Further Constructions of Almost Resilient Functions *

Pinhui Ke[1,2†] , Jie Zhang[1], Qiaoyan Wen[1]
1.School of Science,Beijing University of Posts and Telecommunications
Beijing, 100876, P. R. China
2.School of Mathematics and Computer Science,Fujian Normal University
Fujian 350007, P. R. China
keph@eyou.com

Abstract

Almost resilient function is the generalization of resilient function and have important applications in multiple authenticate codes and almost security cryptographic Boolean functions.In this paper,some secondary constructions are provided.In particular, the theorem 3 in [6] is improved. As ε-almost$(n, 1, k)$-CI functions plays an important role in the secondary constructions, we concluded some properties and constructions. Specially we presented a spectrum characterization of balanced almost CI function, which can be used to identify a balanced almost CI function by computing its walsh spectra.

Keywords: Almost resilient Functions; Resilient function; Almost correlation immune function.

1 Introduction

A Boolean functions is a map from F_2^n to F_2 and by a multi-output Boolean functions we mean a map from F_2^n to F_2^m.They are used as basic primitives for designing ciphers. In order to resist known attacks, several criteria of Boolean functions have been developed. However there are some tradeoffs between these criteria. Strict fulfillment in one criterion may lead to weaken another one. For example, bent functions have the best nonlinearity, but they are never balanced and correlation-Immune. So we may relax the definition's conditions and functions with better parameters could be obtained.

The concept of a resilient function was first introduced by Chor et al.[1], which have been found to be applicable in fault-tolerant distribute computing,quantum cryptographic key distribution and so on.K.Kurosawa et al.[2] generalized the concept and introduced the definition of almost resilient function. An ε-almost(n, m, k)-resilient function is an n-input m-output function f with the property that the deviation of output's distribution from uniform distribution is not great than ε when k arbitrary inputs are fixed and the remaining $n-k$ inputs run through all the 2^{n-k} input tuples. It was showed to have parameters superior to resilient functions. The notations of independent sample space was introduced by Naor and Naor [3], which has been proved to have many cryptographic applications, such as multiple authentication codes[4], almost security cryptographic boolean functions[5] and so on. In [2], the relations between the almost resilient functions and the large sets of almost independent sample spaces were established. Recently,the relation between almost resilient function and its component functions was investigated in [6].

*Research supported by the National Natural Science Foundation of China(60373059), the National Research Foundation for the Doctoral Program of Higher Education of China(20040013007) and the research foundation of the State Key Laboratory of Information Security..

†Corresponding author.

They proved that if each nonzero linear combination of $f_1, f_2, \cdots, f_m$ is an ε-almost$(n, 1, k)$-resilient function, then $F = (f_1, f_2, \cdots, f_m)$ is a $\frac{2^m-1}{2^{m-1}}\varepsilon$-almost$(n, m, k)$-resilient function.

This paper is organized as follows. Some definitions and preliminaries that will be used later in the paper are described in Sect.2. In Sect.3, more constructions of almost resilient functions are provided. For the relation between almost resilient function and large set of almost independent sample space, the constructions of balanced almost resilient function are concerned. As balanced almost CI function plays an important role in the secondary construction, we conclude some construction methods in Sect.4. Especially we prove it is feasible to determine weather a function is a balanced almost CI function by computing its walsh spectra.

2 Preliminaries

The vector spaces of n-tuples of elements from GF(2) is denoted by F_2^n. Let F be a function from F_2^n to F_2^m.

Definition 2.1 *The function F is called an (n, m, k)-resilient function if*

$$Pr[F(x_1, \cdots, x_n) = (y_1, \cdots, y_m) | x_{i_1} x_{i_2} \cdots_{i_k} = \alpha] = 2^{-m}$$

for any k positions $i_1 < i_2 < \cdots < i_k$, for any k-bit string $\alpha \in F_2^k$, and for any $(y_1, \cdots, y_m) \in F_2^m$, where the values $x_j (j \notin \{i_1, i_2, \cdots, i_k\})$ are chosen independently at random.

Following propositions is well-known and useful in understanding the relationship between a resilient functions and its component functions. It has appeared in many references (see, for example, [8]).

