Non-Asymptotic t-Wise Independence of Substitution-Permutation Networks

by

Angelos Pelecanos

S.B., Computer Science and Engineering, Mathematics
Massachusetts Institute of Technology (2021)

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

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022

Author .....
Department of Electrical Engineering and Computer Science
May 6, 2022

Certified by .....
Vinod Vaikuntanathan
Professor of EECS
Thesis Supervisor

Accepted by .....
Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Non-Asymptotic t-Wise Independence of Substitution-Permutation Networks

by
Angelos Pelecanos

Submitted to the Department of Electrical Engineering and Computer Science
on May 6, 2022, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

In this thesis, we study the t-wise independence of block ciphers following the Substitution-Permutation Network design to prove resilience against cryptanalytic attacks and show non-asymptotic bounds for two widely-used ciphers. There are two main contributions of this thesis.

We study the pairwise independence of AES. Replacing the `INV` S-box with an 'ideal' variant, we are able to compute tight convergence properties and prove that this ideal AES is pairwise independent in 5 rounds. As a corollary, we show how to simulate the ideal AES variant using the true AES, after silencing parts of some AES rounds. We call the resulting construction *censored AES* and we prove that it is pairwise independent in 92 rounds. Since this variant is modeled after AES, but does not perform a significant fraction of the mixing steps, we believe that our result is evidence that the true AES is pairwise independent in less than 100 rounds.

In the second part of this thesis, we study the t-wise independence of the MiMC cipher. In particular, we explore whether the use of exponential sums results from algebraic number theory can show convergence to t-wise independence of the MiMC cipher over a prime order field. Even though we do not achieve any concrete results, we believe that this is a direction worth pursuing further.

Thesis Supervisor: Vinod Vaikuntanathan
Title: Professor of EECS

Acknowledgments

Firstly and most importantly, I would like to thank God for everything. Not everyone has had the opportunities I had.

Additionally, I would like to express my sincere gratitude to my research supervisor, Professor Vinod Vaikuntanathan. I am indebted to Vinod for introducing me to the problems of this thesis, for his valuable insight on interesting and tractable problems, and for supporting me during my grad school applications. I am hugely grateful to him for being patient with me, and I need to apologize for any inconvenience I may have caused him.

The work of this thesis would not have been possible without the help of my two other collaborators, Professors Stefano Tessaro and Tianren Liu. They welcomed me as a collaborator from day one and were always willing to explore my ideas, no matter how incorrect or incomplete.

I had my first experience in theoretical computer science research during undergrad, under the supervision of Professor Virginia Vassilevska Williams, whom I thank for the guidance and engaging discussions on graph algorithms and fine-grained complexity.

Before coming to the United States to study, I spent my high school years training for and competing in many informatics olympiads. During that time, I was lucky to have Mr. Panos Eracleous as a teacher, who has been transforming the budding programmers of Cyprus into new computer scientists.

I have collected great memories during my five years at MIT. Apart from academics, I am indebted to the people I had the honor to meet and who made everyday life worth living. In particular, I want to thank Angelos Assos for being my Cypriot friend away from Cyprus.

Finally, I would like to express my deepest gratitude to my beloved family, from whom I had to travel away and sacrifice time with them, for my education: my respected parents, Panayiotis and Irene, my brothers Rev. Fr. Loizos and Fotis, and their beautiful families Katerina, Panayiotis, Eleni, Christos and Ioanna, Giorgos, Irene. May God Bless you all.

Contents

1 Introduction **11**

1.1 Substitution-Permutation Networks 12

1.1.1 Advanced Encryption Standard (AES) 13

1.1.2 MiMC 14

1.2 t-Wise Independence and Cryptanalysis 14

1.3 Our Contributions 15

2 Pairwise Independence of Ideal AES **17**

2.1 Preliminaries 17

2.1.1 Assumptions 17

2.1.2 Pairwise Independence 17

2.1.3 Ideal S-box 18

2.1.4 Layouts and the Layout Graph 20

2.2 Transition Probabilities of the Layout Graph 22

2.2.1 Exact Transition Probabilities 22

2.2.2 Bounds on the Transition Probabilities 27

2.3 Layout Graph Convergence 30

2.3.1 Overview 30

2.3.2 Technical Details 32

3 Pairwise Independence of Censored AES **41**

3.1 Exact Calculation for Ideal S-box 41

3.2 From Ideal S-box to `INV` S-box 42

3.2.1 Warm-Up: > 100 Rounds 42

3.2.2 Improved Analysis 43

3.3 Transition Probabilities for INV S-box via Exponential Sums 45

4 t-Wise Independence of MiMC over Prime Field **51**
4.1 Overview 51
4.2 Polynomial Decomposition 54
4.3 Solutions to System of Equations and Exponential Sums 55

A Proof of Lemma 5 **59**

B Proof of Claim 9 **63**

C Proof of Lemma 4 - Diagonal Equality **65**

List of Figures

1-1 An example of a Substitution-Permutation Network with k blocks and 2 rounds shown in yellow boxes. Each round contains three sets of colored operations: `AddRoundKey`, S-`Box` and `LinearMixing` 13

List of Tables

2.1 This table shows the probability of transitioning from layout c to layout d for different values of k. We are assuming a mixing layer with maximal branching number. It turns out that the transition probability does not depend on the exact position of the non-zero entries, but rather on their number only. 23

3.1 Statistical distance from stationary distribution after a few steps of the AES layout graph with an ideal S-box. 42
3.2 Statistical distance from the ideal S-box after a few repetitions of the `INV` S-box over $\mathbb{F}_{2^8}$. 43

Introduction

A significant part of the traffic on the Web is encrypted using block ciphers. Many of these block ciphers follow the Substitution-Permutation Network (SPN) design, meaning they are efficient encryption schemes that repeat simple and weak encryptions for multiple rounds. Our intuitive understanding of why they work is that the composition of many weak encryptions gives strong security.

Despite their widespread use, mathematically analyzing their resilience against cryptanalytic attacks remains an open problem. In the literature, cryptographers typically rely on reductions to prove the security of cryptographic schemes, by demonstrating how breaking the scheme would result in also solving a problem we believe to be hard. When it comes to block ciphers, we have no candidate 'hard' problems to reduce to. The inherent nature of SPNs as compositions of simple permutations makes them hard to associate with well-studied mathematical problems.

Additionally, the incremental format of SPNs makes them a great candidate to construct attacks against. For example, over the past years cryptanalysts have been developing attacks for more and more rounds of the Advanced Encryption Standard (AES) cipher [GRR17, Gra19, BODK$^+$18, BCC19]. In the past, cryptanalysts have been very successful in devising attacks against the security of popular block ciphers. One example is the use of differential cryptanalysis [BS91] against Data Encryption Standard (DES), the predecessor of AES.

A recent paper by Liu, Tessaro and Vaikuntanathan [LTV21] set forth a research program whose goal is to study the security of practical and widely-used SPN ciphers. One desirable property that they focused on was t-wise independence, which can rule out several classes of known attacks. In this thesis, we continue their study of the t-wise independence of block ciphers and we obtain concrete non-asymptotic results for two widely-used ciphers today, AES and MiMC. Our results can be extended to ciphers that are designed according to the *Substitution-Permutation Network* framework.

1.1 Substitution-Permutation Networks

Given two positive integers k, b, a Substitution-Permutation Network (SPN) is a class of keyed permutations on the set of (bk)-bit strings $\{0,1\}^n$, for $n = bk$. SPNs divide the n bits of their input into k blocks of b consecutive bits. We typically think of such a string of b bits as an element of the finite field $\mathbb{F}_{2^b}$. Then, SPNs apply a transformation to these blocks according to a random secret key for a predetermined number of repetitions (rounds). The transformation during one round of the cipher is parametrized by three operations that happen in sequence

1. `AddRoundKey`. XOR the input with the secret key. Since the XOR operation happens per-bit, `AddRoundKey` applies to each block separately.

2. S`-Box`. The S-Box is a permutation defined by the function $S : \mathbb{F}_{2^b} \to \mathbb{F}_{2^b}$. S is non-linear, such as the inverse over $\mathbb{F}_{2^b}$: $S(x) = x^{2^b-2}$ or the cube: $S(x) = x^3$ over $\mathbb{F}_{2^b}$ for an odd integer b.

3. `LinearMixing`. The last operation is a linear mixing that happens between all k blocks of the input, which is represented as a $k \times k$ matrix with entries from $\mathbb{F}_{2^b}$

This sequence of operations is then repeated r times to obtain an r-round SPN. We assume that the cipher is using a new random n-bit secret key in each round, for a total of $r \cdot n$ bits of randomness in total. In practice, a *key schedule* is used to expand n random bits to $r \cdot n$ bits. The effect of the key schedule on block ciphers and their pseudorandomness is an important open problem [LTV21].

Although what we can prove about SPNs is limited, our intuition behind their design is that the combination of `AddRoundKey` + S`-Box` generate random-looking blocks. The `LinearMixing` layer is then used to *diffuse* the randomness between all blocks.

If $k = 1$, the cipher contains one large block of n bits and the S-box function applies to all n bits at the same time. This is in many cases inefficient since it requires computing the non-linear S-box over a large domain. For efficiency, many modern ciphers (like AES) split the input to $k > 1$ blocks. It is still not exactly clear how the choice of k affects the security of the cipher, since the $k = 1$ variant of AES with trivial mixing is known not to be 4-wise independent [LTV21].

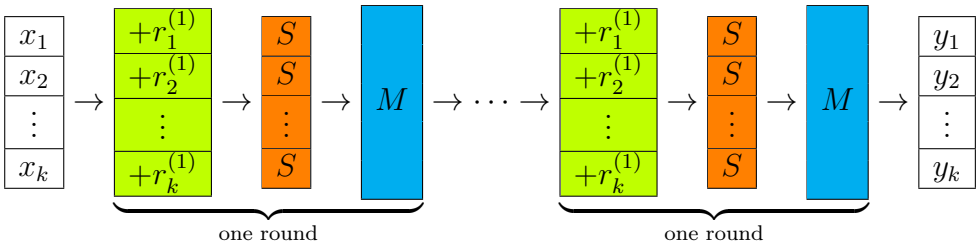


Figure 1-1: An example of a Substitution-Permutation Network with k blocks and 2 rounds shown in yellow boxes. Each round contains three sets of colored operations: `AddRoundKey`, S-`Box` and `LinearMixing`

1.1.1 Advanced Encryption Standard (AES)

The Advanced Encryption Standard (originally known as Rijndael) is the most widely used block cipher in the world. AES became effective as a U.S. federal government standard in May 2002, after approval by the U.S. Secretary of Commerce. It is a particular instantiation of a Substitution-Permutation Network.

The length of the input is $n = 128$ bits, which are divided in $k = 16$ blocks of $b = 8$ bits each. These 16 blocks are usually arranged in a 4×4 array. The S-box is chosen as an affine transformation of the inverse function `INV` over the $\mathbb{F}_{2^8}$ finite field.

The linear mixing layer is a bit more complicated, as it was designed with implementation efficiency and parallelism in mind. In particular, the linear mixing is divided in two steps: `ShiftRows` and `MixColumns`.

`ShiftRows`. This step acts on each row separately. Counting from row 0 down to row 3, each row i is shifted to the left by i positions. This means that the first row remains unchanged, the second row shifts its first block to the left and so on. Below you can find an example on how the `ShiftRows` operation behaves.

α_{11}	α_{12}	α_{13}	α_{14}
α_{21}	α_{22}	α_{23}	α_{24}
α_{31}	α_{32}	α_{33}	α_{34}
α_{41}	α_{42}	α_{43}	α_{44}

`ShiftRows` →

α_{11}	α_{12}	α_{13}	α_{14}
α_{22}	α_{23}	α_{24}	α_{21}
α_{33}	α_{34}	α_{31}	α_{32}
α_{44}	α_{41}	α_{42}	α_{43}

`MixColumns`. The second part of the mixing involves each column separately. If we consider the 4×4 block array as a matrix, then `MixColumns` is performing matrix multiplication with the block array matrix and a known 4×4 matrix `MC`. This means that each element is now the linear combination of the elements in its column.

$$\begin{array}{|c|c|c|c|} \hline \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \hline \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{21} \\ \hline \alpha_{33} & \alpha_{34} & \alpha_{31} & \alpha_{32} \\ \hline \alpha_{44} & \alpha_{41} & \alpha_{42} & \alpha_{43} \\ \hline \end{array} \xrightarrow{\texttt{MixColumns}} \begin{bmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 3 \\ 3 & 1 & 1 & 2 \end{bmatrix} \begin{array}{|c|c|c|c|} \hline \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \hline \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{21} \\ \hline \alpha_{33} & \alpha_{34} & \alpha_{31} & \alpha_{32} \\ \hline \alpha_{44} & \alpha_{41} & \alpha_{42} & \alpha_{43} \\ \hline \end{array}$$

As an example, if the elements of the first column are $\alpha_{11}, \alpha_{22}, \alpha_{33}, \alpha_{44}$ then after `MixColumns` the first element of the first column will equal $2\alpha_{11} + 3\alpha_{22} + \alpha_{33} + \alpha_{44}$. Recall that multiplication and addition is done over $\mathbb{F}_{2^8}$.

1.1.2 MiMC

The name of the MiMC cipher stands for Minimal Multiplicative Complexity and was introduced in [AGR+16] for Succinct Non-interactive Arguments of Knowledge (SNARKS). Other applications include Multi-Party Computation (MPC) and Fully Homomorphic Encryption (FHE).

MiMC is another instantiation of the Substitution-Permutation Network design, whose input is arranged in only one block $k = 1$ and with the cube over $\mathbb{F}_{2^b}$ as its S-box, $S(x) = x^3$. Since it has only one block, there is no need for a linear mixing layer. Two variants of MiMC have been introduced, one defined to work over the extension field $\mathbb{F}_{2^b}$ (for b odd) and the other over prime fields $\mathbb{F}_p$, for p a large prime. For the purposes of this thesis, MiMC will have n independent and uniformly distributed round keys denoted as $r_1, \ldots, r_n$ and its execution on input x can be seen below.

$$x \xrightarrow{\texttt{ARK}} x+r_1 \xrightarrow{S} (x+r_1)^3 \xrightarrow{\texttt{ARK}} (x+r_1)^3+r_2 \xrightarrow{S} \ldots \xrightarrow{S} \Big(\ldots\big((x+r_1)^3+r_2\big)^3+\cdots+r_n\Big)^3$$

1.2 t-Wise Independence and Cryptanalysis

The recent paper of Liu, Tessaro and Vaikuntanathan [LTV21] initiated a research program that studies the t-wise independence of concrete block ciphers. They argue that t-wise independence is a desirable property for a block cipher against various cryptanalytic attacks. Indeed, pairwise independence implies security against differential and linear attacks, whereas extending this property to t-wise independence implies resilience against statistical attacks with at most t inputs and order $\log_2(t)$ differential attacks.

