Tatra Mt. Math. Publ. **33** (2006), 51–69

TATRA MOUNTAINS
Mathematical Publications

IS RIJNDAEL REALLY INDEPENDENT OF THE FIELD POLYNOMIAL?

Otokar Grošek — Spyros S. Magliveras — Ján Ťapuška — Wandi Wei

ABSTRACT. In many papers about the Advanced Encryption Standard (AES) system Rijndael, including papers by its designers J. Daemen, V. Rijmen [*The Rijndal Block Cipher*, AES Proposal: Rijndael, 1999], E. Barkan, E. Biham [*In how many ways can you write Rindael?*, preprint] and [*The book of Rijndaels*, preprint], the authors claim that the generating polynomial of the underlying field *plays no role* in the selection of the power permutation and subsequent affine transformation, whose composition forms the S-box for Rijndael. In this contribution we show that this is not the case if *input/output entropy* and *propagation criteria* are under our consideration.

1. Introduction

Rijndael was adopted as the new Advanced Encryption Standard (AES) in November, 2001, and was published as FIPS197 in the US Federal Register in December, 2001. The selection process for AES was a truly international affair, and Rijndael was adopted by many countries for *sensitive, secret*, and even *top-secret* information, as indicated in CNSS policy No. 15, Fact Sheet 1, June 2003. Rijndael is characterized by the following parameters (whose choice is fixed in Rijndael):

1. The generating polynomial $m(x)$ for $\mathrm{GF}(2^8) \cong \mathbb{F} = \mathrm{GF}(2)[x]/\big(m(x)\big)$.
2. An S-box which is a permutation of the field elements given by $f(X) = X^{-1}$.
3. An affine transformation of $\mathbb{F}$, viewed as a vector space over $\mathrm{GF}(2)$, and given by a matrix A and a translation vector b.

2000 Mathematics Subject Classification: 68P25, 12E99.
Keywords: cryptography, advanced encryption standard, Rijndael, iterative ciphers, block ciphers, S-boxes, Galois fields, generating polynomials, nonlinearity, propagation criterion, correlation immunity.

This material is based upon work supported under Grant No. VEGA 1/0161/03.

4. A “MixColumns” operation which is given by a matrix C [GS03].

In this paper we show that there is indeed a large collection of optimal choices for these parameters.

2. Preliminaries

For p a prime and $q = p^n$, we denote by $\mathbb{F}_q = \mathrm{GF}(q)$ the finite field of order q. Viewed as a vector space over $\mathbb{F}_p$, $\mathbb{F}_q \cong \mathbb{F}_p \oplus \cdots \oplus \mathbb{F}_p$ which we also denote by $\mathbb{F}_p^n$. When we construct a version of $\mathbb{F}_q = \mathbb{F}_p[x]/\big(m(x)\big)$ by selecting a monic, irreducible polynomial $f(x)$ of degree n, we will call $m(x)$ the *generating polynomial* of $\mathbb{F}_q$. It is standard to say that $m(x)$ is *primitive* if polynomial x is a primitive element of $\mathbb{F}_q = \mathbb{F}_p[x]/\big(m(x)\big)$. If $g(x) \in \mathbb{F}_p[x]$ is an irreducible polynomial of degree n, by the *reciprocal* of $g(x)$, denoted by $\overline{g}(x)$, we mean the polynomial $x^n g(1/x)$. The *general linear* group of dimension n over $\mathbb{F}_q$ is denoted by $\mathrm{GL}_n(q)$. Moreover, we denote by $\mathrm{AGL}_n(q)$ the group of all *affine transformations* of n-dimensional vector space $\mathbb{F}_q^n$ into itself, over $\mathbb{F}_q$. Thus $T \in \mathrm{AGL}_n(q)$ if and only if $T\colon x \to Ax + b$, where $A \in \mathrm{GL}_n(q)$ and $b \in \mathbb{F}_q^n$. Of special importance here are the affine transformations in $\mathrm{AGL}_n(2)$ of $\mathbb{F}_{2^n}$ viewed as an n-dimensional vector space over $\mathbb{F}_2$. If n is a natural number, we denoted the set of natural numbers $\{1, 2, \ldots, n\}$ by $[1, n]$.

If $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n) \in \mathbb{F}_2^n$, we denote by $a \odot b$ the inner product of vectors a and b, i.e., the element $\sum\limits_{i=1}^{n} a_i b_i \in \mathbb{F}_2$. Moreover, if $f = (f_1, \ldots, f_n)$ is an n-tuple of functions $f_i\colon \mathbb{F}_2^r \to \mathbb{F}_2^s$, we denote by $g = a \odot f$ the function $g = \sum\limits_{i=1}^{n} a_i f_i$. A Boolean function f of the Boolean variables $x_1, \ldots, x_n$, can be identified with the 2^n-dimensional binary vector corresponding to its truth table description. Moreover, f can be expressed as a linear combination, over $\mathbb{F}_2$, of the Reed-Muller basis of functions $\mathcal{B} = \{\mathbf{1}, x_1, \ldots, x_n, x_1x_2, \ldots, x_{n-1}x_n, \ldots, x_1x_2\cdots x_n\}$, where $\mathbf{1}$ denotes the constant function which takes the value 1 for all $x \in \mathbb{F}_2^n$. Note that $|\mathcal{B}| = 2^n$.

In block ciphers there is an encoding of plaintext blocks into n-tuples of 0's and 1's. This encoding may be viewed as an embedding of plaintexts into the vector space $\mathbb{F}_2^n$, or if more structure is desirable into the field $\mathbb{F}_{2^n} = \mathrm{GF}(2^n)$. An S-Box is a permutation $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$.[1] We denote by $\Sigma = \Sigma_n$ the collection of all permutations of $\mathbb{F}_2^n$ onto itself, i.e., Σ_n is the symmetric group on $\mathbb{F}_2^n$. One of the advantages of using the field $\mathbb{F}_{2^n}$ rather than just the vector space

[1] More generally an S-Box is a function $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$, where $n \geq m$. In this paper, however, for an S-box we will always have $n = m$.

