Construction of Nonlinear Optimal Functions over Finite Fields

July 1, 2022

Introduction

Differential analysis and linear analysis are two important symmetric cryptanalysis methods. Resistance to these two kinds of analysis is a problem that designers must consider. Modern block ciphers usually use the iterative structure of simple round functions, and round functions generally include key mixing layer, substitution layer and diffusion layer. This paper mainly focuses on the construction method of diffusion layer. Branch number is a quantitative index to measure the diffusivity of diffusion layer. Combined with the differential (linear) probability of substitution layer, the ability of the whole block cipher to resist differential and linear analysis can be estimated. The larger the branch number, the stronger the ability of block cipher algorithm to resist differential (linear) cryptanalysis. A function is called optimal diffusion function if its branch number reaches the upper bound.

In 1995, Joan Daemen first proposed a method of constructing optimal linear diffusion function with maximum distance divisible code (MDS)[?] and successfully applied it to the design of AES algorithm. Since then, many methods for constructing optimal linear diffusion function[?, ?, ?, ?] have been put forward. These functions have been widely used in symmetric cryptography algorithm design.

If an optimal diffusion function is nonlinear at the same time, it will undoubtedly improve the security of the algorithm. At the same time, the use of nonlinear optimal diffusion function makes it possible to omit other nonlinear component in the algorithm, ruducing the implementation circuit cost of the algorithm, especially suitable for the design of lightweight cryptography algorithm. But also because of nonlinear, the research is much more difficult than the linear diffusion function, so far the related research results are very few. Theorem B.1.1 in Appendix B of reference[?] gives the necessary and sufficient conditions for a (nonlinear) diffusion function to have the maximum difference branch number, and proves that the difference branch number of the optimal diffusion function is equal to the linear branch number, which has strong theoretical value, but no practical construction method is given. In 2005, Alexander Klimov et al gave a method to construct nonlinear optimal diffusion layer based on T function[**?**], which is the first concrete construction. In 2013, H. Han et al give an example of a nonlinear optimal diffusion function with input 2 output over a finite field[**?**]. In 2017, Qu Chengqin proved There is no linear optimal diffusion function over the ring of integers modulo 2^n[**?**], and a construction method of 2-input 2-output nonlinear optimal diffusion function is given by using orthogonal latin square. In 2018, Liu Yunwen et al used nonlinear codes to construct nonlinear diffusion functions[**?**], and designed a class of nonlinear suboptimal diffusion functions based on modular additions. In 2021 reference[**?**], the necessary and sufficient conditions for the 4-input nonlinear optimal diffusion function are given, and a concrete 4-input nonlinear optimal diffusion function is constructed. The existing construction methods are either limited to the simple case of variable number 2, or too complex to have practical value. For example, the implementation complexity of the proposed nonlinear diffusion function in [**?**] is too high. Only part of the component function in nonlinear optimal diffusion function constructed in reference[**?**] is nonlinear.

In this paper, we propose and prove a necessary and sufficient conditions of nonlinear optimal diffusion function based results in reference [**?**] , and give a concrete construction method, which is the first method of constructing nonlinear optimal diffusion function with practical value so far.

The rest of the paper is organized as follows: Section 2 gives preliminary used throughout the paper. In section 3 a necessary and sufficient conditions is given. We propose a generic construction of a class of nonlinear optimal diffusion function in section 4. Section 5 concludes the paper.

Preliminary

Let $\mathbb{F}_q$ a finite field with q elements and $\mathbb{F}_q^n$ an extending field of $\mathbb{F}_q$. For , $y \in \mathbb{F}_q$, $x + y$ denotes addition of x and $y, x * y$ denotes the multiplication of x and y. For simplicity, we write xy instead of $x * y$ without causing confusion. We denote $x - y = x + (-y)$, where $-y$ is additive inverse of y. An element in $\mathbb{F}_q^n$ is represented as a vector $X = (x_1, x_2, \ldots, x_n)$, where $x_i \in \mathbb{F}_q, i = 1, \ldots, n$. The Hamming weight of X is defined by $w(X) = w(x_1, x_2, \cdots x_n) = \#\{i \mid x_i \neq 0\}$, where $\#\{i \mid x_i \neq 0\}$ means the number of non-zero entry in $(x_1, x_2, \ldots x_n)$.

