ISSN 0032-9460, Problems of Information Transmission, 2008, Vol. 44, No. 3, pp. 266–286. © Pleiades Publishing, Inc., 2008.
Original Russian Text © N.N. Tokareva, 2008, published in Problemy Peredachi Informatsii, 2008, Vol. 44, No. 3, pp. 105–127.

## INFORMATION PROTECTION

# On Quadratic Approximations in Block Ciphers[1]

**N. N. Tokareva**

*Sobolev Institute of Mathematics, Siberian Branch of the RAS, Novosibirsk*
tokareva@math.nsc.ru

Received February 8, 2008; in final form, April 9, 2008

**Abstract**—We consider quadratic approximations (of Boolean functions) of a special form and their potential applications in block cipher cryptanalysis. We show that the use of $k$-bent functions as ciphering functions extremely increases the resistance of ciphers to such approximations. We consider examples of 4-bit permutations recommended for use in S-boxes of the algorithms GOST 28147-89, DES, and $s^3$DES; we show that in almost all cases there exist more probable (than linear) quadratic relations of a special form on input and output bits of these permutations.

**DOI:** 10.1134/S0032946008030083

## 1. INTRODUCTION

**Linear cryptanalysis.** The linear cryptanalysis (LC) method for the FEAL block cipher was proposed in 1992 by Matsui and Yamagishi [1], and for the DES cipher, in 1993 by Matsui [2]. Nowadays, this method (along with differential cryptanalysis [3]) is reputed to be one of the most efficient ones. The idea of the method is as follows. First, for a known ciphering algorithm, a linear relation $L$ on bits of a plaintext, ciphertext, and key is found that holds with probability $p = 1/2+\varepsilon$ far enough from $1/2$. Then, for a fixed unknown key $K$, a cryptanalist collects statistics of $N$ pairs {plaintext; the corresponding ciphertext} and based on it, taking into account the sign of $\varepsilon$, distinguishes between two simple statistical hypotheses: whether or not the relation $L$ holds for this unknown key $K$. As a result, a new probabilistic relation for bits of $K$ is found. For this method to work reliably, $N$ should be proportional to $|\varepsilon|^{-2}$.

There are many works devoted to various generalizations and applications of the LC method. Here are some of them. A detailed analysis of the LC method (in particular, for DES) is given in [4]; see also [5–8]. To improve the efficiency of the LC method, in [9] it was proposed to consider several linear approximations for one combination of key bits; this subject was further developed in [10]. The authors of [11] presented a way to improve the LC method (in particular, for the LOKI91 cipher) by considering probabilistic behavior of some bits in approximation instead of their fixed values. Among recent papers that develop the LC method, we mention [12, 13].

A series of works is devoted to problems of resistance of various ciphering algorithms to linear cryptanalysis. Problems of construction of Feistel-type ciphering schemes resistant to methods of linear and differential cryptanalysis were considered in [14]. In 2001 it was proved [15] that the Russian GOST 28147-89 algorithm (with at least five rounds of ciphering for linear cryptanalysis, and seven rounds for differential cryptanalysis) is resistant to these methods. Resistance of the ciphers RC5, RC6, IDEA, Serpent, AES, Blowfish, and Khufu to the LC method was analyzed in [16–20].

[1] Supported in part by the Siberian Branch of the RAS Integration Project no. 35 "Tree Catalog of Mathematical Internet Resources mathtree.ru," the Russian Foundation for Basic Research, project nos. 07-01-00248 and 08-01-00671, and the Russian Science Support Foundation.

Other works are devoted to the analysis of various classes of approximating functions and construction of functions with the best resistance to such approximations. These works consider *bent functions* [21], i.e., Boolean functions of an even number of variables that have the maximum possible Hamming distance from the set of all linear functions (see the surveys [22, 23]), and their generalizations: *semi-bent functions* [24], *partial bent functions* [22], $\mathbb{Z}$*-bent functions* [25], *homogeneous bent functions* [26], *hyper-bent functions* [27–31], *negabent functions* [32], etc.

**Nonlinear cryptanalysis.** A general approach to the use of nonlinear approximations in linear cryptanalysis was proposed in 1996 [33]. The main idea is simple: extend the class of approximating functions (in $m$ variables) by nonlinear functions, thereby improving the approximation quality. However, here a cryptanalist encounters the following problems.

*How to choose a good nonlinear approximation efficiently*? In the linear case, this can be done by the extensive search of all $2^m$ linear functions. In the general case, exhaustive search of all $2^{2^m}$ Boolean functions is impossible even for small values of $m$.

*How to combine nonlinear approximations of different rounds*? Consider a simple example. Let the $i$th round of ciphering, which takes an intermediate ciphertext $C^{(i-1)}$ to $C^{(i)}$, be organized as follows: $C^{(i)} = S^i(C^{(i-1)} \oplus K^{(i)})$, where $K^{(i)}$ is the subkey of the $i$th round and $S^i$ is a known nonlinear transformation. Assume that a cryptanalist has found an approximation of $S^i$ by a function $f^i$; i.e., the equality $S^i(\boldsymbol{x}) = f^i(\boldsymbol{x})$ holds with high probability for an arbitrary $\boldsymbol{x}$. Then, if $f^i$ is linear, for the $i$th round we have the approximation $C^{(i)} = f^i(C^{(i-1)} \oplus K^{(i)}) = f^i(C^{(i-1)}) \oplus f^i(K^{(i)})$. Since the dependence on the block $C^{(i-1)}$ and on the subkey $K^{(i)}$ is explicit here, this approximation of the $i$th round can be used in the whole chain of round approximations. In the general case, combining round approximations is difficult.

Towards the solution of the first problem, we mention the study [34], which deals with finding quadratic relations for particular permutations used in DES S-boxes, experimental research of [35], and the work [36] concerning application of heuristic algorithms for finding good nonlinear approximations (with examples for S-boxes of the MARS cipher). Probabilistic aspects of approximation of a random Boolean function by the set of all quadratic functions were studied in [37]. Problems of nonlinear approximations of Boolean functions (using their reduced representation) were considered in [38]. The author is unaware of any works aimed at the solution of the second problem.

On the whole, nonlinear cryptanalysis is not yet properly developed.

**Quadratic cryptanalysis.** This paper studies the capabilities of quadratic cryptanalysis of block ciphers based on quadratic approximations of a special form. Namely, in [39] for each integer $k$, $1 \le k \le m/2$, there was defined a binary operation $\langle \cdot, \cdot \rangle_k$ on the set of vectors $\mathbb{Z}_2^m$, which, reasoning from its properties, can be viewed as an analog of the inner product of vectors over $\mathbb{Z}_2$. The definition is given in the framework of the coding-theoretic approach; in essence, it uses the classification of $\mathbb{Z}_4$-linear Hadamard-like codes given in [40, 41]. For a fixed vector $\boldsymbol{u} \in \mathbb{Z}_2^m$, the function $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ in the variables $v_1, \ldots, v_m$ is either linear or quadratic (see Section 2). In this paper, we propose to use approximation by all functions of the form $\langle \boldsymbol{u}, \pi(\boldsymbol{v}) \rangle_k$, where $\pi$ is any permutation on $m$ coordinates and the parameters $\boldsymbol{u}$ and $k$ are arbitrary. The set of such functions consists of $2^m$ (i.e., all) linear functions and at most $2^{m(1+\log_2 m)}$ quadratic functions, so a cryptanalist is able to apply exhaustive search (see Section 3). The choice of these functions is due to the existence of simple formulas to compute the Hamming distance from an arbitrary Boolean function to the class of functions $\langle \boldsymbol{u}, \pi(\boldsymbol{v}) \rangle_k$ for fixed $\pi$ and $k$, and to properties of these functions, which are close to linear.

The paper is of theoretical character. We propose modifications of Matsui's linear cryptanalysis algorithms [2, Algorithms 1 and 2] for an extended class of approximating functions. We give formulas for computing absolute values of biases and reliability of algorithms. We show that using $k$-bent functions as ciphering functions reduces the maximum absolute value of the bias to its minimum

value and hence extremely increases the cipher resistance to these quadratic approximations (see Section 4). We consider examples of 4-bit permutations recommended for use in substitution boxes (S-*boxes*) of the algorithms GOST 28147-89, DES, and $s^3$DES; with the aid of computer, we show that for all (but one) of these permutations there exist more probable (than linear) quadratic relations of a special form on input and output bits of these permutations (see Section 5). Properties of approximating functions that can be used for combining round approximations are considered in Section 6. Practical results on application of quadratic approximations in cryptanalysis will be presented in a subsequent paper of the author.

## 2. $\langle\cdot,\cdot\rangle_k$ OPERATION

In this section, following [39], we give several necessary definitions. Let $m$ be an integer, and let $\mathfrak{F}_m$ be the class of all Boolean functions in $m$ variables. Let $\boldsymbol{v} = (v_1, \ldots, v_m)$ and $\boldsymbol{u} = (u_1, \ldots, u_m)$, where $u_i, v_i \in \mathbb{Z}_2$. For any integer $k$, $1 \le k \le m/2$, a binary operation $\langle\cdot,\cdot\rangle_k \colon \mathbb{Z}_2^m \times \mathbb{Z}_2^m \to \mathbb{Z}_2$ is defined as follows:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle_k = \left( \bigoplus_{i=1}^{k} \bigoplus_{j=i}^{k} (u_{2i-1} \oplus u_{2i})(u_{2j-1} \oplus u_{2j})(v_{2i-1} \oplus v_{2i})(v_{2j-1} \oplus v_{2j}) \right) \oplus \langle \boldsymbol{u}, \boldsymbol{v}\rangle, \tag{1}$$

where $\langle\cdot,\cdot\rangle$ is the standard inner product of binary vectors over $\mathbb{Z}_2$, i.e.,

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = u_1 v_1 \oplus \ldots \oplus u_m v_m,$$

and $\oplus$ denotes addition modulo 2. Note that the coordinates $u_1, \ldots, u_m$ (as well as $v_1, \ldots, v_m$) in the operation $\langle\cdot,\cdot\rangle_k$ are dissimilar. Namely, for a given $k$, precisely $2k$ first coordinates of each of the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ occur in both quadratic and linear terms; the others occur in linear terms only. It follows from the definition that

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle_1 = \langle \boldsymbol{u}, \widehat{\boldsymbol{v}}\rangle, \tag{2}$$

where $\widehat{\boldsymbol{v}}$ is obtained from $\boldsymbol{v}$ by interchanging $v_1$ and $v_2$. It is also clear that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle_k = \langle \boldsymbol{v}, \boldsymbol{u}\rangle_k$ and $\langle a\boldsymbol{u}, \boldsymbol{v}\rangle_k = a\langle \boldsymbol{v}, \boldsymbol{u}\rangle_k$ for any $a \in \mathbb{Z}_2$. As an example, consider the expression for $\langle \boldsymbol{u}, \boldsymbol{v}\rangle_2$ in the case of $m = 4$:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle_2 = (u_1 \oplus u_2)(u_3 \oplus u_4)(v_1 \oplus v_2)(v_3 \oplus v_4) \oplus u_2 v_1 \oplus u_1 v_2 \oplus u_4 v_3 \oplus u_3 v_4. \tag{3}$$

The binary operation $\langle\cdot,\cdot\rangle_k$ was introduced in [39], were its properties were also studied, which allow us to consider this operation to be an analog of the inner product. The integer function $W_f^{(k)}$ defined on $\mathbb{Z}_2^m$ by the equality

$$W_f^{(k)}(\boldsymbol{v}) = \sum_{\boldsymbol{u} \in \mathbb{Z}_2^m} (-1)^{\langle \boldsymbol{u}, \boldsymbol{v}\rangle_k \oplus f(\boldsymbol{u})} \quad \text{for any } \boldsymbol{v} \in \mathbb{Z}_2^m$$

is called the *k-Walsh–Hadamard transform* of a Boolean function $f \in \mathfrak{F}_m$. By [39], for $W_f^{(k)}$ there is an analog of the Parseval equality:

$$\sum_{\boldsymbol{v} \in \mathbb{Z}_2^m} \left( W_f^{(k)}(\boldsymbol{v}) \right)^2 = 2^{2m} \quad \text{for any } f \in \mathfrak{F}_m;$$

hence, for any $k$ and any $f \in \mathfrak{F}_m$ we have

$$\max_{\boldsymbol{v} \in \mathbb{Z}_2^m} |W_f^{(k)}(\boldsymbol{v})| \ge 2^{m/2}. \tag{4}$$

Recall that the *Hamming distance* $\operatorname{dist}(\cdot,\cdot)$ between Boolean functions in $m$ variables is defined to be the number of positions in which their vectors of values differ. If a function $g_{\boldsymbol{u}}^{(k)}$ is defined as $g_{\boldsymbol{u}}^{(k)}(\boldsymbol{v}) = \langle \boldsymbol{u}, \boldsymbol{v}\rangle_k$, then

$$\operatorname{dist}(f, g_{\boldsymbol{u}}^{(k)}) = 2^{m-1} - \frac{1}{2}W_f^{(k)}(\boldsymbol{u}),$$

$$\operatorname{dist}(f, g_{\boldsymbol{u}}^{(k)} \oplus 1) = 2^{m-1} + \frac{1}{2}W_f^{(k)}(\boldsymbol{u}).$$

The distance between a function $f$ and the set of functions $\{\langle \boldsymbol{u}, \boldsymbol{v}\rangle_k \oplus a \mid \boldsymbol{u} \in \mathbb{Z}_2^m,\ a \in \mathbb{Z}_2\}$ in the variables $v_1, \ldots, v_m$ is called the *$k$-nonlinearity* of $f$; we denote it by $N_f^{(k)}$. We have

$$N_f^{(k)} = 2^{m-1} - \frac{1}{2}\max_{\boldsymbol{v}\in\mathbb{Z}_2^m}|W_f^{(k)}(\boldsymbol{v})|;$$

hence, the $k$-nonlinearity of $f$ is not greater than $2^{m-1} - 2^{(m/2)-1}$. For even $m$, a Boolean function $f \in \mathfrak{F}_m$ is called a *$k$-bent function* if all the coefficients $W_f^{(j)}(\boldsymbol{v})$, $j = 1, \ldots, k$, $\boldsymbol{v} \in \mathbb{Z}_2^m$, equal $\pm 2^{m/2}$. In other words, a $k$-bent function is a function for which each nonlinearity parameter $N_f^{(j)}$, $j = 1, \ldots, k$, takes its maximum possible value $2^{m-1} - 2^{(m/2)-1}$. Let $\mathfrak{B}_m^k$ be the class of all $k$-bent functions in $m$ variables. One can show that $\mathfrak{B}_m^1$ coincides with the class of ordinary bent functions. There are strict inclusions $\mathfrak{B}_m^1 \supset \ldots \supset \mathfrak{B}_m^{m/2}$; the set $\mathfrak{B}_m^{m/2}$ is nonempty. For methods of constructing $k$-bent functions, see [39, 42].

