# Primal-dual distance bounds of linear codes with application to cryptography

Ryutaroh Matsumoto, *Member, IEEE,* Kaoru Kurosawa, *Member, IEEE,* Toshiya Itoh, *Nonmember,* Toshimitsu Konno, *Nonmember,* Tomohiko Uyematsu, *Member, IEEE*

***Abstract*— Let $N(d, d^\perp)$ denote the minimum length $n$ of a linear code $C$ with $d$ and $d^\perp$, where $d$ is the minimum Hamming distance of $C$ and $d^\perp$ is the minimum Hamming distance of $C^\perp$. In this paper, we show a lower bound and an upper bound on $N(d, d^\perp)$. Further, for small values of $d$ and $d^\perp$, we determine $N(d, d^\perp)$ and give a generator matrix of the optimum linear code. This problem is directly related to the design method of cryptographic Boolean functions suggested by Kurosawa et al.**

***Index Terms*— Boolean function, dual distance, linear code, minimum distance**

## I. Introduction

One of the fundamental problems in coding theory is to find the minimum length of linear codes for the given minimum Hamming distance $d$ and the given number of codewords $K$, where the length of a linear code means the length of the codewords.

In this paper, we study a variant of this problem: find the minimum length of linear codes $C$ which achieves the given minimum Hamming distance $d$ and the given minimum Hamming distance $d^\perp$ of $C^\perp$, where $C^\perp$ denotes the dual code of $C$. Note that the number of codewords $K$ is replaced by the minimum Hamming distance $d^\perp$ of $C^\perp$ in our new problem. This problem is interesting not only theoretically but also practically: it is directly related to the design of cryptographic Boolean functions as follows.

Block ciphers must be secure against various attacks, in particular against differential attacks [3] and linear attacks [10]. The security of block ciphers is often studied by viewing their S-boxes (or $F$ functions) as a set of Boolean functions. We say that a Boolean function $f(\boldsymbol{x})$ satisfies (propagation criteria) $PC(\ell)$ [12], [13] if $f(\boldsymbol{x}) + f(\boldsymbol{x} + \Delta)$ is uniformly distributed for any $\Delta$ with $1 \leq wt(\Delta) \leq l$, where $wt(\Delta)$ denotes the Hamming weight of $\Delta$.

To appear in IEEE Transactions on Information Theory, Sept. 2006. This work was supported by the Okawa Foundation for Information and Telecommunications, Japan.

R. Matsumoto and T. Uyematsu are with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, 2–12–1 O-okayama, Meguro-ku, Tokyo 152–8552, Japan. {ryutaroh,uematsu}@it.ss.titech.ac.jp

K. Kurosawa is with the Department of Computer and Information Sciences, Ibaraki University, 4-12-1 Nakanarusawa, Hitachi, Ibaraki, 316-8511, Japan. kurosawa@mx.ibaraki.ac.jp

T. Itoh is with the Global Scientific Information and Computing Center (GSIC), Tokyo Institute of Technology, 2–12–1 O-okayama, Meguro-ku, Tokyo 152–8552, Japan. titoh@dac.gsic.titech.ac.jp

T. Konno is with the Department of Computer Science, Tokyo Institute of Technology, 2–12–1 O-okayama, Meguro-ku, Tokyo 152–8552, Japan. tkonno@saturn.sannet.ne.jp

It is clear that $PC(\ell)$ is directly related to the security against differential attacks because $\Delta$ is the input difference and $f(\boldsymbol{x}) + f(\boldsymbol{x} + \Delta)$ is the output difference of $f$. Also, $f(\boldsymbol{x})$ is a bent function [9, Chapter 14] if and only if $f(\boldsymbol{x})$ satisfies $PC(n)$ [13], where a bent function has the largest distance from the set of affine (linear) functions. Hence $PC(n)$ is directly related to the security against linear attacks. The famous strict avalanche criterion (SAC), which was introduced as a criterion of the security of S-boxes [14], is equivalent to $PC(1)$.

More generally, we say that $f(\boldsymbol{x})$ satisfies (extended propagation criteria) $EPC(\ell)$ of order $k$ [12], [13] if $f(\boldsymbol{x})$ satisfies $PC(\ell)$ even if any $k$ bits of $\boldsymbol{x} = (x_1, \cdots, x_n)$ are fixed to any constant bits. (We remark that many authors refer to EPC as just PC, including [8].) For example, SAC$(k)$, which is a generalized version of SAC, is equivalent to $EPC(1)$ of order $k$. As shown above, $EPC(\ell)$ of order $k$ is a more generalized security notion of cryptographic Boolean functions.

