ACTA ARITHMETICA
172.3 (2016)

A large family of Boolean functions

by

HUANING LIU and MIN ZHANG (Xi'an)

1. Introduction. Boolean functions play an important role in stream ciphers, block ciphers and hash functions. Over the years, many researchers have dealt with cryptography criteria for Boolean functions.

Let $\mathbb{F}_2$ be the binary field; then $\mathbb{F}_2^n$ can be visualized as an n-dimensional vector space over $\mathbb{F}_2$. A *Boolean function* $B(x_1,\ldots,x_n)$ of n variables is a mapping from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. Let $\mathbf{x}=(x_1,\ldots,x_n)$, $\mathbf{a}=(a_1,\ldots,a_n)\in\mathbb{F}_2^n$ and let $\langle\mathbf{a},\mathbf{x}\rangle=a_1x_1+\cdots+a_nx_n$ denote the usual inner product. The Fourier coefficients $\widehat{B}(\mathbf{a})$ of $B(x_1,\ldots,x_n)$ are defined as

$$\widehat{B}(\mathbf{a})=\sum_{\mathbf{x}\in\mathbb{F}_2^n}(-1)^{B(\mathbf{x})+\langle\mathbf{a},\mathbf{x}\rangle}.$$

The *nonlinearity* of $B(x_1,\ldots,x_n)$ is defined by

$$\mathrm{nl}(B)=2^{n-1}-\tfrac{1}{2}\max_{\mathbf{a}\in\mathbb{F}_2^n}|\widehat{B}(\mathbf{a})|.$$

A Boolean function has a unique representation as a multivariate polynomial over $\mathbb{F}_2$, named the *algebraic normal form*:

$$B(x_1,\ldots,x_n)=\sum_{I\subseteq\{1,\ldots,n\}}a_I\prod_{i\in I}x_i.$$

The *algebraic degree* $\deg(B)$ is the number of variables in the highest order term with non-zero coefficient, and the *sparsity* $\mathrm{spr}(B)$ is the number of non-zero coefficients of B. The *average sensitivity* $\sigma_{\mathrm{av}}(B)$ is a measure of how the value of $B(x_1,\ldots,x_n)$ changes on average if the ith bit of the argument

---

2010 *Mathematics Subject Classification*: Primary 94C10; Secondary 94A60, 11T71, 11T24.
Key words and phrases: Boolean function, maximum Fourier coefficient, nonlinearity, average sensitivity, sparsity, collision, avalanche effect.
Received 12 March 2015; revised 15 September 2015.
Published online 20 January 2016.

is flipped, i.e.

$$\sigma_{\mathrm{av}}(B) = 2^{-n} \sum_{\mathbf{a}\in\mathbb{F}_2^n} \sum_{i=1}^{n} |B(\mathbf{a}) - B(\mathbf{a}^{(i)})|,$$

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

In recent years many Boolean functions with good cryptographic properties have been constructed by using number-theoretic methods. For example, D. Coppersmith and I. E. Shparlinski [2] constructed a Boolean function by using quadratic residues modulo an odd prime.

Proposition 1.1. *Let $p > 2$ be a prime, and let $s = \lfloor \log_2 p \rfloor$, where $\lfloor x \rfloor$ denotes the maximum integer not greater than x. Define*

$$B(u_1, \dots, u_s) = \begin{cases} 0 & \text{if } u_1 + u_2 \cdot 2 + \cdots + u_s \cdot 2^{s-1} \\ & \text{is a quadratic residue in } \mathbb{F}_p, \\ 1 & \text{if } u_1 + u_2 \cdot 2 + \cdots + u_s \cdot 2^{s-1} \\ & \text{is a quadratic non-residue in } \mathbb{F}_p, \end{cases} \tag{1.1}$$

where $u_j \in \{0, 1\}$ for $1 \le j \le s$. Then

$$\mathrm{spr}(B) \ge 2^{-3/2} p^{1/4} (\log_2 p)^{-1/2} - 1,$$
$$\sigma_{\mathrm{av}}(B) \ge 0.5s + o(s).$$

H. Aly and A. Winterhof [1] studied Boolean functions derived from Fermat quotients modulo p by using the Legendre symbol.

Proposition 1.2. *Let $p > 2$ be a prime. For an integer u with $(u, p) = 1$, the Fermat quotient $q_p(u)$ is defined as the unique integer with*

$$q_p(u) \equiv \frac{u^{p-1} - 1}{p} \pmod{p}, \qquad 0 \le q_p(u) \le p - 1.$$

Also define $q_p(kp) = 0$ for $k \in \mathbb{Z}$. Write $s = \lfloor 2 \log p \rfloor$. Set

$$B(u_1, \dots, u_s) = \begin{cases} 0 & \text{if } \left(\frac{q_p(u_1 + u_2\cdot 2 + \cdots + u_s \cdot 2^{s-1})}{p}\right) = 1, \\ 1 & \text{if } \left(\frac{q_p(u_1 + u_2\cdot 2 + \cdots + u_s \cdot 2^{s-1})}{p}\right) \ne 1, \end{cases} \tag{1.2}$$

where $u_j \in \{0, 1\}$ for $1 \le j \le s$, and $\left(\frac{\cdot}{p}\right)$ denotes the Legendre symbol. Then

$$\max_{\mathbf{a}\in\mathbb{F}_2^s} |\widehat{B}(\mathbf{a})| \ll p^{15/8} (\log p)^{1/4},$$
$$\mathrm{spr}(B) \gg p^{1/4} (\log p)^{-1/2},$$
$$\sigma_{\mathrm{av}}(B) \ge 0.5s + o(s).$$

