Security of Substitution-Permutation Network

by

Cheng Chen

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

Signature redacted

Author ....
Department of Electrical Engineering and Computer Science
August 26, 2015

Signature redacted

Certified by ....
Professor Vinod Vaikuntanathan
Thesis Supervisor

Signature redacted

Accepted by ....
Professor Leslie Kolodziejski
Chair, Department Committee on Graduate Theses

Security of Substitution-Permutation Network

by

Cheng Chen

Submitted to the Department of Electrical Engineering and Computer Science
on August 26, 2015, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

In this thesis, we study the security of a block cipher design called substitution-permutation network (SPN). We prove that when S-box is chosen uniformly at random as a permutation, the resulting SPN is a strong pseudorandom permutation even against an adversary having oracle access to that S-box. We then examine some special cases of SPN for a fixed S-box and prove two special cases of SPN inspired by AES are 2-wise independent.

Thesis Supervisor: Professor Vinod Vaikuntanathan

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Pr. Vinod Vaikuntanathan, for introducing me to this interesting question and all helpful discussion and guidance.

I would also like to thank my co-advisor Pr. Shafi Goldwasser for continuous support for my study and research.

Last but not the least, I want to thank my family, for their patience, support, and encouragement.

Contents

List of Figures

Chapter 1

Introduction

This thesis takes a new step in closing the gap between pseudorandom permutations and their popular bounded-input counterpart, block ciphers, used in practice.

Pseudorandom permutations are collection of permutations that can not be distinguished from random permutations by any efficient adversary. It was first constructed by Ruby and Rackoff [LR86] from pseudorandom functions, and therefore based on one-way functions. Block ciphers are fixed-length permutations that look random in practice. Implementations of block ciphers, such as Data Encryption Standard (DES) and Advanced Encryption Standard (AES), are widespread all over the Internet to secure sensitive data.

Modern block ciphers often follow the substitution-permutation network (SPN) paradigm, e.g., AES. An SPN is computed over several rounds. In each round, the input is first divided into blocks and a substitution function S-box is applied to each block; a diffusion function is then applied to the entire blocks to spread out the local changes; a round key is then combined to the entire blocks to mask the internal computations. In practice, all the round keys are often derived from a shorter master key for efficiency. But throughout the thesis, we assume that all round keys are independently chosen uniformly at random and are hidden from the adversary.

Despite the popularity of SPN, its security is not well understood. In this thesis, we give some asymptotic analyses of SPN.

1.1 Random S-box

DES was once a predominant encryption scheme. Its design was based on Feistel network. In the seminal work of Luby and Rackoff [LR86], they proved the security of Feistel network: the permutation resulted from Feistel network is a pseudorandom permutation if the underlying primitive used in the construction is itself a pseudorandom function. This theoretical result gave people more confidence in using DES.

Due to its popularity, lots of efforts had been devoted to study the security of DES in practice. Linear cryptanalysis was introduced by Matusi in 1993 [Mat94] as a theoretical attack on DES and later successfully used in the practical cryptanalysis of DES; differential cryptanalysis was first presented by Biham and Shamir in 1990 [BS91] and eventually the details of the attack were packaged into a book. Although the early target of both attacks was DES, the wide applicability of both attacks to numerous other block ciphers has solidified the preeminence of both cryptanalysis techniques in the consideration of the security of all block ciphers [Hey02].

As a replacement of DES, AES was designed from the beginning to resist linear and differential cryptanalysis. It worked pretty well in practice and is still widely used all over the Internet. However, we still lack the understanding of the security of SPN.

Recently, a result of Miles and Viola [MV12] showed a similar result on SPN as that was proved on Feistel network. This is the first asymptotic analysis of SPN. In particular, they showed that the function resulted from Feistel network is a pseudorandom function if the S-box is a pseudorandom function. Though interesting itself, the result has two problems:

- In the design and the use of SPN, both the S-box and the resulted function are supposed to be permutations. Besides the structural requirement, this means that the security of SPN should also consider those adversaries that can ask inversion queries to the resulted function.

- In the design and the use of SPN, S-box is supposed to be a public permutation. In other words, the security of SPN should not rely on hiding the S-box from the adversaries.

In this thesis, we strengthen the result of [MV12] in two ways:

- Both S-box function S and the resulted function F are permutations, and the security of SPN considers those adversaries that can make inversion queries.
- In addition to making queries to F, the adversaries are also allowed to make oracle queries to S (and S^{-1}).

Theorem 1. *(Informal) If S is a truly random permutation, then F is indistinguishable from a random permutation for any (even computationally unbounded) adversaries given polynomially-many oracle access to F, F^{-1}, S, S^{-1}.*

1.2 Fixed S-box

One caveat of the theorem above is that: in reality, the S-box is a public function, while in the theorem above, the adversaries can only have oracle access to S. So a natural question to ask is the security of SPN when S is public, or fixed.

There are several notions of security in cryptography. The standard notion is computationally indistinguishability. That is, a permutation is pseudorandom if no computationally bounded adversary can distinguish the permutation from a truly random one. Another useful notion is almost k-wise independence. In the context of permutation, it means that the outputs of any k distinct inputs are distributed close to uniform. Although there is no obvious linkage between these two notions, one can say a permutation that is almost k-wise independent is secure against any (even computationally unbounded) adversaries making only k non-adaptive queries. By non-adaptive, it means that the adversaries can't make later queries depending on the answers to the queries before. Or in the other word, one can think that the adversaries ask all the queries all together.

The second part of the thesis focuses on the almost k-wise independence of some special cases of SPN with a fixed S box. When the S-box is fixed, the only variable here is the round keys. In each round, an independent uniformly at random round key is picked, and is used to update the internal state of SPN. The problem can then be formulated as a random walk on the graph whose vertices are k-tuples of distinct elements in the input domain of SPN. The number of rounds that is required to make the result of SPN close to k-wise independence can be related to the mixing rate of that graph.

This technique was first introduced by Hoory et al. [HMMR04] to study how well the composition of simple Feistel network round permutations resembles a truly random permutation. And we apply them here on two special cases of SPN. The two cases we study here are inspired by the design of AES. For simplicity, the number of blocks is one and the diffusion function is the identity function. The fixed S-box is the patched field inversion function, as that is used in AES, which has been shown to have good resistance against linear and differential cryptanalysis. The first case works on prime order field and the second case works on characteristic 2 field. We prove that both two cases are almost 2-wise independent after constant rounds.

1.3 Organization

In Chapter 2, we review some preliminaries on algebra and random walks which will be used in the later chapters.

In Chapter 3, we define the main object, the SPN structure.

In Chapter 4, we prove the main theorem that SPN is a strong pseudorandom permutation when S is truly random (and accessible as a black box).

In Chapter 5, we study the case when the S-box is fixed and prove the 2-wise independence of SPN for two special cases of S-boxes inspired by the AES construction.

Chapter 2

Preliminaries

2.1 Algebra

Let $\mathbb{F}$ be a finite field of size p^n. An element $y \in \mathbb{F}$ is called a *quadratic residue* over $\mathbb{F}$ if there exists an element $x \in \mathbb{F}$ such that

$$y = x^2$$

Otherwise, y is called a *quadratic nonresidue* over $\mathbb{F}$.

