Universal Hashing Based on Field Multiplication and (Near-)MDS Matrices

Koustabh Ghosh*[1][0000−0002−1820−2247], Jonathan Fuchs[1][0000−0001−5468−7846], Parisa Amiri Eliasi[1][0009−0003−2881−8314], and Joan Daemen[1][0000−0002−4102−0775]

Digital Security Group, Radboud University, Nijmegen, the Netherlands
firstname.lastname@ru.nl

Abstract. In this paper we propose a new construction for building universal hash functions, a specific instance called *multi-265*, and provide proofs for their universality. Our construction follows the key-then-hash parallel paradigm. In a first step it adds a variable length input message to a secret key and splits the result in blocks. Then it applies a fixed-length public function to each block and adds their results to form the output. The innovation presented in this work lies in the public function: we introduce the *multiply-transform-multiply*-construction that makes use of field multiplication and linear transformations. We prove upper bounds for the universality of key-then-hash parallel hash functions making use of a public function with our construction provided the linear transformation are maximum-distance-separable (MDS). We additionally propose a concrete instantiation of our construction multi-265, where the underlying public function uses a near-MDS linear transformation and prove it to be 2^{-154}-universal. We also make the reference code for multi-265 available.

Keywords: Primitive · Keyed hashing · Parallel · Forgery · Multi-265

1 Introduction

Message authentication code (MAC) functions strive to provide protection against forgery where forgery is defined according to the following scenario. An adversary gains access to a generation oracle and a verification oracle, where the generation oracle returns a tag given an input of a message (and nonce) and the verification oracle for a given tag, message (and nonce) returns whether the tag is valid or not. Forgery consists of a successfull verification query where the message (and nonce) was not used in any generation query.

There are two mainstream approaches to build MAC functions: the nonce based Wegman-Carter-Shoup (WCS) construction[18][20] or hash-then-encrypt construction[11]. In both approaches MAC functions consist of two phases: a

* Corresponding author

compression phase that converts a variable-length input into a fixed-size state under a secret key and a scrambling phase that takes this state and turns it into an output by the use of a pseudorandom function (PRF) or permutation (PRP). In this paper we consider only the compression phase and will refer to it as a *keyed hash function*.

The security of the compression phase of a hash-then-encrypt MAC function $\mathrm{F}_{\mathbf{K}}$ depends on the success probability, taken over the key space of an optimal attacker, to generate collisions at the output of F_K: Finding $\mathbf{M}$ and $\mathbf{M}^*$ such $\mathrm{F}_K(\mathbf{M}) = \mathrm{F}_K(\mathbf{M}^*)$. The ε-universality[19] of $\mathrm{F}_{\mathbf{K}}$ upper bounds the probability of obtaining such a collision. This can further be generalised to ε-Δuniversality[19], which is an upper bound for the success probability, taken over the key space of an optimal attacker, to find a particular output difference at the output of $\mathrm{F}_{\mathbf{K}}$: Finding $\mathbf{M}$ and $\mathbf{M}^*$ such $\mathrm{F}_K(\mathbf{M}) - \mathrm{F}_K(\mathbf{M}^*) = \Delta$. The latter is relevant for the security of WC(S) MAC functions.

In the literature we see three main categories of keyed hash function constructions. The first category builds them as modes of strong cryptographic primitive, like constructions based on cryptographic hash functions such as HMAC[2] and NMAC[2], or block ciphers such as CBC-MAC[3], CMAC[6] and PMAC[7]. The second category builds more efficient functions by applying simple algebraic constructions using multiplication and addition in a finite field such as GHASH[16] and Poly1305[4]. The third category does the same but in a different way: by using public permutations with a relatively small number of rounds.

In keyed hash functions of this third category, a message of variable length is parsed into blocks of a fixed size and added block-by-block to a long key. The latter is typically generated from a short key by means of a stream cipher or a key-schedule like computation. The resulting string can be processed in essentially two ways: parallel or serial. The parallel construction applies the public permutation to the blocks in parallel and adds the corresponding results to form the output. We see this construction in the compression phases of Kravatte[5] and Xoofff[8]. The serial construction applies the permutation serially to each block with the permutation result of the previous block added to it, much like CBC-MAC. This construction is an idealized version of the compression phase of Pelican-MAC[9], in the sense that in Pelican-MAC there is no key added to the message prior to compression, but rather it starts from a secret IV.

