ON THE QUESTION OF NONLINEARITY OF VECTORIAL FUNCTIONS OVER FINITE FIELDS

Vladimir RYABOV

CTCrypt 2022

Novosibirsk
June 6-9, 2022

NOTATION AND DEFINITIONS

- $\boldsymbol{F}_q$ - a finite field of q elements, where $q = p^m$, $p \in \boldsymbol{P}$, $m \in \boldsymbol{N}^*$;
- $\boldsymbol{F}_q^{\,n}$ - an n-dimensional vector space over the field $\boldsymbol{F}_q$, where $n \in \boldsymbol{N}^*$;
- $\boldsymbol{P}_q^{\,n,k}$ - the set of all mappings of $\boldsymbol{F}_q^{\,n}$ into $\boldsymbol{F}_q^{\,k}$ or q-valued vectorial functions;
- $\boldsymbol{A}_q^{\,n,k}$ and $\boldsymbol{L}_q^{\,n,k}$ the subsets of all affine fnd linear mappings from the set $P_q^{\,n,k}$.

In the case $k \leq n$, a vectorial function $F \in P_q^{\,n,k}$ is called *balanced* if for any $y \in \boldsymbol{F}_q^{\,k}$ the condition $F^{\,-1}(y) = q^{n-k}$ is satisfied.

- $\boldsymbol{B}_q^{\,n,k}$ - the subset of all balanced vectorial functions from the set $P_q^{\,n,k}$.

For $k = n$ the set $B_q^{\,n,n}$ coincides with the set of *permutations* of the space $\boldsymbol{F}_q^{\,n}$.

- $\boldsymbol{S}_q^{\,n}$ - the subset of all permutations of the space $\boldsymbol{F}_q^{\,n}$;
- $\boldsymbol{\rho}(F,F')$ - the Hamming distance in the space $\boldsymbol{F}_{qk}^{\;q^n}$ between the functions F and F' from $P_q^{\,n,k}$.

ON APPROXIMATION OF DISCRETE FUNCTIONS BY LINEAR ANALOGUES

Two approaches to determining the measure of closeness of a discrete function to linear ones:

- based on an estimate of the Hamming distance from a function to a set of linear mappings;
- based on the difference properties of the function.

[Glukhov M.M., “On the approximation of discrete functions by linear functions”, Mat. Vopr. Kriptogr., 7:4 (2016), 29–50. In Russian]

Both approaches are widely used in cryptography.

The present paper considers the possibilities of approximating *q*-valued vectorial functions by affine analogs in the framework of the first approach.

Let’s now imagine how this approach was previously used for the cases of Boolean functions, *q*-valued functions and *q*-valued vectorial functions.

BOOLEAN FUNCTIONS

Soviet cryptographers in the 1960s studied the *statistical structure* of Boolean functions, which for the function $f \in P_2^{n,1}$ is the set $\{\Delta_f^l \mid l \in L_2^{n,1}\}$, where $\Delta_f^l = 2^{n-1} - \rho(f,l)$. The value of the maximum modulus Δ_f of such coefficients can be considered as a measure of closeness to the set $A_2^{n,1}$. It was shown that the lower bound $\Delta_f \geq 2^{n/2-1}$ is achievable only for even n for the class of functions called *minimal*. [Glukhov M. M., “On the approximation of discrete functions by linear functions”… + Tokareva N. “Bent functions: results and applications to cryptography”, Academic Press, Elsevier, Global, 2015, 220 p.]

A few years later, for even n, Boolean *bent* functions were defined, which are essentially minimal. [Rothaus O. S., “On "bent" functions”, Journal of Combinatorial Theory, Series A, 20:3 (1976), 300-305]

As a similar measure of closeness of the function $f \in P_2^{n,1}$ to the set $A_2^{n,1}$, the value $N_f = \min\{\rho(f,a) \mid a \in A_2^{n,1}\}$ was defined, which was called the *nonlinearity* of the function f. It was shown that the upper bound

$$N_f \leq 2^{n-1} - 2^{n/2-1} \quad (1)$$

is achievable only for even n for Boolean bent functions. [Meier W., Staffelbach O. “Nonlinearity criteria for cryptographic functions”, EUROCRYPT 1989, LNCS 434, Springer, Berlin, Heidelberg, 1990, 549-562]

Q-VALUED FUNCTIONS

The statistical structure of q-valued functions, as well as its analogue with respect to the set of affine functions, were considered by Ambrosimov A.S. [Ambrosimov A.S., “Approximation of k-ary functions by functions from the given system”, Fundam. Prikl. Mat., 3:3, (1997), 653-674. In Russian] Taking into account the normalization used by him, the latter, which we will call the *extended statistical structure*, for the function $f \in P_q^{n,1}$ has the form $\{\mathfrak{d}(f,a) \mid a \in A_q^{n,1}\}$, where $\mathfrak{d}(f,a) = (q-1)/q - \rho(f,l)/q^n$.