The *quadratic character* is a function of y defined as follows:

$$\left(\frac{y}{\mathbb{F}}\right) = \begin{cases} 1 & y \text{ is a quadratic residue over } \mathbb{F} \\ 0 & y \text{ is a quadratic nonresidue over } \mathbb{F} \end{cases}$$

It is different from the Legendre symbol since 0 is counted as a residue here.

We also denote the set of quadratic residue and nonresidue over $\mathbb{F}$ by

$$\begin{aligned} QR_{\mathbb{F}} &\triangleq \{y \mid y \text{ is a quadratic residue over } \mathbb{F}\} \\ NQR_{\mathbb{F}} &\triangleq \{y \mid y \text{ is a quadratic nonresidue over } \mathbb{F}\} \end{aligned}$$

The *trace map*, denoted by $Tr(x)$, is defined over $\mathbb{F}$

$$Tr(x) = x + x^{p} + x^{p^2} + \cdots + x^{p^{n-1}}$$

where p is the characteristic of $\mathbb{F}$.

We are especially interested in the case when $p = 2$.

Proposition 2. *Let $\mathbb{F}$ be a finite field of size 2^n and let $Tr(x)$ be the trace map over $\mathbb{F}$. Then*

(1) $Tr(x+y) = Tr(x) + Tr(y)$

(2) $Tr(x) = 0$ or 1. Moreover, both $Tr(x)$ and $Tr(x)+1$ have 2^{n-1} roots in $\mathbb{F}$.

Proof. (1) Let $x, y \in \mathbb{F}$, then

$$\begin{aligned} Tr(x+y) &= (x+y) + (x+y)^2 + (x+y)^4 + \cdots + (x+y)^{2^{n-1}} \\ &= (x+y) + \left(x^2+y^2\right) + \left(x^4+y^4\right) + \cdots + \left(x^{2^{n-1}} + y^{2^{n-1}}\right) \\ &= Tr(x) + Tr(y) \end{aligned}$$

(2) Let $x \in \mathbb{F}$, then

$$\begin{aligned} Tr(x)^2 + Tr(x) &= \left(x^2 + x^4 + x^8 + \ldots + x^{2^n}\right) + \left(x + x^2 + x^4 + \ldots + x^{2^{n-1}}\right) \\ &= x^{2^n} + x = 0 \end{aligned}$$

Therefore $Tr(x) = 0$ or 1. Moreover, both $Tr(x)$ and $Tr(x)+1$ are polynomials of degree 2^{n-1}, therefore both of them have exactly 2^{n-1} roots in $\mathbb{F}$.

2.2 Quadratic Equations over Finite Fields

Let $\mathbb{F}$ be a finite field of size p^n. We want to study the roots of a quadratic polynomial

$$f(x) = ax^2 + bx + c \quad \text{with } a \neq 0 \tag{2.1}$$

In the case $p \neq 2$, Equation (2.1) can be solved by reducing to the square function $u^2 + d$, and thereby to inverting the square map $u \mapsto u^2$.

$$\frac{1}{a} \cdot f(x) = \left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2} \quad \text{with } u = x + \frac{b}{2a} \text{ and } d = \frac{4ac - b^2}{4a^2}$$

Proposition 3. *Let $\mathbb{F}$ be a finite field of odd characteristic, and let $f(x) = ax^2+bx+c$ be a polynomial of degree 2. Then*

(1) f has no roots in $\mathbb{F} \Leftrightarrow b^2 - 4ac \in NQR_{\mathbb{F}}$.

(2) f has one root 0 in $\mathbb{F} \Leftrightarrow b^2 - 4ac = 0$.

(3) f has two roots in $\mathbb{F} \Leftrightarrow b^2 - 4ac \in QR_{\mathbb{F}}\backslash\{0\}$.

The case $p = 2$ is somewhat different. If $b = 0$, f has one root since the square map $u \mapsto u^2$ over finite field of characteristic 2 is a bijective map. If $b \neq 0$, Equation (2.1) can be solved by reducing to the Artin-Schreier polynomial $u^2 + u + d$, and thereby to inverting the Artin-Schreier map $u \mapsto u^2 + u$.

$$\frac{a}{b^2} \cdot f(x) = \left(\frac{a}{b}x\right)^2 + \frac{a}{b}x + \frac{ac}{b^2} \quad \text{with } u = \frac{a}{b}x \text{ and } d = \frac{ac}{b^2}$$

Lemma 4. *[Pom12] Let $\mathbb{F}$ be a finite field with 2^n elements. Then the polynomial $g(u) = u^2 + u + d$ has a root in $\mathbb{F}$ if and only if $Tr(d) = 0$.*

Proof. If g has a root u in $\mathbb{F}$, then

$$Tr(d) = Tr(u^2 + u) = u^{2^n} + u = 0$$

If $Tr(d) = 0$, let $\overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}$ and $u \in \overline{\mathbb{F}}$ be a root of g in $\overline{\mathbb{F}}$, which always exists, then

$$0 = Tr(d) = Tr(u^2 + u) = u^{2^n} + u$$

This implies $u \in \mathbb{F}$ since all 2^n roots of $u^{2^n} + u$ in $\overline{\mathbb{F}}$ are in $\mathbb{F}$. Therefore g has a root in $\mathbb{F}$.

Proposition 5. *[Pom12] Let $\mathbb{F}$ be a finite field of characteristic 2, and let $f(x) = ax^2 + bx + c$ be a polynomial of degree 2. Then*

(1) f has no roots in $\mathbb{F} \Leftrightarrow b \neq 0$ and $Tr\left(\frac{ac}{b^2}\right) = 1$.

(2) f has one root in $\mathbb{F} \Leftrightarrow b = 0$.

(3) f has two roots in $\mathbb{F} \Leftrightarrow b \neq 0$ and $Tr\left(\frac{ac}{b^2}\right) = 0$.

2.3 Random Walks on Graphs

Let $G = (V, E)$ be a weighted directed graph with weights $A \in [0,1]^{V \times V}$ such that $\sum_j A(i,j) = 1$ for all $i \in V$. Consider a *random walk* on G: we start at a node v_0; if at the r-th step we are at a node $v_r = i$, we move to node j with probability a_{ij}. Clearly, the sequence of random nodes $v_0, v_1, v_2, \ldots$ is a Markov chain. The node v_0 might be fixed, or it may itself be drawn from some initial distribution V_0. We denote by V_r the distribution of v_r:

$$V_r(i) = \Pr[v_r = i]$$

The rule of the walk can be expressed by the equation

$$V_r = A^T V_{r-1} = (A^T)^r V_0$$

It follows that the probability that, starting at i, we reach j in r steps is $A^r(i,j)$.

From now on, we focus on graphs where $\sum_i A(i,j) = 1$ for all $j \in V$.

G is *undirected* if the transition matrix A is symmetric: the probability of moving to j, given that we are at node i, is the same as the probability of moving to node i, given that we are at node j.

A distribution π is *stationary* for the graph G if $A^T\pi = \pi$. In particular, the uniform distribution on V is stationary if the graph is regular.