Fuchs et al. investigated the security of both constructions in[11]. They show that the universality of the parallel construction is at least as good as that of the serial construction, and can be much better. Moreover, both constructions have the same workload per block but the serial construction cannot be parallelized and therefore the parallel construction is superior.

In this paper we study a variant of the permutation-based parallel keyed hashing: instead of a public permutation we make use of a public function that is not invertible.

1.1 Our Contribution

In this paper we first generalize the results of [11] to the parallelization of a public function. The main innovation in this work lies in the public function: we introduce the *multiply-transform-multiply*-construction that makes use of field multiplication and linear transformations. We prove upper bounds for the universality of key-then-hash parallel hash functions making use of a public function with our construction provided the linear transformations are maximum-distance-separable (MDS)[10]. We prove that they are $2/p^n$-Δuniversal with the multiplication taking place in the field with p elements, where p is a prime, provided that the linear transformations employed are $n \times n$ MDS matrices.

In secure multi-party computation (MPC), fully homomorphic encryption (FHE), zero-knowledge (ZK) schemes, data is typically encoded in large prime fields. In various application, part of such a function or circuit call on symmetric cryptographic primitives such as a PRF, a symmetric encryption scheme, or a collision resistant keyed hash function. In such applications, the main bottleneck comes from the number of field applications in the underlying primitive[1][12]. The computational cost of our construction per input word, where word is an element of the chosen prime field, is only one field multiplication and a small number of field additions depending on the chosen linear transformation. Therefore, our construction, with a very low multiplicative cost, is especially suitable for use as a collision-resistant keyed hash function in such applications.

Furthermore this low multiplicative cost is also beneficial when masking is applied as protection against differential power analysis (DPA)[15]. In masking, each variable x is encoded in a number of shares $x_0, x_1, \ldots, x_{d-1}$, such that $x = x_0 + x_1 + \ldots + x_{d-1}$, where any subset of $d-1$ shares have a random distribution with the addition taking place in the underlying field. The linear parts of the algorithm, such as the MDS matrix and the key additions, can be performed on the shares separately. There are two main approaches for computing multiplications: the Ishai-Sahai-Wagner (ISW) approach[14] and threshold implementations[17]. In both techniques, the total computational cost increases quadratically with the number of shares d and linearly with the number of multiplications. Furthermore, ISW requires randomness, which also increases quadratically with the number of shares and linearly with the number of multiplications.

The additive cost of our construction can be further optimized by using specific types of near-MDS matrices, instead of MDS matrices. To that end we additionally propose a concrete instantiation of our construction called *multi-265*. Its public function makes use of the prime field with p elements, where $p = 2^{26} - 5$ and a 6×6 lightweight circulant near-MDS matrix with branch number 6. We prove it gives rise to a keyed hash function that is 2^{-154}-Δuniversal. Despite the fact that its matrix is not MDS, it is still $2/p^6$-Δuniversal.

1.2 Outline of the Paper

This paper is organised as follows. In Section 2, we remind the readers of key-then-hash functions and their universalities[11]. In Section 3, we generalize the

parallel construction of a public permutation[11] to a public function. In Section 4, we introduce the notations that will be used throughout this paper. In Section 5 we describe the propagation properties of field multiplication. In Section 6 we look at a simple construction for the public function that we call the duplicated field multiplication. In Section 7 we introduce the multiply-transform-multiply and prove bounds on its universality if the underlying linear transformation is MDS. In Section 8 we introduce our proposed keyed hash function multi-265, study its security and report on the implementation aspects.

2 Preliminaries

Security analysis in this work builds upon the results of [11] and to that end we adopt a terminology similar to that paper. We denote a public function as $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ where G and G′ are abelian groups $\langle \mathrm{G}, + \rangle$ and $\langle \mathrm{G}', + \rangle$.

The elements of G are called *blocks*. The set containing ℓ-block string is denoted as G^ℓ, i.e., $\mathrm{G}^\ell = \{(x_0, x_1, \ldots, x_{\ell-1}) \mid x_i \in \mathrm{G} \text{ for each } i = 0, 1, \ldots, \ell-1\}$. The set of strings of length 1 upto κ is denoted as $\mathrm{BS}(\mathrm{G}, \kappa) = \cup_{\ell=1}^{\kappa} \mathrm{G}^\ell$. We denote strings in bold uppercase letters, like $\mathbf{M}$, its blocks by M_i, where indexing starts from 0 and the length of that string by $|\mathbf{M}|$.