Proposition 2.1 *Let $F = (f_1, \cdots, f_m)$ be a function from F_2^n to F_2^m, where n and m are integers with $n \geq m \geq 1$, and each f_i is a function on F_2^n. Then F is an (n, m, k)-resilient function if and only if every nonzero combination of $f_1, \cdots, f_m$*

$$f(x) = \bigoplus_{i=1}^{m} c_i f_i(x)$$

is a $(n, 1, k)$-resilient function,where $c = (c_1, \cdots, c_n) \in F_2^n$.

K.Kurosawa et al. introduced a notation of ε-almost(n, m, k)-resilient function[2].

Definition 2.2 *The function F is called a ε-almost(n, m, k)-resilient function if*

$$|Pr[F(x_1, \cdots, x_n) = (y_1, \cdots, y_m) | x_{i_1} x_{i_2} \cdots_{i_k} = \alpha] - 2^{-m}| \leq \varepsilon$$

for any k positions $i_1 < i_2 < \cdots < i_k$, for any k-bit string $\alpha \in F_2^k$, and for any $(y_1, \cdots, y_m) \in F_2^m$, where the values $x_j (j \notin \{i_1, i_2, \cdots, i_k\})$ are chosen independently at random.

An almost k-wise independent sample spaces is probability space on n-bit tuples such that any k-bits are almost independent. A large set of (ε, k)-independent sample spaces, denoted by $LS(\varepsilon, k, n, t)$, is a set of $2^{m-t}(\varepsilon, k)$-independent sample spaces,each of size 2^t,such their union contains all 2^n binary vectors of length n. For details about k-wise independent sample spaces and $LS(\varepsilon, k, n, t)$, we refer to [2, 3].

The relation between $LS(\varepsilon, k, n, t)$ and almost resilient function is revealed in [2].

Proposition 2.2 *If there exists an $LS(\varepsilon, k, n, t)$,then there exists a δ-almost $(n, n-t, k)$-resilient function,where $\delta = \frac{\varepsilon}{2^{n-t-k}}$.*

A (n,m)-function F is called balanced if

$$Pr[F(x_1,\cdots,x_n)=(y_1,\cdots,y_m)]=2^{-m}$$

for all $(y_1,\cdots,y_m)\in F_2^m$.

Proposition 2.3 *If there exists a balanced ε-almost (n,m,k)-resilient function, then there exists a $LS(\delta,k,n,n-m)$,where $\delta=\frac{\varepsilon}{2^{k-m}}$.*

Using Weil-Carlitz-Uchiyama bound, K.Kurosawa et al.[2] present a construction of t-systematic (ε,k)-independent sample spaces and then extended to large set of almost independent sample spaces. So by proposition 2.2, some almost resilient functions are obtained.

Let $F(X)=(f_1,f_2,\cdots,f_m)$ be an (n,m)-function, the *nonlinearity* of F is defined to be $nl(F)=min\{nl(l\circ f):l$ is a non-constant m-variable linear function$\}$,where $nl(f)$ is the least hamming distance between boolean function f and all affine functions. And the *degree* of F defined to be the minimum of the degree of $l\circ f$, where l ranges over all non-constant m-variable linear function.

Similar to the resilient function, correlation immune function can also be generalized. K. Kurosawa et al.[2] called it the almost correlation immune function. In fact,an earlier generalization version of the single output case has been introduced by Yan Yixian[11].

Definition 2.3 *The function F is called an ε-almost (n,m,k)-correlation immune function if*

$$|Pr[F(x_1,\cdots,x_n)=(y_1,\cdots,y_m)|x_{i_1}x_{i_2}\cdots_{i_k}=\alpha]-Pr[F(x_1,\cdots,x_n)=(y_1,\cdots,y_m)]|\le\varepsilon$$

for any k positions $i_1<i_2<\cdots<i_k$, for any k-bit string $\alpha\in F_2^k$, and for any $(y_1,\cdots,y_m)\in F_2^m$, where the values $x_j(j\notin\{i_1,i_2,\cdots,i_k\})$ are chosen independently at random.

The relation between almost CI function and nonuniform $LS(\varepsilon,k,n,t)$ is given in [2].It is easy to see that an ε-almost(n,m,k)-resilient function is equivalent to an balanced ε-almost(n,m,k)-CI function.

Let f be a function from F_2^n to F_2, then

$$S_f(w)=\sum_{x\in F_2^n}(-1)^{f(x)\oplus w\cdot x}$$

is called a *Walsh transformation* of f. Walsh transform is a useful tool and many cryptographic criteria of a Boolean function can be characterized by it.