$\mathbb{F}_2^n$ is the fact that every permutation of the elements of $\mathbb{F}_{2^n}$ is a polynomial with coefficients in $\mathbb{F}_{2^n}$. For example the mapping

$$F(X) = \sum_{i=0}^{2^n-1} y_i\big(1 + (X + x_i)^{2^n-1}\big)$$

maps x_i to y_j, $x_i, y_j \in \mathbb{F}_{2^n}$. The evolution of different techniques in the design of strong S-boxes has a long history. Nowadays, the so called *power permutations* over the field $\mathbb{F}_{2^n}$, initially investigated by Pieprzyk [Pi89], play a crucial role. Such a mapping $\pi_d\colon X \to X^d$ for an appropriate fixed d is then used as a bijection from the field $\mathbb{F}_{2^n}$ onto itself. Since the multiplicative group of $\mathbb{F}_{2^n}$ is cyclic of order 2^n-1, π_d is a bijection if and only if $\gcd(d, 2^n-1) = 1$. In Rijndael the S-box is formed as the composition of π_{-1} with an affine transformation of $\mathbb{F}_{2^n}$ as a vector space over $\mathbb{F}_2$. In general, a potential S-Box for a Rijndael-like system would be a bijection

$$\sigma_{(A,b,d)}\colon X \to T\big(\pi_d(X)\big) = A\pi_d(X) + b, \tag{2.1}$$

where $A \in \mathrm{GL}_n(2)$, $b \in \mathbb{F}_{2^n}$, and d is a unit in the ring of integers modulo $2^n - 1$.

In many papers, including those by the designers of Rijndael [DaR99], [BiBa04a], [BiBa04b], the authors claim that the generating polynomial $m(x)$ of the underlying field over GF(2) *does not play a role* when an S-box is defined as in equation (2.1). In this contribution we show that this is not the case if *input/output entropy* and *propagation characteristics* are the criteria under consideration.

It is easy to establish that the translation b has no effect on the performance characteristics, so we focus on the case where $b = 0$, i.e., we look at S-boxes of the form

$$\sigma_{(A,d)}\colon X \to T\big(\pi_d(X)\big) = AX^d. \tag{2.2}$$

In particular we focus on special matrices $A \in \mathrm{GL}_n(2)$ yielding optimal performance parameters for the S-box of (2.2). We classify matrices A of $\mathrm{GL}_n(2)$ into classes according to the above criteria, and study the distribution of matrices in these classes with respect to the selected generating polynomial $m(x)$ of $\mathrm{GF}(2^n)$ over GF(2), and the degree d of the power permutation π_d.

Properties of S-Boxes $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$, which have been under careful consideration since DES was cryptanalysed can be divided into two groups:

- nonlinear order, nonlinearity, maximum value of the XOR table, and linear approximation bias.
- propagation criterion, correlation immunity, input/output bit-to-bit entropy.

2.1 Group-1 performance criteria

In this section we remind the reader of performance characteristics for an S-box which we aggregate together as the *Group-1* criteria. These include *nonlinear order*, *nonlinearity*, *XORMAX*, and *linear approximation bias.*

1. Let $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a function such that for $x = (x_1, \dots, x_n) \in \mathbb{F}_2^n$, $\sigma(x) = \bigl(\sigma_1(x), \dots, \sigma_m(x)\bigr)$. Then, each σ_i is a Boolean function of the Boolean variables $x_1, \dots, x_n$, and can be expressed as a linear combination over $\mathbb{F}_2$ of the Reed-Muller basis as described above. We denote the degree of σ_i as a polynomial in $x_1, \dots, x_n$ by $\partial\sigma_i$, and remind the reader that the *nonlinear order* of σ is defined by:

$$\operatorname{ord}(\sigma) = \min_{i \in [1,n]} \partial\sigma_i \,. \tag{2.3}$$

2. Let $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ be an arbitrary Boolean function, and let $\mathcal{A}_n$ denote the collection of all affine Boolean functions, i.e., all functions $g\colon \mathbb{F}_2^n \to \mathbb{F}_2$ of the form:

$$g(x_1, \dots, x_n) = a_0 1 \oplus a_1 x_1 \oplus \cdots \oplus a_n x_n \,, \qquad a_i \in \mathbb{F}_2 \,.$$

$\mathcal{A}_n$ is also known as the *first order* Reed-Muller code of length 2^n. Each of the 2^{n+1} functions $g \in \mathcal{A}_n$ *approximates* f in the sense that g agrees (or otherwise) with f over the various domain points $x \in \mathbb{F}_2^n$. The degree of agreement of g with f is measured by the Hamming distance $d(g,f)$, the complete agreement occurring when $d(g,f) = 0$, and the complete disagreement when $d(g,f) = 2^n$. By the *nonlinearity* of f we mean the parameter:

$$N(f) = \min_{g \in \mathcal{A}_n} d(g,f) \,. \tag{2.4}$$

If $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ is a function so that $\sigma(x) = \bigl(\sigma_1(x), \dots, \sigma_m(x)\bigr)$, by the *nonlinearity* of σ we mean:

$$N(\sigma) = \min_{0 \neq a \in \mathbb{F}_2^m} N(a \odot \sigma) \,. \tag{2.5}$$

3. If $V = \mathbb{F}_2^n$, and $\sigma\colon V \to V$ is a given S-box, the XOR-table, XOR-profile or *difference distribution table* [Sti2] for σ is the $2^n \times 2^n$ matrix whose (a,b) entry is

$$D_\sigma(a,b) = \bigl|\bigl\{(x_1,x_2) \in V \times V : x_1 \oplus x_2 = a \,,\ \sigma(x_1) \oplus \sigma(x_2) = b\bigr\}\bigr| \,,$$

where $a, b \in V$. The parameter $\mathrm{XORMAX}(\sigma)$, defined by

$$\mathrm{XORMAX}(\sigma) = \max_{(a,b) \in V \times V} D_\sigma(a,b) \,, \tag{2.6}$$

is of course related to a possible vulnerability of σ to a differential attack. Ideally an S-box $\sigma \in \Sigma$ would have $\mathrm{XORMAX}(\sigma)$ as close to 0 as possible, but such an S-box would not be optimal with respect to other performance criteria.