Let $X \in \mathrm{G}$ be a discrete random variable that has a value that depends on the key K. We denote the probability that a variable X has value x by $\Pr(X = x)$. In words, $\Pr(X = x)$ is the fraction of the keyspace for which variable X has value x. We call two variables independent if $\Pr(X = x, X' = x') = \Pr(X = x) \Pr(X' = x')$ for all $x, x' \in \mathrm{G}$.

The probability mass function (PMF) of a variable X, denoted as g_X, is the array of values $\Pr(X = x)$ over all values x. We have $g_X(x) = \Pr(X = x)$. Clearly, $\forall x : 0 \leq g_X(x) \leq 1$ and $\sum_x g_X(x) = 1$. As such, a PMF can be seen as a mapping $g\colon \mathrm{G} \to [0, 1]$.

For two independent random variables X and Y, let $Z = X + Y$. The PMF g_Z is given by the convolution of two PMFs g_X, g_Y and is denoted as $g_X * g_Y$.

$$g_Z = g_X * g_Y \iff \forall z : g_Z(z) = \sum_x g_X(x) g_Y(x - z)\,,$$

with $-$ is determined by the group operation of G and the summation done over $\mathbb{R}$. We further let g_X^{*n} with $n \in \mathbb{N}$ denote the convolution of g_X n-times.

2.1 ε and ε-Δuniversality

Let $\mathrm{F}_{\mathbf{K}}$ denote a keyed hash function where the key $\mathbf{K}$ is sampled uniformly at random from the key space. The security of a keyed hash function is measured by the probability of generating a collision at the output of $\mathrm{F}_{\mathbf{K}}$: distinct $\mathbf{M}, \mathbf{M}^*$ such that $\mathrm{F}_{\mathbf{K}}(\mathbf{M}) = \mathrm{F}_{\mathbf{K}}(\mathbf{M}^*)$. This probability is upper-bounded by the so called ε-universality. We further look at an even stronger notion of universality: ε-Δuniversality, which gives an upper-bound on the probability taken over all keys of two distinct inputs strings exhibiting a specific output difference.

Definition 1 (ε-universality[19]). *== A keyed hash function* F *is said to be ε-universal if for any distinct strings* $\mathbf{M}, \mathbf{M}^*$

$$\Pr[\mathrm{F}_{\mathbf{K}}(\mathbf{M}) = \mathrm{F}_{\mathbf{K}}(\mathbf{M}^*)] \leq \varepsilon\,.$$

Definition 2 (ε-Δuniversality[19]). *A keyed hash function* F *is said to be ε-Δuniversal if for any distinct strings* $\mathbf{M}, \mathbf{M}^*$ *and for all* $\Delta \in \mathrm{G}$

$$\Pr[\mathrm{F}_{\mathbf{K}}(\mathbf{M}) - \mathrm{F}_{\mathbf{K}}(\mathbf{M}^*) = \Delta] \leq \varepsilon\,.$$

2.2 Key-then-hash Functions

We study keyed hash functions that take as input elements of $\mathrm{BS}(\mathrm{G}, \kappa)$ and return an element of G'. The keys are elements of G^κ. When processing an input, the key is first added to the input and then an unkeyed function is applied to the result. This is a special case of keyed hash functions and such functions are called key-then-hash functions. A key-then-hash function is defined as: $\mathrm{F}\colon \mathrm{BS}(\mathrm{G}, \kappa) \to \mathrm{G}'$ with $\mathrm{F}_{\mathbf{K}}(\mathbf{M}) := \mathrm{F}(\mathbf{K} + \mathbf{M})$. The addition of two strings $\mathbf{M} = (M_0, M_1, \ldots, M_{|\mathbf{M}|-1})$ and $\mathbf{M}^* = (M_0^*, M_1^*, \ldots, M_{|\mathbf{M}^*|-1}^*)$ with $|\mathbf{M}| \leq |\mathbf{M}^*|$ is defined as $\mathbf{M}' := \mathbf{M} + \mathbf{M}^* = (M_0 + M_0^*, M_1 + M_1^*, \ldots, M_{|\mathbf{M}|-1} + M_{|\mathbf{M}|-1}^*)$ with $|\mathbf{M}'| = |\mathbf{M}|$. In Section 3 we demonstrate how to build such functions using a public function as the underlying primitive.