1.3 Our Contributions

This thesis continues the work of [LTV21] on two fronts. First, we obtain improved pairwise independence guarantees for a variant of AES, which we call 'censored AES'. We do this by first replacing the inverse (`INV`) S-box by an ideal S-box, which is easier to analyze. Then, we can analyze the convergence properties of this ideal S-box exactly and we prove that 5 rounds of AES with the ideal S-box are enough to reach pairwise independence.

Theorem. (Informal). *The statistical distance of a substitution-permutation network with an ideal S-box and a linear mixing layer of maximal branching number from pairwise independence after $2\rho + 1$ steps is*

$$d(2\rho + 1) \leq \frac{2^{\rho k - 1}(2e)^{\rho(k-1)/2}}{(2^b - 1)^{\rho k/2}}$$

Additionally, AES with an ideal S-box and a linear mixing layer of maximal branching number is 2^{-128}-close to pairwise independent after only 5 rounds.

We proceed with showing how to simulate the ideal S-box using consecutive `INV` S-boxes. The way to compose `INV` in sequence is to censor (i.e. silence) the mixing steps between them, which is how we obtain our result for the censored variant of AES.

Theorem. (Informal). *Consider a 'censored' variant of the AES cipher with a maximal branching mixing layer, in which a specific subset of the mixing layers is not performed. Then 'censored' AES is 2^{-128}-close to pairwise independent in 92 rounds, where a round could be normal, or with silenced mixing.*

Since the censored variant performs significantly less mixing operations than the true AES, we expect that its convergence to pairwise independence is severely slowed down. Thus, we believe that our result is evidence that the true AES is pairwise independent in less than 100 rounds.

We note here that analyzing AES using idealized components is a promising research direction, especially since recent reduced-round attacks on AES [GRR17, Gra19, BODK+18] do not use the algebraic nature of the S-box.

The second part of this thesis is dedicated to the MiMC cipher. Prior work [LTV21] proved the existence of SPNs that are t-wise independent after $t + o(t)$ rounds. We hope to extend this result by proving that the MiMC cipher over prime order fields is a concrete construction of a t-wise independent cipher in $O(t)$ rounds for large enough fields.

Due to the algebraic structure of the MiMC cipher, we attempt to obtain such a result using character sums, and in particular using Deligne's proof of the Weil conjectures [Del74, Del80]. The motivation behind this approach is that t distinct plaintexts give t MiMC ciphertexts that can be computed using low-degree polynomials. Deligne's result provides an upper bound on the character sums of these polynomials, which can bound the pointwise distance of the ciphertext distribution to t-wise independence.

Erratum note. A prior version of this thesis claimed to prove the t-wise independence of the MiMC cipher in $O(t)$ rounds using the techniques described above. Since then, an issue with the proof was discovered and we were unable to restore the original result. The current version does not include the unverified statement but contains some preliminary work on the MiMC cipher with the hope that similar techniques can recover the original result.

Pairwise Independence of Ideal AES

2.1 Preliminaries

2.1.1 Assumptions

For this thesis, we will adopt the AES model from [LTV21]. In that paper, the authors assume that each round key is sampled uniformly and identically from $\mathbb{F}_{2^b}^k$. In practice, only one such key is sampled and is then expanded to one key per round using a key scheduler. Another simplification in [LTV21] is to ignore the fixed affine transformation after the S-box, since it does not affect our convergence properties.

The new assumption that we make in this paper will be about the mixing layer. For the remainder of this thesis we will assume that the mixing of AES is represented as a maximal branching number matrix $M \in \mathbb{F}_{2^8}^{k \times k}$. This is not the case in practice, since mixing consists of `ShiftRows` and `MixColumns`. Admittedly, we expect that replacing the mixing layer with full branch mixing will improve the convergence to pairwise independence, with the caveat of sacrificing the efficiency and parallelism that the current mixing provides.

Definition 1. *The branching number of a matrix $M \in \mathbb{F}_{2^b}^{k \times k}$ is defined as*

$$\mathrm{br}(M) = \max_{\alpha \in \mathbb{F}_{2^b}^k} (\mathrm{wt}(\alpha) + \mathrm{wt}(M\alpha))$$

where wt is the Hamming weight. The maximal branching number for any $k \times k$ matrix is $k + 1$ and having a maximal branching number is a desirable property for mixing functions [Dae95, KHL$^+$02].

2.1.2 Pairwise Independence

A random permutation $P : \mathbb{F}_{2^b}^k \to \mathbb{F}_{2^b}^k$ is pairwise independent if for all pairs of distinct input plaintexts (x_1, x_2), the output $(P(x_1), P(x_2))$ is uniformly distributed over all pairs of distinct ciphertexts. Equivalently, we want the linear transformation

$$(P(x_1), P(x_1) + P(x_2))$$

to have the first coordinate uniformly distributed over $\mathbb{F}_{2^b}^k$ and the second coordinate uniformly distributed over $\mathbb{F}_{2^b}^k \backslash \{\mathbf{0}\}$. Since the first `AddRoundKey` operation makes the marginal distribution of $P(x_1)$ uniform, what is left is to bound the distance of $P(x_1) + P(x_2)$ from the uniform distribution over $\mathbb{F}_{2^b}^k \backslash \{\mathbf{0}\}$, $U(\mathbb{F}_{2^b}^k \backslash \{\mathbf{0}\})$.

Lemma 1. *The statistical distance of AES from pairwise independent is equal to the maximum statistical distance of AES*(x_1)*+ AES*(x_2) *from* $U(\mathbb{F}_{2^b}^k \backslash \{\mathbf{0}\})$ *for all* x_1, x_2.

With Lemma 1 we have reduced the dimension of our problem from the joint distribution of two ciphertexts to the marginal distribution of their sum. To make this statistical distance arbitrarily small, we will frequently make use of the Amplification Lemma of [KNR05].

Lemma 2. (Amplification Lemma [KNR05]). *Consider a Markov chain* $G = (V, E)$ *with transition matrix* M *and stationary distribution* $\pi(\cdot)$. *Define the statistical distance from stationary after* t *steps as*

$$d(t) = \max_x \|M^t e_x - \pi\|_{TV}$$

Here e_x *is the vector with* 1 *at position* x *and* 0 *everywhere else. Then the distance after* $s + t$ *steps is related to the distance after* t *and* s *steps.*

$$d(s + t) \leq 2d(s)d(t)$$

We will mostly use the Amplification Lemma in the following form, which can be proved using induction for all positive integers ρ.

$$d(\rho t) \leq 2^{\rho-1} d(t)^\rho$$

2.1.3 Ideal S-box

In this section we will understand how the non-linear S-box affects the sum of the two ciphertexts, which will motivate the design of an ideal S-box. Since the mixing layer M is linear, just like the sum operation, the AES mixing has the same effect on the ciphertext sum. Formally for any two ciphertexts c_1, c_2

$$M(c_1) + M(c_2) = M(c_1 + c_2)$$

Consider now the effect of the `AddRoundKey` and S-box operations on two ciphertexts c_1, c_2 with sum d. The round key $r^{(1)}$ is uniformly sampled from $\mathbb{F}_{2^b}^k$ and we apply the S-box to each block of the ciphertexts separately.

$$\underbrace{\begin{matrix} c_1 \\ c_2 \end{matrix}}_{\text{sum } d} \xrightarrow{\texttt{ARK}} \underbrace{\begin{matrix} c_1 + r^{(1)} \\ c_2 + r^{(1)} \end{matrix}}_{\text{sum } d} \xrightarrow{S} \underbrace{\begin{matrix} S\left(c_1 + r^{(1)}\right) \\ S\left(c_2 + r^{(1)}\right) \end{matrix}}_{\text{sum } S(r)+S(r+d)}$$

where we write $r = c_1 + r^{(1)}$, which is uniformly distributed over $\mathbb{F}_{2^b}^k$. Nyberg [Nyb93] analyzed the scalar version of this distribution

$$\tilde{S}(d) \stackrel{\text{def}}{=} S(r) + S(r+d)$$

for any non-zero $d \in \mathbb{F}_{2^b}$, $r \in \mathbb{F}_{2^b}$ chosen uniformly at random, and S being the `INV` function over $\mathbb{F}_{2^b}$. Below we assume b is even, as is the case for AES, and Tr: $\mathbb{F}_{2^b} \to \mathbb{F}_2$ is the linear trace function over the finite field of characteristic 2.

$$\mathbb{P}_r[\tilde{S}(d) = \gamma] = \begin{cases} \frac{4}{2^b} & \gamma = \frac{1}{d} \\ \frac{2}{2^b} & \text{Tr}\left(\frac{1}{\gamma \cdot d}\right) = 0 \\ 0 & \text{otherwise} \end{cases}$$

As we can see, if the sum of the ciphertexts if d, then after `ARK` and the `INV` S-box, the possible sums are distributed over the inverse of a subspace of $\mathbb{F}_{2^b}$, with one value appearing twice as often as the rest. This complex distribution is hard to work with, and indeed prior work [LTV21] proved pairwise independence by approximating this distribution. What we will do, is we will replace this distribution with $U(\mathbb{F}_{2^b} \backslash \{0\})$. The uniform distribution and $\tilde{S}(d)$ are quite far apart in statistical distance, however we are able to relate them because the composition $\tilde{S}(\dots \tilde{S}(d) \dots)$ approximates $U(\mathbb{F}_{2^b} \backslash \{0\})$. The way we will get this composition of $\tilde{S}$, is by censoring the mixing layers between the `ARK` and S-box operations.

Definition 2. *We call a function $S^* : \mathbb{F}_{2^b} \to \mathbb{F}_{2^b}$ the* ideal *S-box if*

$$S^*(d) = \begin{cases} \gamma \sim U(\mathbb{F}_{2^b} \backslash \{0\}) & d \neq 0 \\ 0 & d = 0 \end{cases}$$

Such an ideal S-box can exist in practice, if we replace `AddRoundKey` and `INV` with a random permutation $P : \mathbb{F}_{2^8} \to \mathbb{F}_{2^8}$. To summarize, we showed that the sum of two ciphertexts after one round of AES is obtained after applying $\tilde{S}$ to each block,

followed by the mixing layer. Since the distribution of $\tilde{S}$ is hard to work with, we have defined the ideal S-box S^*. Using this new distribution, we will obtain tight bounds on the number of rounds required for the sum of two plaintexts to converge to $U(\mathbb{F}_{2^b}^k \backslash \mathbf{0})$.

2.1.4 Layouts and the Layout Graph

Our definition of the ideal S-box now motivates us to define 'equivalence groups'. In particular, after one application of the ideal S-box, any block of the ciphertext sum that has a non-zero value will be uniformly distributed over all non-zero values. On the other hand, any zero block remains zero. Thus the distribution of our sum d after one application of the ideal S-box is uniformly distributed over all sums that have non-zero entries in exactly the same positions as d. We will define this set of sums as the 'layout' of d.

Definition 3. *A layout is the set of ciphertext sums that have zero and non-zero entries in the same positions. If we represent the indices with zero entries and non-zero entries with bits* 0 *and* 1 *respectively, we get a string c of k bits that defines the layout. Thus, we say a ciphertext sum x is in layout c if*

$$x_i \neq 0 \Leftrightarrow c_i = 1$$

We also denote with $|c|$ as the number of set bits in c. Equivalently, $|c|$ is the number of non-zero entries in layout c and we will also refer to it as the weight *of layout c.*

Example. Below is an example, assuming a smaller number of blocks, $k = 4$. Here all α_i are non-zero.

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \\ \alpha_4 \end{bmatrix} \in c,\ c = 1101,\ |c| = 3$$

Definition 4. *For a layout f and an integer i (possibly negative) such that*

$$0 < |f| + i \leq k$$

we define f_i to be an arbitrary layout with $|f_i| = |f| + i$.

Definition 5. *For simplicity, we will denote by* $\mathbf{k}$ *the unique layout with no zero entries, i.e. $|\mathbf{k}| = k$.*

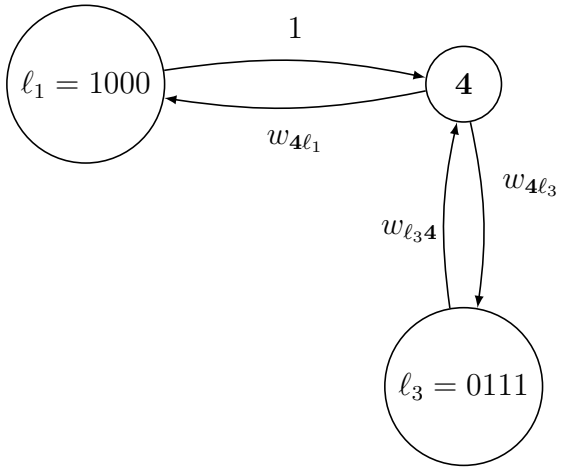
We will analyze the convergence of the ciphertext sum using a Markov chain. The vertices of our Markov chain will be the layouts of ciphertexts sums, since the ideal S-box takes two sums x, y from the same layout to the same distribution $S^*(x) = S^*(y)$. This reduces the total number of vertices in our Markov chain from $2^{kb} - 1$ chiphertext sums to $2^k - 1$ possible layouts.

Definition 6. *The layout graph $G = (V, E)$ for a block cipher consists of $2^k - 1$ vertices, one for each possible layout. We add directed edges from layout c to layout d with weight w_{cd} equal to the probability of starting from any ciphertext sum in c and after one round of the block cipher we end up in layout d.*

Example. Below we show a small part of the layout graph for a block cipher with $k = 4$. On the left side we have one layout ℓ_1 with weight 1 and on the right we have the layout with weight k at the top and a layout ℓ_3 with weight 3 at the bottom. Since the ideal S-box makes any ciphertext sum in these layouts behave similarly, the probability of transitioning from a layout c to another layout d is independent of the starting sum in c. We can compute it to be equal to

$$w_{cd} = \frac{|\{x \in c \mid Mx \in d\}|}{(2^b - 1)^{|c|}}$$

In the case when $|c| = 1$, then the maximal branching number of the mixing layer guarantees that Mx lies in the layout with weight k for any $x \in c$.



In the next section we will compute the rest of the transition probabilities exactly, which will allow us to bound the mixing time of the layout graph.

2.2 Transition Probabilities of the Layout Graph

2.2.1 Exact Transition Probabilities

Lemma 3. *Consider a block cipher with k blocks, the ideal S-box and a full-branch mixing layer. The probability of transitioning from a specific layout c to a particular layout d is*

$$w_{cd} = \underbrace{\frac{\binom{k-1}{k-1}}{(2^b-1)^{k-|d|}} - \frac{\binom{k}{k-1}}{(2^b-1)^{k-|d|+1}} + \frac{\binom{k+1}{k-1}}{(2^b-1)^{k-|d|+2}} - \cdots}_{|c|+|d|-k \text{ terms}}$$
$$= \sum_{i=0}^{|c|+|d|-k-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{k-|d|+i}}$$

The above equality only holds as long as $|c| + |d| \geq k + 1$. In the opposite case, the maximal branching number rules out this transition and hence the probability is equal to 0.

