S \ge (1+o(1))n.$$

For an n-variate Boolean function g, we define its *real degree* $\Delta(g)$ and *real approximate degree* $\delta(g)$ as the smallest possible degree of a real polynomial F in n variables for which

$$g(x_1,\ldots,x_n) = F(x_1,\ldots x_r) \quad \text{and} \quad |g(x_1,\ldots,x_n) - F(x_1,\ldots x_r)| \le 1/3.$$

holds for every $(x_1,\ldots,x_r) \in \mathcal{B}_n$, respectively. Clearly, $\delta(g) \le \Delta(g) \le n$.

By Corollary 2.5 and by Lemma 3.8 of [17], for any Boolean function g, we have

$$\Delta(g) \ge \sigma_{av}(g) \qquad \text{and} \qquad \delta(g) \ge (\sigma_{av}(g)/6)^{1/2},$$

thus Theorem 5 we obtain for the function f, that

$$\Delta(f) \ge (1+o(1))n \qquad \text{and} \qquad \delta(f) \ge (6^{-1/2}+o(1))n^{1/2}.$$

In turn, these bounds imply a lower bound on quantum computational complexity of f, see [7].

3.2 Decision Tree Complexity

We recall that a *decision tree* with input variables $X_1, \ldots, X_n$ is a rooted binary tree in which each edge is labeled with a variable or a negated variable in such a way that labels of edges leaving the same inner node are negations of each other. Further each leaf v of the tree is labeled with some value $\lambda(v) \in \{0, 1\}$.

A decision tree $\mathcal{T}$ defines a Boolean function $g_{\mathcal{T}}$ as follows: Given an input $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{B}_n$, replace each edge label X_i by the induced value, that is, replace each X_i by x_i and each $\neg X_i$ by $\neg x_i$. After the replacement there is exactly one path from the root to some leaf v whose edges are all labeled 1 which is called the *computation path of the input* x. Define $g_{\mathcal{T}}(\mathbf{x})$ to be $\lambda(v)$.

The number of leaves is called the *size* of the decision tree.

We denote by $\mathrm{DT}(g)$ the smallest possible size of a decision tree which computes a Boolean function g.

Theorem 6. *For the function f given by* (1)*, we have*

$$\mathrm{DT}(f) \geq 2^{n/2 + O(n^{0.525})}.$$

Proof. From Lemma 2.2 (taken with S empty) of [11] we obtain

$$\mathrm{DT}(f) \geq \sum_{\mathbf{u} \in \mathcal{B}_n} |\widehat{f}(\mathbf{u})|.$$

On the other hand, from the Parseval identity (2) and the bound of Theorem 4 we see that

$$1 = \sum_{\mathbf{u} \in \mathcal{B}_n} \widehat{f}(\mathbf{u})^2 \leq 2^{-n/2 + O(n^{0.525})} \sum_{\mathbf{u} \in \mathcal{B}_n} |\widehat{f}(\mathbf{u})|.$$

and the desired estimate follows. □

4 Remarks

Clearly the error term $O(n^{0.525})$ in Theorem 4 comes from a result about gaps between consecutive primes [1] and under the Riemann Hypothesis can be reduced to $O(n^{1/2+o(1)})$.

The bound of Theorem 4 implies that the function f has a high *non-linearity*

$$N(f) = 2^{n-1} + O\left(2^{n/2 + O(n^{0.525})}\right)$$

which is defined as the difference

$$N(f) = 2^{n-1} - \frac{1}{2} \max_{\mathbf{u} \in \mathcal{B}_n} |\widehat{f}(\mathbf{u})|.$$

We recall that Boolean functions with large non-linearity play a very important role in cryptography, see [8, 9]. Thus it may be interesting to study some other properties of cryptographic interest for the function f. One can also consider its applicability to stream ciphers, which naturally leads to a question about the period and statistical distribution of sequences $(z_h)_{h=1}^{\infty}$, generated recursively by

$$z_{h+n+1} = f(z_h, \dots z_{h+n}), \qquad h = 1, 2, \dots ,$$

with some initial vector $(z_1, \dots , z_n) \in \mathcal{B}_n$.