3 Parallel Universal Hashing

We first note that ε-universality of a key-then-hash function is upper bounded by the ε-Δuniversality of that function. To that end we now see how a public function can be parallelized to form a key-then-hash function and further prove upper bound on the ε-Δuniversality of such construction.

The analysis of a parallel universal hash construction using public permutations has been presented by Fuchs et al. in [11]. Using a similar approach, we generalize their results to a parallel universal hash construction built on a public function as its underlying primitive.

3.1 Construction

We adapt the parallelization of a public permutation to public functions in Algorithm 1 and depict it in Figure 1. The construction takes as parameters a public function $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ and a maximum string length κ. The inputs to the construction are a key $\mathbf{K} \in \mathrm{G}^\kappa$ and a string $\mathbf{M} \in \mathrm{BS}(\mathrm{G}, \kappa)$. The construction returns a digest $h \in \mathrm{G}'$.

Given any public function f, its parallelization is the key-then-hash function denoted as Parallel [f]. Since these are the key-then-hash functions we study, for the rest of the paper F and Parallel [f] will be used interchangeably to denote parallelized public functions. The key space of Parallel [f] is G^κ and as such we assume the existence of long keys with independent key blocks.

Algorithm 1: The parallelization Parallel [f]

Parameters: A public function $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ and a maximum string length κ
Inputs : A key $\mathbf{K} \in \mathrm{G}^{\kappa}$ and a message $\mathbf{M} \in \mathrm{BS}(\mathrm{G}, \kappa)$
Output : A digest $h \in \mathrm{G}'$

$\mathbf{x} \leftarrow \mathbf{M} + \mathbf{K}$
$h \leftarrow 0$
for $i \leftarrow 0$ **to** $|\mathbf{M}| - 1$ **do**
  $h \leftarrow h + \mathrm{f}(x_i)$
end
return h

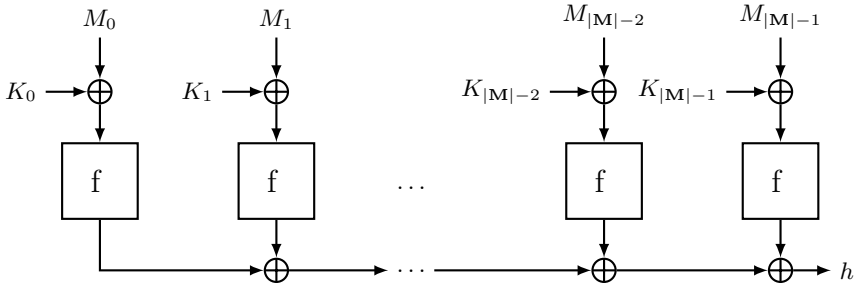

Fig. 1: The parallelization Parallel [f] adapted from [11].

3.2 Propagation Probabilities of Fixed-length Functions

Before we can investigate the universality of Parallel [f], we first look at the differential properties of the underlying fixed length function f.

Classically, a differential defined over the fixed input-length public function $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ is the tuple (A, Δ), where $A \in \mathrm{G}/\{0\}$ is called the input difference and $\Delta \in \mathrm{G}'$ is called the output difference. We now remind the reader of differential probability of a differential over fixed-length public functions.

Definition 3 (Differential probability). *Let* $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ *be a public function. The differential probability of a differential* (A, Δ) *of* f*, denoted as* $\mathsf{DP}_{\mathrm{f}}(A, \Delta)$*, is:*

$$\mathsf{DP}_{\mathrm{f}}(A, \Delta) = \frac{\#\{X \in \mathrm{G} \mid \mathrm{f}(X + A) - \mathrm{f}(X) = \Delta\}}{\#\mathrm{G}}\,.$$

We say that input difference A *propagates to output difference* Δ *with probability* $\mathsf{DP}_{\mathrm{f}}(A, \Delta)$*.*

The universality of a parallelized public permutation depends on the uniformity of the public permutation by Lemma 3[11]. Unlike a public permutation, the relative frequency of outputs of a non-bijective public function f is not constant. Thus in order to generalize the universality of parallelized public permutation to public function, we introduce the definition of *image probability* of f.