By analogy with Boolean case, the Hamming distance from the function $f \in P_q^{n,1}$ to the set $A_q^{n,1}$ was called *nonlinearity* of the function *f*. [Nyberg K. “On the construction of highly nonlinear permutations”, EUROCRYPT 1992, LNCS 658, Springer, Berlin, Heidelberg, 1993, 92-98]

In previous works of the author, the study of the extended statistical structure and nonlinearity of *q*-valued functions was continued. In particular, the upper bound

$$N_f \leq (q-1)\, q^{n-1} - q^{n/2-1} \quad (2)$$

was proved, which, as was reported at the previous symposium, for $q > 2$, even for even *n*, is achieved only for a part of the *q*-valued bent functions.

Q-VALUED VECTORIAL FUNCTIONS

In the recently published work of the author, the study of the extended statistical structure and nonlinearity of q-valued vectorial functions was started. [Ryabov V.G., “Approximation of vector functions over finite fields and their restrictions to linear manifolds by affine analogues”, Diskr. Mat., 34:2 (2022), 83-105. In Russian]

For the mapping $F \in P_q^{n,k}$, the *extended statistical structure* is the set $\{\delta(F,A)\ | A \in A_q^{n,k}\}$, where $\delta(F,A) = (q^k - 1)/q^k - \rho(F,A)/q^n$, and the *nonlinearity*, by analogy with the previous definitions, is given by the formula

$$N_F = \min\{\rho(F,A) \mid A \in A_q^{n,k}\}. \quad (3)$$

The value of the maximum coefficient of the extended statistical structure δ_F and the nonlinearity N_F can be considered as a measure of closeness of the mapping F to the set $A_q^{n,k}$. These parameters are related by the equality $N_F = (q^k - 1)q^{n-k} - \delta_F q^n$. An upper bound was also obtained

$$N_F \leq (q^k - 1)q^{n-k} - q^{n/2-k}. \quad (4)$$

OTHER TYPES
OF NONLINEARITY OF Q-VALUED VECTORIAL FUNCTIONS

To date, other types of nonlinearity of *q*-valued vectorial functions, which are in demand in linear and differential methods of cryptanalysis, have become widespread.

An important place is occupied by the nonlinearity defined for the mapping $F \in P_q^{n,k}$ with a set of coordinate functions $\boldsymbol{f} = (f_1, \ldots, f_k)$ by the formula

$$NL_F = \min\{\, N_{\langle \boldsymbol{w},\boldsymbol{f}\rangle} \mid \boldsymbol{w} \in \boldsymbol{F}_q^{\,k} \setminus \{\boldsymbol{0}\}\,\}. \quad (5)$$

[Nyberg K. "On the construction of highly nonlinear permutations"...]

Using an approach based on the difference properties of a vectorial function from $P_q^{n,k}$ involves studying the structure $\{\delta_F^{\boldsymbol{a},\boldsymbol{b}} \mid \boldsymbol{a} \in \boldsymbol{F}_q^{\,n} \setminus \{\boldsymbol{0}\},\ \boldsymbol{b} \in \boldsymbol{F}_q^{\,k}\}$, where $\delta_F^{\boldsymbol{a},\boldsymbol{b}} = |\{\boldsymbol{x} \in \boldsymbol{F}_q^{\,n} \mid F(\boldsymbol{x} \oplus \boldsymbol{a}) \ominus F(\boldsymbol{x}) = \boldsymbol{b}\}|$. The value of the maximum coefficient of this structure δ_F defines another type of nonlinearity of the vectorial function F.

However, it should be noted that when using NL_F and δ_F to measure the proximity of the mapping F to the set $A_q^{n,k}$, collisions may occur when vectorial functions that are not affine are equated to affine ones.

BENEFITS OF USING NONLINEARITY N_F

The previously used proximity measures for vector functions were not metrics. In contrast, the Hamming distance used here is a metric in a space representing all mappings from $P_q^{n,k}$.

It is directly related to the maximum size of a piecewise affine region and allows one to obtain a lower bound on the number of such regions for all possible representations of a *q*-valued vectorial function in piecewise affine form.