Kurosawa et al. [8] gave the first construction method of $EPC(\ell)$ of order $k$ based on the Maiorana-McFarland construction (see [7]). They showed that there exists an $EPC(\ell)$ of order $k$ function $f(x_1, \cdots, x_n)$ if there exists a linear code $C$ such that $d = k + 1$, $d^\perp = \ell + 1$ and the length of $C$ is $n/2$, where $d$ is the minimum Hamming distance of $C$ and $d^\perp$ is the minimum Hamming distance of $C^\perp$. Carlet generalized this construction to nonlinear codes [5].

We now ask, given $k$ and $\ell$, what is the minimum $n$ for which $EPC(\ell)$ of order $k$ functions $f(x_1, \cdots, x_n)$ exist ? In the design method of Kurosawa et al. [8], this is equivalent to saying that, given $d$ and $d^\perp$, find the minimum length $n$ of a linear code $C$ with $d$ and $d^\perp$. Note that this problem is exactly the same as the one mentioned at the beginning of the introduction.

More formally, let $N(d, d^\perp)$ denote the minimum length $n$ of a linear code $C$ with $d$ and $d^\perp$, where $d$ is the minimum Hamming distance of $C$ and $d^\perp$ is the minimum Hamming distance of $C^\perp$. We then want to find $N(d, d^\perp)$ for given $d$ and $d^\perp$. In this paper, we show lower bounds and upper bounds on $N(d, d^\perp)$. Further, for small values of $d$ and $d^\perp$, we determine $N(d, d^\perp)$ exactly and give a generator matrix of the optimum linear code.

This paper is organized as follows: In Section 2, we introduce relevant concepts and notations. In Section 3, we propose upper bounds on $N(d, d^\perp)$. In Section 4, we propose lower bounds on $N(d, d^\perp)$, show true values of $N(d, d^\perp)$, and compare the proposed bounds with the true values. In Section 5, concluding remarks are given.

## II. Preliminaries

### A. Notation

We use $f$ to denote a Boolean function $\{0,1\}^n \to \{0,1\}$, and $\phi$ to denote a function $\{0,1\}^n \to \{0,1\}^m$, where $m \le n$. We use $\boldsymbol{x}$ to denote $(x_1, \cdots, x_n)$, where $x_i$ is a binary variable.

Let $\cdot$ denote the inner product of two binary vectors over $GF(2)$. For a set $A$, $|A|$ denotes the cardinality of $A$.

Let a linear $[n, m, d]$ code denote a binary linear code $C$ of length $n$, dimension $m$ and the minimum Hamming distance at least $d$. The dual code $C^\perp$ of a linear code $C$ is defined as $C^\perp = \{u \mid u \cdot v = 0 \text{ for all } v \in C\}$ . The dual distance $d^\perp$ of $C$ is defined as the minimum Hamming distance of $C^\perp$.

### B. Resilient Functions

*Definition 1:* We say that $\phi : \{0,1\}^n \to \{0,1\}^m$ is an $(n, m, k)$-resilient function if $\phi(x_1, \cdots, x_n)$ is uniformly distributed even if any $k$ variables $x_{i_1}, \cdots, x_{i_k}$ are fixed into constants. That is,

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

for any $k$ positions $i_1 < \cdots < i_k$, for any $k$-bit string $\alpha \in \{0,1\}^k$ and for any fixed $(y_1, \cdots, y_m) \in \{0,1\}^m$, where the values $x_j$ ($j \notin \{i_1, \ldots, i_k\}$) are chosen independently at random.

### C. $EPC(\ell)$ of order $k$

Define the derivative of $f : \{0,1\}^n \to \{0,1\}$ by

$$D_\Delta f = f(\boldsymbol{x}) + f(\boldsymbol{x} + \Delta)$$

for $\Delta \in \{0,1\}^n$.

*Definition 2:* [12], [13] We say that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ satisfies $EPC(\ell)$ of order $k$ if $D_\Delta f$ is $k$-resilient for any $\Delta \in \{0,1\}^n$ with $1 \le wt(\Delta) \le l$. (We also say that $f$ is an $EPC(\ell)$ of order $k$ function.)

Kurosawa et al. gave a general method to design $EPC(\ell)$ of order $k$ functions by using a linear code [8].

*Proposition 3:* Suppose that there exists a linear $[n, m, k+1]$ code $C$ with the dual distance at least $\ell + 1$. Then there exists an $EPC(\ell)$ of order $k$ function $f : \{0,1\}^{2n} \to \{0,1\}$.

*Remark 4:* The construction of [8] is essentially quadratic in nature with a non-quadratic 'offset' part. After [8], Carlet [5] showed a construction which uses nonlinear Kerdock and Preparata codes as an improvement. It gives non-quadratic Boolean functions not just in their offset part.

