Characterisations of Extended Resiliency and Extended Immunity of S-Boxes

Josef Pieprzyk[1], Xian-Mo Zhang[1], and Jovan Dj. Golić[2]

[1] Centre for Advanced Computing - Algorithms and Cryptography
Department of Computing, Macquarie University
Sydney , NSW 2109, Australia
josef,xianmo@ics.mq.edu.au
[2] Telecom Italia Lab, Telecom Italia
Via G. Reiss Romoli 274, 10148 Turin, Italy
jovan.golic@tilab.com

Abstract. New criteria of extended resiliency and extended immunity of vectorial Boolean functions, such as S-boxes for stream or block ciphers, were recently introduced. They are related to a divide-and-conquer approach to algebraic attacks by conditional or unconditional equations. Classical resiliency turns out to be a special case of extended resiliency and as such requires more conditions to be satisfied. In particular, the algebraic degrees of classically resilient S-boxes are restricted to lower values. In this paper, extended immunity and extended resiliency of S-boxes are studied and many characterisations and properties of such S-boxes are established. The new criteria are shown to be necessary and sufficient for resistance against the divide-and-conquer algebraic attacks by conditional or unconditional equations.

Key Words: Extended Resiliency, Extended Immunity, Divide-and-Conquer Algebraic Attacks

1 Introduction

The concept of divide-and-conquer algebraic attacks by the *conditional* or *unconditional* equations induced from a cipher was introduced recently by Golić in [15]. The basic idea behind the resistance against the new attacks is to design the ciphers so that an attacker cannot induce non-constant equations involving certain subsets of variables within the cipher. For this reason, the notions of extended resiliency along with extended immunity as a special case were introduced in [15].[3] As a special case of extended resiliency, classical resiliency (see for instance [1, 2, 5, 6, 16, 20, 21, 23, 24]) requires additional conditions that are not necessary for resistance against the algebraic attacks by conditional or unconditional equations. Furthermore, the additional conditions restrict the algebraic degrees [3, 21, 22]. Therefore, although possibly useful to resist other

[3] The name 'extended resiliency (immunity)' in this paper corresponds to 'algebraic immunity (resiliency)' in [15]. The name is changed in order to avoid confusion with [4, 12, 13], where 'algebraic immunity' was defined differently.

types of attacks, such as correlation attacks, classically resilient S-boxes are not good candidates for cryptographic blocks that should resist algebraic attacks by conditional or unconditional equations. The extended resiliency (immunity) thus seems to be an appropriate platform for a study of S-boxes that are used in stream or block ciphers. These S-boxes are "provably" immune against the divide-and-conquer algebraic attacks by conditional (unconditional) equations and, unlike classically resilient S-boxes, can have high algebraic degrees.

The aim of the paper is to characterise extended resiliency and extended immunity. More precisely, in Section 2, we recall and describe the divide-and-conquer algebraic attacks by unconditional (conditional) equations induced from an S-box. The algebraic properties of S-boxes to be used in the rest of the paper are provided in Section 3. The extended resiliency and the extended immunity are characterised in Sections 4 and 5, respectively. In Section 6, the extended resiliency (immunity) is characterised in terms of the resistance against algebraic attacks by unconditional (conditional) equations. The relations between the extended immunity, the extended resiliency and the classical resiliency are summarised in Section 7. In Section 8, we demonstrate that algebraic degrees of extended resilient (immune) S-boxes can be as high as $n-1$, where n denotes the input size, and provide the corresponding constructions. In Section 9, an upper bound on extended immunity (resiliency) is analysed. Conclusions and suggestions for future work are given in Section 10. Proofs of mathematical results are provided in the Appendix.

2 Divide-and-Conquer Algebraic Attacks based on Conditional and Unconditional Equations

Algebraic attacks [5, 7, 9–13, 15, 18] have recently been shown to be very powerful against certain types of both stream and block ciphers. Typically, an algebraic attack consists of the following two stages. In the first stage, the attacker finds a collection of equations that holds for some specific input, intermediate, and output variables for the cipher. In the second stage, the attacker observes the accessible variables, fixes the known variables to the observed values, and solves the resulting system of equations. The solution normally reduces the uncertainty of the unknown variables such as the secret key and in some circumstances, the attacker can determine all unknown variables breaking the cipher. The amount of work involved in this attack depends on the algebraic degree of the equations derived by the adversary. The smaller the degree of the equations the more efficient the attack is. To prevent ciphers against algebraic attacks, one would expect that the internal structure of the ciphers does not permit the adversary to derive low degree non-constant equations.

A concept of divide-and-conquer algebraic attacks is recently proposed in [15]. It suggests that algebraic attacks can be based on equations involving only subsets of input or output variables for individual nonlinear components of a cipher such as S-boxes or lookup tables. The equations can be unconditional, involving both input and output variables, or conditioned on the output, involving only the input variables. The conditional scenario is shown to be useful for stream ciphers (e.g., based on linear feedback shift registers), while the unconditional scenario may possibly be used for both block ciphers and stream ciphers. More precisely, in iterated constructions, such equations may possibly reduce the number of intermediate variables involved in the equations and hence also the complexity of algebraic attacks. The two scenarios of induced algebraic equations are described next.

Let $F(x) = y$ be an $n \times m$ S-box, where $x = (x_1, \ldots, x_n) \in GF(2)^n$ and $y \in GF(2)^m$. [4] For a fixed integer t, $0 \leq t \leq n$, let $T = \{j_1, \ldots, j_t\}$ be fixed ordered t-subset of $\{1, \ldots, n\}$ and $\{i_1, \ldots, i_{n-t}\} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_t\}$, where $j_1 < \cdots < j_t$ and $i_1 < \cdots < i_{n-t}$. Set $x' = (x_{j_1}, \ldots, x_{j_t})$ and $x'' = (x_{i_1}, \ldots, x_{i_{n-t}})$. We define a subset $W(F, T)$ of $GF(2)^{t+m}$ as follows.

$$\begin{aligned} W(F,T) = \{&(\alpha', \beta) \mid \alpha' \in GF(2)^t, \beta \in GF(2)^m, \\ &(\exists\, \alpha'' \in GF(2)^{n-t},\ x' = \alpha', x'' = \alpha'', F(x) = \beta)\}. \end{aligned} \tag{1}$$

Let a function h on $GF(2)^{t+m}$ satisfy

$$h(x', y) = 0, \text{ for all } (x', y) \in W(F, T). \tag{2}$$

Then the equation (2), over x' and y, is called an *unconditional algebraic equation* induced from $F(x) = y$ (for the fixed T). Of course, such h always exists as h can be the constant zero function. However the attackers try to find a non-constant h so as to eliminate x'' and involve only the variables in x' and y. To make this divide-and-conquer strategy ineffective, it is desirable that F does not induce non-constant unconditional algebraic equations, e.g., for relatively small values of t.

In particular, when $\beta \in F(GF(2)^n)$ in (1) is fixed, we define a subset $W(F, T, \beta)$ of $GF(2)^t$ as follows.

$$\begin{aligned} W(F,T,\beta) = \{&\alpha' \mid \alpha' \in GF(2)^t, \\ &(\exists\, \alpha'' \in GF(2)^{n-t},\ x' = \alpha', x'' = \alpha'', F(x) = \beta)\}. \end{aligned} \tag{3}$$

Let a function h on $GF(2)^t$ satisfy

$$h(x') = 0, \text{ for all } x' \in W(F, T, \beta). \tag{4}$$

Then the equation (4), over x' only, is called a *conditional algebraic equation* induced from $F(x) = \beta$. Similarly, it may be desirable that F does not induce non-constant conditional algebraic equations, e.g., for relatively small values of t. The extended immunity (resiliency) of vectorial Boolean functions is defined in [15] in order to describe the divide-and-conquer effect of induced algebraic equations. This will be studied in more detail in Section 6 and provides a practical motivation for our work. On the other hand, the extended resiliency (immunity) naturally generalises the well-known notion of classical resiliency (immunity) and it is thus theoretically interesting to investigate its properties and propose new constructions.