3 Construction of almost resilient functions

In the following, if h is a functions from F_2^n to F_2^m or F_2, denote

$$L(h(X)=Y)=\{(x_1,x_2,\cdots,x_n):h(x_1,x_2,\cdots,x_n)=Y\}. \tag{1}$$

Let $X_i,1\le i\le m$, be m independent random variables on F_2. The number of nonzero combination of $X_1,X_2,\cdots,X_m$ is $C_m^1+C_m^2+\cdots+C_m^m=2^m-1$. We divide it into two parts, each contains 2^{m-1} and $2^{m-1}-1$ elements respectively. Denote them as A_1 and A_2.For a fixed $(y_1,y_2,\cdots,y_m)\in F_2^m$ and a nonzero linear combination of $X_1,X_2,\cdots,X_m$, it determine a set $L(\oplus_{i=1}^m c_iX_i=\oplus_{i=1}^m c_iy_i)$. We call the set *determined* by $(y_1,y_2,\cdots,y_m)$. Furthermore we call the set $L(\oplus_{i=1}^m c_iX_i=\oplus_{i=1}^m c_iy_i\oplus 1)$ the *determined complement* set induced by $(y_1,y_2,\cdots,y_m)$. For each nonzero m-bit string $(c_1,c_2,\cdots,c_m)\in F_2^m$ and $a\in F_2$, by (1) it is obvious that

$$|L(\oplus_{i=1}^m c_iX_i=a)|=2^{m-1},L(\oplus_{i=1}^m c_iX_i=0)\cup L(\oplus_{i=1}^m c_iX_i=1)=F_2^m. \tag{2}$$

Lemma 3.1 *[6] Let notations defined as above. For an arbitrary m-bit string* $Y = (y_1, y_2, \cdots, y_m) \in F_2^m$*, then the collection of determined sets of* A_1 *equals to the collection of determined complement sets of* A_2 *added* $2^{m-1}Y$*. Note again that we call the two collections are equal if and only if the elements and its multiplicity in the two collections are identical.*

In [6],relations between almost resilient function and its component functions were presented.

Theorem 3.1 *Let* $F = (f_1, \cdots, f_m)$ *be a function from* F_2^n *to* F_2^m*, where* n *and* m *are integers with* $n \geq m \geq 1$*, and each* f_i *is a function on* F_2^n*. If* F *is an* ε*-almost* (n, m, k)*-resilient function, then each nonzero combination of* $f_1, \cdots, f_m$

$$f(x) = \bigoplus_{i=1}^{m} c_i f_i(x)$$

is an $2^{m-1}\varepsilon$*-almost* $(n, 1, k)$*-resilient function,where* $x = (x_1, \cdots, x_n) \in F_2^n$*.*

Theorem 3.2 *Let* $F = (f_1, \cdots, f_m)$ *be a function from* F_2^n *to* F_2^m*, where* n *and* m *are integers with* $n \geq m \geq 1$*, and each* f_i *is a function on* F_2^n*. If each nonzero combination of* $f_1, \cdots, f_m$

$$f(x) = \bigoplus_{i=1}^{m} c_i f_i(x)$$

is an ε*-almost* $(n, 1, k)$*-resilient function, then* F *is an* $\frac{2^m-1}{2^{m-1}}\varepsilon$*-almost* (n, m, k)*-resilient function , where* $x = (x_1, \cdots, x_n) \in F_2^n$*.*

Remark. By Theorem 3.1 for any ε-almost (n, m, k)-resilient function $F = (f_1, ..., f_m)$ every nonzero linear combination of f_i is an $2^{m-1}\varepsilon$-almost $(n, 1, k)$-resilient function. But by Theorem 3.2 if every nonzero linear combination is an $2^{m-1}\varepsilon$-almost $(n, 1, k)$-resilient function then $(f_1, ..., f_m)$ is $(2^m - 1)\varepsilon$-almost (n, m, k)-resilient function. Thus starting from ε-almost (n, m, k)-resilient function one can obtain an $(2^m - 1)\varepsilon$-almost (n, m, k)-resilient function. This gap between ε and $(2^m - 1)\varepsilon$ implies that both statements are not equally strong.Compared with Proposition 2.1 we could see although the almost resilient function only bias ε from resilient function in definition it is difficult for us to prove the same proposition of resilient function in almost case.

A construction based on a balanced almost $(n, 1, k)$-resilient function was presented in [6].