To better understand the statement of the lemma, we provide an example before jumping into the proof. Let $k = 4$ and α_i, β_i be non-zero elements of $\mathbb{F}_{2^b}$. Consider the transition

$$\alpha \stackrel{\text{def}}{=} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \\ \alpha_4 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ \beta_2 \\ \beta_3 \\ 0 \end{bmatrix} \stackrel{\text{def}}{=} \beta$$

Here α is in the layout $c = 1101$ and β is in the layout $d = 0110$. Since $|c| = 3, |d| = 2$, we can verify from Table 2.1 that the probability of transitioning from α to β after one round of AES is equal to $\frac{1}{2^{2b}}$.

Proof. We will prove this by induction, by fixing the layout c and computing the transition probability for all layouts d, in increasing value of $|d|$ (number of non-zero entries).

Base Case. The maximal branching number of the mixing layer restricts that $|c|+|d| \geq k+1$. Thus our base case will be any layout d with weight $|d| = k-|c|+1$.

# of Blocks k	$\lvert c\rvert$	$\lvert d\rvert$	Probability w_{cd}
4	1	4	1
4	2	3	$\frac{1}{2^b}$
4	2	4	$1-\frac{4}{2^b}$
4	3	2	$\frac{1}{2^{2b}}$
4	3	3	$\frac{1}{2^b}-\frac{4}{2^{2b}}$
4	3	4	$1-\frac{4}{2^b}+\frac{10}{2^{2b}}$
4	4	1	$\frac{1}{2^{3b}}$
4	4	2	$\frac{1}{2^{2b}}-\frac{4}{2^{3b}}$
4	4	3	$\frac{1}{2^b}-\frac{4}{2^{2b}}+\frac{10}{2^{3b}}$
4	4	4	$1-\frac{4}{2^b}+\frac{10}{2^{2b}}-\frac{20}{2^{3b}}$
3	1	3	1
3	2	2	$\frac{1}{2^b}$
3	2	3	$1-\frac{3}{2^b}$
3	3	1	$\frac{1}{2^{2b}}$
3	3	2	$\frac{1}{2^b}-\frac{3}{2^{2b}}$
3	3	3	$1-\frac{3}{2^b}+\frac{6}{2^{2b}}$
5	5	1	$\frac{1}{2^{4b}}$
5	5	2	$\frac{1}{2^{3b}}-\frac{5}{2^{4b}}$
5	5	3	$\frac{1}{2^{2b}}-\frac{5}{2^{3b}}+\frac{15}{2^{4b}}$
5	5	4	$\frac{1}{2^b}-\frac{5}{2^{2b}}+\frac{15}{2^{3b}}-\frac{35}{2^{4b}}$
5	5	5	$1-\frac{5}{2^b}+\frac{15}{2^{2b}}-\frac{35}{2^{3b}}+\frac{70}{2^{4b}}$

Table 2.1: This table shows the probability of transitioning from layout c to layout d for different values of k. We are assuming a mixing layer with maximal branching number. It turns out that the transition probability does not depend on the exact position of the non-zero entries, but rather on their number only.

Layout c has $|c|$ non-zero entries. Let these entries be $|c|$ variables $x_1, \ldots, x_{|c|}$ such that $x_i \neq 0$. The zero entries in d define $k - |d| = |c| - 1$ equations on the x_i variables.

$$\begin{bmatrix} M_{11} & M_{12} & \ldots & M_{1k} \\ \ldots & \ldots & \ldots & \ldots \\ M_{(|c|-1)1} & M_{(|c|-1)2} & \ldots & M_{(|c|-1)k} \\ M_{|c|1} & M_{|c|2} & \ldots & M_{|c|k} \\ \ldots & \ldots & \ldots & \ldots \\ M_{k1} & M_{k2} & \ldots & M_{kk} \end{bmatrix} \begin{bmatrix} x_1 \\ \ldots \\ x_{|c|-1} \\ x_{|c|} \\ \ldots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \ldots \\ 0 \\ \neq 0 \\ \ldots \\ \neq 0 \end{bmatrix}$$

Since we have $|c|$ variables and $|c| - 1$ equations, we need one more equation to solve for x. Without loss of generality, set $x_1 = \alpha$ for all $\alpha \in \mathbb{F}_{2^b} \backslash \{0\}$ and solve the system of equations to obtain the entries of vector x. We show below that these are valid and the only possible solutions to the system.

Claim 1. *Setting $x_1 = \alpha \neq 0$ cannot set another x_i to zero.*

Proof. By contradiction, this would mean that x has $\leq |c| - 1$ non-zero entries and d has $\leq k - |c| + 1$ non-zero entries. In total, their weight is at most $k < k + 1$, which violates the maximal branching number of our mixing M. □

Claim 2. *Setting $x_1 = \alpha \neq 0$ cannot set a non-zero entry of d to zero.*

Proof. By contradiction, this would mean that d has at most $k - |c|$ non-zero entries and x at most $|c|$ entries. Their total Hamming weight is at most $k < k + 1$, which again violates the maximal branching number of M. □

Claim 3. *There can be no solution with $x_1 = 0$.*

Proof. Similar to the first claim, setting $x_1 = 0$ will make the total Hamming weight of x and d less than $k + 1$, violating the maximal branching number. □

As a result, all $2^b - 1$ values of x_1 result in exactly one unique vector in c that transitions to layout d. Out of the $(2^b - 1)^{|c|}$ possible columns in c, exactly $2^b - 1$ of them work, making the probability of such a transition equal to

$$w_{cd} = \frac{1}{(2^b - 1)^{|c|-1}} = \sum_{i=0}^{0} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b - 1)^{|c|-1+i}}$$

Setting up the induction. Recall that layout c has $(2^b - 1)^{|c|}$ possible ciphertext sums. Define the number of such sums from layout c that transition to layout d as

$$T(c, d) = |\{x \mid x \in c \cap Mx \in d\}|$$

The number of sums and the transition probability satisfy

$$T(c,d) = w_{cd} \cdot (2^b - 1)^{|c|}$$

Thus, we will prove the following statement for the number of sums, which will imply our original lemma

$$T(c,d) = \sum_{i=0}^{|c|+|d|-k-1} (-1)^i \binom{k-1+i}{k-1} \cdot (2^b-1)^{|c|+|d|-k-i}$$

Lemma 4. (Diagonal Equality). *The value of $T(c,d)$ only relies on the sum $|c|+|d|$ and not on the exact values of $|c|$ and $|d|$. In particular, 'exchanging' all zero entries of d with non-zero entries of c does not change the value of $T(\cdot,\cdot)$.*

$$T(c_{d-k}, \mathbf{k}) = T(c,d)$$

As a result, instead of computing $T(c,d)$, we will compute $T(c_{d-k}, \mathbf{k})$.

Computing $T(c_{d-k}, \mathbf{k})$. Recall that $|c_{d-k}| = c + d - k$ and $|\mathbf{k}| = k$. Of all possible $(2^b-1)^{|c|+|d|-k}$ sums in c_{d-k}, the number of sums that transition in the $\mathbf{k}$ layout are all *except* the ones that transition in a layout $\mathbf{k}_{-\gamma}$ with $k-\gamma$ non-zeros, for $\gamma > 0$. But the number of sums that end up in a specific layout with $k-\gamma$ non-zeros is $T(c_{d-k}, k_{-\gamma})$, which we have computed using induction, since the total weight is

$$|c_{d-k}| + |k_{-\gamma}| = |c| + |d| - k + k - \gamma = |c| + |d| - \gamma < |c| + |d|$$

Hence we can compute the number of ciphertext sums by subtracting the ones that do not arrive where we want. Note that there are $\binom{k}{\gamma}$ possible layouts with $k-\gamma$ non-zero entries. Thus

$$T(c_{d-k}, \mathbf{k}) = (2^b-1)^{|c|+|d|-k} - \sum_{\gamma=1}^{|c|+|d|-k-1} T(c_{d-k}, \mathbf{k}_{-\gamma}) \cdot \binom{k}{\gamma} \tag{2.1}$$

Intuition. To complete the proof, we will show that the coefficients of each power of $2^b - 1$ are equal in the LHS and the RHS of eq. (2.1). Now we recall our inductive hypothesis, that for positive integer γ

$$T(c_{d-k}, \mathbf{k}_{-\gamma}) = \sum_{i=0}^{|c|+|d|-\gamma-k-1} (-1)^i \binom{k-1+i}{k-1} \cdot (2^b-1)^{|c|+|d|-\gamma-k-i}$$
$$= \binom{k-1}{k-1} \cdot (2^b-1)^{|c|+|d|-\gamma-k} - \binom{k}{k-1} \cdot (2^b-1)^{|c|+|d|-\gamma-k-1} + \dots$$

We want to prove the hypothesis for $\gamma = 0$

$$T(c_{d-k}, \mathbf{k}) = \sum_{i=0}^{|c|+|d|-k-1} (-1)^i \binom{k-1+i}{k-1} \cdot (2^b-1)^{|c|+|d|-k-i}$$
$$= \binom{k-1}{k-1} \cdot (2^b-1)^{|c|+|d|-k} - \binom{k}{k-1} \cdot (2^b-1)^{|c|+|d|-k-1} + \dots$$

Matching Coefficients for Powers of (2^b-1). $T(c_{d-k}, \mathbf{k}_{-\gamma})$ has all powers of 2^b-1 from $(2^b-1)^1$ up to $(2^b-1)^{|c|+|d|-k-\gamma}$.

- $(2^b-1)^{|c|+|d|-k}$. This power only appears outside the summation of eq. (2.1) with coefficient 1, just like our hypothesis.
- $(2^b-1)^{|c|+|d|-k-1}$. This power appears in the first term of the summation with coefficient.
$$-\binom{k}{1} \cdot \binom{k-1}{k-1} = -k$$
This matches our hypothesis.
- $(2^b-1)^{|c|+|d|-k-2}$. This power appears in the first two terms of the summation with coefficient.
$$-\binom{k}{2} \cdot \binom{k-1}{k-1} + \binom{k}{1} \cdot \binom{k}{k-1} = \binom{k+1}{k-1}$$
This matches our hypothesis.
- **Term** $(2^b-1)^{|c|+|d|-k-\alpha}$ for positive integer α. This power appears in the first α terms of the summation with coefficient.
$$-\binom{k}{\alpha} \cdot \binom{k-1}{k-1} + \binom{k}{\alpha-1} \cdot \binom{k}{k-1} - \dots + (-1)^\alpha \binom{k}{1} \cdot \binom{k+\alpha-2}{k-1}$$

$$= \sum_{i=0}^{\alpha-1} (-1)^{i+1} \binom{k}{\alpha} \binom{k-1+i}{k-1}$$

The coefficient of this power in $T(c_{d-k}, \mathbf{k})$ is then

$$(-1)^{\alpha} \binom{k-1+\alpha}{k-1}$$

As a result, we want to prove that

$$\sum_{i=0}^{\alpha-1} (-1)^i \cdot \binom{k}{\alpha-i} \binom{k-1+i}{k-1} = (-1)^{\alpha-1} \binom{k+\alpha-1}{k-1}$$

Which is true by the following lemma, whose proof is included in the appendix.

Lemma 5. *For positive integers α, k with $\alpha \leq k$*

$$\sum_{i=0}^{\alpha-1} (-1)^i \cdot \binom{k}{\alpha-i} \binom{k-1+i}{k-1} = (-1)^{\alpha-1} \binom{k+\alpha-1}{k-1}$$

In particular, when expanding the summation the equality looks as follows

$$\binom{k}{\alpha}\binom{k-1}{k-1} - \binom{k}{\alpha-1}\binom{k}{k-1} + \cdots + (-1)^{\alpha-1}\binom{k}{1}\binom{k+\alpha-2}{k-1} = (-1)^{\alpha-1}\binom{k+\alpha-1}{k-1}$$

As a result, the coefficients of the two sides match. This concludes our proof by induction on the value of $T(\cdot, \cdot)$. Dividing by the total number of values $(2^b-1)^{|c|}$ concludes the proof of the lemma on the transition probabilities.

□

2.2.2 Bounds on the Transition Probabilities

In the previous section we computed the probability of transitioning from any ciphertext sum from a particular layout c to anywhere in a layout d. This probability is not dependent on the exact position of the non-zero entries, but rather it depends on the number of non-zero entries and is equal to

$$w_{cd} = \sum_{i=0}^{|c|+|d|-k-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{k-|d|+i}}$$

In this section we will prove some simple bounds for this probability that will help with our analysis later. Since the exact probability expression is quite technical, we first start with introducing a helpful piece of notation.