For clarification, we use another well-known concept. Let us divide the space $\boldsymbol{F}_q^n$ into sets on which the restrictions of the vectorial function $F \in P_q^{n,k}$ coincide with the restrictions of some affine mappings. The smallest number of such sets under all possible partitions is called the order of affinity and is denoted by ***ard F***. [Fomichev V.M., "Discrete mathematics and cryptology", Dialog-MIFI, Moscow, 2003, 397 p. In Russian].

$$\textbf{\textit{ard F}} \geq [q^n/(q^n - N_F)]. \quad (6)$$

The use of an extended statistical structure creates a basis for studying the possibilities of representing *q*-valued mappings in a piecewise affine form, which can be useful in the case of applying analysis methods that use affine approximations (see, for example, [Gorshkov S.P., Dvinyaninov A.V., "Lower and upper bounds for the affinity order of transformations of Boolean vector spaces", Prikl. Diskr. Mat., 2(20), (2013), 14-18. In Russian]).

PROPERTIES OF NONLINEARITY N_F

[Ryabov V.G., “Approximation of vector functions over finite fields and their restrictions to linear manifolds by affine analogues”…]

- ▷ it was shown that the unordered set of coefficients of the extended statistical structure $\{\eth(F,A)\}$ and the nonlinearity N_F are invariants for **EA**-equivalent q-valued vector functions from $P_q^{n,k}$;
- ▷ $N_F \geq \max\,\{N_{\langle \boldsymbol{w},\boldsymbol{f}\rangle} \mid \boldsymbol{w} \in \boldsymbol{F}_q^{k}\} \;\; \geq NL_F$.

For the balanced mapping $G \in B_q^{n,k}$, we obtain a refinement of the upper bound (4)

$$N_G \leq (q^k - 1)\, q^{n\text{-}k} - \lfloor q^{n/2-k} \rfloor - 1. \quad (7)$$

It follows from (6) that the permutation $S \in S_q^{\,n}$ satisfies the upper bound

$$N_S \leq q^n - 2. \quad (8)$$

Taking into account the correspondence of permutations $S \in S_q^{\,n}$ and $s \in S_{q^n}^{\,1}$, we have a chain of inequalities $0 \leq NL_S \leq N_S \leq N_s \leq q^n - 2$.

THE CASE OF BOOLEAN BALANCED MAPPINGS AND PERMUTATIONS

In the case of balanced mappings and permutations, following from (1) the upper bound of the alternative nonlinearity of the form $2^{n-1} - 2^{n/2-1}$ can be refined using the well-known estimate of the nonlinearity of balanced functions, namely, for $G \in B_q^{n,k}$ ($S \in S_q^n$), $n \geq 4$, we have the inequality of the form

$$NL_G \leq 2^{n-1} - 2^{n/2-1} - 2. \quad (9)$$

[Seberry J., Zhang X.-M., Zheng Y., “Nonlinearity and propagation characteristics of balanced Boolean functions”, Information and Computation, 119:1, (1995), 1-13]

Using the Sidelnikov–Chabaud–Vaudenay’s bound, we also get an upper bound for a permutation $S \in S_q^n$ of the form $NL_S \leq 2^{n-1} - 2^{(n-1)/2}$.

Consider the Boolean permutations from S_2^4. Among the "good" permutations, the researchers include those generated by the power function x^d over the field $\boldsymbol{F}_{16}$ for $d = 7, 11, 13, 14$. Indeed, all these four permutations have, in accordance with (9), the maximum possible alternative nonlinearity equal to 4, a nonlinearity equal to 9, and in accordance with (6) an affinity order greater than or equal to 3.

THE CASE
OF BOOLEAN BALANCED PERMUTATIONS (continued)

Adams C. and Tavares S. proposed "good" permutations in their opinion, the first of which was called the permutation $S' \in S_2^4$ generated the permutation of the form $s' = \{9, 13, 10, 15, 11, 14, 7, 3, 12, 8, 6, 2, 4, 1, 0, 5\} \in S_{16}^1$. [Adams C., Tavares S., "Good S-boxes are easy to find", CRYPTO 1989, LNCS 435, Springer, Berlin, Heidelberg, 1990, 612-615]

Permutation S', like the previous four permutations, has the maximum possible alternative $NL_{S'}$ nonlinearity equal to *4*. However, its nonlinearity $N_{S'}$ is equal to *8* and, as can be seen from the table below, its order of affinity ***ard*** S' is equal to *2*.