Definition 4 (Image probability). *Let* $\mathrm{f}\colon \mathrm{G} \to \mathrm{G}'$ *be a public function. The image probability of an output* $Z \in \mathrm{G}'$ *of* f*, denoted as* $\mathsf{IP}_{\mathrm{f}}(Z)$*, is the number of inputs that* f *maps to* Z *divided by the total number of possible inputs, namely,*

$$\mathsf{IP}_{\mathrm{f}}(Z) = \frac{\#\{X \in \mathrm{G} \mid \mathrm{f}(X) = Z\}}{\#\mathrm{G}}\,.$$

To obtain the ε-Δuniversality of $\mathrm{F} = \mathrm{Parallel}\,[\mathrm{f}]$, we need to obtain an upper bound of the maximum possible value of DP_{f} and IP_{f} over all differentials and outputs of the underlying fixed length public function f respectively. As such we denote them as:

$$\mathrm{MDP}_{\mathrm{f}} = \max_{(A,\Delta)} \mathsf{DP}_{\mathrm{f}}(A, \Delta) \quad \text{and} \quad \mathrm{MIP}_{\mathrm{f}} = \max_{Z} \mathsf{IP}_{\mathrm{f}}(Z)\,.$$

Furthermore we denote by DP_A and IP the probability mass functions $\mathsf{DP}_{\mathrm{f}}(A, Z)$ with $A \in \mathrm{G}$ fixed and $\mathrm{IP}_{\mathrm{f}}(Z)$ respectively.

3.3 Differentials over Parallel [f] and Their Differential Probability

The inputs to $\mathrm{F} = \mathrm{Parallel}\,[\mathrm{f}]$ are of variable lengths and as such the classical definition of differentials no longer work since two distinct strings may now differ in both value and length. With this distinction in mind, a difference between two strings is defined in [11] that is relevant to our approach.

Definition 5 (Difference between two strings [11]). *The difference between two strings* $\mathbf{M}$*,* $\mathbf{M}^*$ *with* $|\mathbf{M}| \leq |\mathbf{M}^*|$ *is defined as the pair* $(\mathbf{A}, \lambda) \in \mathrm{G}^{|\mathbf{M}|} \times \mathbb{Z}_{\geq 0}$*, where* $\mathbf{A} = \mathbf{M} - \mathbf{M}^* = (M_0 - M_0^*, M_1 - M_1^*, \ldots, M_{|\mathbf{M}|-1} - M_{|\mathbf{M}|-1}^*)$ *and* $\lambda = |\mathbf{M}^*| - |\mathbf{M}|$*.*

Now, given two strings $\mathbf{M}$ and $\mathbf{M}^*$, the probability that the strings result in an output difference Δ through F is determined by the difference between the strings.

Proposition 1 (Proposition 1[11]). *Given two strings* $\mathbf{M}, \mathbf{M}^*$ *with* $|\mathbf{M}| \leq |\mathbf{M}^*|$*, the probability that the strings result in an output difference* Δ *through* $\mathrm{F}_{\mathbf{K}}$ *is given by:*

$$\Pr[\mathrm{F}_{\mathbf{K}}(\mathbf{M}) - \mathrm{F}_{\mathbf{K}}(\mathbf{M}^*) = \Delta] = \frac{\#\{\mathbf{K} \in \mathrm{G}^{\kappa} \mid \mathrm{F}(\mathbf{A} + \mathbf{K}) - \mathrm{F}(0^{|\mathbf{A}|+\lambda} + \mathbf{K}) = \Delta\}}{\#\mathrm{G}^{\kappa}}\,,$$

where $(\mathbf{A}, \lambda)$ *is the difference between the strings* $\mathbf{M}$ *and* $\mathbf{M}^*$*.*

This naturally leads to the following definitions of generalised differentials and their DP.

Definition 6 (Generalized differentials and their DP [11]). *Given an input difference* $(\mathbf{A}, \lambda)$ *and output difference* Δ*, the differential probability of the differential* $(\mathbf{A}, \lambda, \Delta)$ *over* F*, denoted as* $\mathsf{DP}_{\mathrm{F}}(\mathbf{A}, \lambda, \Delta)$ *is given by*

$$\mathsf{DP}_{\mathrm{F}}(\mathbf{A}, \lambda, \Delta) = \frac{\#\{\mathbf{K} \in \mathrm{G}^{\kappa} \mid \mathrm{F}(\mathbf{A} + \mathbf{K}) - \mathrm{F}(0^{|\mathbf{A}|+\lambda} + \mathbf{K}) = \Delta\}}{\#\mathrm{G}^{\kappa}}\,.$$