A fundamental result of random walk is that if the graph is strongly connected and non-bipartite, then it has a unique stationary distribution π and, regardless of the initial distribution V_0, the distribution of v_r converges to π as r tends to infinity.

We are especially interested in how fast the random walk converges to the stationary distribution. To start with, we need a way to measure the distance of two distributions. L_2 distance is rather weak in the setting here since it does not require convergence to π everywhere. A strong notion called *total variation distance* (also called *statistical distance*) is often used in the literature.

Definition 6. *(total variation distance) Let P and Q be two probability distributions on a finite space U. The total variation distance between P and Q is*

$$\triangle(P,Q) \triangleq \max_{J \subset U} |X(J) - Y(J)| = \frac{1}{2}\sum_{j \in U} |X(j) - Y(j)|$$

The total variation distance measures the largest possible difference between the probabilities that the two probability distributions can assign to the same event.

After r steps, the total variation distance of the random walk to the stationary distribution π is given by

$$\begin{aligned} \triangle(r) &\triangleq \max_i \max_{J \subset V} \left| \sum_{j \in J} A^r(i,j) - \pi(J) \right| \\ &= \frac{1}{2} \max_i \sum_j |A^r(i,j) - \pi(j)| \end{aligned}$$

The *mixing rate* is a measure of how fast the random walk converges to the stationery distribution. This can be defined as

$$\tau(\epsilon) = \max_{v \in V(G)} \min\{r : \triangle(A^{(r)}(v,\cdot), \pi) < \epsilon\}$$

where $A^{(r)}(v,\cdot)$ is the distribution of the end node after an r-step random walk starting from v.

When the graph G is undirected, the real symmetric transition matrix A is diagonalizable. Let $|\lambda_{|V|}| \leq \cdots \leq |\lambda_2| \leq |\lambda_1| = 1$ be the real eigenvalues of A. A result relating spectral gap, $1 - |\lambda_2|$, to the mixing time is the following:

Theorem 7. *(Theorem 5.1 of [Lov96]) For a random walk starting at any node $i \in V$,* *for any $j \in V$ and $r \geq 1$*

$$|V_r(j) - \pi(j)| \leq \sqrt{\frac{\pi(j)}{\pi(i)}}|\lambda_2|^r$$

More generally, for any $J \subset V$

$$|V_r(J) - \pi(J)| \leq \sqrt{\frac{\pi(J)}{\pi(i)}}|\lambda_2|^r$$

where $V_r(J) = \sum_{j \in J} v_r(j)$.

As a corollary,

$$\triangle(r) = \max_i \max_J \left|V_r(J) - \pi(J)\right| \leq \sqrt{|V|}|\lambda_2|^r$$

Proposition 8.

$$\tau(\epsilon) = O(\frac{1}{1 - |\lambda_2|} \cdot \log(\left|V(G)\right| / \epsilon))$$

The conductance of G is defined as

$$\Phi(G) = \min_{W \subseteq V(G), |W| \leq |W|/2} \frac{\left|E(W, \bar{W})\right|}{d \cdot |W|}$$

where $\bar{W} = V(G) \backslash W$, and $E(W, \bar{W}) = \{(u, v) \in E(G) : u \in W \text{ and } v \notin W\}$. A random walk is lazy if for some constant $\delta > 0$, it holds that $\Pr[v_{t+1} = v \mid v_t = v] \geq \delta$ for all $v \in V(G)$. A fundamental result relating conductance and the mixing time is

Theorem 9. *[SJ89] If the random walk on G is lazy then*

$$\tau(\epsilon) = O(\frac{1}{\Phi^2} \cdot \log(\left|V(G)\right| / \epsilon))$$

2.4 Independence of Permutations

For set U, denote by $U^{\otimes k}$ the set of all k distinct elements in U. For a collection $\mathcal{C} = \{C_s\}_{s \in S}$ of permutations on U, the *output distribution* on $x = (x_1, \ldots, x_k) \in U^{\otimes k}$

is

$$\mathcal{C}(x_1, \ldots, x_k) \triangleq \{(C_s(x_1), \ldots, C_s(x_k)) : s \leftarrow S\}$$

The total variation distance of the output distribution to the uniform distribution is given by

$$\triangle(\mathcal{C}) \triangleq \max_{(x_1, \ldots, x_k) \in U^{\otimes k}} \triangle(\mathcal{C}(x_1, \ldots, x_k), U^{\otimes k})$$

Chapter 3

Substitution-Permutation Network

3.1 Background

A *substitution-permutation network* (SPN) $C_s : \mathbb{F}^m \to \mathbb{F}^m$ is indexed by a key $s = (s_0, \ldots, s_r) \in (\mathbb{F}^m)^{r+1}$, and is specified by the following parameters:

- $r \in \mathbb{N}$, the *number of rounds*
- $\mathbb{F}$, the *working field*
- $S : \mathbb{F} \to \mathbb{F}$, the *S-box*
- $m \in \mathbb{N}$, the *number of S-box invocations per round*
- $M : \mathbb{F}^m \to \mathbb{F}^m$, the *linear transformation*

The *S-box size* is given by $b = \log |\mathbb{F}|$ and the *input/output size* of C_s is $n = mb$ bits.

C_s is computed over r rounds. The ith round is computed over three steps:

(1) m parallel applications of S;

(2) application of M to the entire state;

(3) m parallel field additions with the round key s_i.

On input x, $C_s(x)$ gives $x + s_0$ as input to the first round; the output of round i becomes the input to round $i+1$ (for $1 \leq i < r$), and $C_s(x)$'s output is the output of the rth round.

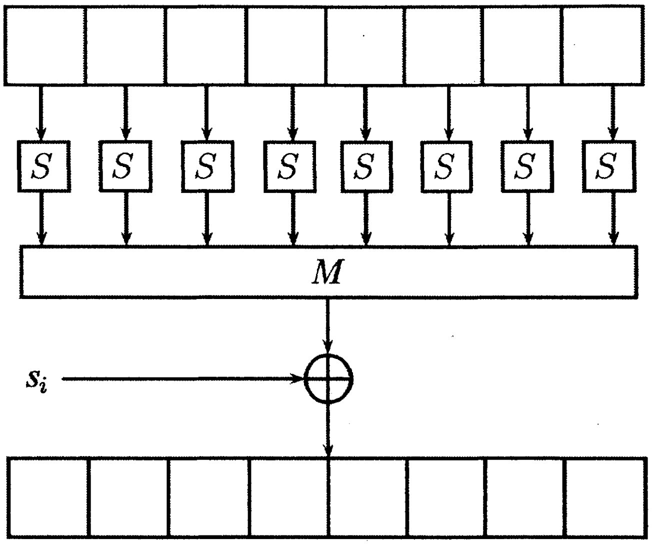

Figure 3-1: One round of an SPN

3.2 Security against linear and differential cryptanalysis

In practice, the security of an SPN is evaluated against two general attacks on block ciphers: linear and differential cryptanalysis. Resistance to these attacks is typically seen as the main security feature of SPNs.

For both attacks, a crucial property in the security proof is that the linear transformation M has maximal *branch number*, defined as follows.