T. Lange and A. Winterhof [4] extended the construction in Proposition 1.1.

PROPOSITION 1.3. *Let $p > 2$ be a prime and $r \ge 1$ be an integer. Let $\mathbb{F}_q$ denote the finite field of order $q = p^r$, and let $\beta_0, \dots, \beta_{r-1}$ be a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. Define $s = \lfloor \log_2 p \rfloor$. Let $u_{ij} \in \{0,1\}$ for $1 \le j \le s$ and $1 \le i \le r$. Write $k_{i-1} = u_{i1} + u_{i2} \cdot 2 + \dots + u_{is} \cdot 2^{s-1}$ for $1 \le i \le r$. Define*

$$(1.3)\quad B(u_{11}, \dots, u_{1s}, \dots, u_{r1}, \dots, u_{rs}) = \begin{cases} 0 & \text{if } k_0\beta_0 + \dots + k_{r-1}\beta_{r-1} \text{ is a square in } \mathbb{F}_q, \\ 1 & \text{if } k_0\beta_0 + \dots + k_{r-1}\beta_{r-1} \text{ is a non-square in } \mathbb{F}_q. \end{cases}$$

Then

$$\operatorname{spr}(B) \ge \big(2^{-3/2}(3^{1/r} + r)^{-1/2} p^{1/4}\big)^r - 1.$$

T. Lange and A. Winterhof [5] further studied the properties of the Boolean function (1.3).

PROPOSITION 1.4. *Let p, r, s, B be as in Proposition 1.3. Then*

$$\max_{\mathbf{a} \in \mathbb{F}_2^{rs}} |\widehat{B}(\mathbf{a})| \le 2^{(2r+3)/4} q^{7/8} (\ln p + 1)^{r/4} + 1,$$
$$\sigma_{\mathrm{av}}(B) \ge 0.5rs + o(rs).$$

Noting that the above constructions produce only a "few" good Boolean functions while in some applications one needs "large" families of Boolean functions, in this paper we construct a large family of Boolean functions by using polynomials over finite fields, and study their cryptographic properties.

THEOREM 1.1. *Let $\mathbb{F}_q$ be the finite field of order $q = p^r$ with p an odd prime and an integer $r \ge 1$, and let $\beta_0, \dots, \beta_{r-1}$ be linearly independent elements of $\mathbb{F}_q$ over $\mathbb{F}_p$. Define $s = \lfloor \log_2 p \rfloor$. Write $k_{i-1} = u_{i1} + u_{i2} \cdot 2 + \dots + u_{is} \cdot 2^{s-1}$ with $u_{ij} \in \{0,1\}$ for $1 \le j \le s$ and $1 \le i \le r$. Assume that $f(x) \in \mathbb{F}_q[x]$ has no multiple zero in $\overline{\mathbb{F}}_q$ and $0 < \deg(f) < p$. Define*

$$(1.4)\quad B(u_{11}, \dots, u_{1s}, \dots, u_{r1}, \dots, u_{rs}) = \begin{cases} 0 & \text{if } f(k_0\beta_0 + \dots + k_{r-1}\beta_{r-1}) \text{ is a square in } \mathbb{F}_q, \\ 1 & \text{if } f(k_0\beta_0 + \dots + k_{r-1}\beta_{r-1}) \text{ is a non-square in } \mathbb{F}_q. \end{cases}$$

Then

$$(1.5)\quad \max_{\mathbf{a} \in \mathbb{F}_2^{rs}} |\widehat{B}(\mathbf{a})| \le 2^{3/4} (\deg(f))^{1/4} q^{7/8} (1 + \log p)^{r/4} + \deg(f),$$

$$(1.6)\quad \operatorname{nl}(B) \ge \frac{q}{2^{r+1}} - 2^{-1/4} (\deg(f))^{1/4} q^{7/8} (1 + \log p)^{r/4} - \tfrac{1}{2}\deg(f).$$

Furthermore, assume that

$$r\big(\deg(f) + r(2(\log_2 p)^{1/2} + 1)\big) < \log_4 p.$$

Then also

$$(1.7)\quad \sigma_{\mathrm{av}}(B) \ge 0.5rs + o(rs).$$

Theorem 1.2. *Let $p > 2$ be a prime, and let $s = \lfloor \log_2 p \rfloor$. Suppose that $f(x) \in \mathbb{F}_p[x]$ has no multiple zero in $\overline{\mathbb{F}}_p$ and $0 < \deg(f) < p$. Define*

$$B(u_1, \dots, u_s) = \begin{cases} 0 & \text{if } f(u_1 + u_2 \cdot 2 + \cdots + u_s \cdot 2^{s-1}) \\ & \text{is a square in } \mathbb{F}_p, \\ 1 & \text{if } f(u_1 + u_2 \cdot 2 + \cdots + u_s \cdot 2^{s-1}) \\ & \text{is not a square in } \mathbb{F}_p. \end{cases} \tag{1.8}$$