[4] Here and throughout, we use a simplified notation $GF(2)^n$ for $(GF(2))^n$.

3 Algebraic Properties of S-boxes

In this section, we provide the necessary background including notations that are used in the next sections. We write all vectors in $GF(2)^n$ as $(0,\dots,0,0)=\alpha_0$, $(0,\dots,0,1)=\alpha_1$, …, $(1,\dots,1,1)=\alpha_{2^n-1}$, and call α_i the *binary representation* of integer i, $i=0,1,\dots,2^n-1$. A Boolean function f is a mapping from $GF(2)^n$ to $GF(2)$. We express f as $f(x)=f(x_1,\dots,x_n)$ where $x=(x_1,\dots,x_n)\in GF(2)^n$. The *truth table* of f is defined as $(f(\alpha_0), f(\alpha_1), \dots, f(\alpha_{2^n-1}))$, and the *sequence* of f is defined as $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \dots, (-1)^{f(\alpha_{2^n-1})})$. The scalar product of $\alpha=(a_1,\dots,a_n)$, $\beta=(b_1,\dots,b_n)\in GF(2)^n$, denoted by $\langle\alpha,\beta\rangle$, is defined as $\langle\alpha,\beta\rangle= a_1b_1\oplus\cdots\oplus a_nb_n$ where $\oplus$ denotes the binary addition. We call $h(x)=\langle\alpha,x\rangle\oplus c$ an *affine function*, where $\alpha,x\in GF(2)^n$ and $c\in GF(2)$. In particular, h is called a *linear function* if $c=0$.

Consider a mapping $F=(f_1,\dots,f_m)$ from $GF(2)^n$ to $GF(2)^m$, where each f_j is a Boolean function on $GF(2)^n$ and is called a coordinate or component function of F. We express F as $F(x)=F(x_1,\dots,x_n)$ where $x=(x_1,\dots,x_n)\in GF(2)^n$. F is also called an *S-box* or a *vectorial Boolean function*. From now, we call F an $n\times m$ S-box. F is said to be *affine* if all its coordinate functions are affine, and in particular, F is said to be *linear* if all its coordinate functions are linear. For any k, $1\le k\le m$, and any k-subset $\{j_1,\dots,j_k\}$ of $\{1,\dots,m\}$, where $j_1<\cdots<j_t$, the mapping $\hat{F}=(f_{j_1},\dots,f_{j_k})$ from $GF(2)^n$ to $GF(2)^k$ is called a *k-subfunction* of F.

Notation 1 *Let $\alpha_i\in GF(2)^n$ be the binary representation of integer i, $i=0,\dots,2^n-1$, and $\gamma_j\in GF(2)^m$ be the binary representation of integer j, $j=0,\dots,2^m-1$. For an $n\times m$ S-box F, we define a $2^n\times 2^m$ $(1,-1)$ matrix $D_F=(d_{ij})$: $d_{ij}=(-1)^{\langle F(\alpha_i),\gamma_j\rangle}$. Also we define a $2^n\times 2^m$ real-valued $(0,1)$ matrix $C_F=(c_{ij})$: $c_{ij}=1$ if and only if $F(\alpha_i)=\gamma_j$.*

Recall that a $k\times k$ $(1,-1)$-matrix M is called a *Hadamard* matrix if $MM^T=kI_k$, where M^T is the transpose of M and I_k is the $k\times k$ identity matrix [17]. A $2^s\times 2^s$ Sylvester-Hadamard matrix, denoted by H_s, is defined by the following recursive relation: $H_0=1$, $H_s=\begin{bmatrix} H_{s-1} & H_{s-1}\\ H_{s-1} & -H_{s-1}\end{bmatrix}$, $s=1,2,\dots$. Clearly H_s is a symmetric matrix. Denote the jth row (column) of H_m by ℓ_j (ℓ_j^T), $j=0,1,\dots.2^m-1$. It is known that ℓ_j is the sequence of a linear function $\psi_j(y)=\langle\gamma_j,y\rangle$ where $y\in GF(2)^m$.

Lemma 1. *Let F be an $n\times m$ S-box. Then $D_FH_m=2^mC_F$.*

Lemma 2. *Let F be an $n\times m$ S-box. Let $diag(\lambda_0,\lambda_1,\dots,\lambda_{2^m-1})$ be a diagonal matrix, where λ_j denotes the number of times that F takes value γ_j. Then $D_F^TD_F=H_m diag(\lambda_0,\lambda_1,\dots,\lambda_{2^m-1})H_m$.*

Recall the basic facts from linear algebra, for instance, from [19]. If an $s\times s$ matrix A with real entries, a nonzero s-dimensional column vector η with real coordinates and a real number κ satisfy $A\eta=\kappa\eta$, then κ is called the eigenvalue of matrix A corresponding to the eigenvector η or, alternatively, η is called an eigenvector of matrix A corresponding to the eigenvalue κ. For a fixed matrix A, each eigenvector corresponds to only one eigenvalue, whereas an eigenvalue does not necessarily correspond to only one eigenvector. Usually, a real square matrix does not necessarily have a real eigenvector. However, any real symmetric $s \times s$ matrix must have s linearly independent real eigenvectors.

Corollary 1. *Let F be an $n \times m$ S-box. Then the jth column ℓ_j^T of H_m is the eigenvector of $D_F^T D_F$ corresponding to the eigenvalue $2^m \lambda_j$, where λ_j denotes the number of times that F takes value γ_j.*

4 Extended Resilient S-boxes

The concept of extended resiliency was originally proposed by Golić [15]. In this section, we derive various characterisations of the extended resiliency.

4.1 Surjective S-boxes

Before defining the extended resiliency, we introduce necessary notations.

Notation 2 *Let F be an $n \times m$ S-box. For a subset S of $GF(2)^n$, we write $\{F(x) \mid x \in S\} = F(S)$.*

Definition 1. *Let F be an $n \times m$ S-box. F is said to be* surjective *(or* onto *$GF(2)^m$) if $F(GF(2)^n) = GF(2)^m$.*

Lemma 3. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is surjective, (ii) all eigenvalues of $D_F^T D_F$ are nonzero, and (iii) the rank of D_F is 2^m.*

Definition 2. *Functions $f_1, \ldots, f_m$ on $GF(2)^n$ are said to be* functionally independent *if for any non-constant Boolean function h on $GF(2)^m$, $h(f_1, \ldots, f_m)$ is non-constant.*

Clearly linear independence is a special case of functional independence.

Lemma 4. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is surjective, (ii) for any integer k, $1 \le k \le m$, and any surjective $m \times k$ S-box P, the $n \times k$ S-box $P(F(x))$ is surjective, and (iii) the coordinate functions of F are functionally independent.*

4.2 Extended Resiliency and its Properties

Notation 3 *For a fixed t-subset $T = \{j_1, \ldots, j_t\}$ of $\{1, \ldots, n\}$ and a fixed vector $\alpha = (a_1, \ldots, a_t) \in GF(2)^t$, we define a subset $S(n, T, \alpha)$ of $GF(2)^n$: $S(n, T, \alpha) = \{x = (x_1, \ldots, x_n) \mid x \in GF(2)^n, x_{j_1} = a_1, \ldots, x_{j_t} = a_t\}$.* [5] *Formally, for $t = 0$, a 0-subset T is the empty set, i.e., $T = \emptyset$ and α is not defined, then $S(n, T, \alpha)$ becomes $GF(2)^n$.*

[5] Here and throughout, a t-subset $\{j_1, \ldots, j_t\}$ is assumed to be ordered so that $j_1 < \cdots < j_t$.

Let $\#X$ denote the number of elements in a set X. Then $\#S(n, T, \alpha) = 2^{n-t}$.