For $X = (x_1, x_2, \dots, x_n), Y = (y_1, y_2, \dots, y_n) \in \mathbb{F}_q^n$, define function $Y = f(X)$ over $\mathbb{F}_q^n$ as

$$f : \mathbb{F}_q^n \to \mathbb{F}_q^n : (x_1, x_2, \dots, x_n) \mapsto (y_1, y_2, \dots, y_n) \tag{1}$$

Where component functions $y_i = f_i(x_1, x_2, \dots, x_n), i = 1, 2, \dots, n$. The inner product of X, Y is defined as $X \cdot Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$. The difference of X, Y is defined as $\Delta = X - Y = (x_1 - y_1, x_2 - y_2, \dots, x_n - y_n)$

Definition 1 *Let* n *be a positive integer, the differential branch number of function $f : \mathbb{F}_q^n \to \mathbb{F}_q^n, X \mapsto Y = f(X)$ is defined as*

$$B_d(f) = \min_{X, Y \neq X} \{w(X - Y) + w(f(X) - f(Y))\}.$$

The linear branch number is defined as

$$B_l(f) = \min_{x, \alpha, \beta \in \mathbb{F}_q, \beta \neq 0, \alpha \cdot x = \beta \cdot f(x)} \{w(\alpha) + w(\beta)\}.$$

Where "+" in the above formula represents the integer addition.

For a function $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$, the upper bound of the differential(linere) branch number is obviously $n + 1$. In general, the difference branch number and the linear branch number of a function are not equal. When one of them achieves the upper bound they are equal, which is proved in [**?**]:

Theorem 1 *[?] A function has the maximum difference branch number if and only if it has the maximum linear branch number.*

We call $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ optimal diffusion function if its branch number equals to $n + 1$. The construction of the first optimal diffusion function based on linear MDS code [**?**] in the coding theory. So it is also commonly called as MDS function in the literature . While constructing a linear optimal diffusion function is very easy, for nonlinear case it is very difficult. There is no practical methed so far.

Conditions for nonlinear optimal diffusion functions

Consider a function $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ as follows:

$$f : \mathbb{F}_q^n \to \mathbb{F}_q^n : (x_1, x_2, \dots, x_n) \mapsto (y_1, y_2, \dots, y_n) \tag{2}$$

Where component functions $y_i = f_i(x_1, x_2, \ldots, x_n), i = 1, 2, \ldots, n$. A necessary and sufficient conditions for nonlinear optimal diffusion function are given in Theorem B.1.1 of [?] from the point of view of equation. Using a equal expression of (2) as:

$$f : \mathbb{F}_q^n \to \mathbb{F}_q^n, (x_1, x_2, \ldots, x_n) \mapsto (x_{n+1}, x_{n+2}, \ldots, x_{2n}) \tag{3}$$

Consider a partition Ξ of the set $\{1, 2, \ldots, 2n\}$ that divides the set of indices in two equally sized subsets ξ and $\bar{\xi}$. Denote the vector with components $x_i \in \xi$ by x_ξ. Define the following set of equations:

$$\left\{ \begin{array}{c} (x_1, x_2, \ldots, x_n) \mapsto (x_{n+1}, x_{n+2}, \ldots, x_{2n}), \\ x_{t_i} = a_i, i \in \xi, a_i \in \mathbb{F}_q. \end{array} \right. \tag{4}$$

For any partition $\Xi = \xi \sqcup \bar{\xi}, i \in \xi, a_i \in \mathbb{F}_q$, If all equations in (4) have unique solutions, they are the optimal differential diffusion function and vice versa. The following theorem is proved in [?]:

Theorem 2 *[?]: functions in the form of (2) have the maximum number of differential branches, i.e. $B_d(f) = n + 1$, iff if any set of equations of form (4) has exactly one solution, whatever the choice of $\xi(\#\xi = n)$ and any $a_i \in \mathbb{F}_q, i \in \xi$*

The authors in [?] did not give construction method of optimal nonlinear diffusion functions from Theorem 2. It is difficult to determine whether a nonlinear function has the maximum branch number using Theorem 2 because finding the inverse of a nonlinear function is hard. So this method is not practical. In the following we will propose a more practical necessary and sufficient conditions of a nonlinear optimal diffusion function. Based on these conditions we give a method to construct a class of nonlinear optimal functions in generally.