If 2 is a primitive root modulo p, then

$$\sigma_{\mathrm{av}}(B) \geq 0.5s + o(s), \tag{1.9}$$

$$\mathrm{spr}(B) \geq \tfrac{1}{4}(\deg(f))^{-1/2} p^{1/4} (\log p)^{-1/2}. \tag{1.10}$$

We further study the properties of our family of Boolean functions. Collision and avalanche effect are important notions in cryptography (see [8]), and can be adapted in the following way.

Assume that $\mathcal{T}$ is a given set (e.g., a set of polynomials) and for each $t \in \mathcal{T}$ we have a unique Boolean function

$$B(x_1, \dots, x_n) = B^{(t)}(x_1, \dots, x_n) \in \{0,1\}^{2^n};$$

let $\mathcal{F} = \mathcal{F}(\mathcal{T})$ be the family of all these functions:

$$\mathcal{F} = \mathcal{F}(\mathcal{T}) = \{B^{(t)} : t \in \mathcal{T}\}. \tag{1.11}$$

Definition 1.1. If $t_1, t_2 \in \mathcal{T}$, $t_1 \neq t_2$ and

$$B^{(t_1)} = B^{(t_2)}, \tag{1.12}$$

then (1.12) is said to be a *collision* in $\mathcal{F} = \mathcal{F}(\mathcal{T})$. If there is no collision in $\mathcal{F} = \mathcal{F}(\mathcal{T})$, then $\mathcal{F}$ is said to be *collision free.*

Definition 1.2. If for any $t \in \mathcal{T}$, changing any value of t changes "many" elements of $B^{(t)}$ (i.e. for $t_1 \neq t_2$ many values of $B^{(t_1)}$ and $B^{(t_2)}$ are different), then we speak about the *avalanche effect*, and we say that $\mathcal{F} = \mathcal{F}(\mathcal{T})$ has the *avalanche property.* If for any $t_1, t_2 \in \mathcal{T}$, $t_1 \neq t_2$, at least $(1/2 - o(1))2^n$ values of $B^{(t_1)}$ and B^{t_2} are different, then $\mathcal{F}$ is said to have the *strict avalanche property.*

To study the collision and avalanche effect, we introduce the following measure (see [9]).

Definition 1.3. If $n \in \mathbb{N}$, $B(\mathbf{x}) \in \{0,1\}^{2^n}$ and $B'(\mathbf{x}) \in \{0,1\}^{2^n}$, then the distance $d(B, B')$ between B and B' is defined by

$$d(B, B') = |\{\mathbf{x} \in \mathbb{F}_2^n : B(\mathbf{x}) \neq B'(\mathbf{x})\}|.$$

If $\mathcal{F} = \mathcal{F}(\mathcal{T})$ is a family of the form (1.11), then the *minimum distance*

$m(\mathcal{F})$ of $\mathcal{F}$ is defined by

$$m(\mathcal{F}) = \min_{\substack{t_1,t_2\in\mathcal{T}\\ t_1\neq t_2}} d(B^{(t_1)}, B^{(t_2)}).$$

It is easy to show that the family $\mathcal{F}$ is collision free if and only if $m(\mathcal{F}) > 0$, and $\mathcal{F}$ has the strict avalanche property if

$$m(\mathcal{F}) \geq (1/2 - o(1))2^n.$$

In Section 6 we will study the collision and avalanche effect of our family of Boolean functions, and prove the following results.

THEOREM 1.3. *Let $\mathbb{F}_q$ be the finite field of order $q = p^r$ with p an odd prime and an integer $r \geq 1$, and let $\beta_0, \ldots, \beta_{r-1}$ be linearly independent elements of $\mathbb{F}_q$ over $\mathbb{F}_p$. Define $s = \lfloor \log_2 p \rfloor$. Write $k_{i-1} = u_{i1} + u_{i2}\cdot 2 + \cdots + u_{is}\cdot 2^{s-1}$ with $u_{ij} \in \{0,1\}$ for $1 \leq j \leq s$ and $1 \leq i \leq r$. Let $\mathcal{T}$ be the set of polynomials $f(x) \in \mathbb{F}_q[x]$ with $1 \leq \deg(x) \leq D$ which do not have multiple zeros. Define*

$$\begin{aligned} B^{(f)} &= B(u_{11}, \ldots, u_{1s}, \ldots, u_{r1}, \ldots, u_{rs}) \\ &= \begin{cases} 0 & \text{if } f(k_0\beta_0 + \cdots + k_{r-1}\beta_{r-1}) \text{ is a square in } \mathbb{F}_q, \\ 1 & \text{if } f(k_0\beta_0 + \cdots + k_{r-1}\beta_{r-1}) \text{ is a non-square in } \mathbb{F}_q, \end{cases} \end{aligned}$$

and $\mathcal{F} = \mathcal{F}(\mathcal{T}) = \{B^{(f)} : f \in \mathcal{T}\}$. Then

$$m(\mathcal{F}) \geq \tfrac{1}{2}\big(2^{rs} - 2Dq^{1/2}\,(1+\log p)^r - 2D\big).$$

COROLLARY 1.1. *If $\mathcal{T}$ and $\mathcal{F}$ are as in Theorem 1.3 and*

$$D < \frac{1}{2^{r+2}} q^{1/2}(1+\log p)^{-r},$$

then $\mathcal{F}$ is collision free.