Definition 7. *For simplicity, write*

$$P^i_{c\to d} = \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{k-d+i}}$$

Then the transition probabilities can be written as

$$w_{cd} = \sum_{i=0}^{c+d-k-1} (-1)^i P^i_{c\to d}$$

Lemma 6. *As long as $k < 2^b - 2$, then the summation terms for the probabilities satisfy*

$$P^i_{c\to d} > P^{i+1}_{c\to d}$$

Proof. We want to show that

$$\frac{\binom{k-1+i}{k-1}}{(2^b-1)^{k-d+i}} > \frac{\binom{k+i}{k-1}}{(2^b-1)^{k-d+i+1}}$$

Exchange numerators with denominators to obtain

$$\Leftrightarrow \frac{(2^b-1)^{k-d+i+1}}{(2^b-1)^{k-d+i}} > \frac{\binom{k+i}{k-1}}{\binom{k-1+i}{k-1}}$$

Expand the binomial coefficient

$$\Leftrightarrow 2^b - 1 > \frac{(k+i)!}{(k-1)!(i+1)!} \cdot \frac{(k-1)!i!}{(k-1+i)!}$$

Cancel out terms

$$\Leftrightarrow 2^b - 1 > \frac{k+i}{i+1}$$

The maximum value of the RHS is achieved when $i = 0$ and the RHS is then equal to $k+1$. As long as $2^b - 2 > k$, the inequality is satisfied. □

Lemma 7. *As long as $k < 2^b - 2$, we can obtain upper and lower bounds for the probability of transition using the truncated summation*

$$\frac{1}{(2^b-1)^{k-d}} - \frac{k}{(2^b-1)^{k-d+1}} = P^0_{c\to d} - P^1_{c\to d} \le w_{cd} \le P^0_{c\to d} = \frac{1}{(2^b-1)^{k-d}}$$

Proof. Consider the summation

$$(-1)^t \sum_{i=0}^{t} (-1)^i P_{c\to d}^i = (-1)^t P_{c\to d}^0 + (-1)^{t+1} P_{c\to d}^1 + \cdots - P_{c\to d}^{t-1} + P_{c\to d}^t$$

And compare it with the next partial sum

$$(-1)^t \sum_{i=0}^{t+1} (-1)^i P_{c\to d}^i = (-1)^t P_{c\to d}^0 + (-1)^{t+1} P_{c\to d}^1 + \cdots - P_{c\to d}^{t-1} + P_{c\to d}^t - P_{c\to d}^{t+1}$$

The two summations are identical, except for the last term of the second summation, which subtracts a non-negative amount. Thus

$$(-1)^t \sum_{i=0}^{t+1} (-1)^i P_{c\to d}^i \leq (-1)^t \sum_{i=0}^{t} (-1)^i P_{c\to d}^i$$

Now we compare the two summations with

$$(-1)^t \sum_{i=0}^{t+2} (-1)^i P_{c\to d}^i = (-1)^t P_{c\to d}^0 + (-1)^{t+1} P_{c\to d}^1 + \cdots - P_{c\to d}^{t-1} + P_{c\to d}^t - P_{c\to d}^{t+1} + P_{c\to d}^{t+2}$$

This summation adds a non-negative amount to the $t+1$ term, so it is not lesser. It also subtracts the term $P_{c\to d}^{t+1} - P_{c\to d}^{t+2}$ from the t expression. Lemma 6 implies that this term is non-negative and thus the above summation cannot be larger than the t expression. Combining everything together we conclude

$$(-1)^t \sum_{i=0}^{t+1} (-1)^i P_{c\to d}^i \leq (-1)^t \sum_{i=0}^{t+2} (-1)^i P_{c\to d}^i \leq (-1)^t \sum_{i=0}^{t} (-1)^i P_{c\to d}^i \tag{2.2}$$

We will now prove the upper and lower bound separately.
Upper Bound. Write the transition probability as

$$w_{cd} = \sum_{i=0}^{c+d-k-1} (-1)^i P_{c\to d}^i$$

From (1), if $c+d-k-1$ is odd, then

$$w_{cd} \leq \sum_{i=0}^{c+d-k-2} (-1)^i P_{c\to d}^i$$

Otherwise, if $c+d-k-1$ is even, then

$$w_{cd} \leq \sum_{i=0}^{c+d-k-3} (-1)^i P^i_{c\to d}$$

In both cases we showed that the probability is no greater than a summation that goes to a smaller even index. We can repeat this procedure until we reach index 0, which is equal to $P^0_{c\to d}$.

Lower Bound. Write the transition probability as

$$w_{cd} = \sum_{i=0}^{c+d-k-1} (-1)^i P^i_{c\to d}$$

From (1), if $c+d-k-1$ is even, then

$$\sum_{i=0}^{c+d-k-2} (-1)^i P^i_{c\to d} \leq w_{cd}$$

Otherwise, if $c+d-k-1$ is odd, then

$$\sum_{i=0}^{c+d-k-3} (-1)^i P^i_{c\to d} \leq w_{cd}$$

In both cases we showed that the probability is no less than a summation that goes to a smaller odd index. We can repeat this procedure until we reach index 1, which is equal to $P^0_{c\to d} - P^1_{c\to d}$. □

2.3 Layout Graph Convergence

2.3.1 Overview

In this section we will prove that the ideal S-box and a mixing matrix with maximal branching number are enough to obtain very fast convergence to pairwise independence.

Recall that working with the layout graph is natural with the ideal S. Indeed, any two ciphertext sums in the same layout are indistinguishable after one application of the ideal S-box. As a result, we will only model the total probability of being inside each layout.

This observation allows us to decrease our state space from the total $2^{kb} - 1$ ciphertext sums to the $2^k - 1$ possible layouts. There are two notable differences that are important to keep in mind:

1. The uniform distribution over the ciphertext sums does not correspond to the uniform distribution over the layouts. Indeed, the all-ones layout has way more sums than a layout with weight 1. The stationary distribution of the layout graph has the following probability for layout c

$$\pi(c) = \frac{(2^b - 1)^{|c|}}{2^{kb} - 1}$$

2. Once we reach ϵ-close to the stationary distribution of the *layout graph*, this does not directly mean that we are ϵ-close to pairwise independent. This is because Lemma 3 does not guarantee anything about the distribution of the sum within the arriving layout d. Fortunately, one more application of the ideal S-box distributes the probability mass of each layout to its ciphertext sums and thus makes our distribution ϵ-close to pairwise independent. What this means, is that if we need R rounds to approximate the stationary distribution in the layout graph, then we need $R + 1$ rounds to reach pairwise independence.

Lemma 8. *If R rounds are enough to reach ϵ-close to the stationary distribution of the layout graph, then $R + 1$ rounds are enough to reach ϵ-close to pairwise independence.*

Proof. Denote by $p_R(\cdot)$ the distribution over the layouts after R rounds in the layout graph. Since we are ϵ-close to the stationary distribution, this means that

$$\sum_{\text{layout } c} |p_R(c) - \pi(c)| \leq \epsilon$$

Now write the distribution over the ciphertext sums after R rounds as $p'_R(\cdot)$. After one more application of the ideal S-box (which we will write as 'round' $R + \frac{1}{2}$) we know the probability of every ciphertext sum x that belongs to layout c

$$p'_{R+1/2}(x) = \frac{p_R(c)}{(2^b - 1)^{|c|}}$$

Then the distance from pairwise independence is equal to

$$\sum_{\text{ciphertext } x} \left| p'_{R+1/2}(x) - \frac{1}{2^{kb} - 1} \right|$$

$$= \sum_{\text{layout } c} \sum_{x \in c} \left| \frac{p_R(c)}{(2^b-1)^{|c|}} - \frac{1}{2^{kb}-1} \right|$$

$$= \sum_{\text{layout } c} \frac{1}{(2^b-1)^{|c|}} \sum_{x \in c} \left| p_R(c) - \frac{(2^b-1)^{|c|}}{2^{kb}-1} \right|$$

$$= \sum_{\text{layout } c} |p_R(c) - \pi(c)| \leq \epsilon$$

We have just showed that one application of the ideal S-box makes the distance from pairwise independence at most ϵ. The remainder of round $R+1$, the mixing step, cannot increase the distance from pairwise independence. Thus after $R+1$ rounds we are ϵ-close to pairwise independence. □

Intuition. To prove convergence to the stationary distribution, we will use two important insights.

1. After one round of the cipher, which will correspond to one step of the walk (ideal S-box + full-branch mixing), the probability mass in a layout with few non-zero entries (will be denoted as a 'sparse' layout) is very low.

2. The ratio of the stationary probabilities of two layouts c, d is equal to $(2^b - 1)^{|c|-|d|}$. For layouts with many non-zero entries ('dense' layouts), the ratio of the probability of transitioning to layouts c or d is very close to $(2^b-1)^{|c|-|d|}$.

The two insights above suggest the following strategy: Wait for one round until the probability of a 'sparse' layout is low. Then during the next round, the transition probabilities will move the probability mass (almost) according to their stationary probability.

A tradeoff arises from our definition of 'sparse' layout. The higher the threshold (more non-zero entries) for a layout to be sparse, the more probability mass ends up in a 'sparse' layout and is not accounted for during step 2. On the other hand, the lower the weight threshold for a 'sparse' layout, the more error we incur in the transition probabilities, as they resemble less and less the stationary probabilities.

2.3.2 Technical Details

Definition 8. *Consider an integer ℓ between $\lfloor \frac{k-1}{2} \rfloor$ and $k-1$. We call a layout with at most ℓ non-zero entries to be a* sparse *layout. In contrast, a layout with more than ℓ non-zero entries is said to be a* dense *layout.*

Lemma 9. (Few Sparse Layouts). *After one step of the random walk, the probability that we end up in a sparse layout is at most*

$$\delta(\ell) = \frac{2^k}{(2^b-1)^{k-\ell}}$$

Proof. From the bounds of Lemma 7

$$w_{cd} \leq \frac{1}{(2^b-1)^{k-d}}$$

Let X^t denote the probability distribution over all $2^k - 1$ valid layouts after t steps. Then the probability of arriving at a fixed sparse layout x with d non-zero entries after $t+1$ steps is

$$\begin{aligned}
\mathbb{P}[X^{t+1} = x] &= \sum_y \mathbb{P}[X^t = y] \cdot w_{yx} \\
&\leq \sum_y \mathbb{P}[X^t = y] \cdot \frac{1}{(2^b-1)^{k-d}} \\
&= \frac{1}{(2^b-1)^{k-d}}
\end{aligned}$$

Note that the above probability is independent of the number of steps in our random walk. Summing the above probability over all sparse layouts gives

$$\begin{aligned}
&\sum_{d=1}^{\ell} \frac{1}{(2^b-1)^{k-d}} \cdot \binom{k}{d} \\
\leq\ &\frac{2^k}{(2^b-1)^{k-\ell}}
\end{aligned}$$

□

Lemma 10. (Dense Layout Ratios are Close to Stationary). *For a fixed dense layout c, the probabilities of transitioning to a dense layout d decreases geometrically with the number of zero entries*

$$\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} - \frac{\binom{k-1+c-\alpha}{k-1}}{(2^b-1)^c} \leq w_{c\mathbf{k}_{-\alpha}} \leq \frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} + \frac{\binom{k-1+c-\alpha}{k-1}}{(2^b-1)^c}$$

Since c is at least $\ell+1$, we get that

$$\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} - \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \leq w_{c\mathbf{k}_{-\alpha}} \leq \frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} + \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

Proof. Write the probability of transitioning to $\mathbf{k}_{-\alpha}$

$$w_{c\mathbf{k}} = \sum_{i=0}^{c-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^i}$$

Multiplying by $(2^b-1)^\alpha$ gives

$$\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} = \sum_{i=0}^{c-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{\alpha+i}}$$

By splitting the summation at $c-\alpha-1$ we get

$$\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} = \sum_{i=0}^{c-\alpha-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{\alpha+i}} + \sum_{i=c-\alpha}^{c-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{\alpha+i}}$$

The first term is exactly equal to $w_{c\mathbf{k}_{-\alpha}}$

$$\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha} = w_{c\mathbf{k}_{-\alpha}} + \sum_{i=c-\alpha}^{c-1} (-1)^i \frac{\binom{k-1+i}{k-1}}{(2^b-1)^{\alpha+i}}$$

The second term is upper and lower bounded by $\pm\frac{\binom{k-1+c-\alpha}{k-1}}{(2^b-1)^c}$, which implies the statement of the claim.

□

Lemma 11. *After two steps of the random walk on the layout graph, the statistical distance from the stationary distribution is at most*

$$\frac{2\cdot 2^k}{(2^b-1)^{k-\ell}} + \frac{2^{k+1}(2e)^{k-1}}{(2^b-1)^{\ell+1}}$$

Proof. We will consider the first two steps of the random walk.

Step 1. The walk distribution after 1 step is X^1 and there is at most $\delta(\ell)$ probability mass in a sparse layout. We consider the case we end up in a dense layout, which

happens with probability at least $1 - \delta(\ell)$.

Step 2. For the second step, our walk distribution is X^2. Even if we start from a dense layout, we could end up in a sparse layout with at most $\delta(\ell)$ probability

However, we know that the probability distribution of each dense layout is approximately distributed geometrically according to the number of zero entries as the following claim suggests.

Claim 4. *Fix a dense layout $\mathbf{k}_{-\alpha}$ with $k - \alpha$ non-zero entries. Conditioned on the event that we start from a dense layout in step* 1*.*

$$\frac{1}{(2^b-1)^\alpha}\cdot\mathbb{P}[X^2=\mathbf{k}]-\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \le \mathbb{P}[X^2=\mathbf{k}_{-\alpha}] \le \frac{1}{(2^b-1)^\alpha}\cdot\mathbb{P}[X^2=\mathbf{k}]+\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

Proof. **Upper Bound**.

$$\begin{aligned}
\mathbb{P}[X^2=\mathbf{k}_{-\alpha}] &= \sum_{\text{dense layout } c} \mathbb{P}[X^1=c]\cdot w_{ck_{-\alpha}} \\
&\le \sum_{\text{dense layout } c} \mathbb{P}[X^1=c]\cdot\left(\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha}+\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}\right) \\
&= \frac{1}{(2^b-1)^\alpha}\sum_{\text{dense layout } c}\mathbb{P}[X^1=c]\cdot w_{c\mathbf{k}} + \sum_{\text{dense layout } c}\mathbb{P}[X^1=c]\cdot\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \\
&= \frac{1}{(2^b-1)^\alpha}\cdot\mathbb{P}[X^2=\mathbf{k}]+\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}
\end{aligned}$$

Lower Bound.

$$\begin{aligned}
\mathbb{P}[X^2=\mathbf{k}_{-\alpha}] &= \sum_{\text{dense layout } c} \mathbb{P}[X^1=c]\cdot w_{c\mathbf{k}_{-\alpha}} \\
&\ge \sum_{\text{dense layout } c} \mathbb{P}[X^1=c]\cdot\left(\frac{w_{c\mathbf{k}}}{(2^b-1)^\alpha}-\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}\right) \\
&= \frac{1}{(2^b-1)^\alpha}\sum_{\text{dense layout } c}\mathbb{P}[X^1=c]\cdot w_{c\mathbf{k}} - \sum_{\text{dense layout } c}\mathbb{P}[X^1=c]\cdot\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \\
&= \frac{1}{(2^b-1)^\alpha}\cdot\mathbb{P}[X^2=\mathbf{k}]-\frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}
\end{aligned}$$

□

Since the probability that we end up in a sparse layout after 2 steps is at most $\delta(\ell)$, we can deduce that

$$\sum_{\text{dense layout } c} \mathbb{P}[X^2 = c] \geq 1 - \delta(\ell)$$

By using the upper bounds of Claim 2, we can get a lower bound on the probability of $\mathbf{k}$ after 2 steps.

$$\mathbb{P}[X^2 = \mathbf{k}] + \sum_{\text{dense layout } \mathbf{k}_{-\alpha}} \left(\frac{1}{(2^b-1)^\alpha} \cdot \mathbb{P}[X^2 = \mathbf{k}] + \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \right) \geq 1 - \delta(\ell)$$

$$\Rightarrow \mathbb{P}[X^2 = \mathbf{k}] \left(\sum_{\alpha=0}^{k-\ell-1} \frac{\binom{k}{\alpha}}{(2^b-1)^\alpha} \right) \geq 1 - \delta(\ell) - \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

$$\Rightarrow \mathbb{P}[X^2 = \mathbf{k}] \geq \frac{1 - \delta(\ell) - \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}}{\sum_{\alpha=0}^{k-\ell-1} \frac{\binom{k}{\alpha}}{(2^b-1)^\alpha}}$$

