On a Conjecture of O'Donnell

Qichun Wang*

Abstract

Let $f : \{-1,1\}^n \to \{-1,1\}$ be with total degree d, and $\widehat{f}(i)$ be the linear Fourier coefficients of f. The relationship between the sum of linear coefficients and the total degree is a foundational problem in theoretical computer science. In 2012, O'Donnell Conjectured that

$$\sum_{i=1}^{n} \widehat{f}(i) \leq d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{1-d}.$$

In this paper, we prove that the conjecture is equivalent to a conjecture on the cryptographic Boolean function. We then prove that the conjecture is true for $d = 1, n-1$. Moreover, we count the number of f's such that the upper bound is achieved.

Keywords: Boolean function, Linear coefficient, Total degree, Resiliency.

1 Introduction

Let $f : \{-1,1\}^n \to \{-1,1\}$. Then it can be written as

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_i,$$

where $[n] = \{1, 2, \dots, n\}$ and $\widehat{f}(S)$ are the Fourier coefficients of f given by

$$\frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x) \prod_{i \in S} x_i.$$

*School of Computer Science and Technology, Nanjing Normal University, Nanjing, P.R.China 210046. E-mail: qcwang@fudan.edu.cn.

The total influence of f, denoted by $Inf[f]$, is defined by

$$Inf[f] = \sum_{S \subseteq [n]} |S| \widehat{f}(S)^2.$$

The total degree of f, denoted by $\deg(f)$, is defined by

$$\deg(f) = \max\{|S| : \widehat{f}(S) \neq 0\}.$$

It is well-known that

$$\sum_{i=1}^{n} \widehat{f}(\{i\}) \leq Inf[f] \leq \deg(f).$$

For simplicity, we use $\widehat{f}(i)$ to denote $\widehat{f}(\{i\})$. In 2009, Parikshit Gopalan and Rocco Servedio conjectured that

$$\sum_{i=1}^{n} \widehat{f}(i) \leq \sqrt{\deg(f)}.$$

More ambitiously, in [5], O'Donnell proposed the following Conjecture.

Conjecture 1.1. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be with total degree d. Then*

$$\sum_{i=1}^{n} \widehat{f}(i) \leq d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{1-d}.$$

It is known that the conjecture is trivial for $d = n$ [6], since

$$\sum_{i=1}^{n} \widehat{f}(i) \leq 2^{-n} |x_1 + x_2 + \ldots + x_n| = n \cdot \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor} 2^{1-n}.$$

It should be noted that

$$\sum_{i=1}^{n} \widehat{f}(i) \geq -d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{1-d},$$

if Conjecture 1.1 holds.

2 An equivalent conjecture

Let $\mathbb{F}_2^n$ be the n-dimensional vector space over the finite field $\mathbb{F}_2 = \{0, 1\}$ and $\mathcal{B}_n$ be the set of all n-variable Boolean functions from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. Let $a = (a_1, a_2, \ldots, a_n) \in \mathbb{F}_2^n$. The Hamming weight of a, denoted by $wt(a)$, is defined by $\sum_{i=1}^n a_i$.

Let $g \in \mathcal{B}_n$. g is called t-resilient if [13]

$$\sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} = 0,$$

for any $v = (v_1, \ldots, v_n) \in \mathbb{F}_2^n$ satisfying $0 \le wt(v) \le t$, where "$\oplus$" is the XOR operator and $v \cdot x = v_1x_1 \oplus \cdots \oplus v_nx_n$ is the usual inner product.

If t is small, and g is not t-resilient, then a nonlinear combiner model of stream cipher using g as combining function can be attacked using the divide-and-conquer attack [12]. For more results on resilient Boolean functions, we refer to e.g. [2, 3, 4, 7, 8, 9, 10, 14, 15].

Conjecture 2.1. *Let $g \in \mathcal{B}_n$ be $(n - d - 1)$–resilient, where $1 \le d \le n - 1$. Then*

$$\sum_{\substack{v\in\mathbb{F}_2^n\\ wt(v)=n-1}} \sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} \le d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{n+1-d}.$$