COROLLARY 1.2. *If $\mathcal{T}$ and $\mathcal{F}$ are as in Theorem 1.3 and*

$$D = o\big(q^{1/2}(1+\log p)^{-r}\big),$$

then $\mathcal{F}$ has the strict avalanche property.

2. The maximum Fourier coefficient and nonlinearity. First we list the following lemmas.

LEMMA 2.1 ([10, Theorem 2]). *Suppose that $q = p^n$, χ is a multiplicative character on $\mathbb{F}_q$ of order $d > 1$, $v_1, \ldots, v_n \in \mathbb{F}_q$ are linearly independent over the prime field of $\mathbb{F}_q$, $f \in \mathbb{F}_q$ is a non-constant polynomial which is not a dth power and which has m distinct zeros in its splitting field over $\mathbb{F}_q$, and $t_1, \ldots, t_n$ are non-negative integers with $t_1 < p, \ldots, t_n < p$. Define*

$$B = \Big\{ \sum_{i=1}^{n} j_i v_i : 0 \leq j_i \leq t_i \text{ for } i = 1, \ldots, n \Big\}.$$

Then

$$\Big|\sum_{z\in B}\chi(f(z))\Big| < mq^{1/2}(1+\log p)^n.$$

Lemma 2.2 ([7, Lemma 2 and Theorem 2]). *Let $q=p^n$, $z_1,\dots,z_k$ be distinct elements of $\mathbb{F}_q$, $h(x)\in\mathbb{F}_q[x]$ with $h(x)=ah_1(x)$, where $a\in\mathbb{F}_q$ and $h_1(x)$ is a monic polynomial. Define $H(x)=h_1(x+z_1)\cdots h_1(x+z_k)$. If $h(x)$ has no multiple zero in $\overline{\mathbb{F}}_q$, $0<\deg(h)<p$, and $k=2$ or $4^{n(\deg(h)+k)}<p$, then $H(x)$ has at least one zero in $\overline{\mathbb{F}}_q$ whose multiplicity is odd.*

Now we study the maximum Fourier coefficient of our Boolean functions. Write $k_{i-1}=u_{i1}+u_{i2}\cdot 2+\cdots+u_{is}\cdot 2^{s-1}$ with $u_{ij}\in\{0,1\}$ for $1\le j\le s$ and $1\le i\le r$. Let χ be the quadratic character of $\mathbb{F}_q$. It is obvious that $(-1)^{B(u_{11},\dots,u_{1s},\dots,u_{r1},\dots,u_{rs})}=\chi(f(k_0\beta_0+\cdots+k_{r-1}\beta_{r-1}))$ for $f(k_0\beta_0+\cdots+k_{r-1}\beta_{r-1})\neq 0$.

Define

$$\mathcal{H}_{2^s}=\{k_0\beta_0+\cdots+k_{r-1}\beta_{r-1}: 0\le k_{i-1}\le 2^s-1 \text{ for } i=1,\dots,r\}.$$

For any $\mathbf{a}\in\mathbb{F}_2^{rs}$, we have

$$\widehat{B}(\mathbf{a})=\sum_{\mathbf{x}\in\mathbb{F}_2^{rs}}(-1)^{B(\mathbf{x})+\langle\mathbf{a},\mathbf{x}\rangle}=\sum_{\substack{z\in\mathcal{H}_{2^s}\\ f(z)\neq 0}}\chi(f(z))(-1)^{\langle z,\mathbf{a}\rangle}+\sum_{\substack{z\in\mathcal{H}_{2^s}\\ f(z)=0}}(-1)^{\langle z,\mathbf{a}\rangle},$$

where $z=k_0\beta_0+\cdots+k_{r-1}\beta_{r-1}$, $k_{i-1}=u_{i1}+u_{i2}\cdot 2+\cdots+u_{is}\cdot 2^{s-1}$, $1\le i\le r$, and

$$\langle z,\mathbf{a}\rangle=\langle(u_{11},\dots,u_{1s},\dots,u_{r1},\dots,u_{rs}),\mathbf{a}\rangle.$$

Denote

$$S(\mathbf{a})=\sum_{z\in\mathcal{H}_{2^s}}\chi(f(z))(-1)^{\langle z,\mathbf{a}\rangle}.$$

We get

$$\begin{aligned}|\widehat{B}(\mathbf{a})| &\le \Big|\sum_{z\in\mathcal{H}_{2^s}}\chi(f(z))(-1)^{\langle z,\mathbf{a}\rangle}\Big|+\deg(f)\\ &=|S(\mathbf{a})|+\deg(f).\end{aligned}\tag{2.1}$$