4. If $\mathbb{F} = \mathbb{F}_{2^n}$, $\sigma\colon \mathbb{F} \to \mathbb{F}$ an S-box, and $B = \{\omega_1, \dots, \omega_n\}$ a basis for $\mathbb{F}$ over $\mathbb{F}_2$, the XOR-profile for σ with respect to B can be defined by means of the above definition, where for $a, b \in \mathbb{F}$ we define $D_\sigma(a,b)$ to be the value of $D_{\sigma'}(a',b')$, where a', b' are the n-tuples representing a and b with respect to B, and σ' the permutation induced by σ on the field viewed as a vector space over $\mathbb{F}_2$.
5. For an S-box σ as above, the *linear approximation table* [Sti2] for σ is the $2^n \times 2^n$ matrix whose (a,b) entry is

$$L_\sigma(a,b) = \left|\left\{x \in \mathbb{F}_2^n : (a \odot x) \oplus \big(b \odot \sigma(x)\big) = 0\right\}\right|$$

where $a, b \in \mathbb{F}_2^n$. From L_σ one can ascertain whether there are entries (a,b) corresponding to random variable $(a \odot x) \oplus (b \odot \sigma(x))$ of bias significantly different from zero. The expected value of entries in L_σ is 2^{n-1}. The value of $|L_\sigma(a,b) - 2^{n-1}|$ significantly greater than zero indicates a significant linear bias for σ, suggesting susceptibility of σ to linear cryptanalysis. Just as in the XOR-profile case, the target for $\max\left|L_\sigma(a,b) - 2^{n-1}\right|$ is 0.

2.2 Group-2 performance criteria

In this section we review the *Group-2* performance criteria for S-boxes.

1. Let $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an S-box with $\sigma(x) = \big(\sigma_1(x), \dots, \sigma_n(x)\big)$. Then for each $i, j \in [1,n]$ we compute the $(i,j)^{th}$ *input/output bit-to-bit* entropy $H\big(x_i/\sigma_j(x)\big)$, and define the parameter:

$$H = \min_{i,j \in [1,n]} H\big(x_i/\sigma_j(x)\big)\,. \tag{2.7}$$

This parameter represents an amount of information about the value of input x_i if the value of output $\sigma_j(x)$ is known. Low values of entropy H indicate high dependence between at least one input bit, and one or more output bits.
2. Recall that a Boolean function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is said to be *balanced* if its truth table has exactly 2^{n-1} 1's. We say that f satisfies the *propagation criterion of degree* k if for each $\alpha \in \mathbb{F}_2^n$ of weight $w(\alpha) \in [1,k]$ the function $f(x) \oplus \big(f(x \oplus \alpha)\big)$ is balanced. More precisely, f satisfies the propagation criterion of degree k if

$$\sum_{x\in\mathbb{F}_2^n} f(x)\oplus f(x\oplus\alpha) = 2^{n-1}\,. \tag{2.8}$$

The propagation criterion of degree 1 is also called the *strict avalanche criterion* (SAC) [Pi89].

3. Let $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function, and denote by $\mathcal{F}$ the set of $\binom{n}{m}2^m$ Boolean functions obtained from f by fixing any m input bits. Let $\mathcal{A} = \{\alpha \in \mathbb{F}_2^n : w(\alpha) \in [1,k]\}$. Then the *propagation criterion bias of degree* k and order m is defined by:

$$\mathrm{PCB}_f(k,m) = \max_{\substack{a\in\mathcal{A}\\ g\in\mathcal{F}}} \left| \sum_{x\in\mathbb{F}_2^n} \big(g(x)\oplus g(x\oplus\alpha)\big) - 2^{n-m-k} \right|. \tag{2.9}$$

4. The *correlation immunity bias* of order m for the Boolean function f is defined by:

$$\mathrm{CIB}_f(m) = \max_{g\in\mathcal{F}} \left|2^m, w(g) - w(f)\right|. \tag{2.10}$$

5. Now, suppose that there is such a general S-box $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^s$ given that $\sigma(x) = \big(\sigma_1(x),\dots,\sigma_s(x)\big)$. Then we define the parameter $\mathrm{PCB}_\sigma(k,m)$ for σ by:

$$\mathrm{PCB}_\sigma(k,m) = \max_{i\in[1,s]} \{\mathrm{PCB}_{\sigma_i}(k,m)\}. \tag{2.11}$$

Similarly, for $\sigma\colon \mathbb{F}_2^n \to \mathbb{F}_2^s$ we define the parameter $\mathrm{CIB}_\sigma(m)$ by:

$$\mathrm{CIB}_\sigma(m) = \max_{i\in[1,s]} \{\mathrm{CIB}_{\sigma_i}(m)\}. \tag{2.12}$$

In this contribution we examine the Group-2 performance criteria:

i) $\mathrm{PCB}_\sigma(1,0)$, ii) $\mathrm{PCB}_\sigma(1,1)$ and iii) $\mathrm{CIB}_\sigma(1)$.

3. Known facts about the Group-1 criteria

The following theorems are known:

THEOREM 3.1. (See [BeDi93].) Let B be a basis of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$, and π_d: $\mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ the power mapping defined by $\pi_d(X) = X^d$, where d is a positive integer, $d < 2^n - 1$. Then, $\mathrm{ord}\big([\pi_d]_B\big) = w_b(d)$, i.e., the nonlinear order of π_d is the Hamming weight of the binary representation of integer d, and is consequently independent of B.