## 3. A CLASS OF APPROXIMATING FUNCTIONS

Consider the following class of Boolean functions in the variables $v_1, \ldots, v_m$, where $m$ is even. For any $k$, $1 \le k \le m/2$, and any permutation $\pi \in S_m$ on the $m$ variables, let

$$\mathfrak{A}_{m,0}^k(\pi) = \{\langle \boldsymbol{u}, \pi(\boldsymbol{v})\rangle_k \mid \boldsymbol{u} \in \mathbb{Z}_2^m\}.$$

Note that for any $\pi \in S_m$, the set $\mathfrak{A}_{m,0}^1(\pi)$ consists of all linear functions. We approximate Boolean functions in the variables $v_1, \ldots, v_m$ used in ciphering by functions of the set

$$\Delta_m = \bigcup_{1\le k\le m/2}\ \bigcup_{\pi\in S_m} \mathfrak{A}_{m,0}^k(\pi);$$

throughout what follows, we refer to this set as the *class of approximating functions*. Informally speaking, owing to arbitrary permutations $\pi$ of the variables $v_1, \ldots, v_m$, we eliminate the "dissimilarity" of these variables in the function $\langle \boldsymbol{u}, \boldsymbol{v}\rangle_k$.

Let us find the cardinality of the class $\Delta_m$ and a way to enumerate its elements. The main difficulty here is due to the fact that in general the sets $\mathfrak{A}_{m,0}^{k'}(\pi')$ and $\mathfrak{A}_{m,0}^{k''}(\pi'')$ are not disjoint.

For a Boolean function $f \in \mathfrak{F}_m$, let the set $\operatorname{ANF}(f)$ consist of all monomials of its algebraic normal form (also called a *polynomial form* or *Zhegalkin polynomial* of the function). For example, for the function $g(v_1, v_2, v_3, v_4) = v_1v_2 \oplus v_1v_3 \oplus v_1v_4 \oplus v_2v_3 \oplus v_3v_4 \oplus v_2 \oplus v_3 \oplus 1$ we have $\operatorname{ANF}(g) = \{v_1v_2,\ v_1v_3,\ v_1v_4,\ v_2v_3,\ v_3v_4,\ v_2,\ v_3,\ 1\}$. For a fixed permutation $\pi \in S_m$, denote by $f^\pi$ the Boolean function defined as $f^\pi(\boldsymbol{v}) = f(\pi(\boldsymbol{v}))$. We divide the variables of a Boolean function $f \in \mathfrak{F}_m$ into pairs; the pair $\{v_{2i-1}, v_{2i}\}$ is given the number $i$. By $\operatorname{Act}(f)$, we denote the subset of $\{1, 2, \ldots, m/2\}$ of the maximum possible cardinality such that for any two distinct elements $i, j$ in $\operatorname{Act}(f)$ the monomials $v_{2i-1}v_{2j-1}$, $v_{2i-1}v_{2j}$, $v_{2i}v_{2j-1}$, and $v_{2i}v_{2j}$ belong to $\operatorname{ANF}(f)$. We say that a pair of variables with number $i$ is *active* for $f$ whenever $i \in \operatorname{Act}(f)$. Note that the cardinality of $\operatorname{Act}(f)$ for any $f$ is either zero or not less than two. By $\rho = \rho(f)$ we denote any permutation in $S_m$ such that $|\operatorname{Act}(f^\rho)| = \max\limits_{\pi\in S_m}|\operatorname{Act}(f^\pi)|$. For instance, consider the set $\operatorname{Act}(g)$ for the function $g$ defined above. Since the monomials $v_1v_3$, $v_1v_4$, and $v_2v_3$ belong to $\operatorname{ANF}(g)$ and the monomial $v_2v_4$ does not, we have $\operatorname{Act}(g) = \varnothing$. However, for $\rho = (1, 3, 2, 4)$ we have $\operatorname{Act}(g^\rho) = \{1, 2\}$.

**Theorem 1.** *A Boolean function $f \in \mathfrak{F}_m$ of degree at most two such that $f(\mathbf{0}) = 0$ belongs to the class $\Delta_m$ if and only if $f$ satisfies the following conditions*:

1. *For any two distinct numbers $i, j$ $(1 \le i, j \le m/2)$, the monomials*

$$v_{2i-1}v_{2j-1},\ v_{2i-1}v_{2j},\ v_{2i}v_{2j-1},\ v_{2i}v_{2j}$$

   *either all belong to* $\mathrm{ANF}(f^\rho)$ *or all do not*;
2. *The set* $\mathrm{ANF}(f^\rho)$ *does not contain monomials of the form* $v_{2i-1}v_{2i}$;
3. *If a pair $i$ is active for $f^\rho$, then exactly one of the variables $v_{2i-1}$ and $v_{2i}$ belongs to* $\mathrm{ANF}(f^\rho)$.

**Proof.** ($\Leftarrow$) Let a function $f$ of degree at most two with $f(\mathbf{0}) = 0$ satisfy conditions 1–3 of the theorem. If the set $\mathrm{Act}(f^\rho)$ is empty, then $f$ is linear according to conditions 1 and 2 and hence belongs to $\Delta_m$.

Now assume that $\mathrm{Act}(f^\rho)$ is nonempty and has the form $\mathrm{Act}(f^\rho) = \{i_1, \dots, i_k\}$, where $2 \le k \le m/2$. Let $j_1, \dots, j_{(m/2)-k}$ be the numbers of inactive pairs of variables of the function $f^\rho$. Consider a permutation $\tau \in S_m$ such that $\tau(i_s) = s$ for any $s = 1, \dots, k$ and $\tau(j_s) = k + s$ for any $s = 1, \dots, (m/2) - k$. Permute pairs of variables of $f^\rho$ according to $\tau$. Namely, consider the function $f^{\rho\circ\pi}$ (here and in what follows, the notation $\rho \circ \pi$ means that we first apply the permutation $\rho$ and then $\pi$), where $\pi \in S_m$ is defined via $\tau$ as follows: $\pi(2s-1) = 2\tau(s) - 1$ and $\pi(2s) = 2\tau(s)$, for any $s = 1, \dots, m/2$. It is easily seen that conditions 1–3 remain valid upon replacing $f^\rho$ with $f^{\rho\circ\pi}$ in each of them, and the set $\mathrm{Act}(f^{\rho\circ\pi}) = \{1, \dots, k\}$, as well as $\mathrm{Act}(f^\rho)$, is of cardinality $k$. Therefore, hereafter we assume without loss of generality that $\mathrm{Act}(f^\rho) = \{1, \dots, k\}$.

Note that by conditions 1 and 2, the number $k$ uniquely determines the quadratic part of $f^\rho$, which is of the form

$$\bigoplus_{i=1}^{k-1} \bigoplus_{j=i+1}^{k} (v_{2i-1}v_{2j-1} \oplus v_{2i-1}v_{2j} \oplus v_{2i}v_{2j-1} \oplus v_{2i}v_{2j}). \tag{5}$$

We show that $f^\rho$ belongs to the set $\mathfrak{A}^k_{m,0}$. Consider a vector $\boldsymbol{u} \in \mathbb{Z}_2^m$ such that

$$u_t = 1 \quad \Longleftrightarrow \quad \begin{cases} v_t \notin \mathrm{ANF}(f^\rho) & \text{for } t = 1, \dots, 2k, \\ v_t \in \mathrm{ANF}(f^\rho) & \text{for } t = 2k+1, \dots, m. \end{cases}$$

Then $f^\rho(\boldsymbol{v}) = \langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$. Indeed, by the definition of $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ we have

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k = \left( \bigoplus_{i=1}^{k-1} \bigoplus_{j=i+1}^{k} Y_i Y_j \right) \oplus \left( \bigoplus_{i=1}^{k} (u_{2i}v_{2i-1} \oplus u_{2i-1}v_{2i}) \right) \oplus \left( \bigoplus_{i=k+1}^{m/2} (u_{2i-1}v_{2i-1} \oplus u_{2i}v_{2i}) \right), \tag{6}$$

where $Y_i = (u_{2i-1} \oplus u_{2i})(v_{2i-1} \oplus v_{2i})$. Using condition 3 and the definition of $\boldsymbol{u}$, we obtain $u_{2i-1} \oplus u_{2i} = 1$ for $i = 1, \dots, k$; hence, $Y_iY_j = v_{2i-1}v_{2j-1} \oplus v_{2i-1}v_{2j} \oplus v_{2i}v_{2j-1} \oplus v_{2i}v_{2j}$, where $1 \le i < j \le k$. Thus, the quadratic part of the function $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ coincides with (5). From (6) we immediately obtain that the linear parts of the functions $f^\rho$ and $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ also coincide. Therefore, since $f^\rho(\mathbf{0}) = \langle \boldsymbol{u}, \mathbf{0} \rangle_k = 0$ and both functions $f^\rho$ and $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ are of degree 2, they coincide. Thus, we have shown that $f^\rho$ belongs to the class $\mathfrak{A}^k_{m,0}(\mathrm{id})$, where id denotes the identity permutation. It remains to note that we have

$$f^\sigma \in \mathfrak{A}^k_{m,0}(\mathrm{id}) \quad \Longleftrightarrow \quad f \in \mathfrak{A}^k_{m,0}(\sigma^{-1}), \quad \text{for any permutation } \sigma \in S_m,$$

which follows from the equivalence

$$\exists \boldsymbol{u} :\ f^\sigma(\boldsymbol{v}) = \langle \boldsymbol{u}, \boldsymbol{v} \rangle_k \quad \Longleftrightarrow \quad \exists \boldsymbol{u} :\ f(\boldsymbol{v}) = \langle \boldsymbol{u}, \sigma^{-1}(\boldsymbol{v}) \rangle_k.$$

Hence we finally conclude that $f$ belongs to $\mathfrak{A}^k_{m,0}(\rho^{-1})$ and therefore to $\Delta_m$.

($\Rightarrow$) If $f$ is a linear function, conditions 1–3 obviously hold. Let $f$ have a nontrivial quadratic part. Then, since $f$ belongs to some class $\mathfrak{A}^k_{m,0}(\pi)$, the cardinality of the quadratic part of $f$ is $4\binom{s}{2}$ for an appropriate $s$, $2 \le s \le k$, which directly follows from the definition of $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$. Since $f^\rho$ is also contained in the class $\Delta_m$, we have $|\mathrm{Act}(f^\rho)| = s$ (for instance, we may take the permutation $\pi^{-1}$ for $\rho$). Then the quadratic part of $\mathrm{ANF}(f^\rho)$ consists only of monomials of the form $v_{2i-1}v_{2j-1}$, $v_{2i-1}v_{2j}$, $v_{2i}v_{2j-1}$, and $v_{2i}v_{2j}$ for any distinct $i, j \in \mathrm{Act}(f^\rho)$; hence, conditions 1 and 2 are fulfilled. The definition of $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ implies the validity of condition 3. $\triangle$

**Corollary.** *For any even $m$, we have*

$$|\Delta_m| = 2^m \left( 1 + \sum_{k=2}^{m/2} \binom{m}{2k} \frac{(2k-1)!!}{2^k} \right).$$

**Proof.** The class $\Delta_m$ contains exactly $2^m$ linear Boolean functions. Using Theorem 1, for any fixed $k$, $2 \le k \le m/2$, we find the number of quadratic functions $f$ in $\Delta_m$ such that $|\mathrm{Act}(f^\rho)| = k$. Each of these functions $f$ is uniquely determined by a set of $k$ unordered pairs of variables (upon the action of the corresponding permutation $\rho$, all these pairs become active) and its linear part. A set of $k$ unordered pairs can be chosen in

$$\frac{1}{k!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2k+2}{2} = \frac{m!}{2^k k!(m-2k)!}$$

ways. For each chosen pair of variables, exactly one variable of this pair belongs to $\mathrm{ANF}(f)$ by condition 3 of Theorem 1. Variables that are not contained in the chosen pairs enter or not $\mathrm{ANF}(f)$ freely. Thus, the number of functions $f \in \Delta_m$, $|\mathrm{Act}(f^\rho)| = k$, is

$$\frac{m!}{2^k k!(m-2k)!} 2^k 2^{m-2k} = \binom{m}{2k} 2^{m-k}(2k-1)!!.$$

Summing over all $k$, $2 \le k \le m/2$, and taking account of linear functions, we obtain the desired expression for the cardinality of $\Delta_m$. $\triangle$

For example, $|\Delta_4| = 28$, $|\Delta_6| = 904$, and $|\Delta_8| = 28\,816$; the number of linear functions in each of these classes is 16, 64, and 256, respectively. It easily follows from the corollary that $|\Delta_m|$ is not greater than $e2^m m!$, which is certainly less than $2^{m(1+\log_2 m)}$. Note that the number of all quadratic functions in $m$ variables is proportional to $2^{m^2}$, and functions of the form $\langle \boldsymbol{u}, \pi(\boldsymbol{v}) \rangle_k$ constitute an extremely small part of them as $m \to \infty$.