Definition 10. *Let $M : \mathbb{F}^m \to \mathbb{F}^m$ be a linear transformation acting on vectors over a field $\mathbb{F}$. The branch number of M is*

$$Br(M) = \min_{\alpha \neq 0^m} (w(\alpha) + w(M(\alpha))) \leq m + 1$$

where $w(\cdot)$ denotes the number of non-zero elements.

Lemma 11. *[MV12] Let $M : \mathbb{F}^m \to \mathbb{F}^m$ be any matrix with maximal branch number $m + 1$. Then all entries of M are non-zero and M is invertible.*

Proof. Assume for contradiction that $M_{i,j} = 0$ for some $1 \leq i, j \leq m$. Let $\alpha \in \mathbb{F}^m$

be the vector such that $\alpha_j = 1$ and $\alpha_{j'} = 0$ for $j' \neq j$. Then $(M\alpha)_i = 0$, and so $Br(M) \leq w(\alpha) + w(M\alpha) \leq m$.

Assume for contradiction that M is not invertible. Let $\alpha, \beta \in \mathbb{F}^m$ such that $\alpha \neq \beta$ and $M\alpha = M\beta$, then $\alpha + \beta \neq 0$ and $M(\alpha + \beta) = 0^m$, so $Br(M) \leq w(\alpha + \beta) + w(M(\alpha + \beta)) \leq m$. □

Lemma 12. *Let $M : \mathbb{F}^m \to \mathbb{F}^m$ be any matrix with maximal branch number $m + 1$. Then M^{-1} also has maximal branch number.*

Proof. Because M has maximal branch number, both M and M^{-1} are invertible. For any $\alpha \neq 0^m$, $M^{-1}\alpha \neq 0^m$. Then, $w(\alpha) + w(M^{-1}\alpha) = w(M^{-1}\alpha) + w(M(M^{-1}\alpha)) \geq m + 1$. Therefore M^{-1} has maximal branch number. □

Chapter 4

Random S-box

In this chapter, we will prove the security of SPNs with a random S-box. Our candidate F_1 is a 2-round SPN when S is a truly random permutation. An adversary can make oracle queries to F_1, F_1^{-1}, S, S^{-1}. We make the assumption here that all of them are only accessible as black boxes. We show that any adversary A has small advantage in distinguishing F_1 from a random permutation F.

Theorem 13. *If computationally unbounded A makes at most q total queries to its oracles, then*

$$\left| \Pr_{S,F_1}[A^{S,S^{-1},F_1,F_1^{-1}} = 1] - \Pr_{S,F}[A^{S,S^{-1},F,F^{-1}} = 1] \right| < O(m^2q^2) \cdot 2^{-b}$$

We note that Miles and Viola [MV12] proved that F_1 is a pseudorandom function when S is a truly random permutation. Theorem 13 strengthens their result in two ways:

- Instead of only given access to F_1 (or F), the adversary can also make queries to S and S^{-1}. In practice, S is public, and therefore proving the security against adversaries that has access to S and S^{-1} makes more sense.

- Instead of only given access to F_1 (or F), the adversary can make inverse queries to F_1^{-1}. Therefore, we prove that F_1 is strong pseudorandom permutation. This can be interpreted as secure against adaptive chosen-ciphertext attack, where the adversary has the additional power to ask for the decryption of ciphertexts of his choice.

The bound achieved here is similar to the classical result of Luby and Rackoff [LR86] on Feistel network in the sense that the advantage is exponentially small in the size of the random permutation, with a polynomial loss in the number of queries. The proof of the theorem follows the framework of Naor and Reingold [NR99].

Our analysis in this chapter holds for SPNs in which the linear transformation M has maximal branch number, which is also essential in the security proof against linear and differential cryptanalysis.

For the remainder of this section, fix any $M \in \mathbb{F}^{m \times m}$ with maximal branch number. For any permutation $S : \mathbb{F} \to \mathbb{F}$ and any set of round keys $(s_0, s_1, s_2) \in (\mathbb{F}^m)^3$, let $F_1 = F_1(S, s_0, s_1, s_2)$ be the 2-round SPN defined by these components. We will show that F_1 is strong pseudorandom permutation, i.e. Theorem 13.

Similar to [MV12], the proof proceeds in two stages. In the first stage, we prove the theorem against non-adaptive adversaries. This is equivalent to saying that F_1 is almost q-wise independent. In particular, we show that with all but negligible probability, $\{S(u_1), \ldots, S(u_{q_1}), F_1(x_1), \ldots, F_1(x_{q_2})\}$ is uniformly distributed.

In the second stage, we follow the framework of [NR99]. In particular, we consider the distribution over transcripts of A's interaction with its oracle, and uses the result of the first stage in a probability argument to show that the transcripts are distributed nearly identical in either setting, and thus that A's distinguishing advantage is small.

Note that 2-round SPN can be written as

$$y = M \cdot S^*(M \cdot S^*(x + s_0) + s_1) + s_2$$

where for any $x = (x[1], \ldots, x[m])$ we define $S^*(x) = (S(x[1]), \ldots, S(x[m]))$. Then,

$$M^{-1}(y - s_2) = S^*(M \cdot S^*(x + s_0) + s_1)$$

4.1 Stage 1: the non-adaptive case

Fix any distinct $u_1, \ldots, u_{q_1} \in \mathbb{F}$, any $v_1, \ldots, v_{q_1} \in \mathbb{F}$, any distinct $x_1, \ldots, x_{q_2} \in \mathbb{F}^m$ and $y_1, \ldots, y_{q_2} \in \mathbb{F}^m$.

Let D_0 be the uniform distribution on (S, s_0, s_1, s_2). Consider another distribution D_1 on (S, s_0, s_1, s_2):

1. Uniformly choose the output of S on $u_1, \ldots, u_{q_1}$.

2. Uniformly choose s_0, s_1.

3. Compute $a_i = x_i + s_0$ and $b_i = M \cdot S^*(a_i) + s_1$ for $1 \le i \le q_2$, and each time the S-box is evaluated on a previously-unseen input, choose the output uniformly. Let H be the set of at most mq_2 S-inputs whose output is determined after this step.

4. Uniformly choose s_2 and compute $c_i = M^{-1}(y_i - s_2)$ for $1 \le i \le q_2$.

5. Compute $S^*(b_i)$ for $1 \le i \le q_2$, and each time the S-box is evaluated on a previously-unseen input, choose the output uniformly.

6. Uniformly choose the output of S on all remaining inputs from all remaining outputs such that different inputs have different outputs.

Conditioned on the event that different inputs of S have different outputs, i.e. S is a permutation, the above distribution is uniform. By union bound, the statistical distance between these two distributions is bounded by

$$\triangle(D_0, D_1) \le \binom{q_1 + 2mq_2}{2} \cdot 2^{-b} = O(m^2 q^2) \cdot 2^{-b}$$

where $q = q_1 + q_2$.

We now define several bad events. The idea is that when none of these bad events happen, all blocks of b_i and u_j's make fresh evaluations of S, i.e. they don't collide either with each other or with other evaluations.

- BAD_1: $\exists i, l, j$ s.t. $a_i[l] = u_j$.

- BAD_2: $\exists i, l, j$ s.t. $b_i[l] = u_j$.

- BAD_3: $\exists i, l, j, l'$ s.t. $b_i[l] = a_j[l']$.