Define $N(d, d^\perp)$ as the minimum $n$ such that there exists a linear $[n, m, d]$ code $C$ with the dual distance at least $d^\perp$. Then $N(k+1, \ell+1)$ is the minimum $n$ such that there exists a $EPC(\ell)$ of order $k$ function $f : \{0,1\}^{2n} \to \{0,1\}$ in the design method of Kurosawa et al. We will consider the upper and lower bounds on $N(d, d^\perp)$, and also determine the true values of $N(d, d^\perp)$ for small $d$ and $d^\perp$.

## III. Upper Bound

In this section, we show upper bounds on $N(d, d^\perp)$. The first bound is based on a Gilbert-Varshamov type argument [9, pp. 557–558].

*Definition 5:*

$$\begin{aligned} S_{n,m} &= \{C \mid C \text{ is an } [n, m] \text{ linear code}\}, \\ S_{n,m}(\boldsymbol{v}) &= \{C \in S_{n,m} \mid C \ni \boldsymbol{v}\}, \\ S^\perp_{n,m}(\boldsymbol{v}) &= \{C \in S_{n,m} \mid C^\perp \ni \boldsymbol{v}\} \end{aligned}$$

*Lemma 6:* For a nonzero vector $\boldsymbol{v} \in GF(2)^n$, we have

$$\frac{|S_{n,m}(\boldsymbol{v})|}{|S_{n,m}|} = \frac{2^m - 1}{2^n - 1}, \tag{1}$$

$$\frac{|S^\perp_{n,m}(\boldsymbol{v})|}{|S_{n,m}|} = \frac{2^{n-m} - 1}{2^n - 1}. \tag{2}$$

Proof is given in Appendix I.

*Theorem 7:* There exists an $[n, m, d]$ binary code with the dual distance $d^\perp$ if

$$\frac{2^m - 1}{2^n - 1} \sum_{i=1}^{d-1} \binom{n}{i} + \frac{2^{n-m} - 1}{2^n - 1} \sum_{i=1}^{d^\perp - 1} \binom{n}{i} < 1.$$

$N(d, d^\perp)$ is upper bounded by the minimum $n$ satisfying the above inequality.

*Proof:* The required code exists iff

$$S_{n,m} \neq \bigcup_{1 \le wt(\boldsymbol{v}) \le d-1} S_{n,m}(\boldsymbol{v}) \cup \bigcup_{1 \le wt(\boldsymbol{v}) \le d^\perp - 1} S^\perp_{n,m}(\boldsymbol{v}).$$

The cardinality of the right hand side is less than or equal to

$$\begin{aligned} &\sum_{1 \le wt(\boldsymbol{v}) \le d-1} |S_{n,m}(\boldsymbol{v})| + \sum_{1 \le wt(\boldsymbol{v}) \le d^\perp - 1} |S^\perp_{n,m}(\boldsymbol{v})| \quad (3) \\ \le &\left( \frac{2^m - 1}{2^n - 1} \sum_{i=1}^{d-1} \binom{n}{i} + \frac{2^{n-m} - 1}{2^n - 1} \sum_{i=1}^{d^\perp - 1} \binom{n}{i} \right) |S_{n,m}| \end{aligned}$$

by Lemma 6. Thus, if the assumption of the theorem is satisfied, the required code exists. ∎

We also introduce another upper bound.

*Proposition 8:*

$$N(d-1, d^\perp) \le N(d, d^\perp) - 1 \text{ (for } d \ge 2), \tag{4}$$

$$N(d, d^\perp - 1) \le N(d, d^\perp) - 1 \text{ (for } d^\perp \ge 2). \tag{5}$$

*Proof:* Let $C$ be a linear code attaining $N(d, d^\perp)$, and $C'$ be the punctured code of $C$. Then $C'$ has the minimum distance at least $d - 1$ and the dual distance at least $d^\perp$, which proves Eq. (4). Equation (5) is proved by considering the punctured code of $C^\perp$. ∎

## IV. Lower bounds

In this section, we give four lower bounds on $N(d, d^\perp)$. The first two are immediate applications of the Griesmer bound and a well-known fact of MDS codes. The third is based on an improvement to the Hamming bound. The fourth is an improvement to Brouwer's bound [4] based on the solvability of a system of linear inequalities [6].