THEOREM 3.2. (See [BeDi93].) If $\Pi = \{\pi_d : gcd(d, 2^n - 1) = 1\}$, then

$$\min_{\pi_d \in \Pi} \operatorname{xormax}(\pi_d) = \begin{cases} 2 & \text{for } n \text{ odd}, \\ 4 & \text{for } n \text{ even}, \end{cases} \tag{3.1}$$

where the minimum is achieved if d has the form $d = 2^n - 1 - 2^t$.

THEOREM 3.3. (See [Pi04].) *If $\Sigma = \Sigma_n$ is the collection of all permutations $\sigma \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, then*

$$\max_{\sigma \in \Sigma} N(\sigma) = \begin{cases} \sum\limits_{i=(n-3)/2}^{n-3} 2^{i+1} & \text{for } n \text{ odd}, \\ \sum\limits_{i=(n-4)/2}^{n-4} 2^{i+2} & \text{for } n \text{ even}. \end{cases} \tag{3.2}$$

The maximum is attained for any $\pi_d \in \Pi$ if d is of the form $d = 2^n - 1 - 2^t$.

We thus see that power permutations $\pi_d \in \Pi$ attain the best possible values for the Group-1 performance criteria. Moreover, it is easy to see that if $\sigma \in \Sigma$, its Group-1 characteristics are preserved when σ is composed with an affine transformation, thus γ defined by

$$\gamma(x) = A\sigma(x) + b$$

has the same characteristics as σ, where $A \in \mathrm{GL}_n(2)$ and $b \in \mathbb{F}_{2^n}$. In particular, the above is true when $\sigma = \pi_d$, with $w_b(d) = n - 1$.

Unfortunately, little is known about the Group-2 characteristics of S-boxes of the form

$$\sigma_{(A,d,f)} \colon x \to Ax^d + b$$

in the representation of the field as $\mathbb{F} = \mathbb{F}_2[x]/(m(x))$.

4. Classes of matrices for Group-2 characteristics

We wish to consider variants of Rijndael in which the S-box is of the form

$$\sigma_{(A,d,f)} \colon x \to Ax^d + b$$

where $m = m(x)$ is an irreducible polynomial of degree n, specifying the representation of the field in the form $\mathbb{F} = \mathbb{F}_2[x]/(m(x))$.

We start to examine which transformations $A \in \mathrm{GL}_n(2)$ are optimal for $d = -1$, and whether the generating polynomial $m(x)$ has an effect. We show in particular that the distribution characteristics are identical for $m(x)$ and its reciprocal $\overline{m}(x)$. We warn the reader that although it may appear counterintuitive that the Group-2 characteristics will depend on the choice of $m(x)$, this is indeed the case as it has already become clear when $n = 4$. One should remember that the Group-2 characteristics are not field invariants, but depend on the representation of the particular field. In Section 8 we generalize our results to arbitrary d, $w(d) = n - 1$.

It is known that

Theorem 4.1.

1. *The Group-2 characteristics are independent of the choice of b in the S-box $x \to Ax^d + b$.*
2. *The Group-2 characteristics are preserved if the rows of A are permuted.*

Hence, for a given generating polynomial $m(x)$, so that $\mathbb{F} = \mathbb{F}_2[x]/\big(m(x)\big)$ we would like to examine the Group-2 characteristics for S-box variants of the form

$$\sigma_{(A,d,m)} \colon x \to Ax^d.$$

5. Results of computer search

In this section we present the results of exhaustive searches with $d = -1$ for each $n = 4, 5, 6$ and all possible irreducible polynomials. We note that $\left|\mathrm{GL}_4(2)\right| = 20{,}160$, $\left|\mathrm{GL}_5(2)\right| = 9{,}999{,}360$, and $\left|\mathrm{GL}_6(2)\right| = 20{,}158{,}709{,}760$, respectively. When $n = 8$, since $\left|\mathrm{GL}_8(2)\right| > 5.34 \times 10^{18}$, we only carry out a random search with irreducible polynomial $x^8 + x^4 + x^3 + x^2 + 1$, and with the AES polynomial $x^8 + x^4 + x^3 + x + 1$.

By Theorem 4.1, part 2., the group-2 characteristics are invariant under row permutations of matrix A. Thus we examine matrices in $\mathrm{GL}_n(2)$ up to row permutations. Consequently, the cardinalities of the resulting classes in the tables below add up to $|\mathrm{GL}_n(2)|/n!$. Target values for the group-2 characteristics are: $H \to 1$, $\mathrm{PCB}(1,0) \to 0$, $\mathrm{PCB}(1,1) \to 0$, and $\mathrm{CIB}(1) \to 0$. We observe that reciprocal irreducible polynomials give rise to identical distributions of matrices and prove this fact as Theorem 7.3 in Section 7. Hence, a single table is presented for each such pair.

5.1 Classes of matrices in $\mathrm{GL}_4(2)$

Remark 5.1. We note that it is already obvious from the two tables below that the choice of irreducible polynomial $m = m(x)$, used to define the field, is significant in the group-2 characteristics of the S-box $\sigma_{(A,d,m)}$.