Theorem 1 and the corollary suggest a way to enumerate all elements of $\Delta_m$ without repetitions.

## 4. QUADRATIC APPROXIMATIONS IN BLOCK CIPHERS

The main idea of our approach is to extend the domain of search for the most probable relations on bits of a plaintext, ciphertext, and key: from the set of linear relations to a set of linear and quadratic relations of a special form. We mainly follow the notation of [22].

Consider a block cipher with $r$ rounds of ciphering. Let

$m = m_{\mathrm{text}}$ be the length of a plaintext and ciphertext;
$P$ be a plaintext, $P \in \mathbb{Z}_2^m$;
$m_{\mathrm{key}}$ be the key length;
$K$ be a ciphering key, $K \in \mathbb{Z}_2^{m_{\mathrm{key}}}$;
$F\colon \mathbb{Z}_2^m \times \mathbb{Z}_2^{m_{\mathrm{key}}} \to \mathbb{Z}_2^m$ be a transformation which is one-to-one for any fixed value of the second argument;

$C = F(P, K)$ be a ciphertext, $C \in \mathbb{Z}_2^m$;
$m'_{\text{key}}$ be the length of a round subkey;
$K^{(i)}$ be a subkey of the $i$th encryption round, $K^{(i)} \in \mathbb{Z}_2^{m'_{\text{key}}}$, $1 \le i \le r$, which is determined by a key $K$;
$F_i \colon \mathbb{Z}_2^m \times \mathbb{Z}_2^{m'_{\text{key}}} \to \mathbb{Z}_2^m$ be a transformation of the $i$th encryption round, $1 \le i \le r$, which is one-to-one for any fixed value of the second argument;
$C^{(0)} = P$;
$C^{(i)} = F_i(C^{(i-1)}, K^{(i)})$ be an intermediate ciphertext, $C^{(i)} \in \mathbb{Z}_2^m$, $1 \le i \le r$;
$C = C^{(r)}$ be a final ciphertext.

We assume that all plaintexts $P$ (as well as keys $K$) are equiprobable. Throughout what follows, the numbers $m$, $m_{\text{key}}$, and $m'_{\text{key}}$ are assumed to be even.

### *4.1. First Algorithm*

The algorithm is based upon the following equality:

$$\langle \boldsymbol{a}, \pi(P) \rangle_i \oplus \langle \boldsymbol{b}, \sigma(C) \rangle_j = \langle \boldsymbol{d}, \tau(K) \rangle_k, \tag{7}$$

where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}_2^m$ and $\boldsymbol{d} \in \mathbb{Z}_2^{m_{\text{key}}}$ are vectors chosen in some way; $\pi, \sigma \in S_m$ and $\tau \in S_{m_{\text{key}}}$ are fixed permutations; and $i$, $j$, and $k$ are integers such that $1 \le i, j \le m/2$ and $1 \le k \le m_{\text{key}}/2$.

We assume that (7) holds with probability $p = 1/2 + \varepsilon$ such that $0 < |\varepsilon| \le 1/2$. We call $\varepsilon$ the *bias* of equality (7). A separate problem for each particular ciphering algorithm is to choose values of $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{d}$, $\pi$, $\sigma$, $\tau$, $i$, $j$, and $k$ so that to make $|\varepsilon|$ as large as possible. In the present paper, we do not deal with this problem. Note that a choice of $i$, $j$, and $k$ influences the form of relation (7) in the following way. If a given parameter ($i$, $j$, or $k$) equals 1, then bits of the corresponding block (plaintext $P$, ciphertext $C$, or key $K$) enter relation (7) linearly, which can be used if we add this relation to a linear system of equations. As the parameter ($i$, $j$, or $k$) grows, the number of bits of a block that enter the nonlinear part of the relation grows proportionally.

Let a ciphering key $K$ be fixed. Consider the set

$$\{(P_t, C_t) \mid t = 1, \ldots, N\}$$

of known pairs (plaintext, ciphertext), $C_t = F(P_t, K)$. The following algorithm is a modification of the Matsui algorithm [2] for finding one bit of the key, which is based on the maximum likelihood principle.

Algorithm 1

- Compute $N_0 = \big|\{t : \langle \boldsymbol{a}, \pi(P_t) \rangle_i \oplus \langle \boldsymbol{b}, \sigma(C_t) \rangle_j = 0\}\big|$;
- Set $\langle \boldsymbol{d}, \tau(K) \rangle_k = \begin{cases} 0 & \text{if } \left(N_0 - \dfrac{N}{2}\right)\varepsilon > 0, \\ 1 & \text{otherwise;} \end{cases}$
- Using the obtained relation, try to find the key.

End of Algorithm

Recall that the *reliability* $\xi_0$ of an algorithm based on a statistical classification procedure is the expectation of the probability that it works correctly. In our case, to work correctly means to find a correct relation on bits of the key. Here we assume that the sought-for key is chosen in the whole

key space randomly, equiprobably, and independently of the set of plaintexts (for details, see [22]). Thus,

$$\xi_0 = \mathbf{E}\{\xi(K)\} = \frac{1}{2^{m_{\text{key}}}} \sum_{K \in \mathbb{Z}_2^{m_{\text{key}}}} \xi(K),$$

where $\xi(K)$ is the probability of choosing plaintexts $P_1, \dots, P_N$ such that the relation for the bits of $K$ will be found correctly. If $p(K) = 1/2 + \varepsilon(K)$, where $\varepsilon(K) \neq 0$ is the probability that (7) holds for a fixed key $K$, then

$$\xi(K) = \sum_{s=0}^{N/2} \binom{N}{s} \left(\frac{1}{2} - |\varepsilon(K)|\right)^s \left(\frac{1}{2} + |\varepsilon(K)|\right)^{N-s}.$$

The reliability $\xi_0$ of Algorithm 1 can be estimated in the same way as in the case of linear cryptanalysis (under additional cryptographic assumptions; see [2, 22] for details) with the use of the normal distribution function; namely,

$$\xi_0 \simeq \Phi_{0,1}(-2|\varepsilon|\sqrt{N}) = \int\limits_{-2|\varepsilon|\sqrt{N}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy. \tag{8}$$

Now we give formulas for computing absolute values of biases and indicate properties of Boolean functions used in ciphering whose presence makes a cipher resistant to the quadratic approximations in question.

For a fixed key $K$, any integers $i$ and $j$ such that $1 \le i, j \le m/2$, and arbitrary permutations $\pi, \sigma \in S_m$, we denote by $\varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0)$ a real number in the interval $[-1/2, 1/2]$ such that the probability that the equality

$$\langle \boldsymbol{a}, \pi(P)\rangle_i \oplus \langle \boldsymbol{b}, \sigma(F(P,K))\rangle_j = 0 \tag{9}$$

holds is $1/2 + \varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0)$.

**Proposition 1.** *For any map $F(\cdot\,, K)\colon \mathbb{Z}_2^m \to \mathbb{Z}_2^m$ and any permutations $\pi, \sigma \in S_m$, we have*

$$2^{m+1}\varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0) = W^{(i)}_{\langle \boldsymbol{b},\sigma(F(\pi^{-1}(\cdot),K))\rangle_j}(\boldsymbol{a}).$$

**Proof.** Let $\mathbb{Z}_2^m = M_0 \cup M_1$, where

$$M_x = \{\boldsymbol{u} \in \mathbb{Z}_2^m \mid \langle \boldsymbol{a}, \pi(\boldsymbol{u})\rangle_i \oplus \langle \boldsymbol{b}, \sigma(F(\boldsymbol{u},K))\rangle_j = x\}$$

for $x = 0, 1$. The definition of the $i$-Walsh–Hadamard coefficient $W^{(i)}_{\langle \boldsymbol{b},\sigma(F(\pi^{-1}(\cdot),K))\rangle_j}(\boldsymbol{a})$ implies that

$$W^{(i)}_{\langle \boldsymbol{b},\sigma(F(\pi^{-1}(\cdot),K))\rangle_j}(\boldsymbol{a}) = \sum_{\boldsymbol{u} \in \mathbb{Z}_2^m} (-1)^{\langle \boldsymbol{a},\pi(\boldsymbol{u})\rangle_i \oplus \langle \boldsymbol{b},\sigma(F(\boldsymbol{u},K))\rangle_j} = |M_0| - |M_1|.$$

Using (9), we obtain $|M_0| = 2^m(1/2 + \varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0))$; hence,

$$|M_0| - |M_1| = 2^{m+1}\varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0). \quad \triangle$$

Recall that $\varepsilon(K)$ denotes the bias of equality (7) for a fixed key $K$. Note that for any $k$, $\boldsymbol{d}$, and $\tau$, we have

$$|\varepsilon(K)| = |\varepsilon_{j,\boldsymbol{b},\sigma}^{i,\boldsymbol{a},\pi}(K;0)|. \tag{10}$$

**Theorem 2.** *Let a key $K \in \mathbb{Z}_2^{m_{\text{key}}}$ be fixed. If a vector $\boldsymbol{b} \in \mathbb{Z}_2^m$, permutations $\pi, \sigma \in S_m$, and a parameter $j$, $1 \le j \le m/2$, are such that the function*

$$\langle \boldsymbol{b}, \sigma(F(\pi^{-1}(\cdot), K))\rangle_j : \mathbb{Z}_2^m \to \mathbb{Z}_2^m$$

*is an $(m/2)$-bent function, then*

$$\max_{i,k,\boldsymbol{a},\boldsymbol{d},\tau} |\varepsilon(K)| = \min_{i,k,\boldsymbol{a},\boldsymbol{d},\tau} |\varepsilon(K)| = 2^{-(m/2)-1}.$$

**Proof.** Since the function $\langle \boldsymbol{b}, \sigma(F(\pi^{-1}(\cdot), K))\rangle_j$ belongs to the class $\mathfrak{B}_m^{m/2}$, we have the equality $|W^{(i)}_{\langle \boldsymbol{b},\sigma(F(\pi^{-1}(\cdot),K))\rangle_j}(\boldsymbol{a})| = \pm 2^{m/2}$ for any $\boldsymbol{a} \in \mathbb{Z}_2^m$ and any $i$, $1 \le i \le m/2$. Then Proposition 1 and equation (10) immediately imply that for any parameters $k$, $\boldsymbol{d}$, and $\tau$, all values of $|\varepsilon(K)|$ equal $2^{-(m/2)-1}$, whence the desired result follows. $\triangle$

Inequality (4) implies that $2^{-(m/2)-1}$ is the minimum possible value of $\max\limits_{\boldsymbol{a} \in \mathbb{Z}_2^m} |\varepsilon(K)|$ for any fixed $i$, $j$, $k$, $\boldsymbol{b}$, $\boldsymbol{d}$, $\pi$, $\sigma$, and $\tau$. By Theorem 2, this minimum value can be attained only when using $(m/2)$-bent functions. The problem of constructing such functions seems to be highly complicated.

### *4.2. Second Algorithm*

Now we consider a modification of the improved Matsui algorithm [2], which is based on the analysis of intermediate ciphertexts. Let integer numbers $s_1$ and $s_2$ with $0 \le s_1 < s_2 \le r$ be chosen. Consider the equality

$$\langle \boldsymbol{a}, \pi(C^{(s_1)})\rangle_i \oplus \langle \boldsymbol{b}, \sigma(C^{(s_2)})\rangle_j = \langle \tau(\boldsymbol{d}), K\rangle_k, \tag{11}$$

where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{Z}_2^m$ and $\boldsymbol{d} \in \mathbb{Z}_2^{m_{\text{key}}}$ are fixed vectors; $\pi, \sigma \in S_m$ and $\tau \in S_{m_{\text{key}}}$ are given permutations; and $i$, $j$, and $k$ are integers such that $1 \le i, j \le m/2$ and $1 \le k \le m_{\text{key}}/2$. We assume that (11) holds with probability $\widetilde{p} = 1/2 + \widetilde{\varepsilon}$ such that $0 < |\widetilde{\varepsilon}| \le 1/2$. Denote by $\widetilde{K}$ a part of bits of a key $K$ which it is sufficient to know to find the values of $\langle \boldsymbol{a}, \pi(C^{(s_1)})\rangle_i$ and $\langle \boldsymbol{b}, \sigma(C^{(s_2)})\rangle_j$ given the vectors $P$ and $C$. Let $m_{s_1,s_2}$ be the number of bits in $\widetilde{K}$.

Algorithm 2

- For each $\widetilde{K} \in \mathbb{Z}_2^{m_{s_1,s_2}}$, find

$$N_0(\widetilde{K}) = \big|\{t : \langle \boldsymbol{a}, \pi(C_t^{(s_1)})\rangle_i \oplus \langle \boldsymbol{b}, \sigma(C_t^{(s_2)})\rangle_j = 0\}\big|;$$

- Arrange all vectors in $\mathbb{Z}_2^{m_{s_1,s_2}}$: $\widetilde{K}_1, \ldots, \widetilde{K}_{2^{m_{s_1,s_2}}}$ so that

$$\left|\frac{N}{2} - N_0(\widetilde{K}_1)\right| \ge \ldots \ge \left|\frac{N}{2} - N_0(\widetilde{K}_{2^{m_{s_1,s_2}}})\right|;$$

- For all $q$ from 1 to $2^{m_{s_1,s_2}}$,
  - ▷ set $\langle \boldsymbol{d}, \tau(K)\rangle_k = \begin{cases} 0 & \text{if } \left(N_0(\widetilde{K}_q) - \dfrac{N}{2}\right)\widetilde{\varepsilon} > 0, \\ 1 & \text{otherwise;} \end{cases}$
  - ▷ using the obtained relation, try to find the key.