Theorem 3.3 *Let* f *be a balanced* ε*-almost* $(n, 1, k)$*-resilient function, then* $g(X, Y, Z) = (f(X) \oplus f(Y), f(Y) \oplus f(Z))$ *is a balanced* $\frac{9}{2}\varepsilon$*-almost* $(3n, 2, 2k + 1)$*-resilient function.*

But the proof of the theorem is tedious. Here we present a direct proof and improve the result.

Theorem 3.4 *Let* f *be a balanced* ε*-almost* $(n, 1, k)$*-resilient function, then* $g(X, Y, Z) = (f(X) \oplus f(Y), f(Y) \oplus f(Z))$ *is a balanced* $\frac{3}{2}\varepsilon$*-almost* $(3n, 2, 2k + 1)$*-resilient function.*

Proof. Denote $h(X, Y) = f(X) + f(Y)$.It is obvious that $h(X, Y)$ is balanced. We first prove that

$$|Pr(h(X, Y) = 1|x_{i_1} \cdots x_{i_r} y_{i_{r+1}} \cdots y_{i_{2k+1}}) - \frac{1}{2}| \leq \varepsilon$$

for any $2k + 1$ positions $x_i, 1 \leq i \leq r$ and $y_i, r + 1 \leq i \leq 2k + 1$.

Without loss of generality, we may assume $r \leq k$. Then by definition

$$|Pr(f(X) = 1|x_{i_1} \cdots x_{i_r}) - \frac{1}{2}| \leq \varepsilon.$$

By notation (1), we have

$$2^{n-r-1} - 2^{n-r}\varepsilon \leq |L(f(X) = 1|x_{i_1} \cdots x_{i_r})| \leq 2^{n-r-1} + 2^{n-r}\varepsilon.$$

So

$$2^{n-(2k+1-r)}(2^{n-r-1}-2^{n-r}\varepsilon) \le |L(f(X)+f(Y)=1|x_{i_1}\cdots x_{i_r}y_{i_{r+1}}\cdots y_{i_{2k+1}})| \le 2^{n-(2k+1-r)}(2^{n-r-1}+2^{n-r}\varepsilon).$$

i.e.

$$|Pr(h(X,Y)=1|x_{i_1}\cdots x_{i_r}y_{i_{r+1}}\cdots y_{i_{2k+1}})-\frac{1}{2}| \le \varepsilon.$$

So $h(X,Y)$ is a balanced ε-almost $(2n,1,2k+1)$ resilient function. The case $f(Y)\oplus f(Z)$ and $f(X)\oplus f(Z)$ can be similarly proved. And each of them is balanced, so g is also balanced. By theorem 3.2, the proof is complete.

We can generalize above result as following.

Theorem 3.5 *Let f_i be a balanced ε_i-almost $(n_i,1,k_i)$-resilient function,$1\le i\le l$,G be a $[l,m,d]$ linear code. Then $F(X_1,X_2,\cdots,X_l)=(f_1(X_1),f_2(X_2),\cdots,f_l(X_l))G^T$ is a balanced $\frac{2^m-1}{2^{m-1}}\varepsilon$-almost $(\sum_{i=1}^{l}n_i,m,dk+d-1)$-resilient function, where $k=min_{1\le i\le l}k_i$ and $\varepsilon=max_{1\le i\le l}\varepsilon_i$.*

Proof. Assume that $G=[a_{ij}],1\le i\le m,1\le j\le l$. Then

$$F=(\oplus_{i=1}^{l}a_{1i}f_i,\oplus_{i=1}^{l}a_{2i}f_i,\cdots,\oplus_{i=1}^{l}a_{mi}f_i).$$

For any nonzero linear combination of its component functions of F,we have

$$\oplus_{j=1}^{m}c_j(\oplus_{i=1}^{l}a_{ji}f_i)=\oplus_{i=1}^{l}f_i(\oplus_{j=1}^{m}c_ja_{ji}).$$

where $c=(c_1,\cdots,c_m)\in F_2^m$ is a nonzero vector. And note that minimum weight of code of G is d. So at least d functions of $f_1,\cdots,f_l$ appear in above formulation.Similar to the proof in theorem 3.4, we known that any nonzero linear combination of its component functions of F is ε-almost $(\sum_{i=1}^{l}n_i,1,dk+d-1)$-resilient function.By theorem 3.2,we complete the proof.