Polynomials $x^4 + x + 1$ **and** $x^4 + x^3 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.954	4	4	2	16
0.811	4	4	4	824

Polynomial $x^4 + x^3 + x^2 + x + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.811	0	2	4	5
0.811	4	4	4	835

5.2 Classes of matrices in $\mathrm{GL}_5(2)$

Polynomials $x^5 + x^2 + 1$ **and** $x^5 + x^3 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.896	4	4	6	15
0.954	4	6	6	32207
0.896	4	6	6	51106

Polynomials $x^5 + x^3 + x^2 + x + 1$ **and** $x^5 + x^4 + x^3 + x^2 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.954	4	4	4	40
0.896	4	4	6	337
0.954	4	6	4	32167
0.896	4	6	6	50784

Polynomials $x^5 + x^4 + x^2 + x + 1$ **and** $x^5 + x^4 + x^3 + x + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.954	4	4	4	4
0.896	4	4	6	35
0.954	4	6	4	32187
0.896	4	6	6	51101

5.3 Classes of matrices in $\mathrm{GL}_6(2)$

Polynomials $x^6 + x + 1$ and $x^6 + x^5 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.974	4	6	6	19
0.954	4	6	8	32
0.974	4	8	6	148834
0.954	4	8	8	147271
0.974	8	6	6	207
0.954	8	6	8	542
0.974	8	8	6	2370935
0.954	8	8	8	15898048
0.974	8	10	6	841925
0.954	8	10	8	8590395

Polynomial $x^6 + x^3 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0.974	4	8	6	21621
0.974	8	6	6	3560
0.974	8	8	6	452995
0.954	8	8	8	6930101
0.974	8	10	6	2884020
0.954	8	10	8	17705911

Polynomials $x^6 + x^4 + x^3 + x + 1$ and $x^6 + x^5 + x^3 + x^2 + 1$.

H	PCB(1, 0)	PCB(1, 1)	CIB(1)	Count
0,974	4	8	6	7647
0.954	4	8	8	45601
0.974	8	6	6	1
0.989	8	10	4	35
0.974	8	8	6	796640
0.954	8	8	8	3575309
0.974	8	10	6	2117545
0.954	8	10	8	10133798
0.974	8	12	6	1508977
0.954	8	12	8	9812655

Polynomials $x^6 + x^5 + x^2 + x + 1$ **and** $x^6 + x^5 + x^4 + x + 1$**.**

H	PCB(1,0)	PCB(1,1)	CIB(1)	Count
0.974	4	8	6	5616
0.954	4	8	8	17344
0.974	8	6	6	19
0.974	8	8	6	550746
0.954	8	8	8	3850082
0.974	8	10	6	3875386
0.954	8	10	8	19699015

Polynomial $x^6 + x^4 + x^2 + x + 1$ **and** $x^6 + x^5 + x^4 + x^2 + 1$**.**

H	PCB(1,0)	PCB(1,1)	CIB(1)	Count
0.974	4	8	6	10804
0.954	4	8	8	30102
0.974	8	8	6	553032
0.989	8	10	4	3
0.954	8	8	8	6803622
0.974	8	10	6	2812796
0.954	8	10	8	17787849

5.4 Classes of matrices in $\mathrm{GL}_8(2)$

Polynomial $x^8 + x^4 + x^3 + x^2 + 1$ **a random sample.**

H	PCB(1,0)	PCB(1,1)	CIB(1)	Count
0.994	12	16	12	27
0.991	12	16	14	673
0.994	12	18	12	3497
0.989	12	16	16	596
0.991	12	18	14	34124
0.994	12	20	12	14632
0.994	16	16	12	69
0.989	12	18	16	27673
0.991	12	20	14	94627
0.991	16	16	14	1914
0.994	16	18	12	5090
0.989	12	20	16	73277
0.989	16	16	16	5330
0.991	16	18	14	101927
0.994	16	20	12	51276
0.989	16	18	16	225745
0.991	16	20	14	2418768
0.989	16	20	16	7426515

Note that our program is skipping all matrices which display at least one parameter worse than those in the table. Thus, in the above case the sum of frequencies in the "Count" column is not exactly the size of the random sample taken, and is much smaller than $\left|\mathrm{GL}_8(2)\right|/8! \approx 1.32 \times 10^{14}$. It is extremely likely that certain types of matrices, possibly optimal ones, have been missed altogether, as the sample size in the table is approximately 1.05×10^7.

The AES polynomial $x^8 + x^4 + x^3 + x + 1$.

H	PCB(1,0)	PCB(1,1)	CIB(1)	Count
0.991	8	14	14	791
0.989	8	14	16	24282
0.994	8	16	12	7375
0.991	8	16	14	2212060
0.989	8	16	16	12084638
0.994	8	18	12	38663
0.991	8	18	14	16374085
0.989	8	8	16	111228559
0.994	12	14	12	506
0.991	12	14	14	261329
0.989	12	14	16	1780188
0.994	12	16	12	4678434
0.991	12	16	14	736097419
0.989	12	16	16	3707806910
0.997	12	18	8	214
0.996	12	18	10	68257
0.994	12	18	12	7214173311

In this case the sample size in the table is 1.18×10^{10} approximately. Thus from our two examples for matrices in $\mathrm{GL}_8(2)$ we can conclude that the probability p to find a matrix belonging to the best case we found is in the range of $3 \times 10^{-6} \leqslant p \leqslant 8 \times 10^{-8}$.

The matrix A for AES with parameters $H = 0.989$, $\mathrm{PCB}(1,0) = 16$, $\mathrm{PCB}(1,1) = 20$, $\mathrm{CIB}(1) = 16$ belongs to the class of skipped ones.