### A. Bounds based on the Griesmer bound and the result in MDS codes

*Proposition 9 (Griesmer):* [9, Section 17.§6] If there exists an $[n, m, d]$ linear code, then

$$n \geq d + \sum_{i=1}^{m-1} \left\lceil \frac{d}{2^i} \right\rceil.$$

*Theorem 10:*

$$N(d, d^\perp) \geq \min[n : 2n \geq d + d^\perp + \min_{m=1,\ldots,n-1} \left\{ \sum_{i=1}^{m-1} \left\lceil \frac{d}{2^i} \right\rceil + \sum_{i=1}^{n-m-1} \left\lceil \frac{d^\perp}{2^i} \right\rceil \right\}] \quad (6)$$

*Proof:* If there exists an $[n, m, d]$ code with dual distance $d^\perp$, then by the Griesmer bound we have

$$2n \geq d + d^\perp + \sum_{i=1}^{m-1} \left\lceil \frac{d}{2^i} \right\rceil + \sum_{i=1}^{n-m-1} \left\lceil \frac{d^\perp}{2^i} \right\rceil. \quad (7)$$

Since $N(d, d^\perp)$ is the minimum $n$ such that there exists a linear code of length $n$, minimum distance $d$ and dual distance $d^\perp$, $2N(d, d^\perp)$ is lower bounded by the minimum of the right hand side of Eq. (7) over possible $n$ and $m$. ■

*Remark 11:* It is well-known that the simplex codes attain the Griesmer bound. However, they do not attain Eq. (6).

The Singleton bound is a corollary to the Griesmer bound and has a simpler expression. It states that if there exists an $[n, m, d]$ code then $m \leq n-d+1$. When the code is binary and $d \geq 3$, it can be tightened to $m \leq n - d$ [11]. The first part of the following result can be seen as a corollary to Theorem 10.

*Theorem 12:*

$$N(d, d^\perp) \geq d + d^\perp - 2. \quad (8)$$

When $d \geq 3$ and $d^\perp \geq 3$, we have[1]

$$N(d, d^\perp) \geq d + d^\perp. \quad (9)$$

*Proof:* Adding $m \leq N(d, d^\perp) - d + 1$ and $N(d, d^\perp) - m \leq N(d, d^\perp) - d^\perp + 1$ shows Eq. (8). A similar argument shows Eq. (9). ■

### B. Bound based on an improved Hamming bound

In this subsection, we will introduce an improvement to the Hamming bound, and derive a lower bound on $N(d, d^\perp)$ as a corollary.

*Definition 13:* For positive integers $d$ and $n$, we define the function $\ell(n, d)$ by

$$\ell(n, d) = \begin{cases} \displaystyle\sum_{i=0}^{(d-1)/2} \binom{n}{i} & \text{for odd } d, \\ \displaystyle\sum_{i=0}^{d/2-1} \binom{n}{i} + \binom{n-1}{d/2-1} & \text{for even } d. \end{cases}$$

Discrete random variables $X_1, \ldots, X_n$ are said to be $d$-wise independent if

$$\Pr[X_{i_1} = x_{i_1}, \ldots, X_{i_d} = x_{i_d}] = \prod_{j=1}^{d} \Pr[X_{i_j} = x_{i_j}]$$

for all $d$-tuples of indices $(i_1, \ldots, i_d)$ and all realizations $(x_{i_1}, \ldots, x_{i_d})$ of random variables.

*Lemma 14:* [1, Proposition 6.4] Let $X_1, \ldots, X_n$ be $(d-1)$-wise independent nonconstant random variables that map the sample space $\Omega$ to $\{0, 1\}$. Then we have $|\Omega| \geq \ell(n, d)$.

*Theorem 15:* For an $[n, m, d]$ linear code $C$, we have $2^{n-m} \geq \ell(n, d)$.

*Proof:* Let $H$ be a parity check matrix for $C$, and $h_i$ be its $i$-th column. Consider the sample space $\Omega = GF(2)^{n-m}$ and the random variable $X_i$ that maps $v \in \Omega$ to the inner product of $v$ and $h_i$. Since any $(d-1)$ columns in $H$ are linearly independent, the random variables $X_1, \ldots, X_n$ are $(d-1)$-wise independent with the uniform probability distribution on $\Omega$. By Lemma 14, $2^{n-m} = |\Omega| \geq \ell(n, d)$. ■

Observe that Theorem 15 is an improvement to the Hamming bound when $d$ is even.