Lemma 5. *For fixed subsets $T = \{j_1, \ldots, j_t\}$, $T' = \{j_1, \ldots, j_{t-1}\}$ and fixed vectors $\alpha = (a_1, \ldots, a_t) \in GF(2)^t$ and $\alpha' = (a_1, \ldots, a_{t-1}) \in GF(2)^{t-1}$, we have $S(n, T, \alpha) \subseteq S(n, T', \alpha')$.*

Definition 3. *Let F be an $n \times m$ S-box. Then F is said to be (n, m, t)-*extended resilient *if for any t-subset T of $\{1, \ldots, n\}$ and any $\alpha \in GF(2)^t$, we have $F(S(n, T, \alpha)) = GF(2)^m$. An (n, m, t)-extended resilient S-box is also said to be t-*extended resilient *if we ignore the particular values of n and m.*

It follows that any (n, m, t)-extended resilient S-box is surjective, in particular, any $(n, m, 0)$-extended resilient S-box is equivalent to a surjective $n \times m$ S-box.

Proposition 1. *For any (n, m, t)-extended resilient S-box, it is necessary that $t \leq n - m$.*

The following claim directly follows from Lemma 5.

Lemma 6. *Let F be an $n \times m$ S-box. Then F is (n, m, t)-extended resilient if and only if F is (n, m, k)-extended resilient for $k = 0, \ldots, t$.*

Due to Lemma 6, we are able to introduce the following definition.

Definition 4. *If F is an (n, m, t)-extended resilient S-box, but is not $(n, m, t+1)$-extended resilient, then t is called the* extended resiliency order *of F.*

Proposition 2. *Let F be an (n, m, t)-extended resilient S-box. Then $F(x)$ runs through each vector in $GF(2)^m$ at least 2^t times while x runs through $GF(2)^n$ once.*

4.3 Characterisations of Extended Resilient S-boxes

Definition 5. *Let S be a subset of $GF(2)^n$. Then the* characteristic function *of S, denoted by χ_S, is a Boolean function on $GF(2)^n$ defined as $\chi_S(\alpha) = 1$ if and only if $\alpha \in S$.*

Theorem 1. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is an (n, m, t)-extended resilient, (ii) for any fixed t-subset T of $\{1, \ldots, n\}$ and any fixed $\alpha \in GF(2)^t$, all eigenvalues of $D_F^T diag(b_0, b_1, \ldots, b_{2^n-1})D_F$ are nonzero, where $(b_0, b_1, \ldots, b_{2^n-1})$ denotes the truth table of the characteristic function of $S(n, T, \alpha)$ and each b_j is regarded a real number, and (iii) the rank of $diag(b_0, b_1, \ldots, b_{2^n-1})D_F$ is 2^m.*

The next claim follows from Lemma 4 and Definition 3.

Theorem 2. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended resilient, (ii) for any integer k, $1 \leq k \leq m$, and any surjective $m \times k$ S-box P, the $n \times k$ S-box $P(F(x))$ is (n, k, t)-extended resilient, and (iii) for any t-subset $T = \{j_1, \ldots, j_t\}$ of $\{1, \ldots, n\}$ and any $\alpha = (a_1, \ldots, a_t) \in GF(2)^t$, the coordinate functions of $F(x)|_{x_{j_1}=a_1, \ldots, x_{j_t}=a_t}$, i.e., the restriction of F to $S(n, T, \alpha)$, are functionally independent.*

The necessity in the following statement holds due to Theorem 2 and the sufficiency is obvious.

Corollary 2. *Let F be an $n \times m$ S-box. Then F is (n, m, t)-extended resilient if and only if for any k, $1 \le k \le m$, every k-subfunction $\hat{F}$ of F is (n, k, t)-extended resilient.*

Theorem 3. *Let F be an $n \times m$ S-box. Then F is (n, m, t)-extended resilient if and only if for any fixed r, $1 \le r \le t$, any fixed r-subset $T = \{j_1, \ldots, j_r\}$ of $\{1, \ldots, n\}$ and every nonzero Boolean function g on $GF(2)^r$, $g(x_{j_1}, \ldots, x_{j_r})F(x)$ is surjective, i.e., $\{g(x_{j_1}, \ldots, x_{j_t})F(x) | x \in GF(2)^n\} = GF(2)^m$.*

The next theorem is helpful for understanding the extended resiliency.

Theorem 4. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended resilient, (ii) for any integer t_0, $0 \le t_0 \le t$, any t_0-subset T_0 of $\{1, \ldots, n\}$ and any $\alpha \in GF(2)^{t_0}$, the restriction of $F(x)$ to $S(n, T_0, \alpha)$ is $(t - t_0)$-extended resilient, and (iii) for any integer t_0, $0 \le t_0 \le t$, any t_0-subset T_0 of $\{1, \ldots, n\}$ and any $\alpha \in GF(2)^{t_0}$, $F(x)$ runs through each vector in $GF(2)^m$ at least 2^{t-t_0} times while x runs through $S(n, T_0, \alpha)$ once.*

The following statement follows from Theorem 4.

Corollary 3. *Let F be an (n, m, t)-extended resilient S-box. For any integer $k \ge 1$, define an $(n+k) \times m$ S-box F^* as $F^*(\alpha, \beta) = F(\alpha)$ for each $\alpha \in GF(2)^n$ and $\beta \in GF(2)^k$. Then F^* is $(n + k, m, t)$-extended resilient.*

5 Extended Immune S-boxes

The extended immunity proposed by Golić [15] is more general than the extended resiliency. In this section, we derive characterisations of the extended immunity.

5.1 Extended Immunity and its Properties

Definition 6. *Let F be an $n \times m$ S-box. Then F is said to be (n, m, t)-*extended immune *if for any t-subset T of $\{1, \ldots, n\}$ and any $\alpha \in GF(2)^t$, we have $F(S(n, T, \alpha)) = F(GF(2)^n)$. An (n, m, t)-extended immune S-box is also said to be t-*extended immune *if we ignore the particular values of n and m.*

An (n, m, t)-extended immune S-box is (n, m, t)-extended resilient if and only if $F(GF(2)^n) = GF(2)^m$. Recall that $S(n, T, \alpha)$ with $\#T = 0$ denotes $GF(2)^n$. Thus any $n \times m$ S-box is $(n, m, 0)$-extended immune.

Proposition 3. *For any (n, m, t)-extended immune S-box F, it is necessary that $t \le n - \log_2 \#F(GF(2)^n)$.*

In particular, if F is an (n, m, t)-extended immune S-box, then the inequality $t \le n - \log_2 \#F(GF(2)^n)$ becomes $t \le n - m$, as in Proposition 1.

Lemma 5 and Definition 6 imply the following lemma.

Lemma 7. *Let F be an $n \times m$ S-box. Then F is (n, m, t)-extended immune if and only if F is (n, m, k)-extended immune for $k = 0, \dots, t$.*

According to Lemma 7, we are able to introduce the following definition.

Definition 7. *If F is an (n, m, t)-extended immune S-box, but is not $(n, m, t+1)$-extended immune, then t is called the* extended immunity order *of F.*

Similarly to Proposition 2, we have the following more general statement.

Proposition 4. *Let F be an (n, m, t)-extended immune S-box. Then $F(x)$ runs through each vector in $F(GF(2)^n)$ at least 2^t times while x runs through $GF(2)^n$ once.*

5.2 Characterisations of Extended Immune S-boxes

We start with the following simple result.

Theorem 5. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended immune and (ii) for any fixed t-subset T of $\{1, \dots, n\}$ and any two vectors $\alpha, \alpha' \in GF(2)^t$, we have $F(S(n, T, \alpha)) = F(S(n, T, \alpha'))$.*

By using a similar argument in the proof of Theorem 1, we can prove Theorem 6, whereas Theorem 7 and Corollary 4 correspond to Theorem 2 and Corollary 2, respectively.