Below we present matrices having the best properties with respect to the AES polynomial $x^8 + x^4 + x^3 + x + 1$ and $d = -1$. The values of the Group-2 characteristics are $H = 0.991$, $\mathrm{PCB}(1,0) = 8$, $\mathrm{PCB}(1,1) = 14$ and $\mathrm{CIB}(1) = 14$.

$$\begin{bmatrix} 1&0&1&0&0&1&1&1\\ 1&1&0&1&0&1&1&1\\ 1&0&1&0&1&1&1&1\\ 1&1&1&0&1&1&1&1\\ 0&1&1&1&1&0&1&0\\ 0&1&1&0&0&0&0&0\\ 0&1&1&0&1&1&0&0\\ 1&0&0&0&0&0&1&0 \end{bmatrix} \begin{bmatrix} 1&0&1&0&0&1&1&1\\ 1&1&0&1&0&1&1&1\\ 1&0&1&0&1&1&1&1\\ 1&1&1&0&1&1&1&1\\ 0&1&1&1&1&0&1&0\\ 0&1&1&0&0&0&0&0\\ 0&1&1&0&1&1&0&0\\ 1&0&1&1&0&1&1&0 \end{bmatrix}.$$

6. Analysis of $\mathrm{PCB}(1,0)$

If σ is an S-box, the parameter $p = \mathrm{PCB}(1,0)$ is defined by

$$p = \max_k \max_{w(\alpha)=1} \left| \sum_{x\in\mathbb{F}} \big(\sigma_k(x) \oplus \sigma_k(x\oplus\alpha)\big) - 2^{n-1} \right| , \tag{6.1}$$

where $\sigma(x) = \big(\sigma_1(x), \dots , \sigma_k(x), \dots , \sigma_n(x)\big)$, and the maximum is taken over all coordinates $1 \le k \le n$, and all weight 1 vectors α. The sum is taken over all elements of the underlying field, each represented in binary notation, i.e., as elements of the associated vector space. Moreover, the sum is the integer sum of these vectors, i.e., we are interested in the extent to which each coordinate function $\sigma_i(x) \oplus \sigma_i(x\oplus\alpha)$ is balanced.

Here, $\sigma(x) = Ax^{-1}$, where inversion is done in the multiplicative group of the field, setting $0^{-1} = 0$, $A \in \mathrm{GL}_n(2)$, and the operation Ax^{-1} is done after x^{-1} has been represented as a binary vector. Hence, for α of weight 1 we are interested in the multiset

$$Ax^{-1} + A(x+\alpha)^{-1} = A\big(x^{-1} + (x+\alpha)^{-1}\big) . \tag{6.2}$$

Independently of the representation of a field $\mathbb{F}$ of order 2^n we denote by $\mathcal{M}(\alpha)$ the multiset[2]

$$\mathcal{M}(\alpha) = \big[x^{-1} + (x+\alpha)^{-1} \colon x \in \mathbb{F}_{2^n}\big] .$$

It is easy to show the following:

Lemma 6.1. $\mathcal{M}(\alpha) = \alpha^{-1}\mathcal{M}(1)$.

Let $m(x)$ be an irreducible polynomial of degree n over $\mathbb{F}_2$. A table whose rows are indexed by the elements $\{\alpha \in \mathbb{F} \colon \alpha = x^i ,\ 0 \le i \le n-1\}$, and columns

[2] A multiset is a collection of elements with multiplicities, denoted as $[\cdots]$ instead of $\{\cdots\}$. The union of multisets including multiplicities is denoted as $\sqcup$.

by the elements of $\mathbb{F}$ whose (α, z) entry is $z^{-1} + (z+\alpha)^{-1}$ will be called the *basic synthem*[3] with respect to $m(x)$. An example of a basic synthem for $n = 4$ and irreducible polynomial $m(x) = x^4 + x^3 + x^2 + x + 1$ is the following:

$$\mathcal{M}_m =$$

$\alpha \backslash x$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	5	5	14	14	12	12	3	3	13	13	1	1	9	9
2	15	11	15	11	13	15	13	15	7	9	7	9	6	14	6	14
4	8	7	10	3	8	7	10	3	9	11	8	12	9	11	8	12
8	4	6	12	4	5	10	14	11	4	6	12	4	5	10	14	11

In the above an integer entry corresponds to the vector representation of the field element with respect to $m(x)$. For example, 7 represents the vector (0,1,1,1) which corresponds to the element $x^2 + x + 1 \pmod{m(x)}$. It is important to note that the entries in synthems are to be viewed as binary vectors, hence it is clear that the Rijndael basic synthem depends on the choice of the irreducible polynomial $m(x)$. Multiplying the basic synthem $\mathcal{M}_m$ on the left by a fixed $A \in \mathrm{GL}_n(2)$ yields $\sigma_\alpha \colon x \to A\big(x^{-1} + (x+\alpha)^{-1}\big)$, and, is called the *synthem related to* A and the chosen irreducible $m(x)$.

7. Reciprocal polynomials

Let $\mathbb{F}$ be a finite field of order 2^n, and $m(x)$ be the irreducible polynomial over $\mathbb{F}_2$ used to construct $\mathbb{F}$. Let $\beta = \mathcal{M}_m$ be the basic synthem of $\mathbb{F}$ with respect to m. Now, let $\mathrm{GL}_n(2)$ act on the collection of all synthems by $\beta^A = A(\beta)$ where a typical element $A \in G$ acts on β elementwise. Moreover, the matrix A operates on the elements of β, each viewed as a vector over $\mathbb{F}_2$.

We are now in the position to prove

THEOREM 7.1. *Let $g(x)$ be an irreducible polynomial of degree n over $\mathbb{F}_2$ and $h(x) = \overline{g}(x) = x^n g(1/x)$ its reciprocal. Then, the basic synthems $\mathcal{M}_g$ and $\mathcal{M}_h$ lie in the same orbit under* $\mathrm{GL}_n(2)$.

P r o o f. Note that as multisets of elements of $\mathbb{F}$, element $\mathcal{M}_g(1)$ is identical to $\mathcal{M}_h(1)$. Let α be a primitive element of $\mathbb{F}$, then $\mathcal{M}(1) = [1,1] \sqcup \big[1/\big(\alpha^j(\alpha^j+1)\big) \colon j \in [1, 2^n - 2]\big]$. Let ω be a root of $g(x)$, then $\mu = \omega^{-1}$ is a root of $h(x)$. When represented as tables of vectors with respect to the bases $\mathcal{B}_g = \{1, \omega, \dots, \omega^{n-1}\}$, and $\mathcal{B}_h = \{1, \mu, \dots, \mu^{n-1}\}$, $\mathcal{M}_g(1)$ and $\mathcal{M}_h(1)$ are

[3] The word *synthem* is from Greek, and means "compressed message".

related by a transfer matrix $T \in \mathrm{GL}_n(2)$ that is for all entries x in $\mathcal{M}_g(1)$, $[x]_{\mathcal{B}_h} = T[x]_{\mathcal{B}_g}$, where $[x]_{\mathcal{B}_g}$ and $[x]_{\mathcal{B}_h}$ denote the vector representations of field elements x with respect to the bases $\mathcal{B}_g$ and $\mathcal{B}_h$, respectively.