End of Algorithm

The reliability of Algorithm 2 can be estimated in the same way as in the case of linear cryptanalysis (see [2, 22]). To provide the required reliability, $N$ must be proportional to $|\widetilde{\varepsilon}|^{-2}$.

As in the case of Algorithm 1, there is a relation between the absolute value of the bias $\widetilde{\varepsilon}$, $k$-Walsh–Hadamard coefficients, and $k$-bent functions.

We denote a set of subkeys $K^{(s_1+1)},\dots,K^{(s_2)}$ by $K^{(s_1+1,\dots,s_2)}$. Let a map $F_{s_1+1,s_2}\colon \mathbb{Z}_2^m \times \big(\mathbb{Z}_2^{m'_{\text{key}}}\big)^{s_2-s_1} \to \mathbb{Z}_2^m$ be defined as the superposition of functions $F_{s_1+1},\dots,F_{s_2}$:

$$F_{s_1+1,s_2}(C^{(s_1)},K^{(s_1+1,\dots,s_2)}) = F_{s_2}(F_{s_2-1}(\dots(F_{s_1+1}(C^{(s_1)},K^{(s_1+1)}),K^{(s_1+2)})\dots),K^{(s_2)}).$$

Then we have

$$C^{(s_2)} = F_{s_1+1,s_2}(C^{(s_1)},K^{(s_1+1,\dots,s_2)}).$$

Similarly to what was done for the first algorithm, consider the equality

$$\langle \boldsymbol{a},\pi(C^{(s_1)})\rangle_i \oplus \langle \boldsymbol{b},\sigma(F_{s_1+1,s_2}(C^{(s_1)},K^{(s_1+1,\dots,s_2)}))\rangle_j = 0 \tag{12}$$

for a fixed set of subkeys $K^{(s_1+1,\dots,s_2)}$. Let it hold with probability $1/2+\widetilde{\varepsilon}^{\,i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}(K^{(s_1+1,\dots,s_2)};0)$, where $-1/2 \le \widetilde{\varepsilon}^{\,i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}(K^{(s_1+1,\dots,s_2)};0) \le 1/2$.

Similarly to Proposition 1, one can easily prove the following statement.

**Proposition 2.** *For any map $F_{s_1+1,s_2}(\cdot\,,K^{(s_1+1,\dots,s_2)})\colon \mathbb{Z}_2^m \to \mathbb{Z}_2^m$, we have*

$$2^{m+1}\widetilde{\varepsilon}^{\,i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}(K^{(s_1+1,\dots,s_2)};0) = W^{(i)}_{\langle \boldsymbol{b},\sigma(F_{s_1+1,s_2}(\pi^{-1}(\cdot),K^{(s_1+1,\dots,s_2)}))\rangle_j}(\boldsymbol{a}).$$

By $\widetilde{\varepsilon}(K)$, we denote the bias in (11) for a fixed $K$. Then for any parameters $k$, $\boldsymbol{d}$, and $\tau$, we have

$$|\widetilde{\varepsilon}(K)| = |\widetilde{\varepsilon}^{\,i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}(K^{(s_1+1,\dots,s_2)};0)| \tag{13}$$

if $K^{(s_1+1,\dots,s_2)}$ is a set of subkeys of $K$.

**Theorem 3.** *Let a key $K \in \mathbb{Z}_2^{m_{\text{key}}}$ and integers $s_1$ and $s_2$ with $0 \le s_1 < s_2 \le r$ be fixed. Let $K^{(s_1+1,\dots,s_2)}$ be a set of subkeys of $K$. Let a vector $\boldsymbol{b} \in \mathbb{Z}_2^m$, permutations $\pi,\sigma \in S_m$, and a parameter $j$, $1 \le j \le m/2$, be such that the function*

$$\langle \boldsymbol{b},\sigma(F_{s_1+1,s_2}(\pi^{-1}(\cdot),K^{(s_1+1,\dots,s_2)}))\rangle_j\colon \mathbb{Z}_2^m \to \mathbb{Z}_2^m$$

*is an $(m/2)$-bent function. Then*

$$\max_{i,k,\boldsymbol{a},\boldsymbol{d},\tau} |\widetilde{\varepsilon}(K)| = \min_{i,k,\boldsymbol{a},\boldsymbol{d},\tau} |\widetilde{\varepsilon}(K)| = 2^{-(m/2)-1}.$$

As in the case of the first algorithm, Theorem 3 implies that using $(m/2)$-bent functions as intermediate ciphering functions allows one to make $\max\limits_{\boldsymbol{a}\in\mathbb{Z}_2^m} |\widetilde{\varepsilon}(K)|$ as small as possible.

## 5. ANALYSIS OF 4-BIT PERMUTATIONS IN S-BOXES OF GOST, DES, AND $s^3$DES

It is well known that the resistance of a block cipher depends straightforwardly on the resistance of substitution boxes (S-boxes) used in it. In this section we consider examples of 4-bit permutations for S-boxes of GOST, DES, and $s^3$DES; with the aid of computer, we show that in almost all cases there exist more probable (than linear) quadratic relations of a special form on input and output bits of these permutations.

**Table 1.** 4-bit permutations with the maximum nonlinearity $NL = 4$

$S^1 = (0, 13, 11, 8, 3, 6, 4, 1, 15, 2, 5, 14, 10, 12, 9, 7)$
$S^2 = (0, 1, 9, 14, 13, 11, 7, 6, 15, 2, 12, 5, 10, 4, 3, 8)$
$S^3 = (0, 1, 11, 13, 9, 14, 6, 7, 12, 5, 8, 3, 15, 2, 4, 10)$
$S^4 = (0, 1, 2, 4, 3, 5, 8, 10, 7, 9, 6, 13, 11, 14, 12, 15)$
$S^5 = (0, 1, 11, 2, 8, 6, 15, 3, 14, 10, 4, 9, 13, 5, 7, 12)$
$S^6 = (0, 1, 11, 2, 8, 3, 15, 6, 14, 10, 4, 9, 13, 5, 7, 12)$
$S^7 = (0, 4, 11, 2, 8, 6, 10, 1, 14, 15, 3, 9, 13, 5, 7, 12)$
$S^8 = (0, 4, 11, 2, 8, 3, 15, 1, 14, 10, 6, 9, 13, 5, 7, 12)$
$S^9 = (0, 11, 15, 9, 1, 5, 6, 8, 3, 10, 4, 12, 14, 13, 7, 2)$
$S^{10} = (0, 7, 10, 14, 9, 1, 13, 8, 12, 2, 11, 15, 3, 5, 4, 6)$
$S^{11} = (4, 10, 9, 2, 13, 8, 0, 14, 6, 11, 1, 12, 7, 15, 5, 3)$
$S^{12} = (8, 2, 11, 13, 4, 1, 14, 7, 5, 15, 0, 3, 10, 6, 9, 12)$
$S^{13} = (10, 5, 3, 15, 12, 9, 0, 6, 1, 2, 8, 4, 11, 14, 7, 13)$
$S^{14} = (5, 10, 12, 6, 0, 15, 3, 9, 8, 13, 11, 1, 7, 2, 14, 4)$
$S^{15} = (3, 9, 15, 0, 6, 10, 5, 12, 14, 2, 1, 7, 13, 4, 8, 11)$
$S^{16} = (15, 0, 10, 9, 3, 5, 4, 14, 8, 11, 1, 7, 6, 12, 13, 2)$
$S^{17} = (12, 6, 3, 9, 0, 5, 10, 15, 2, 13, 4, 14, 7, 11, 1, 8)$
$S^{18} = (13, 10, 0, 7, 3, 9, 14, 4, 2, 15, 12, 1, 5, 6, 11, 8)$

**Table 2.** Permutation $S^2$

| Inputs | | | | Outputs | | | |
|---|---|---|---|---|---|---|---|
| $p_1$ | $p_2$ | $p_3$ | $p_4$ | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |

*Example 1.* In [43] there is listed a series of 4-bit permutations recommended for use in S-boxes of the Russian GOST 28147-89 standard (see permutations $S^1, \ldots, S^{10}$ in Table 1). From each permutation, multiplying it by affine permutations, one gets a class of extremal permutations. All of them are chosen so that to maximally increase the cipher resistance to methods of linear and differential cryptanalysis. Consider their quadratic approximations by functions of the class $\Delta_4$.

To each vector $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$, we assign an integer $\widetilde{x} = 8x_1 + 4x_2 + 2x_3 + x_4$ from 0 to 15. Let $P = (p_1, p_2, p_3, p_4)$ and $C = (c_1, c_2, c_3, c_4)$ be binary inputs and outputs of some 4-bit permutation $S$; i.e., $S(\widetilde{P}) = \widetilde{C}$. For example, the action of $S^2$ is presented in Table 2. Let us find the most probable quadratic and linear relations between input and output bits of a permutation $S$ using the class of functions $\Delta_4$. By the corollary, the number of functions in $\Delta_4$ is 28. Among them, there are 16 linear and 12 quadratic functions; the latter can be listed as follows:

$$\begin{array}{llll}
\langle 0101, v_1v_2v_3v_4\rangle_2, & \langle 0110, v_1v_2v_3v_4\rangle_2, & \langle 1001, v_1v_2v_3v_4\rangle_2, & \langle 1010, v_1v_2v_3v_4\rangle_2, \\
\langle 0101, v_1v_3v_2v_4\rangle_2, & \langle 0110, v_1v_3v_2v_4\rangle_2, & \langle 1001, v_1v_3v_2v_4\rangle_2, & \langle 1010, v_1v_3v_2v_4\rangle_2, \\
\langle 0101, v_1v_4v_2v_3\rangle_2, & \langle 0110, v_1v_4v_2v_3\rangle_2, & \langle 1001, v_1v_4v_2v_3\rangle_2, & \langle 1010, v_1v_4v_2v_3\rangle_2.
\end{array}$$

To this end, we have taken all different sets of two unordered pairs of variables: $\{\{v_1, v_2\}, \{v_3, v_4\}\}$, $\{\{v_1, v_3\}, \{v_2, v_4\}\}$, $\{\{v_1, v_4\}, \{v_2, v_3\}\}$; then for each set we have composed four quadratic functions which differ by their linear parts only.

Consider the relations

$$\langle \boldsymbol{a}, \pi(P)\rangle_i \oplus \langle \boldsymbol{b}, \sigma(C)\rangle_j = 0, \tag{14}$$

where for $i = 1$ a vector $\boldsymbol{a}$ corresponds to the numbers $0, \ldots, 15$ and the identity permutation $\pi$; for $i = 2$, $\boldsymbol{a}$ corresponds to the numbers $5, 6, 9, 10$ and the permutations $\pi = \mathrm{id}, (1, 3, 2, 4), (1, 3, 4, 2)$ (the same for $\boldsymbol{b}$ and $\sigma$ when $j = 1$ or $j = 2$). Under these conditions, the functions $\langle \boldsymbol{a}, \pi(\cdot)\rangle_i$ and $\langle \boldsymbol{b}, \sigma(\cdot)\rangle_j$ run over the whole set $\Delta_4$ without repetitions. For a permutation $S$, consider a table whose rows are enumerated by triples $(i, \widetilde{a}, \pi)$ and columns by triples $(j, \widetilde{b}, \sigma)$; in the intersection of a row and column there is the bias $\varepsilon^{i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}$ of the corresponding equality (14) multiplied by 16 (i.e., the deviation of the number of cases where (14) is fulfilled from a half).