Theorem 6. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended immune and (ii) for any fixed t-subset T of of $\{1, \dots, n\}$ and any fixed $\alpha \in GF(2)^t$, the eigenvalue corresponding to the eigenvector ℓ_j^T of $D_F^T diag(b_0, b_1, \dots, b_{2^n-1}) D_F$ is nonzero if and only if the the eigenvalue corresponding to the eigenvector ℓ_j^T of $D_F^T D_F$ is nonzero, where ℓ_j^T is the jth column of H_m, $j = 0, 1, \dots, 2^m - 1$, $(b_0, b_1, \dots, b_{2^n-1})$ denotes the truth table of the characteristic function of $S(n, T, \alpha)$ and each b_j is regarded as a real number.*

Theorem 7. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended immune, (ii) for any integer k, $1 \le k \le m$, and any $m \times k$ S-box P (not necessarily surjective), $P(F(x))$ is (n, k, t)-extended immune, and (iii) for any t-subset $T = \{j_1, \dots, j_t\}$ of $\{1, \dots, n\}$, any $\alpha = (a_1, \dots, a_t) \in GF(2)^t$, and any Boolean function h on $GF(2)^m$, if $h(F)$ is non-constant, then $h(G)$ is non-constant, where $G(x) = F(x)|_{x_{j_1}=a_1, \dots, x_{j_t}=a_t}$.*

Corollary 4. *Let F be an $n \times n$ S-box. Then F is (n, m, t)-extended immune if and only if for any k, $1 \le k \le m$, every k-subfunction $\hat{F}$ of F is (n, k, t)-extended immune.*

By using a similar argument as in the proof of Theorem 3, we can prove the following characterisation.

Theorem 8. *Let F be an $n \times m$ S-box. Then F is (n, m, t)-extended immune if and only if for any fixed r, $1 \leq r \leq t$, any fixed r-subset $T = \{j_1, \ldots, j_r\}$ of $\{1, \ldots, n\}$ and every nonzero Boolean function g on $GF(2)^r$,*
$\{g(x_{j_1}, \ldots, x_{j_r})F(x) | x = (x_1, \ldots, x_n) \in GF(2)^n\} = F(GF(2)^n)$.

By the same reasoning as for Theorem 4, we can also prove the following theorem, whereas Corollary 5 corresponds to Corollary 3.

Theorem 9. *Let F be an $n \times m$ S-box. Then the following statements are equivalent: (i) F is (n, m, t)-extended immune, (ii) for any integer t_0, $0 \leq t_0 \leq t$, any t_0-subset T_0 of $\{1, \ldots, n\}$ and any $\alpha_0 \in GF(2)^{t_0}$, the restriction of $F(x)$ to $S(n, T_0, \alpha_0)$ is $(t - t_0)$-extended immune, and (iii) for any integer t_0, $0 \leq t_0 \leq t$, any t_0-subset T_0 of $\{1, \ldots, n\}$ and any $\alpha_0 \in GF(2)^{t_0}$, $F(x)$ runs through each vector in $F(GF(2)^n)$ at least 2^{t-t_0} times while x runs through $S(n, T_0, \alpha)$ once.*

Corollary 5. *Let F be an (n, m, t)-extended immune S-box. For any integer $k \geq 1$, define an $(n + k) \times m$ S-box F^* as $F^*(\alpha, \beta) = F(\alpha)$ for each $\alpha \in GF(2)^n$ and $\beta \in GF(2)^k$. Then F^* is $(n + k, m, t)$-extended immune.*

6 Characterisation of Extended Resiliency (Immunity) in Terms of Existence of Unconditional (Conditional) Equations

In this section, we characterise the extended resiliency (immunity) in terms of the resistance against divide-and-conquer algebraic attacks by unconditional (conditional) equations. For completeness, we first give a new proof for a result from [15] about the existence of conditional algebraic equations. Then, we prove a new result about the existence of unconditional algebraic equations.

Lemma 8. *For a fixed t-subset $T = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, n\}$, there is no non-constant conditional algebraic equation over $x' = (x_{j_1}, \ldots, x_{j_t})$ induced from $F(x_1, \ldots, x_n) = \beta$ for any value of $\beta \in F(GF(2)^n)$ if and only if $F(S(n, T, x')) = F(GF(2)^n)$, for every $x' \in GF(2)^t$.*

Theorem 10. *There is no non-constant conditional algebraic equation over $x' = (x_{j_1}, \ldots, x_{j_t})$ induced from $F(x_1, \ldots, x_n) = \beta$ for any t-subset $T = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, n\}$ and any value of $\beta \in F(GF(2)^n)$ if and only if F is (n, m, t)-extended immune.*

Accordingly, for a (n, m, t)-extended immune S-box F, if the attackers want to establish a non-constant conditional algebraic equation induced from F, they have to choose x' with dimension higher than the extended immunity order defined in Definition 7.

Lemma 9. *For a fixed t-subset $T = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, n\}$, there is no non-constant unconditional algebraic equation over $x' = (x_{j_1}, \ldots, x_{j_t})$ and y induced from $F(x_1, \ldots, x_n) = y$ if and only if $F(S(n, T, x')) = GF(2)^m$, for every $x' \in GF(2)^t$.*

Theorem 11. *There is no non-constant unconditional algebraic equation over* $x' = (x_{j_1}, \ldots, x_{j_t})$ *and* y *induced from* $F(x_1, \ldots, x_n) = y$ *for any* t*-subset* $T = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, n\}$ *if and only if* F *is* (n, m, t)*-extended resilient.*

Therefore, for an (n, m, t)-extended resilient S-box F, if the attackers want to establish a non-constant unconditional algebraic equation induced from F, they have to choose x' with dimension higher than the extended resiliency order defined in Definition 4.

7 Relations between Extended Immunity, Extended Resiliency and Classical Resiliency

Classically resilient S-boxes (see for instance [1, 2, 5, 6, 8, 16, 20, 21, 23, 24]) were studied previously. The following is the definition of classical resiliency.

Definition 8. *Let* F *be an* $n \times m$ *S-box. If for any* t*-subset* T *of* $\{1, \ldots, n\}$ *and any* $\alpha \in GF(2)^t$, $F(x)$ *runs through each vector in* $GF(2)^m$ *exactly* 2^{n-m-t} *times while* x *runs through* $S(n, T, \alpha)$ *once, then* F *is said to be* (n, m, t)-classically resilient.

Classical resiliency in Definition 8 was usually called resiliency. In this paper, we call it classical resiliency so as to avoid confusion. Summarily, classical resiliency is a special case of extended resiliency and extended resiliency is a special case of extended immunity. We next establish some relations among the three.

Proposition 5. *Any affine* (n, m, t)*-extended resilient S-box is also* (n, m, t)*-classically resilient.*

Theorem 12. *Let* F *be an affine* (n, m, t)*-extended immune S-box. Then* $\#F(GF(2)^n) = 2^k$*, where* k *is an integer, and there exist a vector* $\beta \in GF(2)^m$ *and an* $m \times k$ *matrix* B *over* $GF(2)$ *such that the mapping* $P(x) = (F(x) \oplus \beta)B$ *is an* (n, k, t)*-classically resilient S-box.*

According to Theorem 12, any affine extended immune S-box can be used to construct a classically resilient S-box by an appropriate affine transformation of the output.

Theorem 13. *Let* Q *be an* (n, m, t)*-extended immune S-box and let* k *be an integer satisfying* $k = \lfloor \log_2 \#Q(GF(2)^n) \rfloor$*, where* $\lfloor r \rfloor$ *denotes the maximum integer less than or equal to* r*. For any mapping* P *from* $Q(GF(2)^n)$ *onto* $GF(2)^k$*, define a mapping* $F = P(Q(x))$ *where* $x \in GF(2)^n$*. Then* F *is an* (n, k, t)*-extended resilient S-box.*

According to Theorem 13, any extended immune S-box can be transformed into an extended resilient S-box by an appropriate mapping of the output.