Lemma 1 (DP of differentials over Parallel [f]). *The differential probability of a differential* $(\mathbf{A}, \lambda, \Delta)$ *over Parallel*[f] *is given by*

$$\mathsf{DP}_{\mathrm{F}}(\mathbf{A}, \lambda, \Delta) = \mathsf{DP}_{A_0} * \mathsf{DP}_{A_1} * \ldots * \mathsf{DP}_{A_{|\mathbf{A}|-1}} * \mathrm{IP}^{*\lambda}(\Delta)\,.$$

Proof. Since the keys that are added to each of the blocks are mutually independent, the difference in the outputs of the first $|\mathbf{A}|$ blocks can be seen as the outcomes of independent stochastic variables whose distributions are given by the PMFs DP_{A_i} respectively, while the outputs of the last λ blocks can be seen as the outcomes of independent stochastic variables whose distribution is given by the PMF IP. Naturally the PMF of the input difference (A, λ) to F denoted as $\mathsf{DP}_{\mathrm{F}}(A, \lambda)$ is given by

$$\mathsf{DP}_{\mathrm{F}}(A, \lambda) = \mathsf{DP}_{A_0} * \mathsf{DP}_{A_1} * \ldots * \mathsf{DP}_{A_{|\mathbf{A}|-1}} * \mathrm{IP}^{*\lambda}\,. \qquad \square$$

Theorem 1 (ε-Δuniversality of Parallel [f]). *The parallelization of a public function* f*, Parallel*[f]*, is* $\max\{\mathrm{MDP_f}, \mathrm{MIP_f}\}$*-*$\Delta$*universal.*

Proof. We first note that if for independent random variables g_X, g_Y and g_Z, $g_Z = g_X * g_Y$, then it follows directly from the definition of g_Z, that

$$\max_z g_Z(z) \leq \max\left(\max_x g_X(x), \max_y g_Y(y)\right)\,.$$

By applying this relation to Lemma 1, we can upper bound DP_{F} as

$$\begin{aligned}\max_{\mathbf{A},\lambda,\Delta} \mathsf{DP}_{\mathrm{F}}(\mathbf{A}, \lambda, \Delta) &\leq \max\left\{\max_{A,\Delta\in\mathrm{G}} \mathsf{DP}_{\mathrm{f}}(A, \Delta), \max_{Z\in\mathrm{G}'} \mathsf{IP}_{\mathrm{f}}(Z)\right\} \\ &= \max\{\mathrm{MDP_f}, \mathrm{MIP_f}\}\,.\end{aligned}$$

$\square$

The tightness of the ε-Δuniversality bound in Theorem 1 depends solely on the tightness of the bounds for $\mathrm{MDP_f}$ and $\mathrm{MIP_f}$ of the underlying public function f since for single block inputs, Parallel [f] = f. Furthermore the bound obtained for the ε-Δuniversality in Theorem 1 is consistent with the results obtained by Fuchs et al. for the ε-Δuniversality of a parallelized public permutation in Theorem 2[11]. Indeed when f is a public permutation, the PMF IP for f is simply the uniform distribution and as such, the ε-Δuniversality of a parallelized public permutation is determined solely by its $\mathrm{MDP_f}$.

4 Notations

$\mathbb{F}_p$ denotes the prime field with p elements and $\mathbb{F}_p^{2n}$ denotes the cartesian product of $\mathbb{F}_p$ $2n$-times. For the public functions proposed in this paper, $\mathrm{G} = \mathbb{F}_p^{2n}$ and $\mathrm{G}' = \mathbb{F}_p^{2n}$ for some prime p and integer $n \geq 1$. As such the input and the output to our public function are both $2n$-tuples, where all elements of the tuple are elements from the finite field $\mathbb{F}_p$ and the block string space is $\mathrm{BS}(\mathbb{F}_p^{2n}, \kappa)$.