This directly implies a lower bound for the probability of being in any dense layout

$$\Rightarrow \mathbb{P}[X^2 = \mathbf{k}_{-\alpha}] \geq \frac{1}{(2^b-1)^\alpha} \cdot \mathbb{P}[X^2 = \mathbf{k}] - \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

$$\geq \frac{1 - \delta(\ell) - \sum_{i=1}^{k-\ell-1} \binom{k}{i} \cdot \frac{\binom{k+\ell-i}{k-1}}{(2^b-1)^{\ell+1}}}{(2^b-1)^\alpha \sum_{i=0}^{k-\ell-1} \frac{\binom{k}{i}}{(2^b-1)^i}} - \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

This lower bound is lower than the stationary probability of layout $\mathbf{k}_{-\alpha}$. Hence we can upper bound the statistical distance by summing over lower bound values and subtracting them from the stationary probability. The sum of the lower bounds above is equal to

$$\mathbb{P}[X^2 = \mathbf{k}] + \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \mathbb{P}[X^2 = \mathbf{k}_{-\alpha}]$$

$$\geq \frac{1 - \delta(\ell) - \sum_{i=1}^{k-\ell-1} \binom{k}{i} \cdot \frac{\binom{k+\ell-i}{k-1}}{(2^b-1)^{\ell+1}}}{\sum_{i=0}^{k-\ell-1} \frac{\binom{k}{i}}{(2^b-1)^i}} + \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \left(\frac{1 - \delta(\ell) - \sum_{i=1}^{k-\ell-1} \binom{k}{i} \cdot \frac{\binom{k+\ell-i}{k-1}}{(2^b-1)^{\ell+1}}}{(2^b-1)^\alpha \sum_{i=0}^{k-\ell-1} \frac{\binom{k}{i}}{(2^b-1)^i}} - \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}} \right)$$

We will collect the terms similar to the first term

$$= \frac{1 - \delta(\ell) - \sum_{i=1}^{k-\ell-1} \binom{k}{i} \cdot \frac{\binom{k+\ell-i}{k-1}}{(2^b-1)^{\ell+1}}}{\sum_{\alpha=0}^{k-\ell-1} \frac{\binom{k}{\alpha}}{(2^b-1)^\alpha}} \left(\sum_{\alpha=0}^{k-\ell-1} \frac{\binom{k}{\alpha}}{(2^b-1)^\alpha} \right) - \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

The denominator of the first term cancels with the factor next to it

$$= 1 - \delta(\ell) - \sum_{i=1}^{k-\ell-1} \binom{k}{i} \cdot \frac{\binom{k+\ell-i}{k-1}}{(2^b-1)^{\ell+1}} - \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

Notice the last two terms are actually equal

$$= 1 - \delta(\ell) - 2 \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

As a result, the maximum statistical distance from the stationary distribution for the dense layouts is at most the total probability of all layouts, which is at most 1, minus the sum of the lower bounds, so at most

$$\delta(\ell) + 2 \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

The total statistical distance (including the loss we incurred during step 1 of the walk) is at most

$$2\delta(\ell) + 2 \sum_{\alpha=1}^{k-\ell-1} \binom{k}{\alpha} \cdot \frac{\binom{k+\ell-\alpha}{k-1}}{(2^b-1)^{\ell+1}}$$

Simplifying the expression. We now focus on simplifying the second term of the statistical distance expression. The fraction in the summation is maximized when $\alpha = 1$

$$\leq \frac{2\binom{k+\ell-1}{k-1}}{(2^b-1)^{\ell+1}} \sum_{\alpha=1}^{k-\ell-1} \cdot \binom{k}{\alpha}$$

The sum of binomials is at most 2^k

$$\leq \frac{2\binom{k+\ell-1}{k-1}}{(2^b-1)^{\ell+1}} \cdot 2^k$$

We now upper bound the binomial coefficient using known bounds

$$\leq \frac{2^{k+1}}{(2^b-1)^{\ell+1}} \cdot \left(\frac{e(k+\ell-1)}{k-1}\right)^{k-1}$$

Since ℓ is at most $k-1$, the fraction with the exponent is at most $2e$

$$\leq \frac{2^{k+1}(2e)^{k-1}}{(2^b-1)^{\ell+1}}$$

As a result, the total statistical distance from uniform is the distance from transition from a dense to a dense plus the distance above

$$\frac{2 \cdot 2^k}{(2^b-1)^{k-\ell}} + \frac{2^{k+1}(2e)^{k-1}}{(2^b-1)^{\ell+1}}$$

□

Theorem 1. *The statistical distance of a substitution-permutation network with the ideal S-box and a linear mixing layer of maximal branching number from pairwise independence after* $2\rho+1$ *steps is*

$$d(2\rho+1) \leq \frac{2^{\rho k-1}(2e)^{\rho(k-1)/2}}{(2^b-1)^{\rho k/2}}$$

Proof. We will choose the best value of ℓ in Lemma 11 to achieve the minimum distance from the stationary distribution and drive it down using the [KNR05] Amplification Lemma. Finally, we will translate our bound from the stationary distribution of the layout graph to a pairwise independence result using Lemma 8.

We start by optimizing the value of ℓ. Note that the product of the two terms is constant in ℓ, so by setting the two terms equal will give us the optimal value of ℓ.

$$\frac{2 \cdot 2^k}{(2^b-1)^{k-\ell}} = \frac{2^{k+1}(2e)^{k-1}}{(2^b-1)^{\ell+1}}$$

Exchange numerators and denominators to get

$$(2^b-1)^{2\ell-k+1} = (2e)^{k-1}$$

Multiply both sides with $(2^b-1)^{k+1}$

$$(2^b-1)^{2(\ell+1)} = (2e)^{k-1}(2^b-1)^{k+1}$$

Invert and take the square root gives

$$\frac{1}{(2^b-1)^{\ell+1}} = \frac{1}{(2e)^{(k-1)/2}(2^b-1)^{(k+1)/2}}$$

Multiply with $2^k(2e)^{k-1}$ to match the expression for the statistical distance

$$\frac{2^k(2e)^{k-1}}{(2^b-1)^{\ell+1}} = \frac{2^k(2e)^{(k-1)/2}}{(2^b-1)^{(k+1)/2}}$$

Hence the total distance is twice that value

$$\leq \frac{2^{k+1}(2e)^{(k-1)/2}}{(2^b-1)^{(k+1)/2}}$$

Actually, recall that the value of ℓ has to be an integer. By rounding our optimal value of ℓ to the closest integer, we might lose a factor of at most $(2^b-1)^{1/2}$. Hence the true statistical distance is at most

$$d(2) \leq \frac{2^{k+1}(2e)^{(k-1)/2}}{(2^b-1)^{k/2}}$$

The KNR Amplification Lemma implies that 2ρ steps on the layout graph satisfy $d(2\rho) \leq 2^{\rho-1} \cdot d(2)^\rho$. Substituting our value for $d(2)$ and applying Lemma 8 completes the proof. □

Pairwise Independence of Censored AES

3.1 Exact Calculation for Ideal S-box

In the previous section we gave a result for the almost-pairwise independence for any block cipher with an ideal S-box and a full-branch mixing layer after $2\rho + 1$ rounds.

We can directly apply Theorem 1 to the specific parameters of AES, which has $k = 16$ blocks, with $b = 8$ bits in each block. Then we get that 11 rounds are enough to reach almost pairwise independence, that is

$$d(11) < 2^{-128}$$

Turns out that we can do better, as we show below.

Theorem 2. *The AES cipher with a maximal branching mixing layer and the ideal S-box is 2^{-128}-close to pairwise independent in 5 rounds.*

Proof. The proof will be computational, meaning we will compute the transition matrix of the layout graph after a few steps and calculate the distance from stationary exactly. An important observation, which will make our calculations more efficient, is the fact that the transition probabilities of the layout graph are only related to the number of non-zero entries in each layout and not their exact position. Thus, we can define a new smaller layout graph $G' = (V', E')$, whose vertex u corresponds to all layouts of weight u. Indeed, since after the first step any mass that goes to layout c will also transition to layout d, as long as $|c| = |d|$, we can also coalesce the edges together and sum their weights. This will restrict the size of this new layout graph from $2^k - 1$ vertices to a mere $k = 16$! In the light of this new grouping, the edge weights correspond to the probability of transitioning from any layout c of weight u to any layout of weight v and are equal to

$$w'_{uv} = \sum_{d:\ |d|=v} w_{cd} = \binom{k}{d} \cdot w_{cd}$$

Number of Steps	Distance from Stationary Distribution
2	$2^{-87.2785}$
3	$2^{-119.700}$
4	$2^{-180.9897}$

Table 3.1: Statistical distance from stationary distribution after a few steps of the AES layout graph with an ideal S-box.

Now, using Lemma 3, we can construct the 16×16 transition matrix and compute the exact distance for a few steps using matrix exponentiation. The results are shown in Table 3.1, where we can see that 3 steps are really close to the distribution that we want. However, to reach sufficient distance from the stationary distribution we need to perform 4 steps of the random walk. Lemma 8 implies that one extra application of the ideal S-box is enough to bring our cipher close to pairwise independent. Thus, 5 rounds suffice for pairwise independence.

□

3.2 From Ideal S-box to `INV` S-box

3.2.1 Warm-Up: > 100 Rounds

The ideal S-box was a great abstraction of the true S-box that allowed a direct and exact analysis of the block cipher. But, what can we say about the `INV` S-box?

The first observation is that we can 'simulate' the ideal S-box by repeating the `INV` S-box multiple times, without mixing in between. Indeed, since the S-box only applies within each layout, applying it more times will make each non-zero block look uniform

It is easy to compute the distance of a few consecutive `INV` S-boxes over $\mathbb{F}_{2^8}$ from the ideal S-box over the same field. These distances can be found in Table 3.2.

Lemma 12. *The 'censored' variant of AES repeated for* 120 *rounds is* 2^{-128}*-close to pairwise independent*

Proof. Our direct analysis of Table 3.1 showed that 2 steps on the layout graph are 2^{-87}-close to the stationary distribution. This implies that 3 rounds of AES with the ideal S-box are $\epsilon_{ideal} \stackrel{\text{def}}{=} 2^{-87}$-close to pairwise independence. We schematically represent these 3 rounds below, with S^* being the ideal S-box and M the mixing layer.

INV S-box Repetitions	Statistical Distance to Ideal S-box
1	$2^{-0.99}$
2	2^{-7}
3	$2^{-8.9}$
11	2^{-40}
20	2^{-72}

Table 3.2: Statistical distance from the ideal S-box after a few repetitions of the INV S-box over $\mathbb{F}_{2^8}$.

$$\boxed{S^*} \to \boxed{M} \to \boxed{S^*} \to \boxed{M} \to \boxed{S^*} \to \boxed{M}$$

We can now replace each S^* by consecutive INV S-boxes. We can think of these consecutive S-boxes as normal AES rounds, with the mixing layers 'censored'. We will call this variant of AES as the 'censored' AES, in which we skip some of the mixing rounds. It is our understanding that mixing helps the block cipher achieve pseudorandomness, so we expect that silencing some of these crucial operations can only hurt our convergence to pairwise independence.

$$\boxed{S^*} \equiv \boxed{\texttt{INV}} \to \boxed{M} \to \cdots \to \boxed{\texttt{INV}} \to \boxed{M}$$

To simulate the ideal S-box sufficiently well, we need enough repetitions of the INV S-box. We will use $r = 20$ repetitions. Table 3.2 shows that 20 repetitions are 2^{-72}-close to the ideal S-box. Recall that the S-box is applied to $k = 16 = 2^4$ blocks in parallel, hence we can use the Union Bound inequality to deduce that we can simulate the ideal S-box with an error of at most $\epsilon_{sim} = 2^{-68}$. The total distance of 'censored AES' from pairwise independence is at most the sum of the ideal AES distance and the error from simulating the three ideal S-boxes.

$$\epsilon_{id} + 3\epsilon_{sim} = 2^{-87} + 3 \cdot 2^{-68} < 2^{-65}$$

We conclude that $3r = 60$ rounds of this censored cipher are 2^{-65} close to pairwise independent. Using the Amplfication Lemma [KNR05] once we get that 120 'censored AES' rounds are $< 2^{-128}$-close to pairwise independent. □

3.2.2 Improved Analysis

From the previous analysis, there are a number of inefficiencies. We focus on simulating the first ideal S-box. Recall from the analysis of Lemma 11 that the first

round was used to bound the probability of a sparse layout. But we actually do not need the ideal S-box to do this, even the INV S-box can give us a bound on the sparse layouts after one INV S-box and one full-branch mixing step. In particular we can use exponential sums to obtain an upper bound for the transition probabilities after one application of the INV S-box and one mixing step.

Lemma 13. *For any two layouts c and d, the probability of transitioning from a fixed input of layout c to any input of layout d after one application of the INV S-box and one mixing step is at most*

$$\mathbb{P}_{INV}[c \to d] \leq \frac{3}{2} \cdot 2^{(b-2)(|d|-k)}$$

If we replace the ideal S-box with the INV one, the ciphertext sums in the same layout do not need to behave in the same way. Indeed, two inputs in the same layout c can reach any layout d with different probabilities. Still, we can obtain an upper bound of these probabilities that depends only on the weight of d. This will allow us to avoid simulating the first of the three ideal S-boxes of Lemma 12.

Theorem 3. *Consider a 'censored' variant of the AES cipher with a maximal branching mixing layer, in which a specific subset of the mixing layers is not performed. Then 'censored' AES is 2^{-128}-close to pairwise independent in 92 rounds, where a round could be normal, or without mixing.*

Proof. We prove the following two Claims.

Claim 5. *After one round of AES (one application of the INV S-box and one full-branch mixing), the probability that we are in a layout with less than 8 non-zero entries is*

$$\delta_{INV}(7) \leq 2^{-42.5}$$

Proof. The proof is the same as in Lemma 9, but now using the probability bound from Lemma 13. □

Claim 6. *The distance of the layout graph from the stationary distribution after two rounds, one with the INV S-box and one with the ideal S-box is at most*

$$d_{INV,\, ideal}(2) \leq 2^{-37.9}$$

Proof. The proof is the same as the proof of Lemma 11. To get this tight bound we use $\delta_{\mathrm{INV}}(7)$ from the claim above and substitute the parameters $k = 16, b = 8$ without any simplifications. □

This means that adding another ideal S-box and mixing step makes the following block cipher $\epsilon_{id} = 2^{-37.9}$-close to pairwise independent from Lemma 8.

$$\boxed{\texttt{INV}} \to \boxed{M} \to \boxed{S^*} \to \boxed{M} \to \boxed{S^*} \to \boxed{M}$$

We again have to figure out how many times to repeat the `INV` S-box to simulate the ideal S-boxes. We will use $r = 11$ repetitions and we can see from Table 3.2 that they are 2^{-40}-close to the ideal S-box. The S-box is applied to $k = 16 = 2^4$ blocks in parallel, which means that the error from simulating the ideal S-box is at most $\epsilon_{sim} = 2^{-36}$. The total distance of 'censored AES' from pairwise independence is at most the sum of the ideal AES distance and the error from simulating the two ideal S-boxes.

$$\epsilon_{id} + 2\epsilon_{sim} < 2^{-34}$$

We conclude that $1 + 2r = 23$ rounds of this censored cipher are 2^{-34}-close to pairwise independent. Using the Amplification Lemma of [KNR05] we repeat the cipher 4 times and we get that 92 rounds of 'censored AES' are $< 2^{-128}$-close to pairwise independent. □

3.3 Transition Probabilities for INV S-box via Exponential Sums

In this section we will bound the transition probabilities between layouts for the `INV` S-box, instead of the ideal S-box.