By Lemma 6.1, for each $i \in [0, n-1]$ we have $\mathcal{M}_g(\omega^i) = \omega^{-i}\mathcal{M}_g(1)$, and $\mathcal{M}_h(\mu^{n-1-i}) = \mu^{-(n-1-i)}\mathcal{M}_h(1)$. Therefore we have the following equalities among multisets of elements of $\mathbb{F}_2[x]/(h(x))$, or equivalently, of the elements of $\mathbb{F}_2[x]/(g(x))$:

$$\begin{aligned}\omega^{n-1}\mathcal{M}_g(\omega^i) &= \omega^{n-1-i}\mathcal{M}_g(1)\\ &= \mu^{-(n-1-i)}\mathcal{M}_g(1)\\ &= \mu^{-(n-1-i)}\mathcal{M}_h(1)\\ &= \mathcal{M}_h(\mu^{n-1-i}).\end{aligned}$$

This shows that as multisets of field elements $\omega^{n-1}\mathcal{M}_g(\omega^i)$ and $\mathcal{M}_h(\mu^{n-1-i})$ are the same. Note that the mapping

$$\phi(x) = \omega^{n-1}x$$

from $\mathbb{F}$ to itself is a non-singular linear transformation. Therefore if synthems are regarded as $n \times 2^n$ tables of vectors, there is a matrix $L \in \mathrm{GL}_n(2)$, such that

$$L\mathcal{M}_g(\omega^i) = \mathcal{M}_h(\mu^{n-1-i}),$$

i.e., for each $x \in \mathbb{F}$:

$$[\omega^{n-1}x]_{\mathcal{B}_h} = L[x]_{\mathcal{B}_h} = LT[x]_{\mathcal{B}_g},$$

where multiplying a matrix by a multiset of vectors is done element-wise, and T represents transformation $T\colon [x]_{\mathcal{B}_h} \to [x]_{\mathcal{B}_g}$.

When i goes over $[0, n-1]$, so does $n-1-i$. We conclude that the synthems of vectors $\mathcal{M}_g$ and $\mathcal{M}_h$ lie in the same $\mathrm{GL}_n(2)$ orbit. $\square$

Note that in the case where $g(x) = \overline{g}(x)$ the above linear transformation would fix the synthem multiset-wise.

Example 7.2. Consider the case $n = 4$ with $g(x) = x^4 + x^3 + 1$ and $h(x) = x^4 + x + 1$. The two representations of the field with respect to $g(x)$ and $h(x)$ are shown in the table below:

	$g(x) = x^4 + x^3 + 1$		$h(x) = x^4 + x + 1$	
1	1	0001	1	0001
α	ω	0010	μ	0010
α^2	ω^2	0100	μ^2	0100
α^3	ω^3	1000	μ^3	1000
α^4	ω^4	1001	μ^4	0011
α^5	ω^5	1011	μ^5	0110
α^6	ω^6	1111	μ^6	1100
α^7	ω^7	0111	μ^7	1011
α^8	ω^8	1110	μ^8	0101
α^9	ω^9	0101	μ^9	1010
α^{10}	ω^{10}	1010	μ^{10}	0111
α^{11}	ω^{11}	1101	μ^{11}	1110
α^{12}	ω^{12}	0011	μ^{12}	1111
α^{13}	ω^{13}	0110	μ^{13}	1101
α^{14}	ω^{14}	1100	μ^{14}	1001

Here, we see that

$$L = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Hence,

$$LT = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

and synthem

$$\mathcal{M}_g =$$

1	1	4	4	9	9	10	10	14	14	1	1	11	11	2	2
12	9	12	9	2	1	2	1	8	7	8	7	5	12	5	12
6	14	8	6	6	14	8	6	1	4	12	15	1	4	12	15
3	12	7	2	4	6	3	11	3	12	7	2	4	6	3	11

is carried by LT onto synthem:

$$\mathcal{M}_h =$$

1	1	7	7	6	6	1	1	13	13	9	9	14	14	11	11
9	15	9	15	10	13	10	13	3	7	3	7	9	12	9	12
13	10	14	8	13	10	14	8	5	6	15	13	5	6	15	13
15	3	5	11	7	15	4	14	15	3	5	11	7	15	4	14

Note that the transformation LT viewed as a permutation of the symbols in $\mathbb{Z}_{16}$ can be written as:

$$\lambda = (0)(1\ 15\ 12\ 13\ 2\ 14\ 3)(4\ 7\ 6\ 9\ 5\ 8\ 10)(11).$$

Moreover, note that LT carries $\mathcal{M}_g(1)$ to $\mathcal{M}_h(4)$, $\mathcal{M}_g(2)$ to $\mathcal{M}_h(3)$, $\mathcal{M}_g(3)$ to $\mathcal{M}_h(2)$ and $\mathcal{M}_g(4)$ to $\mathcal{M}_h(1)$.