- BAD_4: $\exists i, l, l'$ s.t. $l \neq l'$ and $b_i[l] = b_i[l']$.

- BAD_5: $\exists i, l, i', l'$ s.t. $i \neq i'$ and $b_i[l] = b_j[l']$.

We then bound each of these bad events.

- By union bound, $\Pr[BAD_1] \leq q_1 \cdot mq_2 \cdot 2^{-b} = O(mq^2) \cdot 2^{-b}$.

- By union bound, $\Pr[BAD_2] \leq q_1 \cdot mq_2 \cdot 2^{-b} = O(mq^2) \cdot 2^{-b}$.

- By union bound, $\Pr[BAD_3] \leq mq_2 \cdot mq_2 \cdot 2^{-b} = O(m^2q^2) \cdot 2^{-b}$.

- By union bound, $\Pr[BAD_4] < mq_2 \cdot m \cdot 2^{-b} = O(m^2q) \cdot 2^{-b}$.

- By definition, $b_i[l] = b_j[l']$ iff

  $$s_1[l] + \sum_k M[l, k] \cdot S(x_i[k] + s_0[k]) = s_1[l'] + \sum_k M[l', k] \cdot S(x_j[k] + s_0[k])$$

  Let k be such that $x_i[k] \neq x_j[k]$, which must exist because $x_i \neq x_j$. Arbitrarily fix $s_1[l]$, $s_1[l']$, $s_0[k']$ for $k' \neq k$ and the outputs of S on the input sets $I = \{x_i[k'] + s_0[k'], x_j[k'] + s_0[k']\}_{k' \neq k}$. This fixes an $\alpha \in \mathbb{F}$ such that

  $$M[l, k] \cdot S(x_i[k] + s_0[k]) - M[l', k] \cdot S(x_j[k] + s_0[k]) = \alpha$$

  If neither $x_i[k] + s_0[k]$ nor $x_j[k] + s_0[k]$ are in $\{u_1, \ldots, u_{q_1}\} \cup H \cup I$, then the probability it holds is 2^{-b} over the choice of S on these two inputs because all entries of M are non-zero. Further, by union bound, these two inputs fall inside $\{u_1, \ldots, u_{q_1}\} \cup H \cup I$ with probability at most $2 \cdot (q_1 + mq_2 + 2(m-1)) \cdot 2^{-b} = O(mq) \cdot 2^{-b}$. Thus, $\Pr[BAD_5] \leq 2^{-b} + O(mq) \cdot 2^{-b} = O(mq) \cdot 2^{-b}$.

Finally, we have

$$\begin{aligned}
\Pr[BAD] &= \Pr[BAD_1 \cup BAD_2 \cup BAD_3 \cup BAD_4 \cup BAD_5] \\
&\leq \Pr[BAD_1] + \Pr[BAD_2] + \Pr[BAD_3] + \Pr[BAD_4] + \Pr[BAD_5] \\
&= O(mq^2) \cdot 2^{-b} + O(mq^2) \cdot 2^{-b} + O(m^2q^2) \cdot 2^{-b} + O(m^2q) \cdot 2^{-b} + O(mq) \cdot 2^{-b} \\
&= O(m^2q^2) \cdot 2^{-b}
\end{aligned}$$

Conditioned on $\neg BAD$, the evaluations of S on $u_1, \ldots, u_{q_1}$ and S^* on $b_1, \ldots, b_{q_1}$ are fresh. Then

$$\begin{aligned}
& \Pr_{D_1}[S(u_1) = v_1 \cap \cdots \cap S(u_{q_1}) = v_{q_1} \cap F_1(x_1) = y_1 \cap \cdots \cap F_1(x_{q_2}) = y_{q_2} \mid \neg BAD] \\
= \; & \Pr[S(u_1) = v_1 \mid \neg BAD] \times \cdots \times \Pr[S(u_{q_1}) = v_{q_1} \mid \neg BAD] \\
& \times \Pr[S^*(b_1) = c_1 \mid \neg BAD] \times \cdots \times \Pr[S^*(b_{q_2}) = c_2 \mid \neg BAD] \\
= \; & 2^{-(q_1+mq_2)b}
\end{aligned}$$

Therefore, conditioned on $\neg BAD$, and S and F_1 are permutations, i.e. different inputs have different outputs, D_1 is the truly random permutation. Thus, we have

$$\begin{aligned}
\triangle(D_0, U) &\leq \triangle(D_0, D_1) + \triangle(D_1, U) \\
&\leq O(m^2q^2) \cdot 2^{-b} + \Pr_{D_1}[BAD] + \Pr_{D_1}[S \text{ and } F_1 \text{ are not permutations}] \\
&\leq O(m^2q^2) \cdot 2^{-b} + O(m^2q^2) \cdot 2^{-b} + \binom{q_1 + mq_2}{2} \cdot 2^{-b} \\
&= O(m^2q^2) \cdot 2^{-b}
\end{aligned}$$

Therefore, F_1 is $O(m^2q^2) \cdot 2^{-b}$-close to q-independent, and any non-adaptive adversary has only $O(m^2q^2) \cdot 2^{-b}$ advantage distinguishing F_1 from F given oracle access either to S, S^{-1}, F_1, F_1^{-1} or to S, S^{-1}, F, F^{-1}.

4.2 Stage 2: the adaptive case

We now show that even adversaries that make adaptive queries have small distinguishing advantage, i.e. we prove Theorem 13.

Let S, S^{-1}, P, P^{-1} be the oracle that is accessible to the adversary A (P is either F_1 or F). There are 4 types of queries A can make: (S, u) for "what is $S(u)$", (S^{-1}, v) for "what is $S^{-1}(v)$", (P, x) for "what is $P(x)$", and (P^{-1}, y) for "what is $P^{-1}(y)$". For the ith query A makes, define the query/answer pair $\langle s, t\rangle$: $\langle s, t\rangle = \langle u, v\rangle$ if either A's query is (S, u) and the answer is v or A's query is (S^{-1}, v) and the answer is u, and $\langle s, t\rangle = \langle x, y\rangle$ if either A's query is (P, x) and the answer is y or A's query is (P^{-1}, y) and the answer is x. We assume that A makes exactly q queries and refer to the sequence $\{\langle s_1, t_1\rangle, \ldots, \langle s_q, t_q\rangle\}$ of these pairs as the *transcript* of A's computation.

Because A is computationally unbounded, we can make the standard assumption that A is deterministic by fixing the random tape that maximizes the advantage of A. This implies that the transcript of A interacting with S, S^{-1}, P, P^{-1} is determined given S, P. Let $T_{S,P}$ be that transcript.

For a transcript σ, denote its prefix by $\sigma_i = \{\langle s_1, t_1\rangle, \ldots, \langle s_i, t_i\rangle\}$. A transcript is said to be *possible* if for every $i < q$, A's next query is (S, u) or (S^{-1}, v) if $\langle s_{i+1}, t_{i+1}\rangle = \langle u, v\rangle$, and (P, x) or (P^{-1}, y) if $\langle s_{i+1}, t_{i+1}\rangle = \langle x, y\rangle$.