Lemma 13. *For any two layouts c and d, the probability of transitioning from a fixed input of layout c to any input of layout d after one application of the `INV` S-box and one mixing step is at most*

$$\mathbb{P}_{\texttt{INV}}[c \to d] \leq \frac{3}{2} \cdot 2^{(b-2)(|d|-k)}$$

Note. The lemma above is useless for bounding the probability of transitioning to a dense layout. Indeed, setting $|d| = k$ gives us a probability bound of $\frac{3}{2}$. However, when it comes to sparse layouts we can use the lemma to bound the probability of a sparse layout after one round of the true AES.

Proof. To do this, consider two layouts c, d and the transition between them. Without loss of generality, we will represent c to have the first $|c|$ entries to be non-zero.

$$\begin{bmatrix} x_1 \\ \dots \\ x_{|c|} \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

After one round of S-box, we will sample random keys $r_1, \dots, r_k$ and in combination with the `INV` S-box we will get

$$\begin{bmatrix} \tilde{S}(x_1) \\ \dots \\ \tilde{S}(x_{|c|}) \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

For the sake of brevity, we write S as the inverse over $\mathbb{F}_{2^b}$ and

$$\tilde{S}(x_i) := S(x_i + r_i) + S(r_i)$$

Now for our ciphertext sum to transition to layout d, it must hold that after applying the mixing

1. we get zero entries where d has zeros

2. we get non-zero entries where d has non-zero entries

Since we are interested in the maximum probability of transitioning, we will only consider condition 1, which is a weaker condition and can only increase the probability of going from c to d.

For condition 1 to hold, we can apply the multiplication with the mixing matrix and obtain $k-|d|$ equations for our $|c|$ variables. Let the number of possible solutions $(r_1, \dots, r_{|c|})$ to this system of $k - |d|$ equations be at most T for any ciphertext sum in c (meaning for any non-zero $x_1, \dots, x_{|c|}$). Then the transition probability is at most the number of solutions divided by the possible choice of random keys.

$$\mathbb{P}_{\texttt{INV}}[c \to d] \leq \frac{T}{2^{b|c|}}$$

As we will show in Lemma 14 below, the value of T can be bounded using the theory of exponential sums

$$T \leq 6 \cdot 2^{b(|c|+|d|-k)} \cdot 2^{2(k-|d|-1)}$$

This directly implies a bound for the probability of transitioning from layout c to layout d

$$\begin{aligned}\mathbb{P}_{\texttt{INV}}[c \to d] &\leq \frac{6 \cdot 2^{b(|c|+|d|-k)} \cdot 2^{2(k-|d|-1)}}{2^{b|c|}} \\ &= 6 \cdot 2^{b(|d|-k)} \cdot 2^{2(k-|d|-1)} \\ &= \frac{3}{2} \cdot 2^{(b-2)(|d|-k)}\end{aligned}$$

□

Lemma 14. *Consider the following full-rank system of equations with c variables taken from the image of the `INV` S-box and $k - d = \bar{d}$ equations.*

$$\begin{cases} P_1 : A_{11}\tilde{S}(x_1) + A_{12}\tilde{S}(x_2) + \cdots + A_{1c}\tilde{S}(x_c) = 0 \\ P_2 : A_{21}\tilde{S}(x_1) + A_{22}\tilde{S}(x_2) + \cdots + A_{2c}\tilde{S}(x_c) = 0 \\ \cdots \\ P_{\bar{d}} : A_{\bar{d}1}\tilde{S}(x_1) + A_{\bar{d}2}\tilde{S}(x_2) + \cdots + A_{\bar{d}c}\tilde{S}(x_c) = 0 \end{cases}$$

The maximum number of random keys that satisfy this system for any non-zero $(x_1, \ldots, x_c)$ is at most

$$6 \cdot 2^{b(c+d-k)} \cdot 2^{2(k-d-1)}$$

Note. If the variables $\tilde{S}(\cdot)$ were uniform over $\mathbb{F}_{2^b}$, the number of solutions would be equal to $2^{b(c+d-k)}$. We can see that the number of solutions using the S-box has increased by a factor of $6 \cdot 2^{2(k-d-1)}$ approximately, due to the skewed distribution of the S-box.

Proof. Consider the following system of equations with c variables and $k - d = \bar{d}$ equations.

$$\begin{cases} P_1 : A_{11}\tilde{S}(x_1) + A_{12}\tilde{S}(x_2) + \cdots + A_{1c}\tilde{S}(x_c) = 0 \\ P_2 : A_{21}\tilde{S}(x_1) + A_{22}\tilde{S}(x_2) + \cdots + A_{2c}\tilde{S}(x_c) = 0 \\ \cdots \\ P_{\bar{d}} : A_{\bar{d}1}\tilde{S}(x_1) + A_{\bar{d}2}\tilde{S}(x_2) + \cdots + A_{\bar{d}c}\tilde{S}(x_c) = 0 \end{cases}$$

Note that the variables here are the random keys r_i that we are hiding inside the $\tilde{S}(x_i) = S(x_i + r_i) + S(r_i)$. We can use Gaussian Elimination to reduce the system to

$$\Rightarrow \begin{cases} R_1 : A_{11}\tilde{S}(x_1) + A_{12}\tilde{S}(x_2) + \cdots + A_{1c}\tilde{S}(x_c) = 0 \\ R_2 : A_{22}\tilde{S}(x_2) + \cdots + A_{2c}\tilde{S}(x_c) = 0 \\ \cdots \\ R_{\bar{d}} : A_{\bar{d}\bar{d}}\tilde{S}(x_{\bar{d}}) + \cdots + A_{\bar{d}c}\tilde{S}(x_c) = 0 \end{cases}$$

For every solution $(r_{\bar{d}}, \ldots, r_c) \Rightarrow (\tilde{S}(x_{\bar{d}}), \ldots, \tilde{S}(x_c))$ of the last equation $R_{\bar{d}}$, each other equation can reduced by substitution to $\tilde{S}(x_i) = \alpha_i$ for $i < \bar{d}$ and some α_i. The number of possible solutions to this equation is at most 4 from the analysis of $\tilde{S}(\cdot)$ by Nyberg [Nyb93].

As a result, the number of solutions to this system of equations is at most

$$(\#\text{ of solutions to } R_{\bar{d}}) \cdot 4^{\bar{d}-1}$$

$$= (\#\text{ of solutions to } R_{\bar{d}}) \cdot 2^{2(k-d-1)}$$

Solving $R_{\bar{d}}$. We are now interested in the number of random keys $(r_{\bar{d}}, \ldots, r_c)$ that satisfy

$$A_{\bar{d}\bar{d}}\tilde{S}(x_{\bar{d}}) + \cdots + A_{\bar{d}c}\tilde{S}(x_c) = 0$$

We can write the number of solutions T using exponential sums. Let ψ be the additive character of $\mathbb{F}_{2^b}$

$$T = \sum_{r_{\bar{d}},\ldots,r_c} \frac{1}{2^b} \left(\sum_z \psi_z \left(A_{\bar{d}\bar{d}}\tilde{S}(x_{\bar{d}}) + \cdots + A_{\bar{d}c}\tilde{S}(x_c) \right) \right)$$

Since ψ is an additive character

$$T = \sum_{r_{\bar{d}},\ldots,r_c} \frac{1}{2^b} \left(\sum_z \psi_z(A_{\bar{d}\bar{d}}\tilde{S}(x_{\bar{d}})) \ldots \psi_z(A_{\bar{d}c}\tilde{S}(x_c)) \right)$$

Rearranging

$$T = \frac{1}{2^b} \sum_z \left(\sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}}\tilde{S}(x_{\bar{d}})) \right) \ldots \left(\sum_{r_c} \psi_z(A_{\bar{d}c}\tilde{S}(x_c)) \right)$$

The value of $z = 0$ makes every one of the $c - \bar{d} + 1 = c + d - k + 1$ summations equal to 2^b. Hence

$$T = 2^{b(c+d-k)} + \frac{1}{2^b} \sum_{z \neq 0} \left(\sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}} \tilde{S}(x_{\bar{d}})) \right) \dots \left(\sum_{r_c} \psi_z(A_{\bar{d}c} \tilde{S}(x_c)) \right)$$

The value of $z = 0$ is equal to the expected number of solutions. We will now compute the maximum deviation from the mean

$$\left| T - 2^{b(c+d-k)} \right| \leq \left| \frac{1}{2^b} \sum_{z \neq 0} \left(\sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}} \tilde{S}(x_{\bar{d}})) \right) \dots \left(\sum_{r_c} \psi_z(A_{\bar{d}c} \tilde{S}(x_c)) \right) \right|$$

We will move the absolute values inside the summation using the triangle inequality

$$\leq \frac{1}{2^b} \sum_{z \neq 0} \left| \sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}} \tilde{S}(x_{\bar{d}})) \right| \dots \left| \sum_{r_c} \psi_z(A_{\bar{d}c} \tilde{S}(x_c)) \right|$$

We will use the following exponential sums result from [EHN94] to bound the absolute values.

Lemma 15. *([EHN94]) Let* $\mathbf{d} \in \mathbb{F}_{2^b}^s$ *with* $\mathbf{d} \neq \mathbf{0}$ *and let* $\mathbf{e} = (e_1, \dots, e_s) \in \mathbf{F}_{2^b}^s$ *be such that* $e_1, \dots, e_s$ *are distinct. If* ψ *is a non-trivial additive character of* $\mathbb{F}_{2^b}$*, then*

$$\left| \sum_{n \in \mathbb{F}_{2^b}} \psi \left(\sum_{j=1}^{s} d_j S(n + e_j) \right) \right| \leq (2s - 2) 2^{b/2} + s + 1$$

To apply the above lemma, we will rewrite one absolute value. First, since A is a full-rank system of equations, we know that $A_{\bar{d}i} \neq 0$. Thus

$$\left| \sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}} \tilde{S}(x_{\bar{d}})) \right| = \left| \sum_{r_{\bar{d}}} \psi_z(A_{\bar{d}\bar{d}} S(x_{\bar{d}} + r_{\bar{d}}) + A_{\bar{d}\bar{d}} S(r_{\bar{d}})) \right|$$

To use the lemma, we will rewrite $r_{\bar{d}}$ with n, $s = 2$ and $\mathbf{d} = (A_{\bar{d}\bar{d}}, A_{\bar{d}\bar{d}})$. Also $\mathbf{e} = (x_{\bar{d}}, 0)$. Since $A_{\bar{d}\bar{d}} \neq 0$ and $x_{\bar{d}} \neq 0$ and $z \neq 0$, the conditions of the lemma are satisfied, which means that the absolute value is bounded by $2 \cdot 2^{b/2} + 3$.

We can now write the deviation from the expected value

$$\left| T - 2^{b(c+d-k)} \right| \leq \frac{1}{2^b} \sum_{z \neq 0} \left(2 \cdot 2^{b/2} + 3 \right)^{c+d-k+1}$$

$$\leq \left(2 \cdot 2^{b/2} + 3\right)^{c+d-k+1}$$

For $b \geq 8$, we can upper bound the expression as

$$\leq 5 \cdot 2^{b(c+d-k)}$$

Putting everything together, we know that the number of solutions to equation $R_{\bar{d}}$ is

$$\left|T - 2^{b(c+d-k)}\right| \leq 5 \cdot 2^{b(c+d-k)}$$

We are mostly interested in the upper bound, which gives

$$T \leq 6 \cdot 2^{b(c+d-k)}$$

As a result, the number of solutions to a system of equations with c variables and $k - d$ equations is at most

$$\left(6 \cdot 2^{b(c+d-k)}\right) \cdot 2^{2(k-d-1)}$$

□

t-Wise Independence of MiMC over Prime Field

4.1 Overview

In the second part of this thesis, we will extend our study to the t-wise independence of the MiMC cipher. The MiMC cipher is a substitution-permutation network with one block $k = 1$ and the cube function as the S-box.

MiMC was introduced to work over large fields of prime order $\mathbb{F}_p$ or a power of 2, $\mathbb{F}_{2^b}$. We will restrict our attention to prime field case.

We will make a similar assumption with AES. That is, we will assume that we have independently and uniformly sampled keys for each round of MiMC. Below is a schematic representation of n rounds of MiMC, with input x and output y, for $x, y \in \mathbb{F}_p$.

$$x \xrightarrow{\texttt{ARK}} u_1 \xrightarrow{\texttt{Cube}} v_1 \xrightarrow{\texttt{ARK}} u_2 \xrightarrow{\texttt{Cube}} \dots \xrightarrow{\texttt{ARK}} u_n \xrightarrow{\texttt{Cube}} v_n = y$$

To show that n rounds of MiMC are close to t-wise independent, we will show that for any t-tuple of distinct inputs $(x_1, \dots, x_t)$, the distribution of the t-tuple of outputs is close to the uniform distribution over all t-tuples of distinct outputs.

Representing intermediate values. Since total variation distance is convex, without loss of generality we can assume that our input is deterministic. Let the random key in round i be r_i and uniformly sampled from $\mathbb{F}_p$. Then the trail of t distinct plaintexts in MiMC looks like below

$$\begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{matrix} \rightarrow \begin{matrix} (x_1+r_1)^3 \\ (x_2+r_1)^3 \\ \vdots \\ (x_t+r_1)^3 \end{matrix} \rightarrow \cdots \rightarrow \begin{matrix} \left(\ldots\left((x_1+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 \\ \left(\ldots\left((x_2+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 \\ \vdots \\ \left(\ldots\left((x_t+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 \end{matrix}$$

For simplicity, we will write the intermediate steps as $F_i^{(j)}$ where

$$F_i^{(j)} = \left(\ldots\left((x_i+r_1)^3+r_2\right)^3+\cdots+r_j\right)^3$$

Note that each $F_i^{(j)}$ is a j-variate polynomial. To quantify what it means to be 'close' to t-wise independence, we will use pointwise distance. Note that this is a stronger notion and implies total variation distance.