Corollary 3.1 *If there exit an $[l,m,d]$ linear code and ε-almost $(n,1,k)$-resilient function, then an $\frac{2^m-1}{2^{m-1}}\varepsilon$-almost $(ln,m,dk+d-1)$-resilient function must exist.*

If we take $f_1=f_2=f_3=f$ and

$$G=\begin{bmatrix}1&1&0\\0&1&1\end{bmatrix},$$

then theorem 3.4 may regard as a corollary of theorem 3.5.

Theorem 3.6 *Let $F=(f_1,\cdots,f_m)$ be ε_1-almost (n_1,m,t_1)-resilient function and $G=(g_1,\cdots,g_m)$ be ε_2-almost (n_2,m,t_2)-resilient function. Then $F(X)\oplus G(Y)=(f_1(x)\oplus g_1(y),\cdots,f_m(x)\oplus g_m(y))$ is ε-almost (n_1+n_2,m,t_1+t_2+1)-resilient function,where $\varepsilon=max(\varepsilon_1,\varepsilon_2)$.*

Proof. By definition, we need to prove that

$$|Pr(F(X)\oplus G(Y)=\eta|x_{i_1}\cdots x_{i_r}y_{i_{r+1}}\cdots y_{i_{t_1+t_2+1}})-\frac{1}{2^m}| \le \varepsilon$$

holds for arbitrary chosen $\eta\in F_2^m$ and for any t_1+t_2+1 positions $x_i,1\le i\le r$ and $y_i,r+1\le i\le t_1+t_2+1$.

Without loss of generality, assume that $r\le t_1$. Then for arbitrary $G(Y)=\alpha$, there exist exactly one $\beta\in F_2^m$ such that $F(X)=\beta$ and $F(X)+\alpha=\eta$.For $r\le t_1$, we have

$$|Pr(F(X)=\beta|x_{i_1}\cdots x_{i_r})|-\frac{1}{2^m}| \le \varepsilon.$$

i.e.

$$2^{n_1-r-m}-2^{n_1-r}\varepsilon \le |L(F(X)=\beta|x_{i_1}\cdots x_{i_r})| \le 2^{n_1-r-m}+2^{n_1-r}\varepsilon.$$

So we have

$$2^{n_2-(t_1+t_2+1-r)}(2^{n_1-r-m}-2^{n_1-r}\varepsilon) \le |L(F(X)\oplus G(Y)=\eta|x_{i_1}\cdots x_{i_r}y_{i_{r+1}}\cdots y_{i_{t_1+t_2+1}})| \le 2^{n_2-(t_1+t_2+1-r)}(2^{n_1-r-m}+2^{n_1-r}\varepsilon).$$

That is

$$2^{n_1+n_2-(t_1+t_2+1)-m}-2^{n_1+n_2-(t_1+t_2+1)}\varepsilon \le |L(F(X)\oplus G(Y)=\eta|x_{i_1}\cdots x_{i_r}y_{i_{r+1}}\cdots y_{i_{t_1+t_2+1}})| \le 2^{n_1+n_2-(t_1+t_2+1)-m}+2^{n_1+n_2-(t_1+t_2+1)}\varepsilon.$$

Thus we know that $F\oplus G$ is ε-almost (n_1+n_2,m,t_1+t_2+1)-resilient function.

The conclusion of theorem 3.6 can be slightly generalized with a similar proof.

Theorem 3.7 *Let $F_i(X)$,$1\le i\le l$,be ε_i-almost (n_i,m,t_i)-resilient function. Then $\oplus F_i(X_i)$ is an ε-almost $(\sum_{i=1}^{l}n_i,m,\sum_{i=1}^{l}t_i+l-1)$resilient function,where $\varepsilon=max_{1\le i\le l}\varepsilon_i$.*

The theorem 4.1 in [7] could be generalized to almost case.

Theorem 3.8 *Let $F(X)$ be an ε-almost (n,m,t) -resilient function and G be a $[N,k,d]$ linear code.Then*

$$H(X_1,X_2,\cdots,X_N)=(F(X_1),F(X_2),\cdots,F(X_N))G^T$$

is an $\frac{2^{km}-1}{2^{km-m-2}}\varepsilon$-almost $(nN,mk,d(t+1)-1)$-resilient function.

Proof. The proof is similar to that of [7]. The only difference is that any nonzero linear combination of component functions of $F(X)$ is an $\frac{2^m-1}{2^{m-1}}\varepsilon$-almost (n,m,t) -resilient function by theorem 3.1.By theorem 3.2 again, the proof is completed.