8 Algebraic Degree of Extended Resilient S-boxes

The algebraic degree of $n \times m$ S-box $F = (f_1, \ldots, f_m)$, denoted by $deg(F)$, is defined as $deg(F) = \min_g \{deg(g) | g = \bigoplus_{j=1}^{m} c_j f_j,\ (c_1, \ldots, c_m) \neq (0, \ldots, 0)\}$. S-boxes with high algebraic degrees are desirable for resistance against algebraic attacks.

Lemma 10. *The algebraic degree of any $n \times m$ S-box F is at most $n - 1$ if $m \geq 2$.*

From [24], an $n \times m$ S-box is (n, m, t)-classically resilient if and only if each nonzero linear combination of its coordinate functions is $(n, 1, t)$-classically resilient. Due to [22], the algebraic degree of an $(n, 1, t)$-classically resilient function is less than $n - t$ unless $t = n - 1$ (Siegenthaler's inequality). Therefore, the algebraic degree of any (n, m, t)-classically resilient S-box is less than $n - t$ (when $m \geq 2$). We will show that unlike the classical resiliency order, the extended immunity (resiliency) order does not restrict the algebraic degree.

Notation 4 *Let β_j denote the vector in $GF(2)^n$ whose jth component is zero and all other components are one, and $\beta_0 = (1, \ldots, 1) \in GF(2)^n$. Let $(x_1 \cdots x_n)/x_j$ denote the product $x_1 \cdots x_{j-1} x_{j+1} \cdots x_n$ of $n - 1$ variables.*

Lemma 11. *Let f be a Boolean function on $GF(2)^n$. For any integer j, $1 \leq j \leq n$, let $f'(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) \oplus (x_1 \cdots x_n)/x_j$. Then f' differs from f only for $x = \beta_0$ and $x = \beta_j$. If $deg(f) < n - 1$ then $deg(f') = n - 1$.*

Theorem 14. *Let $F = (f_1, \ldots, f_m)$ be an (n, m, t)-extended resilient S-box, $deg(f_j) < n - 1$, $j = 1, \ldots, m$, and $t > \lfloor \log_2(m + 1) \rfloor + 1$. Let $F' = (f'_1, \ldots, f'_m)$ be an $n \times m$ S-box, where $f'_j(x_1, \ldots, x_n) = f_j(x_1, \ldots, x_n) \oplus (x_1 \cdots x_n)/x_j$. Then F' is an (n, m, t_0)-extended resilient S-box such that $t_0 = t - \lfloor \log_2(m + 1) \rfloor - 1$ and $deg(F') = n - 1$.*

Therefore, an (n, m, t_0)-extended resilient S-box F' achieves the maximum degree $n - 1$. In contrast with F', if there exists an (n, m, t_0)-classically resilient S-box, due to Siegenthaler's inequality, its algebraic degree is less than $n - t_0$ (when $m \geq 2$). Furthermore, there is another problem: it is unknown whether this upper bound on the algebraic degree of classically resilient S-boxes can be reached except for special cases. For these reasons, classically resilient S-boxes may not be desirable with respect to algebraic attacks.

In the proof of Theorem 15, we construct (n, m, t)-extended resilient S-boxes with maximum algebraic degree $n - 1$, for any given m and t, but the number of inputs, n, has to be sufficiently large.

Theorem 15. *For any given m and t and $r \geq t + \lfloor \log_2(m+1) \rfloor + 2$, there exists an (rm, m, t)-extended resilient S-box with algebraic degree $rm - 1$.*

9 Upper Bound on Extended Immunity (Resiliency)

Recall that the classical resiliency t of an (n, m, t)-resilient function is upper-bounded by $t \leq \lfloor \frac{2^{m-1}n}{2^m-1} \rfloor - 1$ [14] and $t \leq 2\lfloor \frac{2^{m-2}(n+1)}{2^m-1} \rfloor - 1$ [2]. In this section, we indicate that the upper bound on extended immunity (resiliency) order is different from the bound on the classical resiliency order. According to Proposition 3, any (n, m, t)-extended immune S-box F satisfies $t \leq n - \log_2 \#F(GF(2)^n)$. We next show that this upper bound is tight for large t. We first provide a new proof of a result from [15] concerning Boolean functions, i.e., $m = 1$. Then we generalise this result to an arbitrary m.

Lemma 12. *Let f be a Boolean function on $GF(2)^n$ with an extended immunity order t. Then (i) $t = n$ if and only if f is constant and (ii) $t = n - 1$ if and only if $f(x) = x_1 \oplus \cdots \oplus x_n \oplus c$, where $c \in GF(2)$ is constant.*

Theorem 16. *Let $F = (f_1, \ldots, f_m)$ be an $n \times m$ S-box with an extended immunity order t. Then (i) $t = n$ if and only if F is constant and (ii) $t = n - 1$ if and only if $f_j(x) = x_1 \oplus \cdots \oplus x_n \oplus c_j$ or c_j, where $c_j \in GF(2)$ is constant, $j = 1, \ldots, m$, and there exists a value $j = j_0$ such that $f_{j_0} = x_1 \oplus \cdots \oplus x_n \oplus c_{j_0}$. (iii) For both $t = n$ and $t = n - 1$, the upper bound $t \leq n - \log_2 \#F(GF(2)^n)$ holds with equality.*

According to Theorem 16, the extended immunity of non-constant $n \times m$ S-boxes can be higher than classical resiliency. However, except for $t \leq n - m$, we do not know any other upper bound on the extended resiliency t of (n, m, t)-extended resilient S-boxes. This is an interesting problem to be studied in the future.

10 Conclusions And Future Work

In this paper, we provided different mathematical characterisations of the extended immunity and its special case - the extended resiliency. A characterisation of the extended resiliency (immunity) in terms of the existence of unconditional (conditional) equations is also provided. Relations between the extended immunity, extended resiliency and classical resiliency are examined. Constructions showing that the extended resiliency does not restrict the algebraic degree if the number of inputs is sufficiently large are given too. More efficient constructions and the nonlinearity of extended resilient (immune) S-boxes will be studied in future work. It is also interesting to derive other, possibly sharper bounds on the extended resiliency and extended immunity. Other criteria related to algebraic attacks, such as the minimum degree of algebraic equations induced from S-boxes, in the conditional or unconditional scenarios, can also be taken into consideration.

Acknowledgment

The first two authors were supported by Australian Research Council grants DP0345366 and DP0451484.