*Corollary 16:*

$$N(d, d^\perp) \geq \min\{n \mid n \geq \log_2 \ell(n, d) + \log_2 \ell(n, d^\perp)\}.$$

*Proof:* If there exists an $[n, m, d]$ linear code with dual distance $d^\perp$, then by Theorem 15

$$\begin{aligned} & 2^{n-m} \cdot 2^m \geq \ell(n, d) \cdot \ell(n, d^\perp) \\ \iff \quad & n \geq \log_2 \ell(n, d) + \log_2 \ell(n, d^\perp). \end{aligned} \quad (10)$$

Since $N(d, d^\perp)$ is the minimum $n$ such that there exists a linear code of length $n$, minimum distance $d$ and dual distance $d^\perp$, $N(d, d^\perp)$ is lower bounded by the minimum of the right hand side of Eq. (10) over possible $n$. ■

### C. Bounds based on linear inequalities

For a linear code $C$, define

$$\begin{aligned} A_w &= |\{c \in C : wt(c) = w\}|, \\ A'_w &= |\{c \in C^\perp : wt(c) = w\}|. \end{aligned}$$

We have [9, Section 5.§2]

$$A'_w = \frac{1}{|C|} \sum_{i=0}^{n} A_i P_w(i) = \frac{1}{|C|} \binom{n}{w} + \frac{1}{|C|} \sum_{i=1}^{n} A_i P_w(i),$$

where $P_w(i)$ is the Krawtchouk polynomial defined by

$$P_w(i) = \sum_{j=0}^{w} (-1)^j \binom{i}{j} \binom{n-i}{w-j}.$$

For $w = 1, \ldots, n$, we must have $A'_w \geq 0$. When the code $C$ has minimum distance $d$, we have $A_1 = A_2 = \cdots = A_{d-1} = 0$. We also have $A'_1 = \cdots = A'_{d^\perp - 1} = 0$ if $C$ has dual distance $d^\perp$. Therefore, if there exists a linear code of length $n$, minimum distance $d$ and dual distance $d^\perp$, then there exists a solution $A_d, \ldots, A_n$ to the following system of linear inequalities:

$$\begin{cases} A_i \geq 0 \text{ for } i = d, \ldots, n, \\ \sum_{i=d}^{n} A_i P_w(i) = -\binom{n}{w} \text{ for } w = 1, \ldots, d^\perp - 1, \\ \sum_{i=d}^{n} A_i P_w(i) \geq -\binom{n}{w} \text{ for } w = d^\perp, \ldots, n. \end{cases} \quad (11)$$

*Theorem 17:* [4] $N(d, d^\perp)$ is greater than or equal to the minimum $n$ such that there exists a solution to the above system of linear inequalities.

[1] This improvement was pointed out by an anonymous reviewer.

We will add other constraints to Eq. (11). Since we consider linear codes, there must exist an *integer* solution $(A_d, \ldots, A_n)$ with $A_d + \cdots + A_n = 2^m - 1$ for some nonnegative integer $m$.

A binary linear code is said to be *even* if all codewords have even weight. We call a code *odd* if it is not even. When the code $C$ is odd, then there is the same number of even weighted codewords and odd weighted ones. Moreover, the dual code $C^\perp$ does not contain the codeword with all 1, otherwise $C$ is even. Therefore, if the code $C$ is odd, then we have

$$\left\{ \begin{array}{rcl} \sum_{i:\text{even}} A_i & = & \sum_{i:\text{odd}} A_i, \\ A'_n & = & 0. \end{array} \right. \tag{12}$$

When the code $C$ is even, then the dual code $C^\perp$ contains the codeword with all 1, and we have $A'_i = A'_{n-i}$, because there is one-to-one correspondence between codewords with weight $i$ and weight $n-i$ by adding the all 1 codeword.

Furthermore, we have the following inequality [4] when $C$ is even:

$$4 \sum_{4|i} A_i \geq \sum_{i=0}^{n} A_i,$$

where $4|i$ denotes that 4 divides $i$. Summing up, the evenness of $C$ implies

$$\left\{ \begin{array}{rcl} A_i & = & 0, \text{ for } i = 1, 3, 5, \ldots, \\ 4\sum_{4|i} A_i & \geq & \sum_{i=0}^{n} A_i, \\ A'_n & = & 1, \\ A'_i & = & A'_{n-i}. \end{array} \right. \tag{13}$$

By exchanging the role of $C$ and $C^\perp$, we see that the oddness of $C^\perp$ implies

$$\left\{ \begin{array}{rcl} \sum_{i:\text{even}} A'_i & = & \sum_{i:\text{odd}} A'_i, \\ A_n & = & 0. \end{array} \right. \tag{14}$$

and that the evenness of $C^\perp$ implies

$$\left\{ \begin{array}{rcl} A'_i & = & 0, \text{ for } i = 1, 3, 5, \ldots, \\ 4\sum_{4|i} A'_i & \geq & \sum_{i=0}^{n} A'_i, \\ A_n & = & 1, \\ A_i & = & A_{n-i}. \end{array} \right. \tag{15}$$

When we estimate $N(d, d^\perp)$ and $d$ is even, the code can be either odd or even, and we search a solution for either Eq. (12) or (13). When $d$ is odd, the code is odd and we search a solution for Eq. (12) only. The same rule applies to $d^\perp$.