Just as we point out in the above remark , the gap between theorem 3.1 and 3.2 answer for the increasing of ε in former secondary constructions. So although theorem 3.5 may be seemed as a special case of theorem 3.8(let m=1), we would prefer to theorem 3.5 in secondary construction of almost resilient function under the present condition.

4 Spectrum characterization of ε-almost $(n,1,k)$- resilient functions

As we have describe in the last paragraph of section 3 and proposition 2.3, we are interest in balanced ε-almost$(n,1,k)$-CI functions , i.e. ε-almost $(n,1,k)$-resilient functions.Furthermore as we known that the algebraic degree and correlation immune order is incompatible, almost CI function is also proposed to avoid this dilemma when it was used as combination or filter function in stream cipher.

Some constructions of almost CI functions had been presented in [11].

Theorem 4.1 *[11] Let f be a k order CI function and g be a functions such that $wt(g)$ is a little number. Then $h=f\oplus g$ is a $\frac{3+2^{m+1}}{2^n}wt(g)$-almost (n,1,k)-CI function.*

Theorem 4.2 *[11] Let f_1 be a balanced ε_1-almost $(n,1,k)$-CI function and f_2 be a balanced ε_2-almost $(n,1,k)$-CI function. Then $f(x_1,\cdots,x_n,x_{n+1})=x_{n+1}f_1\oplus(1\oplus x_{n+1})f_2$ is a balanced ε-almost $(n+1,1,k)$-CI function,where $\varepsilon=max(\varepsilon_1,\varepsilon_2)$.*

We could see that it is easily to obtain an almost balanced CI function by modifying a CI function slightly.In this way we may derive many constructions.

It is well known that f is a $(n,1,k)$-CI function if and only if each $f\bigoplus\oplus_{i=1}^{n}a_ix_i$ is a balanced function for all $1\le wt(\alpha)\le k$, $\alpha=(a_1,a_2,\cdots,a_n)\in F_2^n$. It can be restated in the word of walsh transform , which is the well-known Xiao-Massy theorem. In the almost case, Yan Yixian [11] presented the following result.

Theorem 4.3 *Let f be an ε -almost $(n,1,k)$-CI function, then*

$$|Pr(f(X)\bigoplus \oplus_{i=1}^{n} a_i x_i = 1) - \frac{1}{2}| \leq \varepsilon$$

for any $1 \leq wt(\alpha) \leq k$, $\alpha = (a_1, a_2, \cdots, a_n) \in F_2^n$.

It means that for an ε -almost $(n,1,k)$-CI function the resulted function $f \bigoplus \oplus_{i=1}^{n} a_i x_i$ should be almost balanced for all α,$1 \leq wt(\alpha) \leq k$. Now Let us consider the opposite direction i.e. if a function f such that $f \bigoplus \oplus_{i=1}^{n} a_i x_i$ is almost balanced function for all α,$1 \leq wt(\alpha) \leq k$, is the function f an almost $(n,1,k)$-CI function? It is an interesting problem,because if it holds we will be able to determine a function if it is an almost CI function by computing its walsh spectra. Firstly in the case $k = 1$ and f is balanced, we have the following result.

Lemma 4.1 *Let f be a function from F_2^n to F_2 . If f is balanced and*

$$|Pr(f(X) \oplus x_i = 1) - \frac{1}{2}| \leq \varepsilon$$

holds for any $1 \leq i \leq n$ if and only if f is an ε -almost $(n,1,1)$-CI function.

Proof. From theorem 4.3, the sufficiency is obvious. Let us prove the necessity. Without lost of generality, we assume $i = 1$. Denote $P_{ij} = Pr(f(X) = i | x_1 = j), 0 \leq i, j \leq 1$. It is easy to verify that

$$P_{00} + P_{01} = 2Pr(f(x) = 0), P_{10} + P_{11} = 2Pr(f(x) = 1), \tag{3}$$

$$P_{00} + P_{10} = P_{01} + P_{11} = 1. \tag{4}$$

By the condition of the lemma, we have

$$|Pr(f(X) \oplus x_1 = 1) - \frac{1}{2}| \leq \varepsilon.$$

Furthermore,

$$Pr(f(X) \oplus x_1 = 1) = Pr(f(X) = 1, x_1 = 0) + Pr(f(X) = 0, x_1 = 1) = \frac{1}{2}(P_{10} + P_{01}).$$