Let x be an integer with $1<x<s$. Then

$$\begin{aligned}k_{i-1}&=u_{i1}+u_{i2}\cdot 2+\cdots+u_{is}\cdot 2^{s-1}\\ &=u_{i1}+u_{i2}\cdot 2+\cdots+u_{ix}\cdot 2^{x-1}\\ &\quad+u_{i(x+1)}\cdot 2^x+u_{i(x+2)}\cdot 2^{x+1}+\cdots+u_{is}\cdot 2^{s-1}\\ &=u_{i1}+u_{i2}\cdot 2+\cdots+u_{ix}\cdot 2^{x-1}\\ &\quad+2^x(u_{i(x+1)}+u_{i(x+2)}\cdot 2+\cdots+u_{is}\cdot 2^{s-x-1}).\end{aligned}$$

So any $z \in \mathcal{H}_{2^s}$ can be uniquely written as $z = y + w$, where $y \in \mathcal{H}_{2^x}$, and

$$w \in 2^x \mathcal{H}_{2^{s-x}} = \{2^x(k_0\beta_0 + \cdots + k_{r-1}\beta_{r-1}) : 0 \le k_{i-1} \le 2^{s-x} - 1 \quad \text{for } i = 1, \ldots, r\}.$$

Suppose that

$$\begin{aligned}
\mathbf{a} &= (a_{11}, \ldots, a_{1s}, \ldots, a_{r1}, \ldots, a_{rs}),\\
\mathbf{b} &= (a_{11}, \ldots, a_{1x}, \ldots, a_{r1}, \ldots, a_{rx}),\\
\mathbf{c} &= (a_{1(x+1)}, \ldots, a_{1s}, \ldots, a_{r(x+1)}, \ldots, a_{rs}).
\end{aligned}$$

It is obvious that $\langle z, \mathbf{a}\rangle = \langle y, \mathbf{b}\rangle + \langle w, \mathbf{c}\rangle$. By the Cauchy–Schwarz inequality we have

$$\begin{aligned}
|S(\mathbf{a})|^2 &= \Big| \sum_{y \in \mathcal{H}_{2^x}} \sum_{w \in 2^x \mathcal{H}_{2^{s-x}}} \chi(f(y+w))(-1)^{\langle y,\mathbf{b}\rangle + \langle w,\mathbf{c}\rangle} \Big|^2\\
&\le \Big(\sum_{y \in \mathcal{H}_{2^x}} \Big| \sum_{w \in 2^x \mathcal{H}_{2^{s-x}}} \chi(f(y+w))(-1)^{\langle w,\mathbf{c}\rangle} \Big| \Big)^2\\
&\le 2^{rx} \sum_{y \in \mathcal{H}_{2^x}} \Big| \sum_{w \in 2^x \mathcal{H}_{2^{s-x}}} \chi(f(y+w))(-1)^{\langle w,\mathbf{c}\rangle} \Big|^2\\
&= 2^{rx} \sum_{y \in \mathcal{H}_{2^x}} \sum_{w_1 \in 2^x \mathcal{H}_{2^{s-x}}} \sum_{w_2 \in 2^x \mathcal{H}_{2^{s-x}}} \chi(f(y+w_1)f(y+w_2))\\
&\qquad \times (-1)^{\langle w_1,\mathbf{c}\rangle + \langle w_2,\mathbf{c}\rangle}\\
&\le 2^{rx} \sum_{w_1 \in 2^x \mathcal{H}_{2^{s-x}}} \sum_{w_2 \in 2^x \mathcal{H}_{2^{s-x}}} \Big| \sum_{y \in \mathcal{H}_{2^x}} \chi(f(y+w_1)f(y+w_2)) \Big|\\
&\le 2^{rx+rs} + 2^{rx} \sum_{w_1 \in 2^x \mathcal{H}_{2^{s-x}}} \sum_{\substack{w_2 \in 2^x \mathcal{H}_{2^{s-x}}\\ w_1 \ne w_2}} \Big| \sum_{y \in \mathcal{H}_{2^x}} \chi(f(y+w_1)f(y+w_2)) \Big|.
\end{aligned}$$

Then from Lemmas 2.1 and 2.2 we get

$$|S(\mathbf{a})|^2 < 2^{rx+rs} + 2^{rx} \cdot 2^{2r(s-x)} \cdot 2\deg(f) q^{1/2} (1 + \log p)^r.$$

Taking x such that

$$2^{2rx} = 2^{rs} \cdot 2\deg(f) q^{1/2} (1 + \log p)^r,$$

we have

$$\begin{aligned}
|S(\mathbf{a})|^2 &< 2 \cdot 2^{rs} \cdot \big(2^{rs} \cdot 2\deg(f) q^{1/2} (1+\log p)^r\big)^{1/2} \qquad (2.2)\\
&< 2^{3/2} (\deg(f))^{1/2} q^{7/4} (1 + \log p)^{r/2}.
\end{aligned}$$