Let k be an integer, and $0 \leq k \leq n-1$, For any $a_1, a_1, \ldots, a_k \in \mathbb{F}_q$, we select any k variables and $n-k$ component functions, say $x_{t_1}, x_{t_2}, \ldots, x_{t_k}$ and $y_{s_1}, y_{s_2}, \ldots, y_{s_{n-k}}$. Let $x_{t_1} = a_1, x_{t_2} = a_2, \ldots, x_{t_k} = a_k$. The function defined by formula (1) can be expressed as:

$$\left\{ \begin{array}{c} y_{s_i} = f_{s_i}(x_1, x_2, \ldots, x_n), i = 1, 2, \ldots, n-k \\ x_{t_i} = a_i, i = 1, 2, \ldots, k \end{array} \right. \tag{5}$$

We call (5) parameterized function.

Theorem 3 *Function in (1) is a nonlinear optimal diffusion function iff for any $0 \leq k \leq n-$ 1 , the parameterized function (5) is a permutation over $\mathbb{F}_q^{n-k}$*

Proof: According to Theorem 1, It suffices to show the theorem for differential diffusion function.

Assume that f is not an optimal diffusion function. Then there exists a pair of $(\Delta X, \Delta Y) \neq 0$, such that $W(\Delta X)+W(\Delta Y) < n+1$. Without loss of generality we assume that $W(\Delta X) = h$ and $\Delta X = (\underbrace{\Delta x_1, \dots, \Delta x_h}_{h}, \underbrace{0, \dots, 0}_{n-h})$, $\Delta Y = (\underbrace{0, \dots, 0}_{h}, \underbrace{\Delta \mathrm{y}_{h+1}, \dots, \Delta \mathrm{y}_n}_{n-h})$. For any specified variables $x_{h+1}, \dots, x_n$ and h component functions $f_i, i = 1, 2, \dots, h$, there exists a pair of inputs $X = (\underbrace{\mathrm{x}_1, \dots, \mathrm{x}_h}_{h}, \underbrace{\mathrm{x}_{h+1}, \dots, \mathrm{x}_n}_{n-h})$, $X' = (\underbrace{\mathrm{x}_1 - \Delta\mathrm{x}_1, \dots, \mathrm{x}_h - \Delta\mathrm{x}_h}_{h}, \underbrace{\mathrm{x}_{h+1}, \dots, \mathrm{x}_n}_{m-h})$, such that the outputs $Y = Y'$, where $Y = (y_1, \dots, y_h), Y' = (y'_1, \dots, y'_h)$. So $(y_1, \dots, y_h) = (f_1, \dots, f_h)$ is not a permutation on variables $x_1, \dots, x_h$. This contradicts the premise and hence our initial hypothesis is proven to be false. Therefore, the sufficient conditions are established.

Conversely, Assume for $0 \le k \le n-1$, the function in (5) is not a permutation over $\mathbb{F}_q^{n-k}$, then there exists $X = (\underbrace{x_1, \dots, x_{n-k}}_{n-k}, \underbrace{x_{n-k+1}, \dots, x_n}_{k})$, $X' = (\underbrace{x_1 - \Delta x_1, \dots, x_{n-k} - \Delta x_{n-k}}_{n-k}, \underbrace{x_{n-k+1}, \dots, x_n}_{k})$, and $\Delta X = X - X' = (\underbrace{\Delta x_1, \dots, \Delta x_{n-k}}_{n-k}, \underbrace{0, \dots, 0}_{k}) \neq \mathbf{0}$, such that $\Delta Y = \left(y_1 - y'_1, \dots, y_{n-k} - y'_{n-k}\right) = 0$. This implies that the sum of intput and output differential weight is less than n. This contradicts the prerequisites, and the hypothesis is false. The necessary conditions are proved.