Hence,

$$1 - 2\varepsilon \leq P_{10} + P_{01} \leq 1 + 2\varepsilon. \tag{5}$$

From (3) and (5), we have

$$1 - 2\varepsilon + 2Pr(f(X) = 1) \leq P_{10} + P_{01} + P_{10} + P_{11} \leq 1 + 2\varepsilon + 2Pr(f(X) = 1),$$

$$1 - 2\varepsilon + 2Pr(f(X) = 1) \leq 2P_{10} + P_{01} + P_{11} \leq 1 + 2\varepsilon + 2Pr(f(X) = 1).$$

By (4),

$$-\varepsilon \leq P_{10} - Pr(f(X) = 1) \leq \varepsilon.$$

Thus,

$$|Pr(f(X) = 1 | x_1 = 0) - Pr(f(x) = 1)| \leq \varepsilon.$$

Similarly we can prove

$$|Pr(f(X) = 1 | x_i = a) - Pr(f(x) = 1)| \leq \varepsilon, \text{for any } \ 1 \leq i \leq n, a \in F_2. \tag{6}$$

For f is balanced,

$$|Pr(f(X) = 1 | x_i = a) - \frac{1}{2}| \leq \varepsilon.$$

Thus the proof is completed.

Now we prove the main result.

Theorem 4.4 *Let f be a function from F_2^n to F_2 . If f is balanced and*

$$|Pr(f(X)\bigoplus \oplus c_i x_i = 1) - \frac{1}{2}| \leq \varepsilon$$

holds for any $c = (c_1, c_2, \cdots, c_n) \in F_2^n$ and $1 \leq wt(c) \leq k$. Then f is an $(2^k - 1)\varepsilon$-almost $(n, 1, k)$-CI function.

Proof. We prove the theorem in three steps.

1. For a fixed nonzero vector $c \in F_2^n$,

$$|Pr(f(X)\bigoplus \oplus c_i x_i = 1) - \frac{1}{2}| \leq \varepsilon,$$

then we have

$$|Pr(\oplus c_i x_i = 1|f(X)) - \frac{1}{2}| \leq \varepsilon. \tag{7}$$

If we substitute x_1 by $f(X)$ in the proof of lemma 4.1, the proof of step 1 is similar to that of lemma 4.1.

2. If

$$|Pr(\oplus_{i=1}^{k} c_i x_i|f(X)) - \frac{1}{2}| \leq \varepsilon,$$

then we have

$$|Pr(x_1 \cdots x_k|f(X)) - \frac{1}{2^k}| \leq \frac{2^k - 1}{2^{k-1}}\varepsilon. \tag{8}$$

Divide all the nonzero linear combinations of $x_1, \cdots, x_k$ into two part A_1 and A_2, such that $|A_1| = 2^{k-1}$ and $|A_2| = 2^{k-1} - 1$. For any fixed $\alpha = (a_1, \cdots, a_k) \in F_2^k$, by lemma 3.1, we have

$$\sum_{c \in A_1} |L(\oplus c_i x_i = c_i a_i|f(X))| = \sum_{c' \in A_2} |L(\oplus c_i' x_i = c_i' a_i \oplus 1|f(X))|$$

$$+2^{k-1}|L((x_1, \cdots, x_k) = (a_1, \cdots, a_k)|f(X))|. \tag{9}$$

By condition of theorem and step 1, we know (7) holds for any nonzero vector $c = (c_1, c_2, \cdots, c_n) \in F_2^n$ and $1 \leq wt(c) \leq k$. So (7) holds for any nonzero vector $c \in F_2^k$.

By

$$2^{k-1}(\frac{1}{2} - \varepsilon) \leq \sum_{c \in A_1} Pr(\oplus c_i x_i|f(X)) \leq 2^{k-1}(\frac{1}{2} + \varepsilon),$$

$$(2^{k-1} - 1)(\frac{1}{2} - \varepsilon) \leq \sum_{c' \in A_2} Pr(\oplus c_i' x_i|f(X)) \leq (2^{k-1} - 1)(\frac{1}{2} + \varepsilon),$$

and (9),we have

$$\frac{1}{2} - (2^k - 1)\varepsilon \leq 2^{k-1} Pr((x_1, \cdots, x_k) = (a_1, \cdots, a_k)|f(X)) \leq \frac{1}{2} + (2^k - 1)\varepsilon.$$

That is

$$|Pr((x_1, \cdots, x_k) = (a_1, \cdots, a_k)|f(X)) - \frac{1}{2^k}| \leq \frac{2^k - 1}{2^{k-1}}\varepsilon.$$

3. It is easy to verified that

$$Pr(f(X)|x_1 \cdots x_k) = 2^{k-1} Pr(x_1 \cdots x_k|f(X)).$$

So by (8) we have

$$|Pr(f(X)|x_1 \cdots x_k) - \frac{1}{2}| = 2^{k-1}|Pr(x_1 \cdots x_k|f(X)) - \frac{1}{2}| \leq (2^k - 1)\varepsilon.$$

Thus we have done.