Definition 9. *A distribution P is pointwise ϵ-close to distribution Q if for all elements x of their state space*

$$\mathbb{P}[P=x] \leq (1+\epsilon)\cdot\mathbb{P}[Q=x]$$

We will have Q to be the uniform distribution over all t-tuples with distinct elements and P to be the joint distribution of the t polynomials after n rounds of MiMC. Since Q is uniform over all $\binom{p}{t}t!$ possible tuples with distinct elements we can write the pointwise condition for P as follows.

Corollary 1. *The distribution P is pointwise ϵ-close to t-wise independent if for all t-tuples $(y_1,\ldots,y_t)$*

$$\mathbb{P}[P=(y_1,\ldots,y_t)] \leq \begin{cases} \frac{1+\epsilon}{\binom{p}{t}t!} & \text{all } y_i \text{ are distinct} \\ 0 & \text{otherwise} \end{cases}$$

From distributions to solutions to systems of equations. The randomness of P comes exclusively from the set of random keys $(r_1,\ldots,r_n)$. Thus we can relate the probability of the tuple $(y_1,\ldots,y_t)$ under P to the number of solutions to the following system of t equations and n unknowns.

$$\begin{aligned} F_1^{(n)} &= y_1 \\ F_2^{(n)} &= y_2 \\ &\dots \\ F_t^{(n)} &= y_t \end{aligned}$$

In particular, if there are $N(y_1, \dots, y_t)$ solutions to this system, the probability of the $(y_1, \dots, y_t)$ tuple appearing is equal to

$$\frac{N(y_1, \dots, y_t)}{p^n}$$

So we can again rewrite the condition of pointwise ϵ-close to t-wise independent using the number of solutions to the system of equations above.

Corollary 2. *The distribution P is pointwise ϵ-close to t-wise independent if for all t-tuples $(y_1, \dots, y_t)$*

$$N(y_1, \dots, y_t) \leq \begin{cases} \frac{(1+\epsilon)p^n}{\binom{p}{t}t!} & \text{all } y_i \text{ are distinct} \\ 0 & \text{otherwise} \end{cases}$$

Below we formalize how pointwise distance implies total variation distance.

Claim 7. *If P is pointwise ϵ-close to Q, then*

$$\|P - Q\|_{TV} \leq \epsilon$$

Proof. We can write the total variation distance between two probability distributions as

$$\|P - Q\|_{TV} = \sup_{A \in \Omega} P(A) - Q(A)$$

Where Ω is the set of all t-tuples. If a t-tuple $y = (y_1, \dots, y_t)$ does not have distinct elements, then $P(y) = Q(y) = 0$. So we can consider only subsets A that do not include such tuples.

The pointwise guarantee gives that any t-tuple y with distinct elements satisfies

$$P(y) - Q(y) \leq \frac{\epsilon}{\binom{p}{t}t!}$$

Since $|A| \leq \binom{p}{t} t!$ we deduce that

$$\|P - Q\|_{TV} \leq |A| \frac{\epsilon}{\binom{p}{t} t!} \leq \epsilon$$

□

4.2 Polynomial Decomposition

To use exponential sums, we will first decompose a linear combination of the polynomials that represent our ciphertexts by isolating the random keys.

Lemma 16. (Polynomial Decomposition Lemma). *For any $b_1, \ldots, b_t \in \mathbb{F}_p$, we can write the linear combination of the ciphertexts as*

$$b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)} = C_3(r_n) + C_9(r_{n-1}) + \cdots + C_{3^n}(r_1) + \left(b_1 x_1^{3^n} + \cdots + b_t x_t^{3^n}\right)$$

where each $C_{3^j}(r_{n+1-j})$ is a degree-3^j polynomial in r_{n+1-j}, whose coefficients depend on $r_{n-j}, \ldots, r_1$ and $x_1, \ldots, x_t$ and $b_1, \ldots, b_t$.

Proof. We will expand $F_i^{(n)}$ as a cubic function of r_n. Indeed, $F_i^{(n)}$ is expanded to $\left(F_i^{(n-1)} + r_n\right)^3 = r_n^3 + 3r_n^2 F_i^{(n-1)} + 3r_n \left(F_i^{(n-1)}\right)^2 + \left(F_i^{(n-1)}\right)^3$ and we have.

$$\begin{aligned} b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)} = {} & r_n^3 (b_1 + \cdots + b_t) \\ & + 3r_n^2 \left(b_1 F_1^{(n-1)} + \cdots + b_t F_t^{(n-1)}\right) \\ & + 3r_n \left(b_1 \left(F_1^{(n-1)}\right)^2 + \cdots + b_t \left(F_t^{(n-1)}\right)^2\right) \\ & + \left(b_1 \left(F_1^{(n-1)}\right)^3 + \cdots + b_t \left(F_t^{(n-1)}\right)^3\right) \end{aligned}$$

We will write $C_3(r_n)$ as the cubic function that includes the r_n^3, r_n^2 and r_n terms. The constant term can be expanded to isolate r_{n-1}. The $\left(F_1^{(n-1)}\right)^3$ terms are a polynomial of degree 9 in terms of r_{n-1}

$$
\begin{aligned}
b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)} &= C_3(r_n) + \left(b_1 \left(F_1^{(n-1)} \right)^3 + \cdots + b_t \left(F_t^{(n-1)} \right)^3 \right) \\
&= C_3(r_n) \\
&+ r_{n-1}^9 \left(b_1 + \cdots + b_t \right) \\
&+ 9 r_{n-1}^8 \left(b_1 F_1^{(n-2)} + \cdots + b_t F_t^{(n-2)} \right) \\
&\cdots \\
&+ 9 r_{n-1} \left(b_1 \left(F_1^{(n-2)} \right)^8 + \cdots + b_t \left(F_t^{(n-2)} \right)^8 \right) \\
&+ \left(b_1 \left(F_1^{(n-2)} \right)^9 + \cdots + b_t \left(F_t^{(n-2)} \right)^9 \right)
\end{aligned}
$$

We will write $C_9(r_{n-1})$ as the degree-9 function that includes the $r_{n-1}^9, \ldots, r_{n-1}$ terms. The constant term can be again expanded to isolate r_{n-2}. In fact, we can repeat the following procedure n times to decompose the expression as

$$
b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)} = C_3(r_n) + C_9(r_{n-1}) + \cdots + C_{3^n}(r_1) + \left(b_1 x_1^{3^n} + \cdots + b_t x_t^{3^n} \right)
$$

where each C_k (for k a power of 3) looks like below. Here we write $j = n - \log_3 k$.

$$
\begin{aligned}
C_k(r) = r^k (b_1 + \cdots + b_t) &+ \binom{k}{1} \cdot r^{k-1} \left(b_1 F_1^{(j)} + \cdots + b_t F_t^{(j)} \right) + \ldots \\
&+ r \cdot \binom{k}{k-1} \cdot \left(b_1 \left(F_1^{(j)} \right)^{k-1} + \cdots + b_t \left(F_t^{(j)} \right)^{k-1} \right)
\end{aligned}
$$

□

4.3 Solutions to System of Equations and Exponential Sums

In this section we sketch a possible way to prove t-wise independence of MiMC. A desirable result would be of the following form:

Consider the MiMC cipher over $\mathbb{F}_p$ repeated for n rounds with a new uniform and independent random key sampled for each round. The distribution of the ciphertexts of t distinct plaintexts is pointwise ϵ-close to t-wise independent, for

$$\epsilon \leq p^{-O(n)+t}.$$

Such a statement would imply that in particular, for large enough p, $n = O(t)$ rounds of MiMC are enough to reach p^{-1}-close to t-wise independent.

We hope to achieve this result by bounding the number of solutions to a system of polynomial equations, and this will imply a bound on the distance to t-wise independence.

Conjecture 1. *The number of solutions to the system of equations*

$$\begin{cases} \left(\ldots\left((x_1+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 & = y_1 \\ \left(\ldots\left((x_2+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 & = y_2 \\ & \cdots \\ \left(\ldots\left((x_t+r_1)^3+r_2\right)^3+\cdots+r_n\right)^3 & = y_t \end{cases}$$

over the finite field $\mathbb{F}_p$, for p prime, is at most

$$\left(1+p^{-O(n)+t}\right)\cdot\frac{p^n}{\binom{p}{t}t!}$$

Ideas towards a proof. We will use exponential sums to bound the number of solutions. For a non-trivial additive character $\psi : \mathbb{F}_p \to \mathbb{C}$

$$N(y_1,\ldots,y_t) = \sum_{r_1,\ldots,r_n}\left(\frac{1}{p}\sum_{b_1}\psi\left(b_1(F_1^{(n)}-y_1)\right)\right)\ldots\left(\frac{1}{p}\sum_{b_t}\psi\left(b_t(F_t^{(n)}-y_t)\right)\right)$$

Rearranging gives

$$N(y_1,\ldots,y_t) = \frac{1}{p^t}\sum_{b_1,\ldots,b_t}\psi(-b_1y_1-\cdots-b_ty_t)\sum_{r_1,\ldots,r_n}\psi\left(b_1F_1^{(n)}+\cdots+b_tF_t^{(n)}\right)$$

As is typical in exponential sums, we will compute the summation with $b_1 = \cdots = b_t = 0$ separately. In that case the polynomial in the second summation is identically zero and the second sum is equal to p^n.

$$\Rightarrow N(y_1,\ldots,y_n) - p^{n-t}$$
$$= \frac{1}{p^t} \sum_{b_1,\ldots,b_t \neq (0,\ldots,0)} \psi\left(-b_1 y_1 - \cdots - b_t y_t\right) \sum_{r_1,\ldots,r_n} \psi\left(b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)}\right)$$

Note that $b_1 F_1^{(n)} + \cdots + b_t F_t^{(n)}$ is an n-variate polynomial. Typically, we can bound the character sums of such polynomials using Deligne's theorem [Del74, Del80]. Unfortunately, our polynomial does not necessarily satisfy the smoothness conditions to apply Deligne's theorem.

We will try to use the polynomial decomposition of the previous section to create polynomials that satisfy them.

$$\Rightarrow N(y_1,\ldots,y_n) - p^{n-t} = \frac{1}{p^t}$$
$$\sum_{b_1,\ldots,b_t \neq (0,\ldots,0)} \psi\left(b_1(x_1^{3^n} - y_1) + \cdots + b_t(x_t^{3^n} - y_t)\right) \sum_{r_1} \psi\left(C_{3^n}(r_1)\right) \cdots \sum_{r_n} \psi\left(C_3(r_n)\right)$$

Let us now turn our attention to Deligne's theorem for univariate polynomials

Lemma 17. (Deligne [Del74, Del80]). *Let f be a polynomial in $\mathbb{F}_p[x]$ of positive degree d. For a non-trivial additive character ψ, define the complete exponential sum*

$$S(f;p) = \sum_{x \in \mathbb{F}_p} \psi(f(x))$$

For prime p that satisfies $(p,d) = 1$ we can bound the absolute value of this exponential sum by

$$|S(f;p)| \leq (d-1)p^{1/2}$$

Deligne's theorem provides us with sharp bounds on the exponential sums of $C_{3^j}(r_{n+1-j})$ as long as two conditions are met

- $C_{3^j}(r_{n+1-j})$ is not the zero polynomial
- The total degree of $C_{3^j}(r_{n+1-j})$ and the size of the field are coprime, that is $(3^j, p) = 1$.

Since p is a prime number larger than 3, then $(3^j, p) = 1$ and thus the second condition will hold. When it comes to the first condition, it turns out that for large enough j, $C_{3^j}(r_{n+1-j})$ cannot be the zero polynomial.

Claim 8. *The polynomial $C_{3^j}(r_{n+1-j})$ can never be zero for $t \leq 3^j < p$, unless $b_1 = \cdots = b_t = 0$.*

Proof. For $C_{3^j}(r_{n+1-j})$ to be the zero polynomial, we want the coefficients of all powers of r_{n+1-j} to equal 0. Since $3^j < p$, the binomial coefficients are non-zero. What remains is 3^j equations in terms of $b_1, \ldots, b_t$ and $F_1^{(n-j)}, \ldots, F_t^{(n-j)}$, which can be conveniently written as a matrix vector product.

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ F_1^{(n-j)} & F_2^{(n-j)} & \ldots & F_t^{(n-j)} \\ \left(F_1^{(n-j)}\right)^2 & \left(F_2^{(n-j)}\right)^2 & \ldots & \left(F_t^{(n-j)}\right)^2 \\ \vdots & \vdots & \ddots & \vdots \\ \left(F_1^{(n-j)}\right)^{3^j-1} & \left(F_2^{(n-j)}\right)^{3^j-1} & \ldots & \left(F_t^{(n-j)}\right)^{3^j-1} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The matrix on the left is a $3^j \times t$ Vandermonde matrix. Assuming $3^j \geq t$, we can restrict our attention to the first t equations, which give the square Vandermonde matrix V_t. This matrix is invertible as long as the $F_i^{(j)}$ are distinct. The $F_i^{(j)}$ are the intermediate MiMC ciphertexts for different plaintexts, so they are distinct. Thus, we can multiply both sides by V_t^{-1} on the left to obtain

$$V_t^{-1} V_t \mathbf{b} = V_t^{-1} \mathbf{0}$$
$$\Rightarrow \mathbf{0} = \mathbf{0}$$

As a result, $C_{3^j}(r_{n+1-j})$ is the zero polynomial iff $b_1 = \cdots = b_t = 0$. □

This allows us to bound the individual $C_{3^j}(r_{n+1-j})$ for $j \geq \log_3 t$. The question that remains is whether we can use these to bound the quantity $N(y_1, \ldots, y_n)$.

Proof of Lemma 5

Lemma 5. *For positive integers α, k with $\alpha \leq k$*

$$\sum_{i=0}^{\alpha-1}(-1)^i \cdot \binom{k}{\alpha-i}\binom{k-1+i}{k-1} = (-1)^{\alpha-1}\binom{k+\alpha-1}{k-1}$$

In particular, when expanding the summation the equality looks as follows

$$\binom{k}{\alpha}\binom{k-1}{k-1} - \binom{k}{\alpha-1}\binom{k}{k-1} + \cdots + (-1)^{\alpha-1}\binom{k}{1}\binom{k+\alpha-2}{k-1} = (-1)^{\alpha-1}\binom{k+\alpha-1}{k-1}$$

Proof. We will prove this by induction on α.

Base case $\alpha = 1$. The base case is trivial since

$$LHS = \binom{k}{1}\binom{k-1}{k-1} = k = \binom{k-1+1}{k-1} = RHS$$

Inductive Argument. Assume that the statement holds for $\alpha - 1$. We will call $LHS_{\alpha-1}$ the left-hand side of the lemma statement with $\alpha - 1$.