x	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
S'	1001	1101	1010	1111	1011	1110	0111	0011	1100	1000	0110	0010	0100	0001	0000	0101
A_1	**1001**	**1101**	0011	0111	1101	1001	**0111**	**0011**	**1100**	**1000**	**0110**	**0010**	1000	1100	0010	0110
A_2	1110	1011	**1010**	**1111**	**1011**	**1110**	1111	1010	0001	0100	0101	0000	**0100**	**0001**	**0000**	**0101**

The reverse situation is also possible. For example, for the permutation $S'' \in S_2^4$ generated by the permutation s'' from Appendix "A" of GOST R 34.11-94 of the form $s'' = \{6, 12, 7, 1, 5, 15, 13, 8, 4, 10, 9, 14, 0, 3, 11, 2\} \in S_{16}^1$, the value of the alternative nonlinearity $NL_{S''}$ is only *2*, but its nonlinearity $N_{S''}$ is equal to *9*, and the order of affinity ***ard*** S'' is equal to *3*.

ESTIMATES FOR THE DISTRIBUTION OF NONLINEARITY OF BALANCED MAPPINGS AND PERMUTATIONS

Consider a set of mappings with nonlinearity not exceeding a given value of the form $\{F \in P_q^{n,k} \mid N_F \leq r\}$, where $0 \leq r \leq q^n$ (in what follows, we will use the short notation $\{N_F \leq r\}$). As follows from (4), for $r \geq (q^k - 1)\, q^{n-k} - q^{n/2-k}$ the set $\{N_F \leq r\}$ coincides with the entire set $P_q^{n,k}$.

In the case of $k = 1$, it is shown that for Boolean functions for $0 \leq r < 2^{n-1} - 2^{n/2-1}$ the following inequality holds: $|\{N_f \leq r\}| \leq 2^{n+1} \sum_{i=0}^{r} \binom{2^n}{i}$. [Zubkov A.M., Serov A.A., “Bounds for the number of Boolean functions admitting affine approximations of a given accuracy”, Discrete Math. Appl., 20:5-6 (2010), 467-486] This result was generalized to the case of q-valued functions. [Ryabov V.G., “Approximation of restrictions of q-valued logic functions to linear manifolds by affine analogues”, Discrete Math. Appl., 31:6 (2021), 409–419]

In the case of an arbitrary k, with a random and equally probable choice of the function F from the set $P_q^{n,k}$ and $0 \leq r < (q^k - 1)\, q^{n-k} - q^{n/2-k}$, for the probability of the event $\{N_F \leq r\}$ the following estimate was obtained:

$$\mathbf{P}\,(N_F \leq r) \leq q^{k(n+1-q^n)} \sum_{i=0}^{r} \binom{q^n}{i} (q^k - 1)^i. \quad (10)$$

[Ryabov V.G., “Approximation of vector functions over finite fields and their restrictions to linear manifolds by affine analogues”…]

ESTIMATES FOR THE DISTRIBUTION OF NONLINEARITY OF BALANCED MAPPINGS AND PERMUTATIONS (continued)

The following theorem holds for balanced mappings.

Theorem. *Let the mapping G be chosen randomly and with equal probability from the set* $B_q^{n,k}$. *Then for* $0 \le r < q^n - q^{n-1}$ *for the probability of the event* $\{N_G \le r\}$ *we have the following estimate*

$$\mathbf{P}\,(N_G \le r) \le (q^{n-k}!)^k \prod_{h=1}^{k} (q^{n+1} - q^h) \sum_{i=0}^{r} 1/(q^n - i)! \sum_{j=0}^{i} (-1)^i / j\,!. \quad (11)$$

Corollary. *Let the mapping S be chosen randomly and with equal probability from the set* S_q^n. *Then for* $0 \le r < q^n - q^{n-1}$ *for the probability of the event* $\{N_S \le r\}$ *we have the following estimate*

$$\mathbf{P}\,(N_S \le r) \le \prod_{h=1}^{n} (q^{n+1} - q^h) \sum_{i=0}^{r} 1/(q^n - i)! \sum_{j=0}^{i} (-1)^i / j\,!. \quad (12)$$

Applying formulas (10) and (12) to mappings from $P_2^{4,4}$ and permutations from S_2^4, we obtain the relations $\mathbf{P}\,(N_F \le 7) < 0{,}12$ and $\mathbf{P}\,(N_S \le 7) < 0{,}37$. Thus, the nonlinearity of most Boolean mappings of the space $\boldsymbol{F}_2^4$ into itself and permutations of this space is greater than or equal to *8*.