Corollary 4.1 *Let f be a function from F_2^n to F_2. If $S_f(0) = 0$ and $|S_f(w)| \leq 2^{n+1}\varepsilon$ for any $w \in F_2^n, 1 \leq wt(w) \leq k$, then f is an $(2^k - 1)\varepsilon$-almost $(n, 1, k)$-CI function.*

Proof. Note that $|Pr(f(X) \bigoplus \oplus w_i x_i = 1) - \frac{1}{2}| \leq \varepsilon$ holds if and only if $-2^n\varepsilon + 2^{n-1} \leq wt(f(X) \oplus w \cdot X) \leq 2^n\varepsilon + 2^{n-1}$. And f is balanced if and only if $S_f(0) = 0$. By $S_f(w) = \sum_{x \in F_2^n}(-1)^{f(x) \oplus w \cdot x} = 2^n - 2wt(f(x) \oplus w \cdot x)$ and theorem 4.4, the result is obtained.

With corollary 4.1, it is convenient for us to verify a function if it is a balanced almost CI functions by computing its walsh spectra.

5 Conclusion

In this paper, some secondary constructions of almost resilient function are presented. From the relation between almost resilient function and large set of almost independent sample space, the constructions of balanced almost resilient function are concerned. As balanced almost CI function play an important role in the secondary construction, we conclude some constructions methods. Especially we prove it is feasible to determine weather a function is balanced almost CI function by computing its walsh spectra, which can be regard as the generalization of Xiao-Massy theorem in the almost case to a certain extent.

References

[1] B.Chor,O.Goldreich,J.Hȧstad,J.Friedman,S.Rudich,and R.Smoledsky. The bit extraction problem or t-resilient functions. *IEEE Symp. on Foundations of Computer Science*, 1985,Vol.26,pp.396-407.

[2] K.Kurosawa,T.Johansson,D.Stinson. Almost k-wise independent sample spaces and their applications. *J.Cryptology*,2001, Vol.14,no.4,pp.231-253.

[3] J.Naor,M.Naor. Small bias probality spaces:efficient constructions and applications. *SIAM Journal on Computing* 1993,Vol.22,pp.838-856.

[4] M.N.Wegman,J.L.Carter. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences* ,1981,Vol.22, PP.265-279.

[5] K.Kurosawa,R.Matsumoto. Almost security of cryptographic boolean functions,*IEEE Tran. on Info. theory*,2004, Vol.50, No.11,PP.2752-2761.

[6] Pin-Hui Ke,Tai-Lin Liu,Qiao-Yan Wen. Construction of almost resilient functions. *Cryptology and Network Security: 4th International Conference, CANS 2005*,Yvo G. Desmedt et al. ed. LNCS 3810, Springer-Verlag 2005, pp.236-246.

[7] Chuan-Kun Wu,Ed Dawson. On construction of resilient functions. *Information Security and Privacy,Proceedings of First Australasian Conference*,LNCS 1172,Springer-Verlag 1996,pp.79-86.

[8] Xian-Mo Zhang, Yuliang Zheng. Cryptographically resilient functions. *IEEE Tran. on Info. theory*,1997,Vol.43. No.5, PP.1740-1747.

[9] X.Guo-Zhen,J.Massy. A special characterization of correlation immune combining functions,*IEEE Tran. on Info. theory*,1988, Vol.34,PP.569-571.

[10] K.C.Gupta,P.Sarkar. Improved construction of nonlinear resilient S-Boxes. *IEEE Tran. on Info. theory*,2005,Vol.51,No.1,PP.339-348.

[11] Yan Yixian,Lin Xuduan. Coding theory and cryptography. People post and telecomunication publisher,1992.(in chinese).

[12] T.Johansson,E.Pasalic. A construction of resilient functions with high nonlinearity. *IEEE Tran. on Info. theory*,2003,Vol.49, No.2 ,PP.494-501.