Combining (2.1) and (2.2) we immediately get

$$|\widehat{B}(\mathbf{a})| \le 2^{3/4} (\deg(f))^{1/4} q^{7/8} (1 + \log p)^{r/4} + \deg(f).$$

This proves (1.5). Note that $s = \lfloor \log_2 p \rfloor > \log p - 1$. Thus we have

$$\begin{aligned}\mathrm{nl}(B) &= 2^{rs-1} - \tfrac{1}{2}\max_{\mathbf{a}\in\mathbb{F}_2^{rs}} |\widehat{B}(\mathbf{a})| \\ &> 2^{r\log p - r - 1} - 2^{-1/4}(\deg(f))^{1/4} q^{7/8}(1+\log p)^{r/4} - \tfrac{1}{2}\deg(f) \\ &= \frac{q}{2^{r+1}} - 2^{-1/4}(\deg(f))^{1/4} q^{7/8}(1+\log p)^{r/4} - \tfrac{1}{2}\deg(f).\end{aligned}$$

This proves (1.6).

3. The average sensitivity: Case $\mathbb{F}_q$. The ideas in the proof of (1.7) come from [5, proof of Theorem 1], thus we will omit the details. Write

$$M = \lfloor s^{1/2} \rfloor, \quad H = 2M+1, \quad J = \lfloor s - s^{1/2} \rfloor, \quad K = 2^s - H2^J.$$

Write $B'(k) = B(u_{11}, \dots, u_{1s}, \dots, u_{r1}, \dots, u_{rs})$ if

$$k = k_1 + k_2 p + \cdots + k_r p^{r-1}, \quad 0 \le k_i \le p-1,\ 1 \le i \le r,$$

and

$$k_i = u_{i1} + u_{i2}\cdot 2 + \cdots + u_{is}\cdot 2^{s-1} \quad \text{with } u_{ij} \in \{0,1\}$$

for $1 \le j \le s$, $1 \le i \le r$. Define

$$\mathcal{H}'_K = \{k_1 + k_2 p + \cdots + k_r p^{r-1} : 0 \le k_{i-1} \le K-1 \text{ for } i = 1, \dots, r\}.$$

Note that

$$r\big(\deg(f) + r(2(\log_2 p)^{1/2} + 1)\big) < \log_4 p.$$

Thus from Lemmas 2.1, 2.2 and the methods of [5, Theorem 1] we have

$$\begin{aligned}\sigma_{\mathrm{av}}(B) &= 2^{-rs}\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{\substack{k\in\mathcal{H}'_{2^s}\\ B'(k)\ne B'(k^{(ij)})}} 1 \ge 2^{-rs}\sum_{i=1}^{r}\sum_{j=1}^{J}\sum_{\substack{k\in\mathcal{H}'_{2^s}\\ B'(k)\ne B'(k^{(ij)})}} 1 \\ &= 2^{-rs}M^{-1}\Big(\sum_{i=1}^{r}\sum_{j=1}^{J}\sum_{h=1}^{M}\Big|\sum_{\substack{k\in\mathcal{H}'_K\\ B'(k+h2^jp^{i-1})\ne B'((k+h2^jp^{i-1})^{(ij)})}} 1 - \sum_{\substack{k\in\mathcal{H}'_{2^s}\\ B'(k)\ne B'(k^{(ij)})}} 1\Big| \\ &\qquad + \sum_{i=1}^{r}\sum_{j=1}^{J}\sum_{k\in\mathcal{H}'_K}\sum_{\substack{h=1\\ B'(k+h2^jp^{i-1})\ne B'((k+h2^jp^{i-1})^{(ij)})}}^{M} 1\Big) \\ &\ge 2^{-rs}M^{-1}(o(rJM2^{rs}) + 0.5JK^r rM + o(JK^r rM)) \ge 0.5rs + o(rs).\end{aligned}$$

This completes the proof of (1.7).

4. The average sensitivity: Case $\mathbb{F}_p$. We will need the following lemmas.

LEMMA 4.1 ([6, Theorem 2]). *Suppose that p is a prime number, χ is a non-principal character modulo p of order d, and $f(x) \in \mathbb{F}_p[x]$ has degree k and a factorization $f(x) = b(x-x_1)^{d_1} \cdots (x-x_s)^{d_s}$ (where $x_i \neq x_j$ for $i \neq j$) in $\overline{\mathbb{F}}_p$ with $(d, d_1, \ldots, d_s) = 1$. Let X and Y be real numbers with $0 < Y \le p$. Then*

$$\Big| \sum_{X<n\le X+Y} \chi(f(n)) \Big| < 9kp^{1/2} \log p.$$

LEMMA 4.2 ([3]). *Assume that p is a prime, and $f(x) \in \mathbb{F}_p[x]$ has degree k (> 0) and no multiple zero in $\overline{\mathbb{F}}_p$. Suppose that for $l \in \mathbb{N}$ one of the following assumptions holds:*

(i) $l = 2$; (ii) $l < p$, *and* 2 *is a primitive root modulo* p; (iii) $(4k)^l < p$.