We can further assume that A never asks a query if its answer is determined by a previous query/answer pair. That is, for $i \neq j$ both $s_i \neq s_j$ and $t_i \neq t_j$.

Let D_3 be the uniform distribution on (S, F). That is, S is a uniform random permutation on $\mathbb{F}$ and F is a uniform random permutation on $\mathbb{F}^m$. Consider yet another distribution D_2 on (S, F):

1. On the ith query of A

   (a) If A's query is (S, u) and for some $1 \leq j < i$ the jth query/answer pair is (u, v), then answer v.

   (b) If A's query is (S^{-1}, v) and for some $1 \leq j < i$ the jth query/answer pair is (u, v), then answer u.

(c) If A's query is (F, x) and for some $1 \leq j < i$ the jth query/answer pair is (x, y), then answer y.

(d) If A's query is (F^{-1}, y) and for some $1 \leq j < i$ the jth query/answer pair is (x, y), then answer x.

(e) Otherwise, answer uniformly random from $\mathbb{F}$ if $A's$ query is (S, u) or (S^{-1}, v), or uniformly random from $\mathbb{F}^m$ if A's query is (F, x) or (F^{-1}, y).

2. Uniformly choose the output of S and F on all remaining inputs from all remaining outputs such that different inputs have different outputs.

Conditioned on the event that different inputs of S and F have different outputs, i.e. S and F are permutations, the above distribution is uniform. By union bound, for any computationally unbounded adaptive adversary making at most q oracle queries, we have

$$\begin{aligned} & \left|\Pr_{D_2}[A^{S,S^{-1},F,F^{-1}} = 1] - \Pr_{D_3}[A^{S,S^{-1},F,F^{-1}} = 1]\right| \\ \leq & \binom{q}{2} \cdot 2^{-b} + \binom{q}{2} \cdot 2^{-mb} \\ = & O(q^2) \cdot 2^{-b} \end{aligned}$$

For any distinct $u_1, \ldots, u_{q_1} \in \mathbb{F}$, any $v_1, \ldots, v_{q_1} \in \mathbb{F}$, any distinct $x_1, \ldots, x_{q_2} \in \mathbb{F}^m$ and $y_1, \ldots, y_{q_2} \in \mathbb{F}^m$, we have

$$\begin{aligned} & \Pr_{D_2}[S(u_1) = v_1 \cap \cdots \cap S(u_{q_1}) = v_{q_1} \cap F_1(x_1) = y_1 \cap \cdots \cap F_1(x_{q_2}) = y_{q_2}] \\ = & \Pr_{D_2}[S(u_1) = v_1] \times \cdots \times \Pr_{D_2}[S(u_{q_1}) = v_{q_1}] \times \Pr_{D_2}[S^*(b_1) = c_1] \times \cdots \times \Pr_{D_2}[S^*(b_{q_2}) = c_2] \\ = & 2^{-(q_1 + mq_2)b} \\ = & \Pr_{D_1}[S(u_1) = v_1 \cap \cdots \cap S(u_{q_1}) = v_{q_1} \cap F_1(x_1) = y_1 \cap \cdots \cap F_1(x_{q_2}) = y_{q_2} \mid \neg BAD] \end{aligned}$$

Let Γ be the set of all possible transcripts such that $A(\sigma) = 1$. Then,

$$\begin{aligned}
& \left|\Pr_{D_1}[A^{S,S^{-1},F_1,F_1^{-1}} = 1] - \Pr_{D_2}[A^{S,S^{-1},F,F^{-1}} = 1]\right| \\
= & \left|\sum_{\sigma\in\Gamma}(\Pr_{D_1}[T_{S,F_1} = \sigma] - \Pr_{D_2}[T_{S,F} = \sigma])\right| \\
\leq & \left|\sum_{\sigma\in\Gamma}\Pr_{D_1}[BAD]\cdot(\Pr_{D_1}[T_{S,F_1} = \sigma \mid BAD] - \Pr_{D_2}[T_{S,F} = \sigma])\right| \\
& + \left|\sum_{\sigma\in\Gamma}\Pr_{D_1}[\neg BAD](\Pr_{D_1}[T_{S,F_1} = \sigma \mid \neg BAD] - \Pr_{D_2}[T_{S,F} = \sigma])\right| \\
= & \left|\sum_{\sigma\in\Gamma}\Pr_{D_1}[BAD]\cdot(\Pr_{D_1}[T_{S,F_1} = \sigma \mid BAD] - \Pr_{D_2}[T_{S,F} = \sigma])\right| \\
\leq & (\max_{\sigma\in\Gamma}\Pr_{D_1}[BAD])\cdot\left|\sum_{\sigma\in\Gamma}(\Pr_{D_1}[T_{S,F_1} = \sigma \mid BAD] - \Pr_{D_2}[T_{S,F} = \sigma])\right| \\
\leq & 2\max_{\sigma\in\Gamma}\Pr_{D_1}[BAD] = O(m^2q^2)\cdot 2^{-b}
\end{aligned}$$

By union bound, we have

$$\begin{aligned}
& \left|\Pr_{D_0}[A^{S,S^{-1},F_1,F_1^{-1}} = 1] - \Pr_{D_3}[A^{S,S^{-1},F,F^{-1}} = 1]\right| \\
\leq & \left|\Pr_{D_0}[A^{S,S^{-1},F_1,F_1^{-1}} = 1] - \Pr_{D_1}[A^{S,S^{-1},F_1,F_1^{-1}} = 1]\right| \\
& + \left|\Pr_{D_1}[A^{S,S^{-1},F_1,F_1^{-1}} = 1] - \Pr_{D_2}[A^{S,S^{-1},F,F^{-1}} = 1]\right| \\
& + \left|\Pr_{D_2}[A^{S,S^{-1},F,F^{-1}} = 1] - \Pr_{D_3}[A^{S,S^{-1},F,F^{-1}} = 1]\right| \\
= & O(m^2q^2)\cdot 2^{-b} + O(m^2q^2)\cdot 2^{-b} + O(q^2)\cdot 2^{-b} \\
= & O(m^2q^2)\cdot 2^{-b}
\end{aligned}$$

This completes the proof of Theorem 13.

Corollary 14. *If S is strong pseudorandom permutation, then F_1 is strong pseudorandom permutation against any PPT adversary A.*

The above corollary is a straightforward application of hybrid argument on Theorem 13.

Chapter 5

Fixed S-box

In this chapter, we will study the security of SPNs with a fixed S-box. In particular, we study the k-wise independence of SPN, or equivalently, security against computationally unbounded non-adaptive adversaries that can make at most k queries to the function.

We formulate this question as one of the mixing rate of a random walk on the graph whose vertices are k-tuples of distinct elements in $\mathbb{F}^m$, and whose edges are induced by the operation of one SPN round permutation on the vertices. The mixing rate of this graph is exactly that minimal round of SPN we seek.