Theorem 2.2. *Conjecture 1.1 is equivalent to Conjecture 2.1.*

Proof. "$\Rightarrow$" Let $g \in \mathcal{B}_n$ be $(n - d - 1)$–resilient. Then we have

$$\sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus x_1\oplus x_2\oplus\ldots\oplus x_n\oplus v\cdot x} = 0,$$

for any $v \in \mathbb{F}_2^n$ satisfying $d + 1 \le wt(v) \le n$. Let $G(x) = g(x) \oplus x_1 \oplus x_2 \oplus \ldots \oplus x_n$. We define a function $f : \{-1, 1\}^n \to \{-1, 1\}$ as

$$f(x) = (-1)^{G(\frac{x+1}{2})},$$

where $\frac{x+1}{2} = (\frac{x_1+1}{2}, \frac{x_2+1}{2}, \ldots, \frac{x_n+1}{2})$. Then we have

$$\begin{aligned} \sum_{x\in\{-1,1\}^n} f(x) \prod_{i\in S} x_S &= \sum_{x\in\{-1,1\}^n} (-1)^{G(\frac{x+1}{2})} \prod_{i\in S} (-1)^{\frac{x_i+1}{2}+1} \\ &= (-1)^{|S|} \sum_{y\in\mathbb{F}_2^n} (-1)^{G(y)} \prod_{i\in S} (-1)^{v_iy_i} \\ &= (-1)^{|S|} \sum_{y\in\mathbb{F}_2^n} (-1)^{G(y)\oplus v\cdot y} \\ &= 0,\ for\ |S| \ge d+1, \end{aligned}$$

where $v \in \mathbb{F}_2^n$ and $v_i = 1$ if and only if $i \in S$. Therefore, the total degree of f is at most d. By Conjecture 1.1, we have

$$\begin{aligned}
\sum_{i=1}^{n} \widehat{f}(i) &= \frac{1}{2^n} \sum_{i=1}^{n} \sum_{x\in\{-1,1\}^n} (-1)^{G(\frac{x+1}{2})}(-1)^{\frac{x_i+1}{2}+1} \\
&= \frac{1}{2^n} \sum_{i=1}^{n} \sum_{y\in\mathbb{F}_2^n} (-1)^{G(y)}(-1)^{y_i+1} \\
&= -\frac{1}{2^n} \sum_{i=1}^{n} \sum_{y\in\mathbb{F}_2^n} (-1)^{G(y)\oplus y_i\oplus y_1\oplus y_2\oplus\ldots\oplus y_n} \\
&\geq -d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{1-d},
\end{aligned}$$

and the result follows.

"$\Leftarrow$" It is known that Conjecture 1.1 holds for $d = n$. Let $f : \{-1,1\}^n \to \{-1,1\}$ be with total degree d, where $1 \leq d \leq n-1$. Then we define a function $g \in \mathcal{B}_n$ as

$$g(x) = \frac{f(1-2x)+1}{2} \oplus x_1 \oplus x_2 \oplus \ldots \oplus x_n.$$

It is easy to verify that g is $(n-d-1)$–resilient. Then by Conjecture 2.1, we have

$$\begin{aligned}
\sum_{\substack{v\in\mathbb{F}_2^n \\ wt(v)=n-1}} \sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} &= -\sum_{i=1}^{n} \sum_{y\in\{-1,1\}^n} f(y)y_i \\
&\geq -d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{n+1-d},
\end{aligned}$$

and the result follows. □

3 Proof of the conjecture for two cases

In this section, we will prove that Conjecture 2.1 holds for $d = 1$, $n-1$.

3.1 Case $d = 1$

Any $g \in \mathcal{B}_n$ can be written as a multivariate polynomial

$$g(x) = \bigoplus_{S\subseteq[n]} c_S \prod_{i\in S} x_i,$$

where $c_S \in \{0,1\}$. The algebraic degree of g is defined as the degree of this polynomial. It is well-known that the algebraic degree of an n-variable t-resilient Boolean function is at most $n-t-1$ [1, 11]. We state this as a lemma.

Lemma 3.1. *Let $g \in \mathcal{B}_n$ be t–resilient, where $0 \le t \le n-2$. Then the algebraic degree of g is at most $n-t-1$.*

Theorem 3.2. *Conjecture 2.1 holds for $d=1$. Moreover, the bound is achieved if and only if $g(x) = v \cdot x$, where $v \in \mathbb{F}_2^n$ and $wt(v) = n-1$.*

Proof. If $d=1$, then g is $(n-2)$-resilient. By Lemma 3.1, the algebraic degree of g is at most 1. That is, $g = a_0 \oplus a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n$, where $a_j \in \mathbb{F}_2$ and $0 \le j \le n$. Clearly, $g(x) \oplus v \cdot x$ is not balanced only when $(a_1, \ldots, a_n) = v$. Therefore,

$$\sum_{\substack{v \in \mathbb{F}_2^n \\ wt(v)=n-1}} \sum_{x \in \mathbb{F}_2^n} (-1)^{g(x) \oplus v \cdot x} \le \sum_{x \in \mathbb{F}_2^n} |(-1)^{a_0}| = 2^n.$$

Moreover, the equality holds if and only if $a_0 = 0$ and $(a_1, \ldots, a_n) = v$, and the result follows. □

Clearly, for $d=1$, there are exactly n functions achieving the bound.