References

1. C. H. Bennett, G. Brassard, and J. M. Robert. Privacy amplification by public discussion. *SIAM J. Computing*, 17:210–229, 1988.
2. J. Bierbrauer, K. Gopalakrishnan, and D. R. Stinson. Bounds on resilient functions and orthogonal arrays. In *Advances in Cryptology - CRYPTO '94*, volume 839 of *Lecture Notes in Computer Science*, pages 247–256. Springer-Verlag, Berlin, Heidelberg, New York, 1994.
3. P. Camion, C. Carlet, P. Charpin, and N. Sendrier. On correlation-immune functions. In *Advances in Cryptology - CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 87–100. Springer-Verlag, Berlin, Heidelberg, New York, 1991.
4. C. Carlet. Improving the algebraic immunity of resilient and nonlinear functions and constructing bent functions. (`http://eprint.iacr.org/2004/276/`), 2004.
5. C. Carlet and E. Prouff. Vectorial functions and covering sequences. In *Finite Fields and Applications*, volume 2948 of *Lecture Notes in Computer Science*, pages 215–248. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
6. J. H. Cheon. Nonlinear vector resilient functions. In *Advances in Cryptology - CRYPTO '01*, volume 2139 of *Lecture Notes in Computer Science*, pages 458–469. Springer-Verlag, Berlin, Heidelberg, New York, 2001.
7. J. H. Cheon and D. H. Lee. Resistance of S-boxes against algebraic attacks. In *Proceedings of Fast Software Encryption '04*, volume 3017 of *Lecture Notes in Computer Science*, pages 83–94. Springer-Verlag, Berlin, Heidelberg, New York, 2004.
8. B. Chor, O. Goldreich, J. Håstad, J. Friedman, S. Rudich, and R. Smolensky. The bit extraction problem or t-resilient functions. In *Proc. 26th IEEE Symp. on Foundations of Computer Science*, pages 396–407, 1985.
9. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - CRYPTO '03*, volume 2729 of *Lecture Notes in Computer Science*, pages 176–194. Springer-Verlag, Berlin, Heidelberg, New York, 2003.
10. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - EUROCRYPT '03*, volume 2656 of *Lecture Notes in Computer Science*, pages 345–359. Springer-Verlag, Berlin, Heidelberg, New York, 2003.
11. N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In *Advances in Cryptology - ASIACRYPT '02*, volume 2501 of *Lecture Notes in Computer Science*, pages 267–287. Springer-Verlag, Berlin, Heidelberg, New York, 2002.
12. D. Dalai, K. Gupta, and S. Maitra. Results on algebraic immunity for cryptographically significant Boolean functions. In *Proceedings of INDOCRYPT '04*, volume 3348 of *Lecture Notes in Computer Science*, pages 92–106. Springer-Verlag, Berlin, Heidelberg, New York, 2004.
13. D. Dalai, K. Gupta, and S. Maitra. Results on algebraic immunity for cryptographically significant Boolean functions: Construction and analysis in term of algebraic immunity. In *Proceedings of Fast Software Encryption '05*, volume 3557 of *Lecture Notes in Computer Science*, pages 98–111. Springer-Verlag, Berlin, Heidelberg, New York, 2005.
14. J. Friedman. On the bit extraction problem. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 314–319, 1992.

15. J. Dj. Golić. Vectorial Boolean functions and induced algebraic equations. (http://eprint.iacr.org/2004/225/), 2004.
16. K. C. Gupta and P. Sarkar. Improved construction of nonlinear resilient s-boxes. In *Advances in Cryptology - ASIACRYPT '02*, volume 2501 of *Lecture Notes in Computer Science*, pages 466–483. Springer-Verlag, Berlin, Heidelberg, New York, 2002.
17. M. Hall, Jr. *Combinatorial Theory.* Ginn-Blaisdell, Waltham, 1967.
18. W. Meier, E. Pasalic, and C. Carlet. Algebraic attacks and decomposition of Boolean functions. In *Advances in Cryptology - EUROCRYPT '04*, volume 3027 of *Lecture Notes in Computer Science*, pages 474–491. Springer-Verlag, Berlin, Heidelberg, New York, 2004.
19. M. O'Nan. *Linear Algebra.* Harcourt Brace Jovanovich, New York, 1976.
20. E. Pasalic and S. Maitra. Further constructions of resilient Boolean functions with very high nonlinearity. *IEEE Transactions on Information Theory*, 48(7):1825–1834, 2002.
21. P. Sarkar and S. Maitra. Nonlinearity bounds and constructions of resilient Boolean functions. In *Advances in Cryptology - CRYPTO 2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 515–532. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
22. T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, 30(5):776–779, 1984.
23. D. R. Stinson. Resilient functions and large sets of orthogonal arrays. *Congressus Numerantium*, 92:105–110, 1993.
24. X. M. Zhang and Y. Zheng. Cryptographically resilient functions. *IEEE Transactions on Information Theory*, 43(5):1740–1747, 1997.

Appendix: Proofs of Mathematical Results

Proof of Lemma 1. From the structure of D_F, the ith row vector l_i of D_F is the sequence of linear function $\psi(y) = \langle F(\alpha_i), y\rangle$. On the other hand, the $j$the column of H_m is ℓ_j^T - the sequence of a linear function $\psi_j(y) = \langle \gamma_j, y\rangle$. Note that the sequences of different linear functions are orthogonal. Thus, l_i and ℓ_j will be orthogonal if $F(\alpha_i) \neq \gamma_j$, while l_i and ℓ_j will be identical if $F(\alpha_i) = \gamma_j$. According to the definition of C_F, this proves the lemma. □

Proof of Lemma 2. Due to Lemma 1, we have $H_m D_F^T D_F H_m = 2^{2m} C_F^T C_F$. Due to the definition of C_F, each row of C_F has precisely one nonzero entry. Thus any two columns of C_F are orthogonal. On the other hand, the number of ones in the jth column is equal to the number of times that F takes γ_j. Thus $C_F^T C_F = diag(\lambda_0, \lambda_1, \ldots, \lambda_{2^m-1})$. Summarising the above, we have proved $H_m D_F^T D_F H_m = 2^{2m} diag(\lambda_0, \lambda_1, \ldots, \lambda_{2^m-1})$. Since H_m is a $2^m \times 2^m$ Hadamard matrix, it follows that $D_F^T D_F = H_m diag(\lambda_0, \lambda_1, \ldots, \lambda_{2^m-1}) H_m$. □

Proof of Corollary 1. Since H_m is a $2^m \times 2^m$ Hadamard matrix, due to Lemma 2, we have $D_F^T D_F H_m = 2^m H_m diag(\lambda_0, \lambda_1, \ldots, \lambda_{2^m-1})$ or, equivalently, $D_F^T D_F \ell_j^T = 2^m \lambda_j \ell_j^T$, $j = 0, 1, \ldots, 2^m - 1$. □

Proof of Lemma 3. According to Corollary 1, (i) and (ii) are equivalent. According to the well known fact from linear algebra, (ii) holds if and only if the rank of $D_F^T D_F$ is 2^m. Further, $D_F^T D_F$ and D_F have the same rank. This proves the equivalence of (ii) and (iii). $\square$

Proof of Lemma 4. Assume that (i) holds. Consequently, $P(F(GF(2)^n))$= $P(GF(2)^m)$= $GF(2)^k$. This proves (i) $\Longrightarrow$ (ii). Assume that (ii) holds. Clearly, any non-constant function h on $GF(2)^m$ is a surjective $m \times 1$ S-box. Due to (ii), $h(F)$ is a surjective $n \times 1$ S-box and then non-constant. By virtue of Definition 2, we thus proved (ii) $\Longrightarrow$ (iii). Assume that (iii) holds. We now prove (i) by contradiction. Assume that F does not take a value $\beta \in GF(2)^m$. Define a non-constant function h on $GF(2)^m$ as $h(y) = 1$ if and only if $y = \beta$. Clearly, $h(F(x))$ is the constant zero function. This contradicts (iii), so that (i) is true. This proves (iii) $\Longrightarrow$ (i). $\square$

Proof of Proposition 2. Fix a t-subset T of $\{1, \ldots, n\}$. The 2^t sets $S(n, T, \alpha)$, $\alpha \in GF(2)^t$, are disjoint and partition $GF(2)^n$. When x runs once through $S(n, T, \alpha)$ for any fixed $\alpha \in GF(2)^t$, $F(x)$ runs at least once through each vector in $GF(2)^m$. Accordingly, $F(x)$ runs through each vector in $GF(2)^m$ at least 2^t times when x runs once through $GF(2)^n$. $\square$

Proof of Theorem 1. Let a column vector ζ be the sequence of a Boolean function g on $GF(2)^n$. It is easy to verify that $diag(b_0, b_1, \ldots, b_{2^n-1})\zeta$ is the sequence of the restriction $g(x_1, \ldots, x_n)|_{x_{j_1}=a_1,\ldots,x_{j_t}=a_t}$. For this reason, the matrix, from Notation 1, corresponding to the restriction $F(x_1, \ldots, x_n)|_{x_{j_1}=a_1,\ldots,x_{j_t}=a_t}$ is identical with $diag(b_0, b_1, \ldots, b_{2^n-1})D_F$. The theorem then follows from Lemma 3. $\square$