Let $d_1, \ldots, d_l$ be distinct elements of $\mathbb{F}_p$. Then

$$H(x) = f(x+d_1) \cdots f(x+d_l)$$

has at least one zero in $\overline{\mathbb{F}}_p$ whose multiplicity is odd.

Now we use the methods of [2, Theorem 6] to prove (1.9). Set

$$m = \lfloor s^{1/2} \rfloor, \qquad k = 2m+1, \qquad l = \lfloor s - s^{1/2} \rfloor, \qquad R = 2^s - k2^l.$$

Write $B(x) = B(u_1, \ldots, u_s)$ if $x = u_1 + u_2 \cdot 2 + \cdots + u_s \cdot 2^{s-1}$. Thus from Lemmas 4.1, 4.2 and the methods of [2, Theorem 6] we have

$$\begin{aligned}
\sigma_{\mathrm{av}}(B) &= 2^{-s} \sum_{i=1}^{s} \sum_{\substack{x=0 \\ B(x) \neq B(x^{(i)})}}^{2^s-1} 1 \ge 2^{-s} \sum_{i=1}^{l} \sum_{\substack{x=0 \\ B(x) \neq B(x^{(i)})}}^{2^s-1} 1 \\
&= 2^{-s} m^{-1} \Big(\sum_{i=1}^{l} \sum_{j=1}^{m} \Big| \sum_{\substack{x=0 \\ B(x+j2^{i+1}) \neq B((x+j2^{i+1})^{(i)})}}^{R-1} 1 - \sum_{\substack{x=0 \\ B(x) \neq B(x^{(i)})}}^{2^s-1} 1 \Big| \\
&\qquad + \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{\substack{x=0 \\ B(x+j2^{i+1}) \neq B((x+j2^{i+1})^{(i)})}}^{R-1} 1 \Big) \\
&\ge 2^{-s} m^{-1} (o(lm2^s) + 0.5Rlm + o(Rlm)) \ge 0.5s + o(s).
\end{aligned}$$

This proves (1.9).

5. The sparsity: Case $\mathbb{F}_p$. Define the integer a by $2^a > \mathrm{spr}(B) \ge 2^{a-1}$. For each $m \in \{0, 1, \ldots, 2^a - 1\}$ with

$$m = m_1 + m_2 \cdot 2 + \cdots + m_a \cdot 2^{a-1},$$

we consider the function

$$B_m(u_1,\dots,u_{s-a}) = B(u_1,\dots,u_{s-a},m_1,\dots,m_a).$$

It is obvious that the number of distinct monomials in $u_1,\dots,u_{s-a}$ occurring in all the B_m does not exceed $\mathrm{spr}(B)$. Note that $2^a > \mathrm{spr}(B)$. Thus one can find a non-trivial linear combination

$$\sum_{m=0}^{2^a-1} c_m B_m(u_1,\dots,u_{s-a}), \qquad c_1,\dots,c_{2^a-1}\in\mathbb{F}_2,$$

which vanishes identically.

Let χ be the quadratic character of $\mathbb{F}_q$. Note that

$$(-1)^{B(u_1,\dots,u_s)} = \chi\big(f(u_1+u_2\cdot 2+\dots+u_s\cdot 2^{s-1})\big)$$

for $f(u_1+u_2\cdot 2+\dots+u_s\cdot 2^{s-1})\neq 0$. Thus from Lemmas 4.1 and 4.2 we have

$$\begin{aligned}
2^{s-a} &= \sum_{y=0}^{2^{s-a}-1} (-1)^{\sum_{m=0}^{2^a-1} c_m B_m(y)}\\
&= \sum_{\substack{y=0\\ f(y+2^a m)\neq 0}}^{2^{s-a}-1} \prod_{m=0}^{2^a-1} \chi\,(f(y+2^a m))^{c_m} + \sum_{\substack{y=0\\ f(y+2^a m)=0}}^{2^{s-a}-1} \prod_{m=0}^{2^a-1} 1\\
&\le \Big|\sum_{y=0}^{2^{s-a}-1} \chi\Big(\prod_{m=0}^{2^a-1} f(y+2^a m)^{c_m}\Big)\Big| + \deg(f)\\
&\le 2^a \deg(f) p^{1/2}\log p + \deg(f) \le 2\deg(f)2^a p^{1/2}\log p.
\end{aligned}$$

Noting that $2^s = 2^{\lfloor \log_2 p\rfloor} \ge 2^{\log_2 p-1} = p/2$, we have

$$2^a \ge (4\deg(f)\log p)^{-1/2} p^{1/4}.$$

Therefore

$$\mathrm{spr}(B) \ge 2^{a-1} \ge \tfrac14 (\deg(f))^{-1/2} p^{1/4} (\log p)^{-1/2}.$$

This completes the proof of (1.10).