**Table 3.** Biases for $S^2$

| $16\varepsilon^{i,\boldsymbol{a},\pi}_{j,\boldsymbol{b},\sigma}$ | | $j=1$ id | | | | | | | | | | | | | | | | $j=2$ id | | | | $j=2$ (1,3,2,4) | | | | $j=2$ (1,3,4,2) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 5 | 6 | 9 | 10 | 5 | 6 | 9 | 10 | 5 | 6 | 9 | 10 |
| $i=1$ id | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 0 | −2 | 0 | 2 | −2 | 0 | −2 | **4** | 0 | 2 | 0 | −2 | 2 | 0 | 2 | **4** | −2 | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 0 | 0 | **4** | 0 |
| | 2 | 0 | 0 | 0 | 0 | 0 | **4** | **4** | 0 | 2 | 2 | −2 | −2 | 2 | −2 | 2 | −2 | 2 | ⑥ | 0 | 0 | 2 | **4** | −2 | 0 | **4** | 2 | 0 | 2 |
| | 3 | 0 | 2 | 0 | −2 | −2 | 0 | 2 | 0 | −2 | **4** | 2 | **4** | 0 | −2 | 0 | 2 | 0 | 2 | **4** | 2 | 2 | −2 | 0 | 0 | **4** | −2 | 0 | 2 |
| | 4 | 0 | −2 | 0 | −2 | 0 | 2 | 0 | 2 | 0 | −2 | **4** | 2 | 0 | 2 | **4** | −2 | **4** | −2 | 0 | 2 | **4** | −2 | 0 | 2 | 0 | 0 | −2 | −2 |
| | 5 | 0 | 0 | **4** | 0 | 2 | 2 | −2 | 2 | 0 | **4** | 0 | 0 | −2 | 2 | −2 | −2 | −2 | 2 | 0 | **4** | 0 | 0 | −2 | 2 | 0 | **−4** | 2 | 2 |
| | 6 | 0 | −2 | **4** | 2 | 0 | −2 | 0 | −2 | 2 | 0 | −2 | **4** | 2 | 0 | 2 | 0 | −2 | 0 | 0 | −2 | −2 | 2 | 2 | −2 | 0 | 2 | 2 | 0 |
| | 7 | 0 | **4** | 0 | 0 | 2 | 2 | −2 | 2 | −2 | −2 | −2 | 2 | **4** | 0 | 0 | 0 | 0 | 0 | **−4** | 0 | 2 | **4** | 0 | −2 | 0 | 2 | 2 | **−4** |
| | 8 | 0 | 0 | 2 | −2 | 0 | 0 | 2 | −2 | 0 | 0 | 2 | −2 | **4** | **4** | −2 | 2 | 2 | 0 | 2 | 0 | **4** | 0 | 2 | −2 | **4** | 2 | 0 | −2 |
| | 9 | 0 | 2 | 2 | **4** | −2 | **4** | 0 | −2 | 0 | −2 | 2 | 0 | −2 | 0 | 0 | 2 | **4** | 0 | −2 | 2 | 0 | 2 | 0 | ⑥ | **−4** | 2 | 0 | 2 |
| | 10 | 0 | 0 | −2 | 2 | **4** | 0 | −2 | −2 | 2 | 2 | **4** | 0 | 2 | −2 | 0 | 0 | 0 | −2 | 2 | **4** | 2 | 0 | **4** | 2 | 0 | 0 | **4** | 0 |
| | 11 | 0 | −2 | −2 | **4** | 2 | 0 | **4** | 2 | −2 | 0 | 0 | 2 | 0 | 2 | −2 | 0 | 2 | 2 | 2 | −2 | −2 | 2 | 2 | 2 | 0 | **4** | 0 | **4** |
| | 12 | 0 | 2 | 2 | 0 | 0 | −2 | 2 | **4** | **4** | −2 | 2 | 0 | 0 | −2 | −2 | 0 | 2 | −2 | 2 | −2 | 0 | −2 | 2 | 0 | 0 | 2 | −2 | 0 |
| | 13 | 0 | 0 | −2 | −2 | 2 | 2 | 0 | 0 | **4** | 0 | −2 | 2 | −2 | 2 | 0 | **4** | 0 | 2 | −2 | 0 | 0 | 0 | **−4** | 0 | 0 | −2 | −2 | 0 |
| | 14 | 0 | 2 | 2 | 0 | **4** | −2 | 2 | 0 | −2 | 0 | 0 | −2 | −2 | 0 | **4** | 2 | 0 | 0 | 2 | −2 | −2 | −2 | 0 | 0 | 0 | 0 | −2 | 2 |
| | 15 | 0 | **4** | −2 | 2 | −2 | −2 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | **4** | 2 | −2 | −2 | 0 | 2 | 0 | −2 | 0 | 2 | 0 | 0 | 0 | 2 | 2 |
| $i=2$ id | 5 | 0 | −2 | 2 | 0 | **4** | −2 | −2 | 0 | 2 | **4** | 0 | 2 | 2 | 0 | 0 | −2 | **−4** | 0 | 2 | 2 | 0 | 0 | 2 | −2 | 2 | −2 | **4** | 0 |
| | 6 | 0 | 0 | ⑥ | 2 | −2 | 2 | 0 | 0 | 0 | 0 | −2 | 2 | −2 | 2 | 0 | 0 | 0 | 2 | −2 | 0 | −2 | 2 | −2 | 2 | −2 | 0 | 0 | 2 |
| | 9 | 0 | 0 | 0 | **4** | 0 | 0 | 0 | **−4** | 2 | −2 | 2 | 2 | 2 | −2 | 2 | 2 | 2 | −2 | 0 | 0 | 0 | 2 | **4** | 2 | −2 | **4** | 2 | 0 |
| | 10 | 0 | 2 | 0 | 2 | 2 | **4** | −2 | 0 | 0 | 2 | **4** | −2 | −2 | 0 | −2 | 0 | 2 | 0 | 0 | ⑥ | 2 | 0 | 0 | ⑥ | −2 | −2 | 2 | 2 |
| $i=2$ (1,3,2,4) | 5 | 0 | 0 | 2 | 2 | **4** | 0 | −2 | 2 | **4** | 0 | 2 | −2 | 0 | 0 | −2 | −2 | 0 | −2 | 0 | 2 | 0 | 0 | 2 | 2 | −2 | 0 | 2 | 0 |
| | 6 | 0 | 2 | **4** | −2 | −2 | 0 | 2 | **4** | 0 | 2 | 0 | 2 | −2 | 0 | −2 | 0 | 0 | 2 | 2 | 0 | 0 | −2 | −2 | 0 | 2 | −2 | −2 | 2 |
| | 9 | 0 | 2 | −2 | 0 | 0 | −2 | 2 | 0 | 2 | 0 | **4** | 2 | 2 | **−4** | 0 | 2 | 2 | −2 | **4** | 0 | 2 | −2 | **4** | 0 | 2 | 2 | 0 | 0 |
| | 10 | 0 | 0 | 0 | 0 | 2 | 2 | −2 | −2 | −2 | ⑥ | 2 | 2 | 0 | 0 | 0 | 0 | −2 | 2 | 2 | ⑥ | 2 | 0 | 0 | 2 | 2 | **−4** | **4** | 2 |
| $i=2$ (1,3,4,2) | 5 | 0 | 0 | **4** | **4** | 0 | 0 | 0 | 0 | **4** | **−4** | 0 | 0 | 0 | 0 | 0 | 0 | 2 | −2 | −2 | −2 | −2 | 2 | 2 | 2 | **−4** | **4** | 0 | 0 |
| | 6 | 0 | 0 | 2 | −2 | 0 | **−4** | 2 | 2 | 2 | 2 | 0 | **4** | 2 | −2 | 0 | 0 | −2 | 0 | **4** | −2 | 0 | −2 | 2 | **−4** | **4** | 0 | 0 | 0 |
| | 9 | 0 | **4** | 0 | 0 | −2 | 2 | 2 | 2 | 0 | 0 | **4** | 0 | −2 | −2 | −2 | 2 | **4** | 0 | 2 | 2 | 2 | −2 | 0 | **4** | 0 | 0 | −2 | 2 |
| | 10 | 0 | 0 | 2 | 2 | −2 | 2 | 0 | **−4** | −2 | 2 | 0 | **4** | 0 | 0 | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 0 | 2 | 2 |

Although here we present a way to construct such a table for a 4-bit permutation, we note that it can easily be generalized to the case of an arbitrary $t$-bit permutation or a transformation $P \to C$, where $P$ and $C$ have different number of bits.

By the *nonquadraticity* of a permutation $S$, we call the number

$$NQ(S) = \min_{i,j} \min_{\boldsymbol{a}\neq\mathbf{0},\, \boldsymbol{b}\neq\mathbf{0}} \min_{\delta\in\mathbb{Z}_2,\pi,\sigma} |\{P:\ \langle \boldsymbol{a}, \pi(P)\rangle_i \oplus \langle \boldsymbol{b}, \sigma(C)\rangle_j \neq \delta\}|.$$

In other words, $NQ(S)$ is the difference of 8 and the maximum of the absolute values of elements of the table (except for those in the first row and first column). A relation corresponding to an element of the table with absolute value $8 - NQ(S)$ holds with probability either $\dfrac{NQ(S)}{16}$ or $1 - \dfrac{NQ(S)}{16}$ (i.e., most or least probably). By the *nonlinearity* of $S$, we call the number

$$NL(S) = \min_{\boldsymbol{a}\neq\mathbf{0},\, \boldsymbol{b}\neq\mathbf{0}} \min_{\delta,\pi,\sigma} |\{P:\ \langle \boldsymbol{a}, \pi(P)\rangle_1 \oplus \langle \boldsymbol{b}, \sigma(C)\rangle_1 \neq \delta\}|.$$

The parameter $NL(S)$ can be obtained as the difference of 8 and the maximum of the absolute values of elements of the part of the table that corresponds only to linear relations between input and out bits, i.e., to the part with $i = j = 1$ (except for zero combinations). Clearly, $NQ(S) \le NL(S)$. By [44], we have $NL(S) \le 4$ for any 4-bit permutation $S$.

For the permutation $S^2$ we have $NL(S^2) = 4$ and $NQ(S^2) = 2$ (see Table 3). In Table 3, elements with absolute values 4 and 6 are given in bold and enclosed in circles, respectively. Any linear relation on input and output bits of $S^2$ holds with a probability of at most 3/4, whereas there are seven quadratic relations with probability 7/8. Consider the relation with $i = 2$, $\widetilde{a} = 6$, $\pi = \mathrm{id}$, $j = 1$, $\widetilde{b} = 2$, and $\sigma = \mathrm{id}$, i.e.,

$$\langle (0110), (p_1, p_2, p_3, p_4)\rangle_2 \oplus \langle (0010), (c_1, c_2, c_3, c_4)\rangle_1 = 0.$$

Using formulas (2) and (3), we obtain the equality

$$c_3 = p_1p_3 \oplus p_1p_4 \oplus p_2p_3 \oplus p_2p_4 \oplus p_1 \oplus p_4$$

for input and output bits, which holds with probability $(8+6)/16$, i.e., 7/8. Note that this relation is linear with respect to the bits $c_1$, $c_2$, $c_3$, and $c_4$.

Similarly, if we consider a relation with $i = 1$, $\widetilde{a} = 9$, $\pi = \mathrm{id}$, $j = 2$, $\widetilde{b} = 10$, and $\sigma = (1, 3, 2, 4)$, namely,

$$\langle (1001), (p_1, p_2, p_3, p_4)\rangle_1 \oplus \langle (1010), (c_1, c_3, c_2, c_4)\rangle_2 = 0,$$

then after the transformations (2) and (3) we obtain the relation

$$p_2 \oplus p_4 = c_1c_2 \oplus c_1c_4 \oplus c_2c_3 \oplus c_3c_4 \oplus c_3 \oplus c_4,$$

which is linear in $p_1$, $p_2$, $p_3$, and $p_4$ and holds with probability 7/8.

In Table 4, the most probable relations on $P$ and $C$ for the permutations $S^1, \ldots, S^{10}$ are presented in a compact form, which we explain by an example of the relations that we have obtained for $S^2$. One relation is presented in the table as $C\{3\} = P\{13, 14, 23, 24, 1, 4\}$, the other one, as $C\{12, 14, 23, 34, 3, 4\} = P\{2, 4\}$. Note that for each of the ten permutations we manage to construct more probable (than linear) quadratic relations using functions of the class $\Delta_4$. We have $NL(S^t) = 4$ and $NQ(S^t) = 2$ for any $t = 1, \ldots, 10$.

By this example, we see that using relations of the form (14) in systems of equations with unknown (input or output) bits may lead to more probable approximations of unknown bits, while

**Table 4.** Most probable quadratic relations for input and output bits of the permutations $S^1, \ldots, S^{10}$

| $S$ | Quadratic relations with probability 7/8 |
|---|---|
| $S^1$ | $C\{1,3,4\} = P\{13,14,23,24,1,3\}$,<br>$C\{1,3,4\} = P\{12,13,24,34,1,3\}$,<br>$C\{13,14,23,24,2,3\} = P\{3,4\}$,<br>$C\{12,14,23,34,2,3\} = P\{3,4\}$, |
| $S^2$ | $C\{3\} = P\{13,14,23,24,1,4\}$,<br>$C\{2,4\} = P\{12,14,23,34,3,4\}$,<br>$C\{13,14,23,24,1,4\} = P\{3\}$,<br>$C\{12,14,23,34,3,4\} = P\{2,4\}$,<br>$C\{13,14,23,24,2,4\} = P\{13,14,23,24,2,4\}$,<br>$C\{12,14,23,34,3,4\} = P\{13,14,23,24,2,4\}$,<br>$C\{13,14,23,24,2,4\} = P\{12,14,23,34,3,4\}$, |
| $S^3$ | $C\{2\} = P\{13,14,23,24,1,4\}$,<br>$C\{1,2,3,4\} = P\{12,14,23,34,3,4\}$,<br>$C\{12,14,23,34,3,4\} = P\{4\}$,<br>$C\{13,14,23,24,1,4\} = P\{1,2\}$, |
| $S^4$ | $C\{12,13,24,34,1,3\} = P\{1,2\}$,<br>$C\{12,14,23,34,3,4\} = P\{1,2,3,4\}$,<br>$C\{12,13,24,34,1,3\} = P\{12,13,24,34,1,3\}$,<br>$C\{12,14,23,34,1,4\} = P\{12,13,24,34,1,3\}$,<br>$C\{12,14,23,34,1,4\} = P\{12,14,23,34,1,4\}$, |
| $S^5$ | $C\{1,2\} = P\{12,14,23,34,2,3\}$,<br>$C\{12,14,23,34,2,3\} = P\{2,3\}$, |
| $S^6$ | $C\{12,14,23,34,1,4\} = P\{1,2,3,4\}$, |
| $S^7$ | $C\{1,3\} = P\{13,14,23,24,2,4\}$,<br>$C\{12,13,24,34,1,2\} = P\{1,3\} \oplus 1$,<br>$C\{12,13,24,34,3,4\} = P\{12,14,23,34,1,2\}$,<br>$C\{13,14,23,24,1,4\} = P\{12,13,24,34,3,4\} \oplus 1$, |
| $S^8$ | $C\{1,3\} = P\{13,14,23,24,2,4\}$,<br>$C\{12,13,24,34,1,2\} = P\{1,3\} \oplus 1$,<br>$C\{12,13,24,34,3,4\} = P\{12,14,23,34,1,2\}$,<br>$C\{13,14,23,24,1,4\} = P\{12,13,24,34,3,4\} \oplus 1$, |
| $S^9$ | $C\{1,2\} = P\{12,14,23,34,1,4\}$,<br>$C\{1,2,3\} = P\{13,14,23,24,1,3\}$,<br>$C\{1,2,4\} = P\{12,13,24,34,1,3\}$,<br>$C\{13,14,23,24,2,3\} = P\{1,2,4\}$,<br>$C\{12,14,23,34,1,2\} = P\{12,14,23,34,1,4\}$,<br>$C\{13,14,23,24,1,4\} = P\{12,13,24,34,1,2\}$, |
| $S^{10}$ | $C\{1,4\} = P\{13,14,23,24,1,3\}$,<br>$C\{1,2,3,4\} = P\{12,14,23,34,1,4\}$,<br>$C\{13,14,23,24,1,3\} = P\{1,3\}$,<br>$C\{12,14,23,34,3,4\} = P\{1,3\}$,<br>$C\{13,14,23,24,2,4\} = P\{12,14,23,34,3,4\}$,<br>$C\{13,14,23,24,2,4\} = P\{13,14,23,24,1,3\}$ |

solution of the system does not become more complicated (the system may remain linear in the unknowns).