Remark 3.3. *Naturally, one may generalize Conjecture 2.1 to the case when g is of algebraic degree d. However, the bound does not always hold in this case. For example, $g = x_2x_3 \oplus x_2x_4 \oplus x_3x_4 \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_4$ is a balanced function with algebraic degree 2. However,*

$$\sum_{i=1}^{4} \sum_{x \in \mathbb{F}_2^4} (-1)^{g(x) \oplus x_i \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_4} = 24 > d \cdot \binom{d-1}{\lfloor \frac{d-1}{2} \rfloor} 2^{n+1-d} = 16.$$

3.2 Case $d = n-1$

The following lemma gives three combinatorial formulas, which will be used afterwards.

Lemma 3.4. *The following three expressions are all equal to*

$$n \cdot 2^{n-2} + (n-1)\binom{n-2}{\lfloor \frac{n-2}{2} \rfloor}.$$

(i) for $n \geq 4$ even,

$$\sum_{i=0}^{\frac{n}{2}-1}(n-i)\binom{n}{i}+\frac{n}{4}\binom{n}{\frac{n}{2}};$$

(ii) for $n \geq 9$ and $mod(n,4)=1$,

$$2\sum_{i=0}^{\frac{n-5}{4}}(n-2i)\binom{n}{2i}+\frac{n+1}{2}(2^{n-1}-2\sum_{i=0}^{\frac{n-5}{4}}\binom{n}{2i});$$

(iii) for $n \geq 7$ and $mod(n,4)=3$,

$$2\sum_{i=0}^{\frac{n-3}{4}}(n-2i)\binom{n}{2i}+\frac{n-1}{2}(2^{n-1}-2\sum_{i=0}^{\frac{n-3}{4}}\binom{n}{2i}).$$

Proof. We only prove (i) and the other two formulas can be proved similarly. Since n is even, we have

$$\begin{aligned}\frac{n}{4}\binom{n}{\frac{n}{2}} &= \frac{n}{4}(2\binom{n-2}{\frac{n}{2}-1}+2\binom{n-2}{\frac{n}{2}-2})\\ &= \frac{n}{2}(\binom{n-2}{\frac{n}{2}-1}+\frac{n-2}{n}\binom{n-2}{\frac{n}{2}-1})\\ &= (n-1)\binom{n-2}{\frac{n}{2}-1}.\end{aligned}$$

Since $(1+x)^n=\sum_{i=0}^{n}\binom{n}{i}x^i$, the derivation

$$\frac{d}{dx}((1+x)^n)=n(1+x)^{n-1}=\sum_{i=1}^{n}i\binom{n}{i}x^{i-1}.$$

Therefore, $\sum_{i=1}^{n}i\binom{n}{i}=n\cdot 2^{n-1}$, and

$$\sum_{i=0}^{\frac{n}{2}-1}(n-i)\binom{n}{i}=n\cdot 2^{n-2},$$

and the result follows. □

Lemma 3.5. *Let $A_n=\mathbf{1}_n-I_n$ be the matrix over $\mathbb{F}_2$, where $\mathbf{1}_n$ is the $n\times n$ matrix whose elements are all 1, and I_n is the identity matrix. Then the rank of A_n is*

$$rank(A_n)=\begin{cases} n & \text{if } mod(n,2)=0,\\ n-1 & \text{otherwise,}\end{cases}$$

Proof. If $mod(n,2)=0$, then $A_n^2 = I_n$ and $rank(A_n) = n$. If $mod(n,2) = 1$, then the determinant of A_n is 0 and $rank(A_n) < n$. Since A_{n-1} is a submatrix of A_n, we have $rank(A_n) \geq rank(A_{n-1}) = n-1$, and the result follows. $\square$