Definition 15. *(SPN graph) A k-wise SPN graph $G_k = (V, E)$ is a 2^{mb}-regular directed graph such that*

- *The vertex set V contains every k distinct elements in $\mathbb{F}^m$, i.e. $V = (\mathbb{F}^m)^{\otimes k}$.*
- *For every $s \in \mathbb{F}^m$, there is an edge from x to x' if $M(S^*(x)) + s = x'$.*

In order to prove that the random walk converges to the uniform distribution, we need to argue that G_k is strongly connected and non-bipartite. As it turns out, bipartite is not an absolute obstacle for the convergence of random walk. We can modify our random walk by using a lazy random walk, which we stay at the current node with probability 1/2. The transition matrix of the lazy walk is

$$A' = \frac{A + I}{2}$$

In our setting, one can imagine that the family of permutations is modified to $\{C_{s,t}\}$ with $t = \{0,1\}^r$. At round i, the round permutation P is applied to the current state if and only if $t_i = 1$.

5.1 Special cases

In this section we consider two special cases of SPNs and prove their 2-wise independence. Our two candidates F_2 and F_3 work on prime field $\mathbb{F}_p$ and characteristic-2 field $\mathbb{F}_{2^n}$ respectively, with only single block, the linear transformation M is the identical matrix and S-box is the patched inversion function.

$$S(x) = \begin{cases} 0 & x = 0 \\ 1/x & x \neq 0 \end{cases}$$

The choice of patched inversion function as S-box is motivated by the design of AES, and it is differentially uniform and has high degree.

When the reduced SPN graph $G = (V, E)$ is undirected, we have $\tau(\epsilon) = O((1 - |\lambda_2|)^{-1} \cdot \log(|V|/\epsilon))$, and thus it suffices to bound λ_2.

It is generally hard to compute $|\lambda_2|$ exactly. Fortunately, a result of Chung, Graham and Wilson [CGW89] relates $|\lambda_2|$ to the number of 4-cycles. We generalize their result to the weighted undirected graph case.

Theorem 16. *Suppose $A(i,j) = o(1/\sqrt{|V|})$. If for any fixed $i, k \in V$,*

$$\left| \sum_j A(i,j)A(j,k) - \frac{1}{|V|} \right| = o\left(\frac{1}{|V|}\right)$$

Then $|\lambda_2| = o(1)$.

Proof. On one hand,

$$tr(A^4) = \sum_i \lambda_i^4 \geq 1 + |\lambda_2|^4$$

On the other hand,

$$\begin{aligned} tr(A^4) &= \sum_i A^4(i,i) = \sum_{i,j,k,l} A(i,j)A(j,k)A(k,l)A(l,i) \\ &= \sum_i A(i,i)^4 + 2\sum_{i<j} A(i,j)^4 + 2\sum_{i<k}(\sum_j A(i,j)A(j,k))^2 \\ &= o\left(\frac{1}{|V|}\right) + o(1) + (1+o(1)) = 1 + o(1) \end{aligned}$$

Combining the two facts above,

$$|\lambda_2| = o(1)$$

□

A weighted undirected graph is said to be *quasi-random* if it satisfies the conditions in Theorem 16. As a corollary, a quasi-random graph is connected and non-bipartit, and has rapid mixing time $\tau(\epsilon) = O(\log(|V|/\epsilon))$.

Before going to the these two cases, we observe that each step of a random walk in SPN graph can be think of has two stages: on node (x_1, x_2), it first chooses s, and jumps to $(x_1 + s, x_2 + s)$, then it goes to $(M(S(x_1+s)), M(S(x_2+s)))$. Therefore we can view the SPN graph as having $|\mathbb{F}|$ clusters. The cluster u is the set of all pairs of inputs (x_1, x_2) such that $x_2 - x_1 = u$. The random walk first randomly pick a node from the cluster, then follow the edge defined by S and M. So it is equivalent to consider the random walk in the reduced SPN graph as defined below.

Definition 17. *(reduced SPN graph) A 2-wise reduced SPN graph $RG_k = (V, E)$ is weighted directed graph that*

- *The vertex set V contains elements in $\mathbb{F}\backslash 0$.*

- *For every two nodes u, u', the weight of the edge from u to u' is*

$$\Pr[s \leftarrow \mathbb{F} : (x+u)^{-1} - x^{-1} = u']$$

When $M(S(\cdot))$ is itself's inverse, as in the case when S is patched inversion function and M is the identity function, the reduced graph is undirected.

5.1.1 $\mathbb{F} = \mathbb{F}_p$, $m = 1$, $S(x) = 1/x$, $M(x) = x$ and $k = 2$

We have

$$A(u, u') = \Pr[x \leftarrow \mathbb{F}_p : (x + u)^{-1} - x_1^{-1} = u']$$

When $x = 0, -u$, it holds that $uu' = 1$.

When $x \neq 0, -u$, the equation is equivalent to

$$\frac{4}{u^2}\left(x_1 + \frac{u}{2}\right)^2 = 1 - \frac{4}{uu'}$$

which has two solutions if $1 - 4/uu' \in QR_{\mathbb{F}}\backslash\{0\}$, one solution if $uu' = 4$, and zero solution otherwise.

Lemma 18. *As a corollary of the law of quadratic reciprocity, $-3 \in QR_{\mathbb{F}_p}$ if and only if $p \bmod 3 = 1$.*

Combining the facts above, we have

$$A(u, u') = \begin{cases} 4/p & uu' = 1 \wedge p \bmod 3 = 1 \\ 2/p & (uu' = 1 \wedge p \bmod 3 = 2) \vee (1 - 4/uu' \in QR_{\mathbb{F}_p}\backslash\{0, -3\}) \\ 1/p & uu' = 4 \\ 0 & 1 - 4/uu' \in NQR_{\mathbb{F}_p}\backslash\{-3\} \end{cases}$$

We then prove that the graph is quasi-random. In order to do this, we need the following theorem on the distribution of quadratic residues and nonresidues over $\mathbb{F}_p$.

Theorem 19. *[Per92] Let p be a prime number and $a \in \mathbb{F}_p$. Define the joint distribution of the quadratic characters of $(x, x + a)$ for randomly chosen x as*

$$\left\{(y_1, y_2) : x \leftarrow \mathbb{F}_p, y_1 \leftarrow \left(\frac{x}{\mathbb{F}_p}\right), y_2 \leftarrow \left(\frac{x+a}{\mathbb{F}_p}\right)\right\}$$

We have for all $b_1, b_2 \in \{0,1\}$,

$$\left|\Pr[y_1 = b_1 \wedge y_2 = b_2] - \frac{1}{4}\right| \leq \frac{3+\sqrt{p}}{p}$$

Now for any fixed $i, k \in V = \mathbb{F}_p \backslash \{0\}$, we have

$$\begin{aligned}
& \sum_j A(i,j)A(j,k) \\
= \; & \sum_{j:ij\in\{1,4\}\vee jk\in\{1,4\}} A(i,j)A(j,k) + \sum_{j:1-4/ij,1-4/jk\in QR_{\mathbb{F}_p}\backslash\{0,-3\}} \frac{4}{p^2} \\
= \; & O\left(\frac{1}{p^2}\right) + \frac{4}{p^2}\cdot p \cdot \Pr\left[j \leftarrow \mathbb{F}_p : j \neq 0 \wedge ij, jk \notin \{1,4\} \wedge \left(\frac{1-4/ij}{\mathbb{F}}\right) = 1 \wedge \left(\frac{1-4/jk}{\mathbb{F}}\right) = 1\right] \\
= \; & O\left(\frac{1}{p^2}\right) + \frac{4}{p}\cdot \Pr\left[j \leftarrow \mathbb{F}_p : j \neq 0 \wedge \left(\frac{1/j - i/4}{\mathbb{F}}\right) = \left(\frac{-i/4}{\mathbb{F}}\right) \wedge \left(\frac{1/j - k/4}{\mathbb{F}}\right) = \left(\frac{-k/4}{\mathbb{F}}\right)\right] \\
= \; & O\left(\frac{1}{p^2}\right) + \frac{4}{p}\cdot \left(\frac{1}{4} \pm O\left(\frac{1}{\sqrt{p}}\right)\right) = \frac{1}{p-1} \pm o\left(\frac{1}{p-1}\right)
\end{aligned}$$

as desired.