5 Acknowledgements

The author is grateful to the organisers of the Dagstuhl Workshop "Complexity of Boolean Functions" (March, 2006) for the invitation where this work was essentially done. In fact the author learned about the function f and its role in the complexity theory from a talk given at that workshop by Martin Sauerhoff about his work [21].

This work was supported in part by ARC grant DP0556431.

References

[1] R. C. Baker, G. Harman, and J. Pintz, 'The difference between consecutive primes, II', *Proc. Lond. Math. Soc.*, **83** (2001), 532–562.

[2] A. Bernasconi, 'On the complexity of balanced Boolean functions', *Inform. Proc. Letters*, **70** (1999), 157–163.

[3] A. Bernasconi, C. Damm and I. E. Shparlinski, 'Circuit and decision tree complexity of some number theoretic problems', *Inform. and Comp.*, **168** (2001), 113–124.

[4] R. B. Boppana, 'The average sensitivity of bounded-depth circuits', *Inform. Proc. Letters*, **63** (1997), 257–261.

[5] J. Bruck, 'Harmonic analysis of polynomial threshold functions', *SIAM J. Discr. Math.*, **3** (1990), 168–177.

[6] J. Bruck and R. Smolensky, 'Polynomial threshold functions, $\mathcal{AC}^0$ functions, and spectral norms', *SIAM J. Comp.*, **21** (1992), 33–42.

[7] R. Beals, H. Buhrman, R. Cleve, M. Mosca and R. de Wolf, 'Quantum lower bounds by polynomials', *J. ACM*, **48** (2001), 778–797.

[8] C. Carlet, 'On the degree, nonlinearity, algebraic thickness, and nonnormality of Boolean functions, with developments on symmetric functions', *IEEE Trans. Inform. Theory*, **50** (2004), 2178–2185.

[9] C. Carlet and C. Ding, 'Highly nonlinear mappings', *J. Compl.*, **20** (2004), 205–244.

[10] M. Goldmann, 'Communication complexity and lower bounds for simulating threshold circuits', *Theoretical advances in neural computing and learning*, Kluwer Acad. Publ., Dordrecht, 1994, 85–125.

[11] S. Jukna, A. Razborov, P. Savicky and I. Wegener, 'On **P** versus **NP**$\cap$*co*-**NP** for decision trees and read-once branching programs', *Comp. Compl.*, **8** (1999), 357–370.

[12] J. Kahn, G. Kalai and N. Linial, 'The influence of variables on Boolean functions', *Proc. 29st IEEE Symp. on Found. of Comp. Sci.*, IEEE, 1988, 68–80.

[13] S. V. Konyagin and I. E. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999.

[14] N. Linial, Y. Mansour and N. Nisan, 'Constant depth circuits, Fourier transform, and learnability', *J. ACM*, **40** (1993), 607–620.

[15] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam, 1977.

[16] Y. Mansour, 'Learning Boolean functions via the Fourier transform', *Theoretical advances in neural computing and learning*, Kluwer Acad. Publ., Dordrecht, 1994, 391–424.

[17] N. Nisan and M. Szegedy, 'On the degree of Boolean functions as real polynomials', *Comp. Compl.*, **4** (1994), 301–313.

[18] R. Raz, 'Fourier analysis for probabilistic communication complexity', *Comp. Compl.*, **5** (1995), 205–221.

[19] V. Roychowdhry, K.-Y. Siu and A. Orlitsky, 'Neural models and spectral methods', *Theoretical advances in neural computing and learning*, Kluwer Acad. Publ., Dordrecht, 1994, 3–36.

[20] M. Sauerhoff, 'Randomness versus nondeterminism for read-once and read-k branching programs', *Proc. 20th Symp. on Theor. Aspects in Comp. Sci.*, Lect. Notes in Comp. Sci. Springer-Verlag, Berlin, **2607** (2003), 307–318.

[21] M. Sauerhoff, 'Quantum vs. classical read-once branching programs', *Preprint*, 2005, 1–35 (see `http://arxiv.org/abs/quant-ph/0504198`).

[22] P. Savický and S. Žák, 'A read-once lower bound and a $(1, +k)$-hierarchy for branching programs', *Theor. Comp. Sci.*, **238** (2000), 347–362.