*Remark 18:* We remark on the computational complexity on the bound presented in this subsection. When we require $A_d, \ldots, A_n$ to be integers, we have to solve an integer programming problem for which there is no known polynomial time algorithm in the number of variables [2, Section 11.8]. When we allow $A_d, \ldots, A_n$ to be any real numbers, we solve a linear programming problem that can be solved in roughly $O((n-d)^5)$ arithmetic operations [2, Section 9.3]. In both case, it quickly becomes difficult to compute the lower bound for large $n$.

### D. Numerical Examples

In this subsection, we give numerical examples of the derived bounds in Table I. An entry $x$ in Table I means that $N(d, d^\perp) \geq x$ for the lower bounds, and $N(d, d^\perp) \leq x$ for the upper bound. True values of $N(d, d^\perp)$ are also listed, which are obtained by exhaustive search. Generator matrices of codes attaining $N(d, d^\perp)$ are listed in Appendix II. We could not determine the true values of $N(d, d^\perp)$ by exhaustive search with $(d, d^\perp)$ not listed in Table I. We remark that $N(2, \delta) = N(\delta, 2) = \delta$ because the trivial $[\delta, 1, \delta]$ code has dual distance 2.

From Table I, we can make the following observations. Lower bounds are increasing in order of Corollary 16, Theorem 17, and the improvement of Theorem 17 in Sec. IV-C. Theorems 10 and 12 give smaller lower bounds. The upper bound in Theorem 7 is very loose for small values of $d$ and $d^\perp$. This looseness seems to come from the fact that many elements are counted several times in Eq. (3).

Additional constraints in Sec. IV-C give the true values of $N(d, d^\perp)$ as a lower bound except for $(d, d^\perp) = (5, 5)$. They also improve Theorem 17 in the parameters $(d, d^\perp) = (5, 3)$, $(5, 4)$, $(6, 3)$, $(6, 4)$, $(6, 5)$, $(6, 6)$. These improvements significantly reduced the required time for exhaustive search.

## V. Conclusion

In this paper, we considered the minimum length of linear codes with specified minimum Hamming distances and dual distances, from which cryptographic Boolean functions are constructed. We obtained an upper bound by a Gilbert-Varshamov type argument, and lower bounds by applying the Griesmer, the Hamming, and the linear programming bound. The true values for the minimum length are also determined by exhaustive search for certain range of parameters. These lower bounds and true values are useful for estimating the necessary input length of cryptographic Boolean functions for given cryptographic strength. This paper also demonstrated that the upper bound proposed herein is too loose, and it remains an open problem to derive a tight upper bound.

## Acknowledgment

We would like to thank the reviewers' critical comments which improve this paper a lot. In particular, Theorem 12 is improved by the reviewer's comment.

## Appendix I
## Proof of Lemma 6

*Lemma 19:* For nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in GF(2)^n$, we have

$$|S_{n,m}(\boldsymbol{u})| = |S_{n,m}(\boldsymbol{v})|, \tag{16}$$
$$|S^\perp_{n,m}(\boldsymbol{u})| = |S_{n,n-m}(\boldsymbol{u})|, \tag{17}$$
$$|S^\perp_{n,m}(\boldsymbol{u})| = |S^\perp_{n,m}(\boldsymbol{v})|. \tag{18}$$

*Proof:* We define the group $GL_n$ as the set of bijective linear maps $f$ on $GF(2)^n$. In the following equation, $S_{n,m} \ni$

TABLE I
TRUE VALUES AND ESTIMATES OF $N(d, d^\perp)$ BY THE DERIVED BOUNDS

| $d$ | $d^\perp$ | true value | lower bounds | | | | | upper bound |
|---|---|---|---|---|---|---|---|---|
| | | | Thm. 12 | Thm. 10 | Cor. 16 | Thm. 17 (conventional) | Sect. IV-C | Thm. 7 |
| 3 | 3 | 6 | 6 | 5 | 6 | 6 | 6 | 17 |
| 4 | 3 | 7 | 7 | 6 | 7 | 7 | 7 | 21 |
| 4 | 4 | 8 | 8 | 7 | 8 | 8 | 8 | 25 |
| 5 | 3 | 11 | 8 | 7 | 9 | 10 | 11 | 24 |
| 5 | 4 | 13 | 9 | 8 | 11 | 11 | 13 | 29 |
| 5 | 5 | 16 | 10 | 11 | 14 | 14 | 14 | 34 |
| 6 | 3 | 12 | 9 | 8 | 10 | 11 | 12 | 28 |
| 6 | 4 | 14 | 10 | 9 | 12 | 12 | 14 | 33 |
| 6 | 5 | 17 | 11 | 12 | 15 | 15 | 17 | 38 |
| 6 | 6 | 18 | 10 | 13 | 16 | 16 | 18 | 42 |
| 7 | 3 | 14 | 10 | 9 | 12 | 14 | 14 | 31 |
| 7 | 4 | 15 | 11 | 10 | 14 | 15 | 15 | 37 |
| 8 | 3 | 15 | 11 | 10 | 14 | 15 | 15 | 35 |
| 8 | 4 | 16 | 12 | 11 | 15 | 16 | 16 | 40 |