Theorem 3.6. *Conjecture 2.1 holds for $d = n-1$. Moreover, the number of g's achieving the bound is $\binom{\binom{n}{\frac{n}{2}}}{\frac{1}{2}\binom{n}{\frac{n}{2}}}$, for n even,*

$$\binom{2\binom{n}{\frac{n+1}{2}}}{2^{n-1} - 2\sum_{i=0}^{\frac{n-5}{4}}\binom{n}{2i}}, \; for \; mod(n,4) = 1,$$

and

$$\binom{2\binom{n}{\frac{n+1}{2}}}{2^{n-1} - 2\sum_{i=0}^{\frac{n-3}{4}}\binom{n}{2i}}, \; for \; mod(n,4) = 3.$$

Proof. Since $d = n-1$, g is 0-resilient. That is, g is a balanced function. We use 0_g to denote the set $\{x \in \mathbb{F}_2^n : g(x) = 0\}$. Then $|0_g| = 2^{n-1}$. Clearly, If $v \neq 0$, then

$$\sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} = 2\sum_{x\in 0_g} (-1)^{v\cdot x} = 4|\{x \in 0_g : \; v\cdot x = 0\}| - 2^n.$$

Let $A = \mathbf{1}_n - I_n$, where $\mathbf{1}_n$ is the $n\times n$ matrix whose elements are all 1, and I_n is the identity matrix. Then

$$\begin{aligned}
&\sum_{\substack{v\in\mathbb{F}_2^n\\ wt(v)=n-1}} \sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} \\
= \;& 4 \sum_{\substack{v\in\mathbb{F}_2^n\\ wt(v)=n-1}} |\{x \in 0_g : \; v\cdot x = 0\}| - n\cdot 2^n \\
= \;& 4 \sum_{x\in 0_g} |\{v \in \mathbb{F}_2^n : wt(v) = n-1 \; and \; v\cdot x = 0\}| - n\cdot 2^n \\
= \;& 4 \sum_{b\in\mathbb{F}_2^n} \sum_{\substack{x\in 0_g\\ Ax=b}} (n - wt(b)) - n\cdot 2^n.
\end{aligned}$$

Case 1: n is even. Then by Lemma 3.5, A is invertible and $Ax = b$ has exactly one solution for any $b \in \mathbb{F}_2^n$. Therefore,

$$\begin{aligned}
&\sum_{b\in\mathbb{F}_2^n}\sum_{\substack{x\in 0_g\\ Ax=b}}(n-wt(b))\\
\le\ & n\binom{n}{0}+(n-1)\binom{n}{1}+\ldots+(\frac{n}{2}+1)\binom{n}{\frac{n}{2}-1}+\frac{n}{2}\frac{1}{2}\binom{n}{\frac{n}{2}},
\end{aligned}$$

and the number of g's such that the equality holds is $\binom{\binom{n}{\frac{n}{2}}}{\frac{1}{2}\binom{n}{\frac{n}{2}}}$. Then by Lemma 3.4,

$$\sum_{\substack{v\in\mathbb{F}_2^n\\ wt(v)=n-1}}\sum_{x\in\mathbb{F}_2^n}(-1)^{g(x)\oplus v\cdot x}\le 4(n-1)\cdot\binom{n-2}{\frac{n}{2}-1}.$$

Case 2: n is odd. Then by Lemma 3.5, the rank of A is $n-1$. Clearly, $Ax=b$ has two solutions if $wt(b)$ is even, and no solution otherwise. If $mod(n,4)=1$, then

$$\begin{aligned}
&\sum_{b\in\mathbb{F}_2^n}\sum_{\substack{x\in 0_g\\ Ax=b}}(n-wt(b))\\
\le\ & 2n\binom{n}{0}+2(n-2)\binom{n}{2}+\ldots+2(\frac{n+5}{2})\binom{n}{\frac{n-5}{2}}+\frac{n+1}{2}(2^{n-1}-2\sum_{i=0}^{\frac{n-5}{4}}\binom{n}{2i}),
\end{aligned}$$

and the number of g's such that the equality holds is

$$\binom{2\binom{n}{\frac{n-1}{2}}}{2^{n-1}-2\sum_{i=0}^{\frac{n-5}{4}}\binom{n}{2i}}.$$

If $mod(n,4)=3$, then

$$\begin{aligned}
&\sum_{b\in\mathbb{F}_2^n}\sum_{\substack{x\in 0_g\\ Ax=b}}(n-wt(b))\\
\le\ & 2n\binom{n}{0}+2(n-2)\binom{n}{2}+\ldots+2(\frac{n+3}{2})\binom{n}{\frac{n-3}{2}}+\frac{n-1}{2}(2^{n-1}-2\sum_{i=0}^{\frac{n-3}{4}}\binom{n}{2i}),
\end{aligned}$$

and the number of g's such that the equality holds is

$$\binom{2\binom{n}{\frac{n+1}{2}}}{2^{n-1}-2\sum_{i=0}^{\frac{n-3}{4}}\binom{n}{2i}}.$$