For function (1) there are $\sum_{k=0}^{n-1} q^k \left(C_n^{n-k}\right)^2$ parameterized function with form (5), where $C_{\mathrm{n}}^{n-k} = \frac{n(n-1)\cdots(k+1)}{(n-k)(n-k-1)\cdots 1}$. From the Theorem 3, We only need check whether each of them is a permutation. For example, for a 4 × 4 function over $\mathbb{F}_{2^8}$, the total number of functions that need to be verified is $\sum_{k=0}^{3} 2^{8k} \left(C_4^{4-k}\right)^2 = 1 + 2^{12} + 9 \times 2^{18} + 2^{28}$. If using Theorem 2 we have to verify 2^{32} many 32×32 functions, which has much more computation compare with theorem 3. Next section we propose a practical construction method of a class of nonlinear optimal diffusion function over finite field.

A Construction method of nonlinear optimal diffusion function

Consider following vector function $F : \mathbb{F}_q^n \to \mathbb{F}_q^n, (x_1, \dots, x_n) \mapsto (y_1, \dots, y_n)$ over $\mathbb{F}_q^n$:

$$\left\{\begin{array}{c} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + f\left(x_1 + x_2 + \cdots + x_n\right) \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + f\left(x_1 + x_2 + \cdots + x_n\right) \\ \cdots\cdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + f\left(x_1 + x_2 + \cdots + x_n\right) \end{array}\right. \tag{6}$$

Where $a_{ij} \in \mathbb{F}_q, i, j = 1, \dots, n$, and are constants. For simply function (3) can be expressed as:

$$Y = AX + If.$$

Where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, I = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$. We define A_i, A'_i matrices obtaining by Replacing the i-th column of the matrix A with unit vector I and vector Y respectively, and define $|A|$ as determinant of matrix A. Under the condition $|A| \neq 0$ we can uniquely solve X in the equation (6) from Y :

$$x_i = -\frac{|\mathrm{A}_i|}{|\mathrm{A}|} f\left(\sum_{i=1}^{n} x_i\right) + \frac{|A'_i|}{|\mathrm{A}|}, i = 1, 2, \dots, n, \tag{7}$$

Adding the left and right sides of equation (7) we get :

$$\sum_{i=1}^{n} x_i = -\frac{\sum_{i=1}^{n} |\ \mathrm{A}_i|}{|\mathrm{A}|} f\left(\sum_{i=1}^{n} x_i\right) + \frac{\sum_{i=1}^{n} |A'_i|}{|\mathrm{A}|} \tag{8}$$

Where symbol $\sum$ represent additions in $\mathbb{F}_q$. Furthermore we can get $\sum_{i=1}^{n} x_i$ from equation (8) if certain conditions are met. We distinguish two cases: 1) If $\sum_{i=1}^{n} |A_i| = 0$, then $\sum_{i=1}^{n} x_i = \frac{\sum_{i=1}^{n} |A'_i|}{|\mathrm{A}|}$. Substituting $\sum_{i=1}^{n} x_i$ in equation (7) by $\frac{\sum_{i=1}^{n} |A'_i|}{|\mathrm{A}|}$ we get

$$x_i = -\frac{|\mathrm{A}_i|}{|\mathrm{A}|} f\left(\frac{\sum_{i=1}^{n} |A'_i|}{|\mathrm{A}|}\right) + \frac{|A'_i|}{|\mathrm{A}|}, i = 1, 2, \dots, n. \tag{9}$$

2) If $\sum_{i=1}^{n} |A_i| \neq 0$, and add a condition that the function

$$\mathrm{y} = x + \frac{\sum_{i=1}^{n} |\ \mathrm{A}_i|}{|\mathrm{A}|} f(x) \tag{10}$$

is a permutation over $\mathbb{F}_q$, then we can obtain a unique solution of (8), i.e. $\sum_{i=1}^{n} x_i = g\left(\frac{\sum_{i=1}^{n}|A'_i|}{|A|}\right)$, where g is the inverse of $y = x + \frac{\sum_{i=1}^{n}|A_i|}{|A|}f(x)$. Substituting it to (7) we get:

$$x_i = -\frac{|A_i|}{|A|}f\left(g\left(\frac{\sum_{i=1}^{n}|A'_i|}{|A|}\right)\right) + \frac{|A'_i|}{|A|}, i = 1, 2, \ldots, n. \tag{11}$$

Therefore, for each $(y_1, y_2, \ldots, y_n) \in \mathbb{F}_q^n$, vector function (6) has unique inverse $(x_1, x_2, \ldots, x_n)$. Furthermore its expression is given by formula (9) or (11).