Proof of Theorem 3. We only need to prove the theorem for $r = t$ because a function g on $GF(2)^r$ with $1 \leq r < t$ can be regarded as a function on $GF(2)^t$ that does not depend on some $t-r$ variables. Assume that F is (n, m, t)-extended resilient. For any fixed t-subset $T = \{j_1, \ldots, j_t\}$ of $\{1, \ldots, n\}$ and any given nonzero Boolean function g on $GF(2)^t$, there exists a vector $\alpha \in GF(2)^t$ satisfying $g(\alpha) = 1$. Then
$\{g(x_{j_1}, \ldots, x_{j_t})F(x)|x \in S(n, T, \alpha)\}$ =$F(S(n, T, \alpha))$. Since F is (n, m, t)-extended resilient, $F(S(n, T, \alpha)) = GF(2)^m$. Thus $g(x_{j_1}, \ldots, x_{j_t})F(x)$ is surjective. Conversely, assume that F satisfies the property from the theorem. We now prove that F is (n, m, t)-extended resilient. For a given t-subset $T = \{j_1, \ldots, j_t\}$ of $\{1, \ldots, n\}$ and any given $\alpha = (a_1, \ldots, a_t) \in GF(2)^t$, we define a nonzero Boolean function g on $GF(2)^t$ such that $g(y) = 1$ if and only if $y = \alpha$. Therefore, $\{g(x_{j_1}, \ldots, x_{j_t})F(x)|x \in GF(2)^n\}$ =$\{g(x_{j_1}, \ldots, x_{j_t})F(x)|x \in S(n, T, \alpha)\}$= $F(S(n, T, \alpha))$. Due to the assumption, $\{g(x_{j_1}, \ldots, x_{j_t})F(x)|x \in GF(2)^n\} = GF(2)^m$. Hence $F(S(n, T, \alpha)) = GF(2)^m$. Since both T with $\#T = t$ and $\alpha \in GF(2)^t$ are arbitrary, we have proved that F is (n, m, t)-extended resilient. $\square$

Proof of Theorem 4. By Definition 3, it is easy to see that (i) $\Longrightarrow$ (ii). Due to Proposition 2, (ii) $\Longrightarrow$ (iii). Assume that (iii) holds. We let $t_0 = t$. Then it follows that F is (n, m, t)-extended resilient. This proves (iii) $\Longrightarrow$ (i). $\square$

Proof of Theorem 7. Assume that (i) holds. For any subset $S(n, T, \alpha)$ of $GF(2)^n$ with $\#T = t$, since F is (n, k, t)-extended immune, $P(F(S(n, T, \alpha))) = P(F(GF(2)^n))$. Thus $P(F(x))$ is (n, k, t)-extended immune. We have thus proved (i) $\Longrightarrow$ (ii). Assume now that (ii) holds. Let h be a function on $GF(2)^m$ such that $h(F)$ is non-constant, i.e., $\#h(F(GF(2)^n)) = 2$. As, due to (ii), $h(F)$ is $(n, 1, t)$-extended immune, for any subset $S(n, T, \alpha)$ of $GF(2)^n$ with $\#T = t$ and $\alpha \in GF(2)^t$, we have $\#h(F(S(n, T, \alpha))$=$\#h(F(GF(2)^n)) = 2$. This means that $h(G)$ is non-constant, where G is defined in the Theorem. We have thus proved that (ii) $\Longrightarrow$ (iii). Finally, assume that (iii) holds. We prove (i) by contradiction. Assume that (i) does not hold. Then there exists a subset $S(n, T, \alpha)$ of $GF(2)^n$ with $\#T = t$ and $\alpha \in GF(2)^t$ such that $F(S(n, T, \alpha)) \neq F(GF(2)^n)$. This implies that F is non-constant. Let $\beta \in F(GF(2)^n)\setminus F(S(n, T, \alpha))$. We choose a non-constant function h on $GF(2)^m$ such that $h(y) = 1$ if and only if $y = \beta$. Since F takes value β and F is non-constant, from the definition of h, it follows that $h(F)$ takes both values 1 and 0. However, since $\beta \notin F(S(n, T, \alpha))$, $h(G)$ is the zero function, where G is defined in the Theorem. This contradicts (iii), so that (i) is true. Thus we have proved (iii) $\Longrightarrow$ (i). □

Proof of Lemma 8. For a fixed t-subset T and any given value of $\beta \in F(GF(2)^n)$, due to (4), there exists a non-constant algebraic equation over x' induced from $F(x) = \beta$ if and only if $\#W(F, T, \beta) < 2^t$. Consequently, there is no non-constant algebraic equation over x' induced from $F(x) = \beta$ if and only if $\#W(F, T, \beta) = 2^t$, or equivalently, for each $x' \in GF(2)^t$, there exists $x'' \in GF(2)^{n-t}$ such that $F(x) = \beta$. Since this is true for an arbitrary $\beta \in F(GF(2)^n)$, it then follows that there is no non-constant algebraic equation over x' induced from $F(x) = \beta$ for any $\beta \in F(GF(2)^n)$ if and only if $F(S(n, T, x')) = F(GF(2)^n)$, for every $x' \in GF(2)^t$. □

Proof of Theorem 10. As a t-subset T is arbitrary, the theorem directly follows from Lemma 8 and Definition 6. □

Proof of Lemma 9. For a fixed t-subset T, due to (2), there is no non-constant algebraic equation over x' and y induced from $F(x) = y$ if and only if $\#W(F, T) = 2^{t+m}$, or equivalently, for each $x' \in GF(2)^t$, $\#F(S(n, T, x')) = 2^m$, that is, $F(S(n, T, x')) = GF(2)^m$, for every $x' \in GF(2)^t$. □

Proof of Theorem 11. As a t-subset T is arbitrary, the theorem directly follows from Lemma 9 and Definition 3. □

Proof of Proposition 5. Let F be an affine (n, m, t)-extended resilient S-box. Let $T = \{j_1, \ldots, j_t\}$ be a subset of $\{1, \ldots, n\}$ and $\alpha = (a_1, \ldots, a_t) \in GF(2)^t$. Due to Theorem 2, all the coordinate functions of $F(x)|_{x_{j_1}=a_1,\ldots,x_{j_t}=a_t}$ are functionally independent and then also linearly independent. Since $F(x)|_{x_{j_1}=a_1,\ldots,x_{j_t}=a_t}$ is affine, due to linear algebra, $F(x)|_{x_{j_1}=a_1,\ldots,x_{j_t}=a_t}$ runs through each vector in $GF(2)^m$ exactly 2^{n-m-t} times while x runs through $S(n, T, \alpha)$ once. This proves that F is (n, m, t)-classically resilient. □

Proof of Theorem 12. Since F is affine, there exists a vector $\beta \in GF(2)^m$ such that $F(x) \oplus \beta$ is linear. By linear algebra, $\#F(GF(2)^n) = 2^k$ for an integer k and

$F(GF(2)^n) \oplus \beta$ is a k-dimensional subspace U of $GF(2)^m$. Therefore, there exists an $m \times k$ matrix B over $GF(2)$ such that $\{\gamma B | \gamma \in U\}$ is identical with $GF(2)^k$. Set $P(x) = (F(x) \oplus \beta)B$. According to Theorem 7, P is an (n, k, t)-extended resilient S-box. On the other hand, $P(GF(2)^n) = \{\gamma B | \gamma \in U\} = GF(2)^k$. Thus P is a linear (n, k, t)-extended resilient S-box. According to Proposition 5, P is an (n, k, t)-classically resilient S-box. □

Proof of Theorem 13. Clearly, the condition $\#Q(GF(2)^n) \geq 2^k$ guarantees the existence of a mapping from $Q(GF(2)^n)$ onto $GF(2)^k$. Accordingly, if P is such a mapping, then $P(Q(GF(2)^n))= GF(2)^k$. For any t-subset T of $\{1, \ldots, n\}$ and any $\alpha \in GF(2)^t$, since Q is (n, m, t)-extended immune, $Q(S(n, T, \alpha)) = Q(GF(2)^n)$. As P is a mapping from $Q(GF(2)^n)$ onto $GF(2)^k$, $P(Q(S(n, T, \alpha)))$ $= P(Q(GF(2)^n))= GF(2)^k$. □