CONSTRUCTION OF HIGHLY NONLINEAR VECTORIAL FUNCTIONS

For $q = 2$ and even values of $n \geq 2k$, the alternative nonlinearity reaches the upper bound in (1) only in the case of Boolean vectorial bent functions.

For $q > 2$ and $k = 1$, not all bent functions have maximum alternative nonlinearity (2) [Ryabov V.G., "Nonlinearity of bent functions over finite fields", Mat. Vopr. Kriptogr., 12:4 (2021), 87–98. In Russian]. For k > 1, a similar situation occurs for q-valued vectorial bent functions.

Method. For $q = p^m$, $m \geq 2$ and an even n/m, we take the bent function $f \in P_q^{n/m,1}$ and represent the field $\boldsymbol{F}_q$ as an m-dimensional vector space on the field $\boldsymbol{F}_p$. Then the mapping F whose coordinate functions are the coordinates f in this representation is a vectorial bent function from $P_p^{n,m}$ [Ambrosimov A.S., "Properties of bent functions of q-valued logic over finite fields", Discrete Math. Appl., 4:4 (1994), 341-350; Ryabov V.G., "Criteria for the Maximum Nonlinearity of a Function over a Finite Field", Diskr. Mat., 33:3 (2021), 79-91. In Russian].

Example. Take two bent functions from $P_9^{2,1}$: $f_1 = x_1x_2$ with nonlinearity $N_{f_1} = 64$ and $f_2 = x_1^2 \oplus 3x_2^2$ with maximum nonlinearity $N_{f_2} = 71$ [Ryabov V.G., "Maximally nonlinear functions over finite fields", Diskr. Mat., 33:1 (2021), 47-63. In Russian]. The resulting vector functions from $P_3^{4,2}$ have the form: $F_1 = \{2x_1x_3 \oplus x_1x_4 \oplus x_2x_3, x_1x_3 \oplus x_2x_4\}$ and $F_2 = \{2x_1^2 \oplus 2x_3^2 \oplus x_4^2 \oplus 2x_1x_2 \oplus x_3x_4, x_1^2 \oplus x_2^2 \oplus 2x_3^2 \oplus 2x_3x_4\}$. For their alternative nonlinearity: $NL_{F_1} = 48$ and $NL_{F_2} = 51$. For nonlinearity and order of affinity: $N_{F_1} = 64$ and $N_{F_2} = 68$; $\boldsymbol{ard}\, F_1 \geq 5$ and $\boldsymbol{ard}\, F_1 \geq 7$.

CONSTRUCTION OF HIGHLY NONLINEAR VECTORIAL FUNCTIONS (continued)

In the case $p = 2$, if the alternative nonlinearity of a Boolean vectorial function $F \in P_2^{n,2}$ satisfies the condition $NL_F = 2^{n-1} - 2^{n/2-1}$, then the following inequalities hold:

$$N_F \geq 3 \cdot (2^{n-2} - 2^{n/2-2});\ \mathbf{ard}\, F \geq 3 \text{ if } n = 4, 6 \text{ and } \mathbf{ard}\, F \geq 4 \text{ if } n \geq 8. \quad (12)$$

For n divisible by 4, the relations (12) hold for the Boolean vectorial bent function from $P_2^{n,2}$, which consists of the coordinates of the bent function from $P_4^{n/2,1}$.

[Ryabov V.G. “Approximation of vector functions over finite fields and their restrictions to linear manifolds by affine analogues”…].

In the case $p = 3$, the following statement is true.

Statement. *If the alternative nonlinearity of a Boolean vectorial function* $F \in P_3^{n,2}$ *satisfies the condition* $NL_F = 2 \cdot 3^{n-1} - 3^{n/2-1}$, *then the following inequalities hold:*

$$N_F \geq 4 \cdot (2 \cdot 3^{n-2} - 3^{n/2-2});$$

$$\mathbf{ard}\, F \geq 4, \text{ if } n = 2; \qquad \mathbf{ard}\, F \geq 7, \text{ if } n = 4; \quad (13)$$

$$\mathbf{ard}\, F \geq 8, \text{ if } n = 6; \qquad \mathbf{ard}\, F \geq 9, \text{ if } n \geq 8.$$

Corollary. *For* $q = 9$ *and* n *divisible by 4, for a vector bent function* $F \in P_3^{n,2}$ *consisting of the coordinates of a maximally nonlinear bent function* $f \in P_9^{n,1}$, *the relations (13) hold.*

THANK YOU FOR YOUR ATTENTION!

E-mail: 4vryabov@gmail.com