To sum up, we can get the following theorem:

Theorem 4 *The function (6) is a permutation over $\mathbb{F}_q^n$ if the following to conditions hold:*

1) $|\mathrm{A}| \neq 0$*;*

2) If $\sum_{i=1}^{n} |\ \mathrm{A}_i| \neq 0$, then the function $y = x + \frac{\sum_{i=1}^{n}|\ \mathrm{A}_i|}{|\mathrm{A}|}f(x)$ is a permutation over $\mathbb{F}_q$.

From theorem 4 , determining whether the function (3) is a permutation can be transformed into whether the function (11) is a permutation.

Conditions that function (3) is a optimal diffusion can be deduced easily form theorem 3 and theorem 4:

Theorem 5 *The function (3) is an optimal diffusion function over $\mathbb{F}_q^n$ if the following two conditions hold:*

1) For each $k = 2, 3, \ldots n$, the $k \times k$ submatrix of A is nonsingular.

2) For each $k = 2, 3, \ldots n$, if $\sum_{i=1}^{n} |\mathrm{A}_i| \neq 0$,

then the function $y = x + \frac{\sum_{i=1}^{n}|\mathrm{A}_i|}{|\mathrm{A}|}f(x)$ is a permutation over $\mathbb{F}_q$

Remark 1 *If $\sum_{i=1}^{n} |\mathrm{A}_i| \neq 0$, and $|\mathrm{A}| \neq 0$, then $y = \frac{|\mathrm{A}|}{\sum_{i=1}^{k}|\mathrm{A}_i|}x + f(x)$ is a permutation iff function $y = x + \frac{\sum_{i=1}^{n}|\mathrm{A}_i|}{|\mathrm{A}|}f(x)$ is a permutation. In order to be consistent with the form of orthomorphic permutation[**?**], we use $y = \frac{|\mathrm{A}|}{\sum_{i=1}^{k}|\mathrm{A}_i|}x + f(x)$. Note when $\frac{|\mathrm{A}|}{\sum_{i=1}^{k}|\ \mathrm{A}_i|} = 1$, $y = \frac{|\mathrm{A}|}{\sum_{i=1}^{k}|\mathrm{A}_i|}x + f(x)$ is called orthomorphic permutation.*

It is easy to get example of nonlinear optimal diffusion function with form (3) by computer search. We show two of them here:

Example 1 *The function*

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \mathrm{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \end{pmatrix}$$

is nonlinear optimal diffusion function over finite field $\mathbb{F}_{257}$*, where*

$$A = \begin{bmatrix} 60 & 212 & 212 & 0 \\ 212 & 0 & 60 & 212 \\ 212 & 212 & 0 & 60 \\ 0 & 60 & 212 & 212 \end{bmatrix}, f(\mathrm{x}) = x^{129}$$

Example 2 *Define finite field* $\mathbb{F}_{2^8} = \mathbb{F}_2/\left(x^8+x^4+x^3+x^2+1\right)$*. Let* w *be a prime element. The function*

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \mathrm{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \\ f(x_1+x_2+x_3+x_4) \end{pmatrix}$$

is nonlinear optimal diffusion function over finite field $\mathbb{F}_{2^8}$*, where*

$$A = \begin{bmatrix} w^{19} & w^{91} & w^{145} & w^{19} \\ w^{91} & w^{145} & w^{19} & w^{19} \\ w^{145} & w^{19} & w^{19} & w^{91} \\ w^{19} & w^{19} & w^{91} & w^{145} \end{bmatrix}, f(\mathrm{x}) = x^{86} + w^{91}x$$

Conclusion

In this paper, we propose a practical criterion for nonlinear optimal diffusion function. Using this criterion we construct generally a class of nonlinear optimal diffusion function over finite field. This is the first practical example of optimal nonlinear diffusion function so far. These functions can be used in the design of block cipher, stream cipher, hash functions, etc. It may be used in the construction of nonlinear codes in the field of coding.