*Example 2.* In [45] there are given eight 4-bit permutations that were used in GOST ciphering implemented for the Central Bank of the Russian Federation and also in a GOST one-way hash function. All of them have $NL = 2$, except for one permutation with $NL = 4$ (see the permutation $S^{11}$ in Table 1). For each of them, we have $NQ = 2$, so on the average we add 5–6 new most probable quadratic relations of a special form on input and output bits of each permutation.

**Table 5.** Most probable quadratic relations for input and output bits of the permutations $S^{11}, \ldots, S^{18}$

| $S$ | Quadratic relations with probability 7/8 |
|---|---|
| $S^{11}$ | $C\{2,3\} = P\{13,14,23,24,1,3\} \oplus 1,$<br>$C\{2,3\} = P\{12,13,24,34,1,3\} \oplus 1,$<br>$C\{12,14,23,34,1,4\} = P\{1,4\},$<br>$C\{13,14,23,24,1,3\} = P\{13,14,23,24,2,4\},$<br>$C\{13,14,23,24,1,3\} = P\{12,13,24,34,1,3\} \oplus 1,$ |
| $S^{12}$ | $C\{1,2,3\} = P\{13,14,23,24,1,3\} \oplus 1,$<br>$C\{13,14,23,24,1,3\} = P\{3,4\} \oplus 1,$<br>$C\{12,14,23,34,1,2\} = P\{3,4\} \oplus 1,$<br>$C\{13,14,23,24,2,4\} = P\{1,3,4\},$<br>$C\{12,14,23,34,3,4\} = P\{1,3,4\},$<br>$C\{13,14,23,24,1,3\} = P\{12,14,23,34,2,3\} \oplus 1,$<br>$C\{12,13,24,34,1,3\} = P\{12,14,23,34,2,3\} \oplus 1,$ |
| $S^{13}$ | $C\{12,14,23,34,1,4\} = P\{13,14,23,24,2,3\} \oplus 1,$<br>$C\{12,13,24,34,1,3\} = P\{13,14,23,24,2,3\} \oplus 1,$ |
| $S^{14}$ | $C\{1\} = P\{12,14,23,34,1,4\},$<br>$C\{3\} = P\{12,14,23,34,3,4\},$<br>$C\{1,2\} = P\{12,14,23,34,2,3\} \oplus 1,$<br>$C\{2,3\} = P\{12,14,23,34,1,2\} \oplus 1,$<br>$C\{12,14,23,34,3,4\} = P\{12,14,23,34,1,2\} \oplus 1,$<br>$C\{12,14,23,34,1,4\} = P\{12,14,23,34,2,3\} \oplus 1,$ |
| $S^{15}$ | $C\{1,2,3\} = P\{12,14,23,34,1,2\} \oplus 1,$<br>$C\{12,13,24,34,1,3\} = P\{2,4\},$<br>$C\{12,13,24,34,2,4\} = P\{1,2,4\},$<br>$C\{13,14,23,24,2,4\} = P\{2,4\} \oplus 1,$<br>$C\{13,14,23,24,1,3\} = P\{1,2,4\} \oplus 1,$<br>$C\{12,14,23,34,1,4\} = P\{12,13,24,34,2,4\},$<br>$C\{12,13,24,34,1,3\} = P\{12,13,24,34,2,4\},$ |
| $S^{16}$ | $C\{12,13,24,34,1,2\} = P\{12,13,24,34,1,2\},$<br>$C\{12,13,24,34,1,2\} = P\{13,14,23,24,2,3\},$<br>$C\{12,14,23,34,1,2\} = P\{12,14,23,34,2,3\},$<br>$C\{12,14,23,34,1,4\} = P\{12,14,23,34,1,2\},$<br>$C\{12,14,23,34,1,4\} = P\{13,14,23,24,1,3\},$<br>$C\{13,14,23,24,2,4\} = P\{12,13,24,34,1,3\},$<br>$C\{13,14,23,24,1,4\} = P\{12,14,23,34,1,2\},$<br>$C\{13,14,23,24,2,4\} = P\{12,14,23,34,2,3\},$ |
| $S^{17}$ | $C\{2,3\} = P\{12,13,24,34,1,3\},$<br>$C\{1,3,4\} = P\{13,14,23,24,2,3\} \oplus 1,$<br>$C\{13,14,23,24,2,3\} = P\{2\} \oplus 1,$<br>$C\{12,13,24,34,3,4\} = P\{2\} \oplus 1,$<br>$C\{13,14,23,24,1,4\} = P\{1,2\} \oplus 1,$<br>$C\{12,13,24,34,1,2\} = P\{1,2\} \oplus 1,$<br>$C\{13,14,23,24,1,4\} = P\{12,13,24,34,1,3\} \oplus 1,$ |
| $S^{18}$ | no relations |

*Example 3.* For all 32 permutations on 16 elements used in DES S-boxes (see, e.g., [45]), the parameters $NL$ and $NQ$ coincide and are equal to 2. Note that for each permutation we add from 0 to 11 (on the average, 4–5) new most probable quadratic relations on input and output bits.

*Example 4.* Consider 32 permutations (see, e.g., [45]) used in S-boxes of the modified $s^3$DES algorithm [46,47], which are reputed to be resistant to methods of linear and differential cryptanalysis. Only seven of them (these are the permutations $S^{12}, \ldots, S^{18}$ in Table 1) have nonlinearity $NL = 4$, and the other 25 have $NL = 2$. For six of the seven permutations with $NL = 4$ we have

$NQ = 2$, and on the average there are about six quadratic relations with probability 7/8. For one permutation only, $S^{18}$, do we have $NL = NQ = 4$.

Quadratic relations with probability 7/8 for the permutations $S^{11}, \dots, S^{18}$ are given in Table 5.

## 6. REMARKS AND ADDITIONS

Here we present the properties of the function $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ that can be used for combining round approximations in quadratic cryptanalysis of particular ciphers.

For a vector $\boldsymbol{u} = (u_1, \dots, u_m)$, let $\bar{\boldsymbol{u}}^k = (u_1 \oplus u_2, \dots, u_{2k-1} \oplus u_{2k})$ be a vector of length $k$. By $*$ we denote the standard componentwise multiplication of vectors. Let $|\boldsymbol{u}| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle$. We have the following fact.

**Proposition 3.** *For any vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{Z}_2^m$ and any number $k$, $1 \le k \le m/2$, there is the equality*

$$\langle \boldsymbol{u} \oplus \boldsymbol{v}, \boldsymbol{w} \rangle_k = \langle \boldsymbol{u}, \boldsymbol{w} \rangle_k \oplus \langle \boldsymbol{v}, \boldsymbol{w} \rangle_k \oplus \langle \bar{\boldsymbol{u}}^k, \bar{\boldsymbol{w}}^k \rangle \langle \bar{\boldsymbol{v}}^k, \bar{\boldsymbol{w}}^k \rangle \oplus |\bar{\boldsymbol{u}}^k * \bar{\boldsymbol{v}}^k * \bar{\boldsymbol{w}}^k|.$$

**Proof.** According to (1), we have

$$\begin{aligned}
\langle \boldsymbol{u} \oplus \boldsymbol{v}, \boldsymbol{w} \rangle_k &= \langle \boldsymbol{u}, \boldsymbol{w} \rangle \oplus \langle \boldsymbol{v}, \boldsymbol{w} \rangle \oplus \left( \bigoplus_{i=1}^{k} \bigoplus_{j=i}^{k} (\bar{u}_i^k \oplus \bar{v}_i^k)(\bar{u}_j^k \oplus \bar{v}_j^k) \bar{w}_i^k \bar{w}_j^k \right) \\
&= \langle \boldsymbol{u}, \boldsymbol{w} \rangle_k \oplus \langle \boldsymbol{v}, \boldsymbol{w} \rangle_k \oplus \left( \bigoplus_{i=1}^{k} \bigoplus_{j=i}^{k} \bar{u}_i^k \bar{v}_j^k \bar{w}_i^k \bar{w}_j^k \right) \oplus \left( \bigoplus_{i=1}^{k} \bigoplus_{j=i}^{k} \bar{u}_j^k \bar{v}_i^k \bar{w}_i^k \bar{w}_j^k \right) \\
&= \langle \boldsymbol{u}, \boldsymbol{w} \rangle_k \oplus \langle \boldsymbol{v}, \boldsymbol{w} \rangle_k \oplus \left( \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{k} \bar{u}_i^k \bar{v}_j^k \bar{w}_i^k \bar{w}_j^k \right) \oplus \left( \bigoplus_{i=1}^{k} \bar{u}_i^k \bar{v}_i^k \bar{w}_i^k \right).
\end{aligned}$$

It remains to note that the third term coincides with $\langle \bar{\boldsymbol{u}}^k, \bar{\boldsymbol{w}}^k \rangle \langle \bar{\boldsymbol{v}}^k, \bar{\boldsymbol{w}}^k \rangle$, and the fourth term equals $|\bar{\boldsymbol{u}}^k * \bar{\boldsymbol{v}}^k * \bar{\boldsymbol{w}}^k|$. $\triangle$

It follows from Proposition 3 that the smaller the value of $k$, the less significant is the nonlinear "supplement" arising when passing from $\langle \boldsymbol{u} \oplus \boldsymbol{v}, \boldsymbol{w} \rangle_k$ to the sum $\langle \boldsymbol{u}, \boldsymbol{w} \rangle_k \oplus \langle \boldsymbol{v}, \boldsymbol{w} \rangle_k$. In combining round approximations (see Section 1), this supplement can be estimated with some probability given partial information on unknown bits.

**Analog of linearity.** This property is based upon another approach to the definition of the functions $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$, that from the point of view of coding theory. Omitting details, we may say that the set of vectors of values of all functions $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k \oplus a$, $\boldsymbol{u} \in \mathbb{Z}_2^m$, $a \in \mathbb{Z}_2$, forms a binary Hadamard-like code $A_m^k$, on which a group operation consistent with the Hamming metric can be defined. This operation allows us to speak about an analog of linearity for the functions $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$. In detail, these properties of the functions $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k$ are considered in [39]. Here we briefly outline the main features.

Let $m$ be an arbitrary integer, and let the parameter $k$, $1 \le k \le m/2$, be fixed. Let $\boldsymbol{G}_m^k$ be an $(m-k) \times 2^m$ matrix over $\mathbb{Z}_4$ composed of lexicographically ordered columns $\boldsymbol{z}^T$, where $\boldsymbol{z} \in \mathbb{Z}_4^k \times (2\mathbb{Z}_4)^{m-2k}$. Such matrices were first considered in [40,41] for the construction of $\mathbb{Z}_4$-linear Hadamard-like codes and perfect codes. For instance,

$$\boldsymbol{G}_2^1 = \begin{pmatrix} 0123 \end{pmatrix}, \qquad \boldsymbol{G}_4^1 = \begin{pmatrix} 0000111122223333 \\ 0022002200220022 \\ 0202020202020202 \end{pmatrix}, \qquad \boldsymbol{G}_4^2 = \begin{pmatrix} 0000111122223333 \\ 0123012301230123 \end{pmatrix}.$$

**Table 6.** Vectors for $\boldsymbol{C}_4^1$

| $\widetilde{u}$ | $\boldsymbol{u}$ | $\varphi_1^{-1}(\boldsymbol{u})$ |
|---|---|---|
| 0 | 0000 | 000 |
| 1 | 0001 | 001 |
| 2 | 0010 | 010 |
| 3 | 0011 | 011 |
| 4 | 0100 | 100 |
| 5 | 0101 | 101 |
| 6 | 0110 | 110 |
| 7 | 0111 | 111 |
| 12 | 1100 | 200 |
| 13 | 1101 | 201 |
| 14 | 1110 | 210 |
| 15 | 1111 | 211 |
| 8 | 1000 | 300 |
| 9 | 1001 | 301 |
| 10 | 1010 | 310 |
| 11 | 1011 | 311 |

**Table 7.** Vectors for $\boldsymbol{C}_4^2$

| $\widetilde{u}$ | $\boldsymbol{u}$ | $\varphi_2^{-1}(\boldsymbol{u})$ |
|---|---|---|
| 0 | 0000 | 00 |
| 1 | 0001 | 01 |
| 3 | 0011 | 02 |
| 2 | 0010 | 03 |
| 4 | 0100 | 10 |
| 5 | 0101 | 11 |
| 7 | 0111 | 12 |
| 6 | 0110 | 13 |
| 12 | 1100 | 20 |
| 13 | 1101 | 21 |
| 15 | 1111 | 22 |
| 14 | 1110 | 23 |
| 8 | 1000 | 30 |
| 9 | 1001 | 31 |
| 11 | 1011 | 32 |
| 10 | 1010 | 33 |

Now let

$\beta, \gamma\colon \mathbb{Z}_4^i \to \mathbb{Z}_2^i$ be coordinatewise extensions of the maps $\beta, \gamma\colon \mathbb{Z}_4 \to \mathbb{Z}_2$ such that $\beta\colon 0, 1 \to 0$; $2, 3 \to 1$ and $\gamma\colon 0, 3 \to 0$; $1, 2 \to 1$, for any integer $i$;

$\varphi\colon \mathbb{Z}_4^i \to \mathbb{Z}_2^{2i}$ be the coordinatewise extension of the *Gray map*: $\varphi(a) = (\beta(a), \gamma(a))$ for $a \in \mathbb{Z}_4$ (see [48, 49]);

$\varphi_k\colon \mathbb{Z}_4^k \times \mathbb{Z}_2^{m-2k} \to \mathbb{Z}_2^m$ be a map such that $\varphi_k\colon (\boldsymbol{u}', \boldsymbol{u}'') \to (\varphi(\boldsymbol{u}'), \boldsymbol{u}'')$ for any vectors $\boldsymbol{u}' \in \mathbb{Z}_4^k$ and $\boldsymbol{u}'' \in \mathbb{Z}_2^{m-2k}$ (see [39]).