Back to our statement for α, we will rewrite the first $\alpha-1$ terms using the identity

$$\binom{k}{\alpha} = \frac{k-\alpha+1}{\alpha} \cdot \binom{k}{\alpha-1}$$

The left-hand side becomes

$$\binom{k}{\alpha}\binom{k-1}{k-1} - \binom{k}{\alpha-1}\binom{k}{k-1} + \binom{k}{\alpha-2}\binom{k+1}{k-1} - \cdots + (-1)^{\alpha-1}\binom{k}{1}\binom{k+\alpha-2}{k-1}$$

$$= \frac{k-\alpha+1}{\alpha} \cdot \binom{k}{\alpha-1}\binom{k-1}{k-1} - \cdots + (-1)^{\alpha-2}\frac{k-1}{2} \cdot \binom{k}{1}\binom{k+\alpha-3}{k-1} + (-1)^{\alpha-1}\binom{k}{1}\binom{k+\alpha-2}{k-1}$$

Notice that the first $\alpha - 1$ terms look like $LHS_{\alpha-1}$. The difference is that the coefficients in front of them are not all the same, they are increasing from left to right. The i^{th} coefficient (starting from 0 on the left) is

$$c_i = \frac{k - \alpha + 1 + i}{\alpha - i}$$

It is clear that the first term has the smallest coefficient c_0. We will thus split all other coefficients $c_i = c'_i + c_0$. This way, we can group the c_0 terms to obtain $LHS_{\alpha-1}$ and perform the induction.

For $i > 0$, the value of c'_i is

$$\begin{aligned} c'_i &= \frac{k - \alpha + 1 + i}{\alpha - i} - \frac{k - \alpha + 1}{\alpha} \\ &= \frac{\alpha k - \alpha^2 + \alpha + \alpha i - \alpha k + \alpha^2 - \alpha + ik - \alpha i + i}{\alpha(\alpha - i)} \\ &= \frac{i(k+1)}{\alpha(\alpha - i)} \end{aligned}$$

After replacing the c_0 terms with $LHS_{\alpha-1}$, the left-hand side of our lemma becomes

$$\begin{aligned} =& c_0 \cdot LHS_{\alpha-1} - c'_1 \cdot \binom{k}{\alpha - 2}\binom{k}{k-1} + \dots \\ &+ (-1)^{\alpha-2} c'_{\alpha-2} \cdot \binom{k}{1}\binom{k+\alpha-3}{k-1} + (-1)^{\alpha-1}\binom{k}{1}\binom{k+\alpha-2}{k-1} \end{aligned}$$

The remainder of the proof will coalesce the middle terms with c'. Note that the first and last term are of the same magnitude as our desired result, so we will deal with them at the end.

Middle Terms. We will write these middle terms as f_i with the i^{th} term being the left-most (starting from 1)

$$f_i = (-1)^i \frac{i(k+1)}{\alpha(\alpha - i)} \cdot \binom{k}{\alpha - 1 - i}\binom{k-1+i}{k-1}$$

Turns out that the partial sums of these f_i's follow a specific pattern.

Claim 9. *For positive integers k, α, i such as $i < \alpha$ and $\alpha - 1 - i < k$, define f_i as*

$$f_i = (-1)^i \cdot \frac{i(k+1)}{\alpha(\alpha - i)} \cdot \binom{k}{\alpha - 1 - i}\binom{k-1+i}{k-1}$$

For any $\delta < \alpha$, consider the partial sum of the f_i's from 1 to δ. This partial sum has the following form

$$\sum_{i=1}^{\delta} f_i = (-1)^{\delta} \cdot \frac{\delta(k+\delta)}{\alpha(\alpha-1)} \cdot \binom{k}{\alpha - 1 - \delta} \cdot \binom{k-1+\delta}{k-1}$$

Using Claim 9 we deduce that the sum of all $\alpha - 2$ f_i's is equal to

$$\sum_{i=1}^{\alpha-2} f_i = (-1)^{\alpha-2} \frac{(\alpha-2)(k+\alpha-2)}{\alpha(\alpha-1)} \cdot \binom{k}{1} \cdot \binom{k+\alpha-3}{k-1}$$

Thus, we can write the equation

$$= c_0 \cdot LHS_{\alpha-1} + (-1)^{\alpha-2} \frac{(\alpha-2)(k+\alpha-2)}{\alpha(\alpha-1)} \cdot \binom{k}{1} \cdot \binom{k+\alpha-3}{k-1} + (-1)^{\alpha-1} \binom{k}{1}\binom{k+\alpha-2}{k-1}$$

Remaining Terms. We will now add the three remaining terms together. We start by making their binomial coefficients the same. We can change the binomial of the second term to look like the last term using the fraction in front

$$= c_0 \cdot LHS_{\alpha-1} + (-1)^{\alpha-2} \frac{(\alpha-2)}{\alpha} \cdot \binom{k}{1} \cdot \binom{k+\alpha-2}{k-1} + (-1)^{\alpha-1} \binom{k}{1}\binom{k+\alpha-2}{k-1}$$

We now substitute $LHS_{\alpha-1}$ and c_0

$$= \frac{k-\alpha+1}{\alpha} \cdot (-1)^{\alpha-2} \binom{k+\alpha-2}{k-1} + (-1)^{\alpha-2} \frac{(\alpha-2)}{\alpha} \cdot \binom{k}{1} \cdot \binom{k+\alpha-2}{k-1} + (-1)^{\alpha-1} \binom{k}{1}\binom{k+\alpha-2}{k-1}$$

Factorize their common binomial coefficient

$$= \left[k - \frac{k-\alpha+1+k(\alpha-2)}{\alpha}\right] (-1)^{\alpha-1} \binom{k+\alpha-2}{k-1}$$

$$= \left[\frac{\alpha k - k + \alpha - 1 - \alpha k + 2k}{\alpha}\right] (-1)^{\alpha-1} \binom{k+\alpha-2}{k-1}$$

$$= \left[\frac{k+\alpha-1}{\alpha}\right] (-1)^{\alpha-1} \binom{k+\alpha-2}{k-1}$$

Again, we can transform the binomial coefficient to the one we want using the fraction in front.

$$= (-1)^{\alpha-1}\binom{k+\alpha-1}{k-1}$$

This completes the proof. □

Proof of Claim 9

Claim 9. *For positive integers k, α, i such as $i < \alpha$ and $\alpha - 1 - i < k$, define f_i as*

$$f_i = (-1)^i \cdot \frac{i(k+1)}{\alpha(\alpha - i)} \cdot \binom{k}{\alpha - 1 - i}\binom{k - 1 + i}{k - 1}$$

For any $\delta < \alpha$, consider the partial sum of the f_i's from 1 to δ. This partial sum has the following form

$$\sum_{i=1}^{\delta} f_i = (-1)^{\delta} \cdot \frac{\delta(k+\delta)}{\alpha(\alpha-1)} \cdot \binom{k}{\alpha - 1 - \delta} \cdot \binom{k - 1 + \delta}{k - 1}$$

Proof. We will prove the claim by induction.

Base Case $\delta = 1$. It is trivial to verify that for $\delta = 1$ the claim is correct.

Inductive Argument. Assume that the statement holds for $\delta - 1$. To obtain the value of the summation for δ, we will just add the last term

$$\sum_{i=1}^{\delta} f_i = \sum_{i=1}^{\delta-1} f_i + f_{\delta}$$

From the inductive hypothesis, we can replace the summation on the right hand side

$$= (-1)^{\delta-1} \frac{(\delta-1)(k+\delta-1)}{\alpha(\alpha-1)} \cdot \binom{k}{\alpha - \delta} \cdot \binom{k - 2 + \delta}{k - 1} + (-1)^{\delta} \frac{\delta(k+1)}{\alpha(\alpha-\delta)} \cdot \binom{k}{\alpha - 1 - \delta}\binom{k - 1 + \delta}{k - 1}$$

The binomial coefficients of the first term are a constant factor away from the binomials of the second term.

$$\binom{k}{\alpha - \delta} = \frac{k - \alpha + \delta + 1}{\alpha - \delta} \cdot \binom{k}{\alpha - 1 - \delta}$$

$$\binom{k-2+\delta}{k-1} = \frac{\delta}{k-1+\delta} \cdot \binom{k-1+\delta}{k-1}$$

Hence we can substitute the above equations into our original equation and factorize the binomial coefficients out to get

$$= (-1)^{\delta} \left[-\frac{(\delta-1)(k+\delta-1)}{\alpha(\alpha-1)} \cdot \frac{k-\alpha+\delta+1}{\alpha-\delta} \cdot \frac{\delta}{k-1+\delta} + \frac{\delta(k+1)}{\alpha(\alpha-\delta)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

We can now compute the term in the brackets

$$= (-1)^{\delta} \left[-\frac{\delta(\delta-1)(k-\alpha+\delta+1)}{\alpha(\alpha-1)(\alpha-\delta)} + \frac{\delta(k+1)}{\alpha(\alpha-\delta)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

$$= (-1)^{\delta} \left[-\frac{\delta\left((\delta-1)(k-\alpha+\delta+1) - (k+1)(\alpha-1)\right)}{\alpha(\alpha-1)(\alpha-\delta)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

$$= (-1)^{\delta} \left[-\frac{\delta\left((\delta-1)(k+1) + (\delta-1)(\delta-\alpha) - (k+1)(\alpha-1)\right)}{\alpha(\alpha-1)(\alpha-\delta)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

$$= (-1)^{\delta} \left[-\frac{\delta\left((\delta-\alpha)(k+1) + (\delta-1)(\delta-\alpha)\right)}{\alpha(\alpha-1)(\alpha-\delta)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

Factorizing and cancelling the common terms results in the statement of our claim.

$$= (-1)^{\delta} \left[\frac{\delta(k+\delta)}{\alpha(\alpha-1)} \right] \cdot \binom{k}{\alpha-1-\delta} \binom{k-1+\delta}{k-1}$$

□

Proof of Lemma 4 - Diagonal Equality

Lemma 4. (Diagonal Equality). *The value of $T(c,d)$ only relies on the sum $|c|+|d|$ and not on the exact values of $|c|$ and $|d|$. In particular, 'exchanging' all zero entries of d with non-zero entries of c does not change the value of $T(\cdot,\cdot)$.*

$$T(c_{d-k}, \mathbf{k}) = T(c,d)$$

Proof. We will re-write $T(c,d)$ as a system of k linear equations with $2k$ variables and some additional constraints (such as some variables equal to 0 and some other variables not equal to 0). For now let's focus on the linear equations.

$$M \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}$$

We rearrange everything to the LHS

$$\Rightarrow \begin{bmatrix} M & I \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Now, we will rename the variables according to whether they are equal to 0 or unequal to 0. We will denote z_i to be the i^{th} zero variable (in arbitrary order) and n_i to be the i^{th} non-zero variable. We have a total of $|c|+|d| > k$ non-zero variables. We will also rearrange our variable vector from above

$$\Rightarrow \begin{bmatrix} M' \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ \vdots \\ n_{|c|+k} \\ z_1 \\ \vdots \\ z_{k-|c|} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Note that matrix M' is a permutation of the columns of $[M \ I]$. We can now apply Gaussian Elimination to our system of equations to obtain

$$\Rightarrow \begin{bmatrix} I & M'' \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ \vdots \\ n_{|c|+k} \\ z_1 \\ \vdots \\ z_{k-|c|} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

And now we can move the first identity matrix to the RHS to match the format of our original equation

$$\Rightarrow I \cdot \begin{bmatrix} n_1 \\ \vdots \\ n_k \end{bmatrix} + M'' \cdot \begin{bmatrix} n_{k+1} \\ \vdots \\ z_{k-|c|} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Rightarrow M'' \cdot \begin{bmatrix} n_{k+1} \\ \vdots \\ z_{k-|c|} \end{bmatrix} = \begin{bmatrix} n_1 \\ \vdots \\ n_k \end{bmatrix}$$

As a result, the number of solutions to our original system with constraints is equal to the number of solutions to this new system with constraints. But because we arranged the variables in such a way, all the variables on the RHS are non-zero. Thus, if M'' is a matrix with a full-branching number, then the number of solutions to this system is equal to $T(c_{d-k}, \mathbf{k})$.

But M'' has to be full branch. Otherwise, if we could find two vectors u, v such that $|u|+|v| < k+1$ and $M''u = v$, then we could invert our procedure above to obtain an equality $Mu' = v'$ for some vectors u', v' such that $|u'| + |v'| = |u| + |v| < k + 1$. This contradicts the maximal branching number of M. As a result, M'' is a matrix

with a maximal branching number and thus the number of solutions to this system of equations with additional constraints is $T(c_{d-k}, \mathbf{k})$. This concludes that moving zero entries from the LHS to the RHS does not change the number of values that make the specific transition.

$$T(c_{d-k}, \mathbf{k}) = T(c, d)$$

□

Bibliography

[AGR+16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. Mimc: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. *ASIACRYPT*, 2016.

[BCC19] Christina Boura, Anne Canteaut, and Daniel Coggia. A general proof framework for recent aes distinguishers. *Fast Software Encryption*, 2019.

[BODK+18] Achiya Bar-On, Orr Dunkelman, Nathan Keller, Eyal Ronen, and Adi Shamir. Improved key recovery attacks on reduced-round aes with practical data and memory complexities. *CRYPTO*, 2018.

[BS91] Eli Biham and Adi Shamir. Differential cryptanalysis of des-like cryptosystems. *Journal of Cryptology*, 1991.

[Dae95] Joan Daemen. Cipher and hash function design strategies based on linear and differential cryptanalysis. *Ph.D. Thesis, KU Leuven*, 1995.

[Del74] Pierrre Deligne. La conjecture de weil i. *Publications Mathématiques de l'IHÉS*, 1974.

[Del80] Pierrre Deligne. La conjecture de weil ii. *Publications Mathématiques de l'IHÉS*, 1980.

[EHN94] Jurgen Eichenauer-Herrmann and Harald Niederreiter. Bounds for exponential sums and their applications to pseudorandom numbers. *ACTA ARITHMETICA*, 1994.

[Gra19] Lorenzo Grassi. Mixture differential cryptanalysis: A new approach to distinguishers and attacks on round-reduced aes. *Fast Software Encryption*, 2019.

[GRR17] Lorenzo Grassi, Christian Rechberger, and Sondre Ronjom. A new structural-differential property of 5-round aes. *EUROCRYPT*, 2017.

[KHL+02] Ju-Sung Kang, Seokhie Hong, Sangjin Lee, Okyeon Yi, Choonsik Park, and Jongin Lim. Practical and provable security against differential and linear cryptanalysis for substitution-permutation networks. *Etri Journal*, 2002.

[KNR05] Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. *RANDOM*, 2005.

[LTV21] Tianren Liu, Stefano Tessaro, and Vinod Vaikuntanathan. The t-wise independence of substitution-permutation networks. *CRYPTO*, 2021.

[Nyb93] Kaisa Nyberg. Differentially uniform mappings for cryptography. *EUROCRYPT*, 1993.