Proof of Lemma 10. We prove the lemma by contradiction. Assume for contradiction that there exists an $n \times m$ S-box $F = (f_1, \ldots, f_m)$ with $m \geq 2$ such that $deg(F) = n$. Then we have $deg(f_1) = deg(f_2) = n$ and hence both f_1 and f_2 contain the product term $x_1 \cdots x_n$. This term cancels out in $f_1 \oplus f_2$ and therefore $deg(f_1 \oplus f_2) < n$. Further, by definition, this implies that $deg(F) < n$, so that we have a contradiction. □

Proof of Lemma 11. The first part is straightforward to verify. Secondly, if $deg(f) < n - 1$, then the term $(x_1 \cdots x_n)/x_j$ remains in the algebraic normal form of f' and hence $deg(f') = n - 1$. □

Proof of Theorem 14. Due to Lemma 11, each f'_j contains exactly one term of degree $n-1$, i.e., $(x_1 \cdots x_n)/x_j$. Since $m \leq n$ and $\{(x_1 \cdots x_n)/x_j \mid 1 \leq j \leq m\}$ are linearly independent Boolean functions, any nonzero linear combination of $f'_1, \ldots, f'_m$ must contain a nonzero linear combination of $(x_1 \cdots x_n)/x_j$, $j = 1, \ldots, m$, that cannot be eliminated. This implies that $deg(F') = n - 1$. For any fixed subset $T_0 = \{j_1, \ldots, j_{t_0}\}$ of $\{1, \ldots, n\}$ and any fixed vector $\alpha_0 \in GF(2)^{t_0}$, due to Theorem 4, $F(x)$ runs through each vector in $GF(2)^m$ at least 2^{t-t_0} times while x runs through $S(n, T_0, \alpha_0)$ once. According to Lemma 11, F' is obtained from F by changing exactly $m + 1$ of its values, $F(\beta_j)$, $j = 1, \ldots, m$, and $F(\beta^*)$. Since $t - t_0 = \lfloor \log_2(m+1) \rfloor + 1$, we have $t - t_0 > \log_2(m+1)$. Thus $2^{t-t_0} - (m+1) > 0$ and hence $F'(x)$ runs through each vector in $GF(2)^m$ at least once while x runs through $S(n, T_0, \alpha_0)$ once, or in other words, $F'(S(n, T_0, \alpha_0)) = GF(2)^m$. □

Proof of Theorem 15. Let P be a permutation on $GF(2)^m$ and let $r \geq 1$. Let $F(x) = P(z_1) \oplus \cdots \oplus P(z_r)$, where $z_1, \ldots, z_r \in GF(2)^m$ and $x = (z_1, \ldots, z_r) \in GF(2)^{rm}$. We first prove that F is $(rm, m, r-1)$-extended resilient with $deg(F) \leq m$. We note that F is a surjective $rm \times m$ S-box. Rewrite x as $x = (x_1, \ldots, x_{rm})$. Choose any subset $T = \{j_1, \ldots, j_{r-1}\}$ of $\{1, \ldots, rm\}$ and any $\alpha \in GF(2)^{r-1}$. Then there must exist an index i, $1 \leq i \leq r$, such that the sets of variables in z_i and $(x_{j_1}, \ldots, x_{j_{r-1}})$ are disjoint. Since P is a permutation on $GF(2)^m$, then we have $GF(2)^m \supseteq F(S(rm, T, \alpha)) \supseteq P(GF(2)^m)= GF(2)^m$. This proves that F is $(rm, m, r-1)$-extended resilient. Due to the construction, the algebraic degree of F cannot exceed m.

In particular, we choose r satisfying $r - 1 \geq t + \lfloor \log_2(m+1) \rfloor + 1$. Let F' be an $rm \times m$ S-box obtained from F as in Theorem 14. Then according to Theorem 14, $deg(F') = rm - 1$ and F' is (rm, m, t_0)-extended resilient with $t_0 = (r-1) - \lfloor \log_2(m+1) \rfloor - 1 \geq t$. By virtue of Lemma 6, F' is then also (rm, m, t)-extended resilient. □

Proof of Lemma 12 . If f is constant, then $t = n$ by the definition of extended immunity. If $t = n$, then the upper bound $t \leq n - \log_2 \#f(GF(2)^n)$ implies that $\#f(GF(2)^n) = 1$, which means that f is constant. This proves (i).

As for (ii), to prove the sufficiency, note that the restriction of f to any set $S(n, T_i, \alpha)$, where $T_i = \{1, \ldots, n\} \setminus \{i\}$ and $\alpha \in GF(2)^{n-1}$, is a non-constant affine function of the remaining variable x_i, so that $f(S(n, T_i, \alpha)) = GF(2)$. This means that $t \geq n - 1$. However, as f is non-constant, (i) implies that $t = n - 1$.

To prove the necessity in (ii), assume that $t = n - 1$. Then, firstly, from (i) it follows that $f(GF(2)^n) = GF(2)$. For any fixed i, $1 \leq i \leq n$, f can be expressed as $f(x) = x_i g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \oplus h(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ where both g and h are Boolean functions on $GF(2)^{n-1}$. We next prove that g is the constant one by contradiction. Assume that there exists some vectotr $\alpha \in GF(2)^{n-1}$ such that $g(\alpha) = 0$. Since h does not depend on x_i, we have $\#f(S(n, T_i, \alpha)) = 1$, where T_i is as above. This contradicts the fact that $f(S(n, T_i, \alpha)) = f(GF(2)^n) = GF(2)$ and hence proves that g is the constant one. Thus $f(x) = x_i \oplus h(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, which means that the algebraic normal form of f contains the linear term x_i and x_i appears as a linear term only. Since this holds for each $i \in \{1, \ldots, n\}$, f can be expressed as $f(x)$=$x_1 \oplus \cdots \oplus x_n$ $\oplus c$, where $c \in GF(2)$ is constant. □

Proof of Theorem 16. The claim (i) is proved in the same way as for $m = 1$ in Lemma 12.

As for (ii), to prove the sufficiency, assume that F has the form specified. Since f_j is either constant or identical to l or $l \oplus 1$, where $l(x) = x_1 \oplus \cdots \oplus x_n$, and f_{j_0} is not constant, it follows that, for each $x \in GF(2)^n$, the value of $f_{j_0}(x)$ uniquely determines the values of the remaining functions $f_j(x)$. Therefore, $\#F(GF(2)^n) = 2$. By the same argument, we also obtain that $\#F(S(n, T, \alpha)) = 2$, for any $(n-1)$-subset T of $\{1, \ldots, n\}$, because the restriction of f_{j_0} to $S(n, T, \alpha)$ is a non-constant affine function of x_i. Consequently, $F(S(n, T, \alpha)) = F(GF(2)^n)$ and $\#F(GF(2)^n) = 2$. Hence, in view of (i), we get $t = n - 1$.

To prove the necessity in (ii), assume that $t = n-1$. In view of (i), this means that F is a non-constant $(n, m, n-1)$-extended immune function. Corollary 4 then implies that each f_j is $(n, 1, n-1)$-extended immune, i.e., has an extended immunity order $n - 1$ or n. According to Lemma 12, each f_j has the form $f_j = x_1 \oplus \cdots \oplus x_n \oplus c_j$ or c_j, where $c_j \in GF(2)$ is constant. Since F is non-constant, there must exist a value $j = j_0$ such that f_{j_0} is non-constant.

Finally, as for (iii), (i) and (ii) imply that for $t = n$ and $t = n - 1$, we have $\#F(GF(2)^n) = 1$ and $\#F(GF(2)^n) = 2$, respectively, so that the upper bound holds with equality in both cases. □