Let $\boldsymbol{h}^{\boldsymbol{u}} = \varphi_k^{-1}(\boldsymbol{u})\boldsymbol{G}_m^k$ be a vector of length $2^m$ over $\mathbb{Z}_4$. Consider a square matrix $\boldsymbol{C}_m^k = (c_{\boldsymbol{u},\boldsymbol{v}}^k)$, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{Z}_2^m$, of size $2^m$ over $\mathbb{Z}_4$ whose rows are all possible vectors $\boldsymbol{h}^{\boldsymbol{u}}$ arranged in the ascending lexicographic order of the vectors $\varphi_k^{-1}(\boldsymbol{u})$. We assume that columns of $\boldsymbol{C}_m^k$ are also enumerated by vectors $\boldsymbol{v}$ in the ascending lexicographic order of the vectors $\varphi_k^{-1}(\boldsymbol{v})$. For examples, vectors $\boldsymbol{u}$ for enumeration of rows of the matrices $\boldsymbol{C}_4^1$ and $\boldsymbol{C}_4^2$ are presented in Tables 6 and 7. It is also convenient to assign to a vector $\boldsymbol{u}$ the number $\widetilde{u} = 8u_1 + 4u_2 + 2u_3 + u_4$.

Tables 8 and 9 present the matrices $\boldsymbol{C}_4^1$ and $\boldsymbol{C}_4^2$ together with enumeration of their rows and columns.

It is easy to prove that all matrices $\boldsymbol{C}_m^k$ are symmetric. The matrices $\boldsymbol{C}_m^k$ can be obtained iteratively [39]:

$$\boldsymbol{C}_{m+1}^k = (\boldsymbol{C}_m^k \otimes \boldsymbol{J}_2) + (\boldsymbol{J}_n \otimes \boldsymbol{C}_1^0),$$
$$\boldsymbol{C}_{m+2}^{k+1} = (\boldsymbol{J}_4 \otimes \boldsymbol{C}_m^k) + (\boldsymbol{C}_2^1 \otimes \boldsymbol{J}_n),$$

where $\boldsymbol{J}_s$ is the all-one square matrix of size $s$, $\boldsymbol{C}_1^0 = \begin{pmatrix} 00 \\ 02 \end{pmatrix}$, and $\boldsymbol{A} \otimes \boldsymbol{B}$ is the Kronecker product

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & \dots & a_{1p}\boldsymbol{B} \\ \dots\dots\dots\dots\dots & & \\ a_{p1}\boldsymbol{B} & \dots & a_{pp}\boldsymbol{B} \end{pmatrix}$$

of a square matrix $\boldsymbol{A} = (a_{ij})$, $1 \le i, j \le p$, and a matrix $\boldsymbol{B} = (b_{ij})$, $1 \le i, j \le q$.

**Proposition 4** [39]. *For any integer $m$ and $k$, $1 \le k \le m/2$, and any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{Z}_2^m$, we have $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k = \beta(c_{\boldsymbol{u},\boldsymbol{v}}^k)$.*

**Table 8.** Matrix $\boldsymbol{C}_4^1$

| $c^1_{\boldsymbol{u},\boldsymbol{v}}$ | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **12** | **13** | **14** | **15** | **8** | **9** | **10** | **11** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1** | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 |
| **2** | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 |
| **3** | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 |
| **4** | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| **5** | 0 | 2 | 0 | 2 | 1 | 3 | 1 | 3 | 2 | 0 | 2 | 0 | 3 | 1 | 3 | 1 |
| **6** | 0 | 0 | 2 | 2 | 1 | 1 | 3 | 3 | 2 | 2 | 0 | 0 | 3 | 3 | 1 | 1 |
| **7** | 0 | 2 | 2 | 0 | 1 | 3 | 3 | 1 | 2 | 0 | 0 | 2 | 3 | 1 | 1 | 3 |
| **12** | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 |
| **13** | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 |
| **14** | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 |
| **15** | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 |
| **8** | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| **9** | 0 | 2 | 0 | 2 | 3 | 1 | 3 | 1 | 2 | 0 | 2 | 0 | 1 | 3 | 1 | 3 |
| **10** | 0 | 0 | 2 | 2 | 3 | 3 | 1 | 1 | 2 | 2 | 0 | 0 | 1 | 1 | 3 | 3 |
| **11** | 0 | 2 | 2 | 0 | 3 | 1 | 1 | 3 | 2 | 0 | 0 | 2 | 1 | 3 | 3 | 1 |

**Table 9.** Matrix $\boldsymbol{C}_4^2$

| $c^2_{\boldsymbol{u},\boldsymbol{v}}$ | **0** | **1** | **3** | **2** | **4** | **5** | **7** | **6** | **12** | **13** | **15** | **14** | **8** | **9** | **11** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1** | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 |
| **3** | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 |
| **2** | 0 | 3 | 2 | 1 | 0 | 3 | 2 | 1 | 0 | 3 | 2 | 1 | 0 | 3 | 2 | 1 |
| **4** | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| **5** | 0 | 1 | 2 | 3 | 1 | 2 | 3 | 0 | 2 | 3 | 0 | 1 | 3 | 0 | 1 | 2 |
| **7** | 0 | 2 | 0 | 2 | 1 | 3 | 1 | 3 | 2 | 0 | 2 | 0 | 3 | 1 | 3 | 1 |
| **6** | 0 | 3 | 2 | 1 | 1 | 0 | 3 | 2 | 2 | 1 | 0 | 3 | 3 | 2 | 1 | 0 |
| **12** | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 |
| **13** | 0 | 1 | 2 | 3 | 2 | 3 | 0 | 1 | 0 | 1 | 2 | 3 | 2 | 3 | 0 | 1 |
| **15** | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 |
| **14** | 0 | 3 | 2 | 1 | 2 | 1 | 0 | 3 | 0 | 3 | 2 | 1 | 2 | 1 | 0 | 3 |
| **8** | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| **9** | 0 | 1 | 2 | 3 | 3 | 0 | 1 | 2 | 2 | 3 | 0 | 1 | 1 | 2 | 3 | 0 |
| **11** | 0 | 2 | 0 | 2 | 3 | 1 | 3 | 1 | 2 | 0 | 2 | 0 | 1 | 3 | 1 | 3 |
| **10** | 0 | 3 | 2 | 1 | 3 | 2 | 1 | 0 | 2 | 1 | 0 | 3 | 1 | 0 | 3 | 2 |

This fact, as well as iterative formulas for the matrices $\boldsymbol{C}_m^k$, can be used for fast computation of the coefficients $W_f^{(k)}(\boldsymbol{v})$ and biases.

Let $\star\colon \mathbb{Z}_2^m \times \mathbb{Z}_2^m \to \mathbb{Z}_2^m$ be a binary operation such that $\boldsymbol{u} \star \boldsymbol{v} = \varphi_k(\varphi_k^{-1}(\boldsymbol{u}) \dot{+} \varphi_k^{-1}(\boldsymbol{v}))$ for any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{Z}_2^m$, where $\dot{+}$ denotes addition over $\mathbb{Z}_4$ for the first $k$ coordinates of the vectors $\varphi_k^{-1}(\boldsymbol{u})$ and $\varphi_k^{-1}(\boldsymbol{v})$, and addition over $\mathbb{Z}_2$ for the last $m - 2k$ coordinates.

**Proposition 5** [39]. *For any integer $m$ and $k$ and any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{Z}_2^m$, we have $c^k_{\boldsymbol{u},\boldsymbol{w}} + c^k_{\boldsymbol{v},\boldsymbol{w}} = c^k_{\boldsymbol{u}\star\boldsymbol{v},\boldsymbol{w}}$, where $+$ denotes addition over $\mathbb{Z}_4$.*

Proposition 5 follows from the fact that we have $\boldsymbol{h}^{\boldsymbol{u}} + \boldsymbol{h}^{\boldsymbol{v}} = \boldsymbol{h}^{\boldsymbol{u}\star\boldsymbol{v}}$ for any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{Z}_2^m$. Proposition 4 implies that the vector of values of the Boolean function $\langle \boldsymbol{u}, \cdot \rangle_k\colon \mathbb{Z}_2^m \to \mathbb{Z}_2$ is the image (under the map $\beta$) of the vector of values of a function $\langle\langle \boldsymbol{u}, \cdot \rangle\rangle_k\colon \mathbb{Z}_2^m \to \mathbb{Z}_4$ such that $\langle\langle \boldsymbol{u}, \boldsymbol{v} \rangle\rangle_k = c^k_{\boldsymbol{u},\boldsymbol{v}}$. In other words, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_k = \beta(\langle\langle \boldsymbol{u}, \boldsymbol{v} \rangle\rangle_k)$. Note that $\langle\langle \boldsymbol{u}, \boldsymbol{v} \rangle\rangle_k = \langle\langle \boldsymbol{v}, \boldsymbol{u} \rangle\rangle_k$. By Proposition 5, the functions $\langle\langle \boldsymbol{u}, \boldsymbol{v} \rangle\rangle_k$, $\boldsymbol{u} \in \mathbb{Z}_2^m$, possess the property of linearity over $\mathbb{Z}_4$; i.e., $\langle\langle \boldsymbol{u}', \boldsymbol{v} \rangle\rangle_k + \langle\langle \boldsymbol{u}'', \boldsymbol{v} \rangle\rangle_k = \langle\langle \boldsymbol{u}' \star \boldsymbol{u}'', \boldsymbol{v} \rangle\rangle_k$. This fact can be used in quadratic cryptanalysis. In particular, one can replace the main relation $\langle \boldsymbol{a}, \pi(P) \rangle_i \oplus \langle \boldsymbol{b}, \sigma(C) \rangle_j = \langle \boldsymbol{d}, \tau(K) \rangle_k$ for bits of a plaintext, ciphertext, and key (say, for Algorithm 1) with a relation over $\mathbb{Z}_4$ of the form $\langle\langle \boldsymbol{a}, \pi(P) \rangle\rangle_k + \langle\langle \boldsymbol{b}, \pi(C) \rangle\rangle_k = \langle\langle \boldsymbol{d}, \pi(K) \rangle\rangle_k$ letting $i = j = k$ and $\pi = \sigma = \tau$. In relations of this type, linearity of the functions $\langle\langle \boldsymbol{u}, \boldsymbol{v} \rangle\rangle_k$ over $\mathbb{Z}_4$ can be used directly. However, this case requires additional investigation. In particular, one should describe a way to choose, based on the gathered statistics, a value of $\langle\langle \boldsymbol{d}, \pi(K) \rangle\rangle_k$ among the four possibilities 0, 1, 2, and 3 (instead of two possibilities, as above).

The author is deeply grateful to reviewers for valuable remarks, which helped the author to improve the presentation of results.

## REFERENCES

1. Matsui, M. and Yamagishi, A., A New Method for Known Plaintext Attack of FEAL Cipher, *Advances in Cryptology—EUROCRYPT'92. Proc. Workshop on the Theory and Application of Cryptographic Techniques, Balatonfüred, Hungary, 1992*, Rueppel, R.A., Ed., Lect. Notes Comp. Sci., vol. 658, Berlin: Springer, 1993, pp. 81–91.

2. Matsui, M., Linear Cryptanalysis Method for DES Cipher, *Advances in Cryptology—EUROCRYPT'93. Proc. Workshop on the Theory and Application of Cryptographic Techniques, Lofthus, Norway, 1993*, Helleseth, T., Ed., Lect. Notes Comp. Sci., vol. 765, Berlin: Springer, 1994, pp. 386–397.

3. Biham, E. and Shamir, A., Differential Cryptanalysis of DES-like Cryptosystems, *J. Cryptology*, 1991, vol. 4, no. 1, pp. 3–72.

4. Nyberg, K., Linear Approximation of Block Ciphers, *Advances in Cryptology—EUROCRYPT'94. Proc. Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, 1994*, De Santis, A., Ed., Lect. Notes Comp. Sci., vol. 950, Berlin: Springer, 1995, pp. 439–444.

5. Harpers, C., Kramer, G.G., and Massey, J.L., A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-up Lemma, *Advances in Cryptology—EUROCRYPT'95. Proc. Int. Conf. on the Theory and Application of Cryptographic Techniques, Saint-Malo, France, 1995*, Guillou, L.C. and Quisquater, J.-J., Eds., Lect. Notes Comp. Sci., vol. 921, Berlin: Springer, 1995, pp. 24–38.