Then by Lemma 3.4,

$$\sum_{\substack{v\in\mathbb{F}_2^n \\ wt(v)=n-1}} \sum_{x\in\mathbb{F}_2^n} (-1)^{g(x)\oplus v\cdot x} \leq 4(n-1)\cdot \binom{n-2}{\frac{n-3}{2}},$$

and the result follows. □

4 Conclusion

In this paper, we transformed a problem in theoretical computer science to a problem in cryptography, and proved that the conjecture proposed by O'Donnell is equivalent to a conjecture on the cryptographic Boolean function. We proved that the conjecture is true for $d = 1, n-1$, and counted the number of f's such that the upper bound is achieved. We hope that our work would attract more researchers working on cryptographic Boolean functions to be interested in this conjecture.

Acknowledgment

The author would like to thank Dr. Timothy Foo for useful discussion. This work was supported by the National Natural Science Foundation of China (Grant 61572189).

References

[1] C. Carlet, "Boolean Functions for Cryptography and Error Correcting Codes," Chapter of the monography "Boolean Models and Methods in Mathematics, Computer Science, and Engineering", Cambridge University Press, pp. 257–397, 2010. Available: http://www-roc.inria.fr/secret/Claude.Carlet/pubs.html.

[2] C. Carlet and P. Charpin, "Cubic Boolean functions with highest resiliency", *IEEE Trans. Inform. Theory* 51:2 (2005), pp. 562–571.

[3] A. Canteaut, C. Carlet, P. Charpin and C. Fontaine, "Propagation Characteristics and Correlation-Immunity of Highly Nonlinear Boolean Functions", *Advances in Cryptology – EUROCRYPT 2000*, LNCS 1807, Springer–Verlag, 2000, pp. 507–522.

[4] P. Charpin and E. Pasalic, “Highly Nonlinear Resilient Functions Through Disjoint Codes in Projective Spaces”, *Des. Codes Cryptogr.* 37:2 (2005), pp. 319–346.

[5] R. O’Donnell, “Open problems in analysis of boolean functions”, arXiv preprint, arXiv:1204.6447, 2012.

[6] S. K. Jha, “On the Sum of Linear Coefficients of a Boolean Valued Function”, arXiv preprint, arXiv:1611.01029, 2016.

[7] T. Johansson and E. Pasalic, “A construction of resilient functions with high nonlinearity”, *IEEE Trans. Inform. Theory* 49:2 (2003), pp. 494–501.

[8] S. Maitra and P. Sarkar, “Highly Nonlinear Resilient Functions Optimizing Siegenthaler’s Inequality”, *Advances in Cryptology – CRYPTO 1999*, LNCS 1666, Springer–Verlag, 2000, pp. 198–215

[9] E. Pasalic and S. Maitra, “Linear codes in generalized construction of resilient functions with very high nonlinearity”, *IEEE Trans. Comput.*, 48:8 (2002), pp. 2182–2191.

[10] P. Sarkar and S. Maitra, “Nonlinearity Bounds and Constructions of Resilient Boolean Functions”, *Advances in Cryptology – CRYPTO 2000*, LNCS 1880, Springer–Verlag, 2000, pp. 515–532.

[11] T. Siegenthaler, “Correlation-immunity of nonlinear combining functions for cryptographic applications”, *IEEE Trans. on Inform. Theory*, 30:5 (1984), pp. 776–780.

[12] T. Siegenthaler, “Decrypting a Class of Stream Ciphers Using Ciphertext Only”, *IEEE Trans. Comput.*, 34:1 (1985), pp. 81–85.

[13] G. Z. Xiao and J. L. Massey, “A spectral characterization of correlation-immune combining functions,” *IEEE Trans. Inform. Theory* 34:3 (1988), pp. 569–571.

[14] W. Zhang and E. Pasalic, “Constructions of Resilient S-Boxes With Strictly Almost Optimal Nonlinearity Through Disjoint Linear Codes,” *IEEE Trans. Inform. Theory* 60:3 (2014), pp. 1638–1651.

[15] X. Zhang and Y. Zheng, “Cryptographically resilient functions,” *IEEE Trans. Inform. Theory* 43:5 (1997), pp. 1740–1747.