6. Collision and avalanche effect. Assume that $f,g\in\mathcal{T}$ and $f\neq g$. For $\mathbf{x}\in\mathbb{F}_2^{rs}$, it is easy to show that

$$\tfrac12\big(1-(-1)^{B^{(f)}(\mathbf{x})+B^{(g)}(\mathbf{x})}\big) = \begin{cases} 0 & \text{if } B^{(f)}(\mathbf{x}) = B^{(g)}(\mathbf{x}),\\ 1 & \text{if } B^{(f)}(\mathbf{x}) \neq B^{(g)}(\mathbf{x}).\end{cases}$$

Define

$$\mathcal{H}_{2^s} = \{k_0\beta_0+\dots+k_{r-1}\beta_{r-1} : 0\le k_{i-1}\le 2^s-1 \text{ for } i=1,\dots,r\},$$

and let χ be the quadratic character of $\mathbb{F}_q$. It follows that

$$\begin{aligned} d(B^{(f)}, B^{(g)}) &= \sum_{\mathbf{x}\in\mathbb{F}_2^{rs}} \tfrac{1}{2}\big(1-(-1)^{B^{(f)}(\mathbf{x})+B^{(g)}(\mathbf{x})}\big) \\ &= \frac{1}{2}\Big(2^{rs} - \sum_{\mathbf{x}\in\mathbb{F}_2^{rs}} (-1)^{B^{(f)}(\mathbf{x})+B^{(g)}(\mathbf{x})}\Big) \\ &\geq \frac{1}{2}\Big(2^{rs} - \sum_{\substack{z\in\mathcal{H}_{2s}\\ f(z)g(z)\neq 0}} \chi\big(f(z)g(z)\big) - 2D\Big) \\ &= \frac{1}{2}\Big(2^{rs} - \sum_{z\in\mathcal{H}_{2s}} \chi\big(f(z)g(z)\big) - 2D\Big). \end{aligned}$$

Note that $f\neq g$, and f, g have no multiple zeros. Thus fg is not the constant multiple of the square of a polynomial over $\mathbb{F}_q$. By Lemma 2.1 we immediately get

$$d(B^{(f)}, B^{(g)}) \geq \tfrac{1}{2}\big(2^{rs} - 2Dq^{1/2}(1+\log p)^r - 2D\big).$$

Therefore

$$m(\mathcal{F}) = \min_{\substack{f,g\in\mathcal{T}\\ f\neq g}} d(B^{(f)}, B^{(g)}) \geq \tfrac{1}{2}\big(2^{rs} - 2Dq^{1/2}(1+\log p)^r - 2D\big).$$

This proves Theorem 1.3.

If $D < 2^{-r-2}q^{1/2}\,(1+\log p)^{-r}$, then

$$m(\mathcal{F}) \geq \tfrac{1}{2}\big(2^{rs} - 2Dq^{1/2}(1+\log p)^r - 2D\big) > 0,$$

and thus $\mathcal{F}$ is collision free. Furthermore, if $D = o(q^{1/2}(1+\log p)^{-r})$, then Theorem 1.3 gives

$$m(\mathcal{F}) \geq (1-o(1))\,2^{rs},$$

which means that $\mathcal{F}$ has the strict avalanche property. This completes the proof of Corollaries 1.1 and 1.2.

Acknowledgements. This work is supported by National Natural Science Foundation of China under Grant Nos. 11571277 and 11201370, the Science and Technology Program of Shaanxi Province of China under Grant Nos. 2013JM1017, 2014JM1007, 2014JQ2-1006 and 2014KJXX-61, and the Natural Science Foundation of the Education Department of Shaanxi Province of China under Grant No. 2013JK0558.

The authors express their gratitude to the referee for his/her helpful and detailed comments.

References

[1] H. Aly and A. Winterhof, *Boolean functions derived from Fermat quotients*, Cryptogr. Comm. 3 (2011), 165–174.

[2] D. Coppersmith and I. E. Shparlinski, *On polynominal approximation of the discrete logarithm and the Diffie–Hellman mapping*, J. Cryptology 13 (2000), 339–360.

[3] L. Goubin, C. Mauduit and A. Sárközy, *Construction of large families of pseudorandom binary sequences*, J. Number Theory 106 (2004), 56–69.

[4] T. Lange and A. Winterhof, *Incomplete character sums over finite fields and their application to the interpolation of the discrete logarithm by Boolean functions*, Acta Arith. 101 (2002), 223–229.

[5] T. Lange and A. Winterhof, *Interpolation of the discrete logarithm in $\mathbb{F}_q$ by Boolean functions and by polynomials in several variables modulo a divisor of $q-1$*, Discrete Appl. Math. 128 (2003), 193–206.

[6] C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol*, Acta Arith. 82 (1997), 365–377.

[7] C. Mauduit and A. Sárközy, *On large families of pseudorandom binary lattices*, Unif. Distrib. Theory 2 (2007), 23–37.

[8] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1996.

[9] V. Tóth, *Collision and avalanche effect in families of pseudorandom binary sequences*, Period. Math. Hungar. 55 (2007), 185–196.

[10] A. Winterhof, *Some estimates for character sums and applications*, Des. Codes Cryptogr. 22 (2001), 123–131.

Huaning Liu, Min Zhang
School of Mathematics
Northwest University
Xi'an 710127, Shaanxi, P.R. China
E-mail: hnliumath@hotmail.com
zhangmindada@163.com