6. Buttyan, L. and Vajda, I., Searching for the Best Linear Approximation of DES-like Cryptosystems, *Electron. Lett.*, 1995, vol. 31, no. 11, pp. 873–874.

7. Matsui, M., New Structure of Block Ciphers with Provable Security Against Differential and Linear Cryptanalysis, *Advances in Cryptology—EUROCRYPT'96. Proc. Int. Conf. on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, 1996*, Maurer, U.M., Ed., Lect. Notes Comp. Sci., vol. 1070, Berlin: Springer, 1996, pp. 205–218.

8. Daemen, J., Govaerts, R., and Vandewalle, J., Correlation Matrices, *Proc. 2nd Int. Workshop on Fast Software Encryption (FSE'94), Leuven, Belgium, 1994*, Preneel, B., Ed., Lect. Notes Comp. Sci., vol. 1008, Berlin: Springer, 1995, pp. 275–285.

9. Kaliski, B.S., Jr., and Robshaw, M.J.B., Linear Cryptanalysis Using Multiple Approximations, *Advances in Cryptology—CRYPTO'94. Proc. 14th Ann. Int. Cryptology Conf., Santa Barbara, USA, 1994*, Desmedt, Y., Ed., Lect. Notes Comp. Sci., vol. 839, Berlin: Springer, 1994, pp. 26–39.

10. Biryukov, A., De Cannière, C., and Quisquater, M., On Multiple Linear Approximations, *Advances in Cryptology—CRYPTO 2004. Proc. 24th Ann. Int. Cryptology Conf., Santa Barbara, USA, 2004*, Franklin, M.K., Ed., Lect. Notes Comp. Sci., vol. 3152, Berlin: Springer, 2004, pp. 1–22.

11. Sakurai, K. and Furuya, S., Improving Linear Cryptanalysis of LOKI91 by Probabilistic Counting Method, *Proc. 4th Int. Workshop on Fast Software Encryption (FSE'97), Haifa, Israel, 1997*, Biham, E., Ed., Lect. Notes Comp. Sci., vol. 1267, Berlin: Springer, 1997, pp. 114–133.

12. Baignères, T., Junod, P., and Vaudenay, S., How Far Can We Go beyond Linear Cryptanalysis?, *Advances in Cryptology—ASIACRYPT 2004. Proc. 10th Int. Conf. on the Theory and Application of Cryptology and Information Security, Jeju Island, Korea, 2004*, Lee, P.J., Ed., Lect. Notes Comp. Sci., vol. 3329, Berlin: Springer, 2004, pp. 432–450.

13. Selçuk, A.A., On Probability of Success in Linear and Differential Cryptanalysis, *J. Cryptology*, 2008, vol. 21, no. 1, pp. 131–147.

14. Knudsen, L.R., Practically Secure Feistel Cyphers, *Proc. Cambridge Security Workshop on Fast Software Encryption, Cambridge, UK, 1993*, Anderson, R.J., Ed., Lect. Notes Comp. Sci., vol. 809, Berlin: Springer, 1994, pp. 211–221.

15. Shorin, V.V., Jelezniakov, V.V., and Gabidulin, E.M., Linear and Differential Cryptanalysis of Russian GOST, *Proc. Int. Workshop on Coding and Cryptography (WCC 2001), Paris, France, 2001*, Electron. Notes Discrete Math., vol. 6, Amsterdam: Elsevier, 2001, pp. 538–547.

16. Borst, J., Preneel, B., and Vandewalle, J., Linear Cryptanalysis of RC5 and RC6, *Proc. 6th Int. Workshop on Fast Software Encryption (FSE'99), Rome, Italy, 1999*, Knudsen, L.R., Ed., Lect. Notes Comp. Sci., vol. 1636, Berlin: Springer, 1999, pp. 16–30.

17. Hawkes, P. and O'Connor, L., On Applying Linear Cryptanalysis to IDEA, *Advances in Cryptology—ASIACRYPT'96. Proc. Int. Conf. on the Theory and Applications of Cryptology and Information Security, Kyongju, Korea, 1996*, Kim, K. and Matsumoto, T., Eds., Lect. Notes Comp. Sci., vol. 1163, Berlin: Springer, 1996, pp. 105–115.

18. Biham, E., Dunkelman, O., and Keller, N., Differential-Linear Cryptanalysis of Serpent, *Proc. 10th Int. Workshop on Fast Software Encryption (FSE 2003), Lund, Sweden, 2003*, Johansson, T., Ed., Lect. Notes Comp. Sci., vol. 2887, Berlin: Springer, 2003, pp. 9–21.

19. Mansoori, S.D. and Bizaki, H.K., On the Vulnerability of Simplified AES Algorithm Against Linear Cryptanalysis, *Int. J. Comp. Sci. Network Security*, 2007, vol. 7, no. 7, pp. 257–263.

20. Nakahara, J., Jr., A Linear Analysis of Blowfish and Khufu, *Proc. 3rd Int. Conf. on Information Security Practice and Experience (ISPEC 2007), Hong Kong, China, 2007*, Dawson, E. and Wong, D.S., Eds., Lect. Notes Comp. Sci., vol. 4464, Berlin: Springer, 2007, pp. 20–32.

21. Rothaus, O.S., On "Bent" Functions, *J. Combin. Theory, Ser. A*, 1976, vol. 20, no. 3, pp. 300–305.

22. Logachev, O.A., Sal'nikov, A.A., and Yashchenko, V.V., *Bulevy funktsii v teorii kodirovaniya i kriptologii* (Boolean Functions in Coding Theory and Cryptology), Moscow: Mos. Tsentr Nepreryvnogo Mat. Obrazovaniya (MCCME), 2004.

23. Dobbertin, H. and Leander, G., A Survey of Some Recent Results on Bent Functions, *Proc. 3rd Int. Conf. on Sequences and Their Applications (SETA 2004), Seul, Korea, 2004*, Helleseth, T., Sarwate, D.V., Song, H.-Y., and Yang, K., Eds., Lect. Notes Comp. Sci., vol. 3486, Berlin: Springer, 2005, pp. 1–29.

24. Chee, S., Lee, S., and Kim, K., Semi-bent Functions, *Advances in Cryptology—ASIACRYPT'94. Proc. 4th Int. Conf. on the Theory and Applications of Cryptology, Wollongong, Australia, 1994*, Pieprzyk, J. and Safavi-Naini, R., Eds., Lect. Notes Comp. Sci., vol. 917, Berlin: Springer, 1995, pp. 107–118.

25. Dobbertin, H. and Leander, G., Cryptographer's Toolkit for Construction of 8-Bit Bent Functions, *Cryptology ePrint Archive*, Report no. 2005/089. Available at `http://eprint.iacr.org/2005/089`.

26. Qu, C., Seberry, J., and Pieprzyk, J., Homogeneous Bent Functions, *Discrete Appl. Math.*, 2000, vol. 102, no. 1–2, pp. 133–139.

27. Youssef, A.M. and Gong, G., Hyper-Bent Functions, *Advances in Cryptology—EUROCRYPT 2001. Proc. 20th Int. Ann. Conf. on the Theory and Application of Cryptographic Techniques, Innsbruck, Austria*, Pfitzmann, B., Ed., Lect. Notes Comp. Sci., vol. 2045, Berlin: Springer, 2001, pp. 406–419.

28. Kuz'min, A.S., Markov, V.T., Nechaev, A.A., and Shishkov, A.B., Approximation of Boolean Functions by Monomial Ones, *Diskret. Mat.*, 2006, vol. 18, no. 1, pp. 9–29 [*Discrete Math. Appl.* (Engl. Transl.), 2006, vol. 16, no. 1, pp. 7–28].

29. Carlet, C. and Gaborit, P., Hyper-bent Functions and Cyclic Codes, *J. Combin. Theory, Ser. A*, 2006, vol. 113, no. 3, pp. 466–482.

30. Youssef, A.M., Generalized Hyper-bent Functions over $GF(p)$, *Discrete Appl. Math.*, 2007, vol. 155, no. 8, pp. 1066–1070.

31. Kuz'min, A.S., Markov, V.T., Nechaev, A.A., Shishkin, V.A., and Shishkov, A.B., Bent and Hyper-bent Functions over a Field of $2^\ell$ Elements, *Probl. Peredachi Inf.*, 2008, vol. 44, no. 1, pp. 15–37 [*Probl. Inf. Trans.* (Engl. Transl.), 2008, vol. 44, no. 1, pp. 12–33].

32. Parker, M.G. and Pott, A., On Boolean Functions Which Are Bent and Negabent, *Int. Workshop on Sequences, Subsequences, and Consequences (SSC 2007), Los Angeles, USA, 2007. Revised Invited Papers*, Golomb, S.W., Gong, G., Helleseth, T., and Song, H.-Y., Eds., Lect. Notes Comp. Sci., vol. 4893, Berlin: Springer, 2007, pp. 9–23.

33. Knudsen, L.R. and Robshaw, M.J.B., Non-linear Approximations in Linear Cryptanalysis, *Advances in Cryptology—EUROCRYPT'96. Proc. Int. Conf. on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, 1996*, Maurer, U.M., Ed., Lect. Notes Comp. Sci., vol. 1070, Berlin: Springer, 1996, pp. 224–236.

34. Shimoyama, T. and Kaneko, T., Quadratic Relation of S-box and Its Application to the Linear Attack of Full Round DES, *Advances in Cryptology—CRYPTO'98. Proc. 18th Ann. Int. Cryptology Conf., Santa Barbara, USA, 1998*, Krawczyk, H., Ed., Lect. Notes Comp. Sci., vol. 1462, Berlin: Springer, 1998, pp. 200–211.

35. Nakahara, J., Jr., Preneel, B., and Vandewalle, J., Experimental Non-linear Cryptanalysis, *COSIC Internal Report*, Katholieke Univ. Leuven, 2003.

36. Estévez-Tapiador, J.M., Clark, J.A., and Hernández-Castro, J.C., Non-linear Cryptanalysis Revisited: Heuristic Search for Approximations to S-Boxes, *Cryptography and Coding. Proc. 11th IMA Int. Conf., Cirencester, UK, 2007*, Galbraith, S.D., Ed., Lect. Notes Comp. Sci., vol. 4887, Berlin: Springer, 2007, pp. 99–117.

37. Ryazanov, B.V. and Chechëta, S.I., On the Approximation of a Random Boolean Function by a Set of Quadratic Forms, *Diskret. Mat.*, 1995, vol. 7, no. 3, pp. 129–145 [*Discrete Math. Appl.* (Engl. Transl.), 1995, vol. 5, no. 5, pp. 473–489].

38. Ivanov, A.V., Using Reduced Representations of Boolean Functions in Constructing Their Nonlinear Approximations, *Vestn. Tomsk. Gos. Univ.*, 2007, no. 23, pp. 31–35.

39. Tokareva, N.N., Bent Functions with Stronger Nonlinearity Properties: $k$-Bent Functions, *Diskretn. Anal. Issled. Oper., Ser. 1*, 2007, vol. 14, no. 4, pp. 76–102 [*J. Appl. Indust. Math.* (Engl. Transl.), 2008, vol. 2, no. 4, to appear].

40. Krotov, D.S., $\mathbb{Z}_4$-linear Perfect Codes, *Diskretn. Anal. Issled. Oper., Ser. 1*, 2000, vol. 7, no. 4, pp. 78–90 (Engl. transl. available at `http://arxiv.org/abs/0710.0198`).

41. Krotov, D.S., $\mathbb{Z}_4$-linear Hadamard and Extended Perfect Codes, in *Proc. Int. Workshop on Coding and Cryptography, 2001, Paris, France*, pp. 329–334.

42. Tokareva, N.N., Description of $k$-Bent Functions in Four Variables, *Diskretn. Anal. Issled. Oper., Ser. 1*, 2008, vol. 15, no. 4, pp. 74–83.

43. Rostovtsev, A.G. and Makhovenko, E.B., *Vvedenie v teoriyu iterirovannykh shifrov* (Introduction to the Theory of Iterated Ciphers), St. Petersburg: Mir i Sem'ya, 2003.

44. Heys, H.M. and Tavares, S.E., Substitution-Permutation Networks Resistant to Differential and Linear Cryptanalysis, *J. Cryptology*, 1996, vol. 9, no. 1, pp. 1–19.

45. Schneier, B., *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, New York: Wiley, 1996, 2nd ed. Translated under the title *Prikladnaya kriptografiya. Protokoly, algoritmy, ishodnye teksty na yazyke Si* Moscow: Triumf, 2002.

46. Kim, K., Park, S., and Lee, S., Reconstruction of $s^2$DES S-Boxes and Their Immunity to Differential Cryptanalysis, in *Proc. 1993 Korea–Japan Joint Workshop on Information Security and Cryptography, Seoul, Korea*, 1993, pp. 282–291.

47. Biham, E. and Biryukov, A., How to Strengthen DES Using Existing Hardware, *Advances in Cryptology—ASIACRYPT'94. Proc. 4th Int. Conf. on the Theory and Applications of Cryptology, Wollongong, Australia, 1994*, Pieprzyk, J. and Safavi-Naini, R., Eds., Lect. Notes Comp. Sci., vol. 917, Berlin: Springer, 1995, pp. 398–412.

48. Nechaev, A.A., Kerdock Code in a Cyclic Form, *Diskret. Mat.*, 1989, vol. 1, no. 4, pp. 123–139 [*Discrete Math. Appl.* (Engl. Transl.), 1991, vol. 1, no. 4, pp. 365–384].

49. Hammons, A.R., Jr., Kumar, P.V., Calderbank, A.R., Sloane, N.J.A., and Solé, P., The $\mathbb{Z}_4$-Linearity of Kerdock, Preparata, Goethals, and Related Codes, *IEEE Trans. Inform. Theory*, 1994, vol. 40, no. 2, pp. 301–319.