THEOREM 7.3. *Let $g(x)$ be an irreducible polynomial of degree n over $\mathbb{F}_2$, and $h(x) = \overline{g}(x)$ its reciprocal. Then, the distribution tables with respect to $g(x)$ and $h(x)$ for the second group characteristics of S-boxes $f(x) = Ax^{-1}$, $A \in \mathrm{GL}_n(2)$, are identical.*

Proof. This follows directly from the definitions of the bit-to-bit entropy H, $\mathrm{PCB}(1,0)$, $\mathrm{PCB}(1,1)$, $\mathrm{CIB}(1)$ and Theorem 7.1. □

8. Impact of the degree of the power permutation

Finally we show that there is no impact of the degree of the power permutation $F(X) = X^d$ on a distribution of matrices when $w(d) = n-1$. However, the actual matrices contained in the particular classes differ when changing the value of d. Thus the properties of a given matrix depend on a chosen d. Also recall that there is no computational reason to prefer some d since everything can be realized as a look-up table!

Suppose that $t \neq 0$. Then values for $d = 2^n - 1 - 2^t$ and $d = -2^t$ are the same. Thus $F(X) = X^d = \left(X^{-1}\right)^{2^t}$, which yields

$$\left(X^{-1}\right)^{2^t} + \left((X+\alpha)^{-1}\right)^{2^t} = \beta^{2^t}. \tag{8.1}$$

Since (8.1) is equivalent to

$$\left(X^{-1} + (X+\alpha)^{-1} + \beta\right)^{2^t} = 0\,, \tag{8.2}$$

and there are no divisors of zero, (8.1) is equivalent to (8.2), and has the same number of roots. Thus the synthem $\mathcal{M}(\alpha)$ is the same regardless of d, $w(d) = n-1$ and we get

THEOREM 8.1. *Let $g(x)$ be an irreducible polynomial of degree n over $\mathbb{F}_2$, $\mathbb{F} = \mathbb{F}_2[x]/\big(g(x)\big)$, and $w(d) = n-1$. Then, the distribution tables with respect to $g(x)$ for the second group characteristics of S-boxes $f(x) = Ax^d$, $A \in \mathrm{GL}_n(2)$, are identical regardless of the exponent d.*

9. Conclusions

It has been claimed by various researchers that the choice of the generating polynomial $m(x)$ of the underlying field plays no role in optimally selecting the linear transformation A used in the S-box for Rijndael. In this paper we show that the polynomial used has indeed a profound effect on the *group 2* characteristics of S-box variants for Rijndael, when the matrix A is to be optimally selected. In addition we show that the distributional characteristics, when the matrix A ranges over $\mathrm{GL}_n(2)$ are identical for an irreducible polynomial and its reciprocal. Finally we proved that there is no impact on distribution of matrices A for d with Hamming weight $w(d) = n - 1$.

REFERENCES

[BiBa04a] BARKAN, E.—BIHAM, E.: *In how many ways can you write Rindael?* preprint 2002/157, `http://www.cs.technion.ac.il/~biham`.

[BiBa04b] BARKAN, E.—BIHAM, E.: *The book of Rijndaels*, preprint 2002/58, `http://www.cs.technion.ac.il/~biham`.

[BeDi93] BETH, T.—DING, C.: *On almost perfect nonlinear permutations*, in: Advances in Cryptology, EUROCRYPT '93 (Helleseth, Tor, ed.), Lecture Notes Comput. Sci., Vol. 765, Springer-Verlag, Berlin, 1994, pp. 65–76.

[DaR99] DAEMEN, J.—RIJMEN, V.: *The Rijndael Block Cipher*, AES Proposal: Rijndael, 1999.

[DaR02] DAEMEN, J., RIJMEN, V.: *The Design of Rijndael*, AES—The Advanced Encryption Standard, Springer-Verlag, Berlin, 2002.

[GS03] GROŠEK, O.—ŠIŠKA, J.: *Semigroup of matrices over* $\mathrm{GF}\,2^s$ *and its relation to AES*, Comp. and Informatics **22** (2003), 417–426.

[Jun92] JUNGNICKEL, D.: *Finite Fields*, Structure and Arithmetics, I. Wissenschaftverlag, Meinnhaim, 1992.

[LiN97] LIDL, R.—NIEDERREITER, H.: *Finite Fields*, Cambridge University Press, Cambridge, 1997.

[MuR02] MURPHY, S.—ROBSHAW, M. J. B.: *Essential algebraic structure within the AES*, Advances in Cryptology, CRYPTO '02, Lecture Notes in Comput. Sci., Vol. 2442, Springer-Verlag, Berlin, 2002, pp. 1–16.

[FIPS197] *National Institute of Standards and Technology*, NIST FIPS PUB 197, "Advanced Encryption Standard", U.S. Department of Commerce, 26th november 2001.

[Pi89] PIEPRZYK, J.: *Nonlinearity of exponent permutation*, in: Proceedings of EUROCRYPT '89, Springer-Verlag, Berlin, pp. 80–92.

[Pi04] PIEPRZYK, J.—HARDJONO, T.—SEBERRY, J.: *Fundamentals of Computer Security*, Springer-Verlag, Berlin, 2003.

[Seb95] SEBERRY, J.—XIAN-MO, Z.—YULIANG, Z.: *Nonlinearity and propagation characteristics of balanced Boolean functions*, Inform. and Comput. **119** (1995), 1–13.

[Sti2] STINSON, D. R.: *Cryptography Theory and Practice, second edition*, Chapman and Hall/CRC, Boca-Raton, USA, 2002, p. 339.

[Tap04] ŤAPUŠKA, J.: *Searching for cryptographically strong S-boxes*, MSc. Thesis., STU FEI Bratislava, 2004.

Received October 15, 2004

Otokar Grošek
Ján Ťapuška
Department of Applied Informatics
and Information Technology
Slovak University of Technology
Ilkovičova 3
SK–812 19 Bratislava
SLOVAKIA

E-mail: otokar.grosek@stuba.sk
tapuska@decef.elf.stuba.sk

Spyros S. Magliveras
Wandi Wei
Center for Cryptology and Information Security
Department of Mathematical Sciences
Florida Atlantic University
777 Glades Rd.
Boca Raton, FL. 33431
U. S. A.

E-mail: spyros@fau.edu
wwei@fau.edu