5.1.2 $\mathbb{F} = \mathbb{F}_{2^n}$, $m = 1$, $S(x) = 1/x$, $M(x) = x$ and $k = 2$

We have

$$A(u,u') = \Pr[x \leftarrow \mathbb{F}_{2^n} : (x+u)^{-1} - x^{-1} = u']$$

When $x = 0, u,$, it holds that $uu' = 1$.

When $x \neq 0, u$, the equation is equivalent to

$$\frac{x}{u}(\frac{x}{u} + 1) = \frac{1}{uu'}$$

which has two solutions if $Tr(1/uu') = 0$ and zero solution otherwise.

Note $Tr(1) = 1 + 1^2 + 1^4 + \cdots + 1^{2^{n-1}} = n \bmod 2$.

Combining the facts above, we have

$$A(u,u') = \begin{cases} 4/2^n & uu' = 1 \wedge n \bmod 2 = 0 \\ 2/2^n & (uu' = 1 \wedge n \bmod 2 = 1) \vee (uu' \neq 1 \wedge Tr(1/uu') = 0) \\ 0 & uu' \neq 1 \wedge Tr(1/uu') = 1 \end{cases}$$

We then prove that the graph is quasi-random. In order to do this, we need the following properties of trace.

Lemma 20. *For any $a, b \in \mathbb{F}_{2^n} \backslash \{0\}$ with $a \neq b$, we have*

$$\Pr[x \leftarrow \mathbb{F}_{2^n} : Tr(ax) = Tr(bx) = 0] = \frac{1}{4}$$

Proof. The trace map is linear over $\mathbb{F}_2$, so the solution to $Tr(ax) = 0$ forms a vector space of dimension $n-1$ over $\mathbb{F}_2$. More specifically, think of the canonical mapping from $\mathbb{F}_{2^n}$ to $(\mathbb{F}_2)^n$, $x \in \mathbb{F}_{2^n}$ is a solution to $Tr(ax) = 0$ if and only if the image $(x_0, \ldots, x_{n-1}) \in (\mathbb{F}_2)^n$ satisfies $(Tr(a), Tr(2a), \ldots, Tr(2^{n-1}a))^T(x_1, \ldots, x_n) = 0$. $(Tr(a), Tr(2a), \ldots, Tr(2^{n-1}a))$ and $(Tr(b), Tr(2b), \ldots, Tr(2^{n-1}b))$ are linear independent when $a \neq b$. Therefore the joint solution forms a vector space of dimension $n-2$. □

Now for any fixed $i, k \in V = \mathbb{F}_{2^n} \backslash \{0\}$, we have

$$\begin{aligned} & \sum_j A(i,j)A(j,k) \\ = & \sum_{j: ij=1 \vee jk=1} A(i,j)A(j,k) + \sum_{j: ij \neq 1 \wedge jk \neq 1 \wedge Tr(1/ij) = Tr(1/jk) = 0} \frac{4}{2^{2n}} \\ = & O\left(\frac{1}{2^{2n}}\right) + \frac{4}{2^n} \cdot 2^{2n} \cdot \Pr\left[j \leftarrow \mathbb{F}_{2^n} : j \neq 0 \wedge ij \neq 1 \wedge jk \neq 1 \wedge Tr(1/ij) = Tr(1/jk) = 0\right] \\ = & O\left(\frac{1}{2^{2n}}\right) + \frac{4}{2^n} \cdot \Pr\left[j \leftarrow \mathbb{F}_{2^n} : j \neq 0 \wedge Tr(1/ij) = Tr(1/jk) = 0\right] \\ = & O\left(\frac{1}{2^{2n}}\right) + \frac{4}{2^n} \cdot \left(\frac{1}{4} - \frac{1}{2^n}\right) = \frac{1}{2^n - 1} \pm o\left(\frac{1}{2^n - 1}\right) \end{aligned}$$

as desired.

Bibliography

[BS91] Eli Biham and Adi Shamir. Differential cryptanalysis of des-like cryptosystems. In *Proceedings of the 10th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '90, pages 2–21, London, UK, UK, 1991. Springer-Verlag.

[CGW89] Fan R. K. Chung, Ronald L. Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[Hey02] Howard M. Heys. A tutorial on linear and differential cryptanalysis. *Cryptologia*, 26(3):189–221, July 2002.

[HMMR04] Shlomo Hoory, Avner Magen, Steven Myers, and Charles Rackoff. Simple permutations mix well. In Josep Daz, Juhani Karhumki, Arto Lepist, and Donald Sannella, editors, *Automata, Languages and Programming*, volume 3142 of *Lecture Notes in Computer Science*, pages 770–781. Springer Berlin Heidelberg, 2004.

[Lov96] L. Lovász. Random walks on graphs: A survey. In D. Miklós, V. T. Sós, and T. Szőnyi, editors, *Combinatorics, Paul Erdős is Eighty*, volume 2, pages 353–398. János Bolyai Mathematical Society, Budapest, 1996.

[LR86] Michael Luby and Charles Rackoff. How to construct pseudo-random permutations from pseudo-random functions. In HughC. Williams, editor, *Advances in Cryptology ? CRYPTO ?85 Proceedings*, volume 218 of *Lecture Notes in Computer Science*, pages 447–447. Springer Berlin Heidelberg, 1986.

[Mat94] Mitsuru Matsui. Linear cryptanalysis method for des cipher. In Tor Helleseth, editor, *Advances in Cryptology ? EUROCRYPT ?93*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer Berlin Heidelberg, 1994.

[MV12] Eric Miles and Emanuele Viola. Substitution-permutation networks, pseudorandom functions, and natural proofs. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology - CRYPTO 2012*, volume 7417 of *Lecture Notes in Computer Science*, pages 68–85. Springer Berlin Heidelberg, 2012.

[NR99] Moni Naor and Omer Reingold. On the construction of pseudorandom permutations: Luby-rackoff revisited. *Journal of Cryptology*, 12(1):29–66, 1999.

[Per92] Rene Peralta. On the distribution of quadratic residues and nonresidues modulo a prime number. *Mathematics of Computation*, pages 433–440, 1992.

[Pom12] Klaus Pommerening. Quadratic equations in finite fields of characteristic 2. `http://www.staff.uni-mainz.de/pommeren/MathMisc/QuGlChar2.pdf`, 2012.

[SJ89]. Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Inf. Comput.*, 82(1):93–133, July 1989.