$C_1$ is a fixed linear code, and $g$ is a fixed bijective linear map on $GF(2)^n$ such that $g(\boldsymbol{v}) = \boldsymbol{u}$.

$$\begin{aligned} & |S_{n,m}(\boldsymbol{u})| \\ = \ & |\{C \in S_{n,m} \mid C \ni \boldsymbol{u}\}| \\ = \ & |\{f(C_1) \mid f(C_1) \ni \boldsymbol{u}, f \in GL_n\}| \\ = \ & |\{f(C_1) \mid f(C_1) \ni g(\boldsymbol{v}), f \in GL_n\}| \\ = \ & |\{g^{-1} \circ f(C_1) \mid g^{-1} \circ f(C_1) \ni \boldsymbol{v}, f \in GL_n\}| \\ = \ & |\{f(C_1) \mid f(C_1) \ni \boldsymbol{v}, f \in GL_n\}| \\ = \ & |S_{n,m}(\boldsymbol{v})|. \end{aligned}$$

Equation (16) is proved.

By taking the dual code, we see that there is a one-to-one correspondence between $S_{n,m}$ and $S_{n,n-m}$, and we have

$$\begin{aligned} & |S^\perp_{n,m}(\boldsymbol{u})| \\ = \ & |\{C \in S_{n,m} \mid C^\perp \ni \boldsymbol{u}\}| \\ = \ & |\{C \in S_{n,n-m} \mid C \ni \boldsymbol{u}\}| \\ = \ & |S_{n,n-m}(\boldsymbol{u})|, \end{aligned}$$

which proves Eq. (17). Equation (18) is deduced from Eqs. (16) and (17). ■

*Proof of Lemma 6.* Let $B$ be the set of a pair of a nonzero vector $\boldsymbol{u}$ and $C \in S_{n,m}$ such that $\boldsymbol{u} \in C$. For each $C \in S_{n,m}$, there are $2^m - 1$ nonzero vectors $\boldsymbol{u}$ such that $\boldsymbol{u} \in C$, and we have $|B| = (2^m - 1)|S_{n,m}|$.

For each nonzero vector $\boldsymbol{u}$ there are $|S_{n,m}(\boldsymbol{u})|$ linear codes $C$ such that $\boldsymbol{u} \in C$, and we have

$$|B| = \sum_{\boldsymbol{0} \neq \boldsymbol{u} \in GF(2)^n} |S_{n,m}(\boldsymbol{u})| = (2^n - 1)|S_{n,m}(\boldsymbol{v})|$$

by Eq. (16). Thus Eq. (1) is proved. Equation (2) follows from Eqs. (17) and (1). ■

## APPENDIX II
## LINEAR CODES ATTAINING $N(d, d^\perp)$

In this Appendix, we give the name or generator matrices of linear codes attaining $N(d, d^\perp)$. Matrices are generator matrices of linear codes attaining $N(d, d^\perp)$ unless otherwise specified.

$N(3,3) = 6$: Attained by the $[6, 3, 3]$ shortened Hamming code.
$N(4,3) = 7$: Attained by the $[7, 4, 3]$ Hamming code.
$N(4,4) = 8$: Attained by the $[8, 4, 4]$ extended Hamming code.
$N(5,3) = 11$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

$N(5,4) = 13$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

$N(5,5) = 16$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

$N(6,3) = 12$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

$N(6,4) = 14$: The generator matrix of its dual code is

$$\begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0&1&1&1 \\ 0&1&0&0&0&0&0&0&0&0&1&0&1&1 \\ 0&0&1&0&0&0&0&0&0&0&1&1&0&1 \\ 0&0&0&1&0&0&0&0&0&0&1&1&1&0 \\ 0&0&0&0&1&0&0&0&0&1&0&0&1&1 \\ 0&0&0&0&0&1&0&0&0&1&0&1&0&1 \\ 0&0&0&0&0&0&1&0&0&1&0&1&1&0 \\ 0&0&0&0&0&0&0&1&0&1&1&0&0&1 \\ 0&0&0&0&0&0&0&0&1&1&1&0&1&0 \end{pmatrix}$$

$N(6,5) = 17$: The generator matrix of its dual code is

$$\begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&1&0&0&0&0&0&0&0&0&1&1&0&0&1&1&1 \\ 0&0&1&0&0&0&0&0&0&1&0&1&0&1&0&1&1 \\ 0&0&0&1&0&0&0&0&0&1&1&0&1&0&1&0&1 \\ 0&0&0&0&1&0&0&0&1&0&0&1&0&1&1&0&1 \\ 0&0&0&0&0&1&0&0&1&0&1&0&1&0&1&1&0 \\ 0&0&0&0&0&0&1&0&1&1&0&1&1&0&1&1&1 \\ 0&0&0&0&0&0&0&1&1&1&1&0&1&1&0&1&1 \end{pmatrix}$$

$N(6,6) = 18$:

$$\begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&1&0&0&0&0&0&0&0&0&0&1&1&0&0&1&1&1 \\ 0&0&1&0&0&0&0&0&0&0&1&0&1&0&1&0&1&1 \\ 0&0&0&1&0&0&0&0&0&0&1&1&0&1&0&1&0&1 \\ 0&0&0&0&1&0&0&0&0&1&0&0&1&0&1&1&0&1 \\ 0&0&0&0&0&1&0&0&0&1&0&1&0&1&0&1&1&0 \\ 0&0&0&0&0&0&1&0&0&1&1&0&1&1&0&1&1&1 \\ 0&0&0&0&0&0&0&1&0&1&1&1&0&1&1&0&1&1 \\ 0&0&0&0&0&0&0&0&1&1&1&1&1&0&1&1&1&0 \end{pmatrix}$$

$N(7,3) = 14$: Attained by the $[14,4,7]$ punctured simplex code.
$N(7,4) = 15$: Attained by the $[15,5,7]$ punctured first order Reed-Muller code.
$N(8,3) = 15$: Attained by the $[15,4,8]$ simplex code.
$N(8,4) = 16$: Attained by the $[16,5,8]$ first order Reed-Muller code.

## REFERENCES


[1] N. Alon, L. Babai, and A. Itai, "A fast and simple randomized parallel algorithm for the maximal independent set problem," *J. Algorithms*, vol. 7, no. 4, pp. 567–583, Dec. 1986.

[2] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Nashua, NH, USA: Athena Scientific, 1997.

[3] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," in *Advances in Cryptology – CRYPTO '90*, ser. Lecture Notes in Computer Science, vol. 537. Springer-Verlag, 1991, pp. 2–21.

[4] A. E. Brouwer, "The linear programming bound for binary linear codes," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 677–680, May 1993.

[5] C. Carlet, "On cryptographic propagation criteria for Boolean functions," *Inform. and Comput.*, vol. 151, no. 1-2, pp. 32–56, May 1999.

[6] P. Delsarte, "Bounds for unrestricted codes, by linear programming," *Philips Res. Rep.*, vol. 27, pp. 272–289, 1972.

[7] J. F. Dillon, "Elementary hadamard difference sets," Ph.D. dissertation, Univ. of Maryland, 1974.

[8] K. Kurosawa and T. Satoh, "Design of SAC/PC($l$) of order $k$ boolean functions and three other cryptographic criteria," in *Advances in Cryptology – EUROCRYPTO'97*, ser. Lecture Notes in Computer Science, vol. 1233. Springer-Verlag, 1997, pp. 434–449.

[9] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: Elsevier, 1977.

[10] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Advances in Cryptology – EUROCRYPT '93*, ser. Lecture Notes in Computer Science, vol. 765. Springer-Verlag, 1994, pp. 386–397.

[11] V. S. Pless, W. C. Huffman, and R. A. Brualdi, "An introduction to algebraic codes," in *Handbook of Coding Theory*, V. Pless and W. C. Huffman, Eds. Amsterdam: Elsevier, 1998, pp. 3–139.

[12] B. Preneel, R. Govaerts, and J. Vandewalle, "Boolean functions satisfying higher order propagation criteria," in *Advances in Cryptology — EUROCRYPT '91 Proceedings*, ser. Lecture Notes in Computer Science, vol. 547. Springer-Verlag, 1991, pp. 141–152.

[13] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts, and J. Vandewalle, "Propagation characteristics of Boolean functions," in *Advances in Cryptology — EUROCRYPT '90 Proceedings*, ser. Lecture Notes in Computer Science, vol. 473. Springer-Verlag, 1991, pp. 162–173.

[14] A. F. Webster and S. E. Tavares, "On the design of S-boxes," in *Advances in Cryptology — CRYPTO '85 Proceedings*, ser. Lecture Notes in Computer Science, vol. 218. Springer-Verlag, 1986, pp. 523–534.