We represent $X \in \mathrm{G} = \mathbb{F}_p^{2n}$ as $X = (x_0, x_1, \ldots, x_{n-1}, y_0, y_1, \ldots, y_{n-1})^\intercal$. For simplicity, we slightly abuse the notations to denote $X = (\mathbf{x}, \mathbf{y})$, where $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})^\intercal \in \mathbb{F}_p^n$ and $\mathbf{y} = (y_0, y_1, \ldots, y_{n-1})^\intercal \in \mathbb{F}_p^n$. Similarly input differences and key blocks are denoted as $A = (\mathbf{a}, \mathbf{b})$ and $K = (\mathbf{h}, \mathbf{k})$ respectively. An output $Z \in \mathrm{G}' = \mathbb{F}_p^{2n}$ is denoted as $Z = (z_0, z_1, \ldots, z_{2n-1})^\intercal$ and similarly output difference Δ is given by $\Delta = (\delta_0, \delta_1, \ldots, \delta_{2n-1})$.

The number of non-zero components in a vector $\mathbf{x} \in \mathbb{F}_p^n$ is the hamming weight of $\mathbf{x}$ that is denoted as $\mathrm{w}(\mathbf{x})$ and we denote $(0, 0, \ldots, 0)^\intercal \in \mathbb{F}_p^n$ as 0^n. Since multiplication in $\mathbb{F}_p$ is an integral part of our public function, we first look at its differential properties.

5 Differential Properties of Field Multiplication

We first consider as public function $\mathrm{f}\colon \mathbb{F}_p \times \mathbb{F}_p \to \mathbb{F}_p$ to denote the multiplication in $\mathbb{F}_p$. We remind the reader of the differential properties of field multiplication.

Lemma 2. *When* f *is the field multiplication, the image probabilities of its outputs are given by*

$$\mathsf{IP}_{\mathrm{f}}(Z) = \begin{cases} \frac{2p-1}{p^2} & , \text{ when } Z = 0 \\ \frac{p-1}{p^2} & , \text{ otherwise.} \end{cases}$$

Proof. For $Z = 0$, $f(x, y) = xy = 0$ implies $x = 0$ or $y = 0$. So $\mathsf{IP}_{\mathrm{f}}(0) = \frac{2p-1}{p^2}$. For $Z = z \neq 0$, $xy = z$ implies $x = z/y$ with $y \neq 0$ and thus $\mathsf{IP}_{\mathrm{f}}(Z) = \frac{p-1}{p^2}$. □

So, for field multiplication we have $\mathrm{MIP}_{\mathrm{f}} = \frac{2p-1}{p^2}$ and is achieved only for $Z = 0$.

Lemma 3. *When* f *is the field multiplication,* $\mathsf{DP}_{\mathrm{f}}(A, \Delta) = \frac{1}{p}$ *for any* $A \in \mathbb{F}_p^2$ *and any* $\Delta \in \mathbb{F}_p$.

Proof. An input difference $A = (a, b)$ propagates to the output difference $\Delta = \delta$ under a key $K = (h, k)$ for f if:

$$(a + h)(b + k) - hk = \delta \quad \Longrightarrow \quad bh + ak + ab = \delta \,. \tag{1}$$

(1) describes a line in $\mathbb{F}_p \times \mathbb{F}_p$ with p points and thus $\mathrm{DP}(A, \Delta) = 1/p$. □

So, when f is the field multiplication, $\mathrm{MDP}_{\mathrm{f}} = 1/p$ and $\mathrm{MIP}_{\mathrm{f}} = (2p - 1)/p^2$. Thus by Theorem 1, we see that Parallel [f] is ε-Δuniversal, where $\varepsilon = \frac{2p-1}{p^2}$.

6 Duplicated Multiplication as Public Function

In our quest to build a public function based on field multiplication, we now look at a slightly more complicated public function, that we call *duplicated field multiplication*. It is defined as follows:

$$\mathrm{f}\colon \mathbb{F}_p^2 \to \mathbb{F}_p^2 : \mathrm{f}(x, y) = (xy, (x + u)(y + v)) \,. \tag{2}$$

for some constants $u, v \in \mathbb{F}_p \setminus \{0\}$. Here instead of computing only one multiplication, we compute two multiplications in parallel, where the input of the second multiplication is offset by (u, v).

Lemma 4. *When* f *is the duplicated multiplication,* $\max_{Z \in \mathbb{F}_p^2} \mathsf{IP}_{\mathrm{f}}(Z) = \frac{2}{p^2}$

Proof. Let $Z = 0^2 = (0, 0)$. $\mathrm{f}(x, y) = (0, 0)$ implies $(xy, (x + u)(y + v)) = (0, 0)$. Thus

$$xy = 0 \quad \text{and} \quad (x + u)(y + v) = 0 \; .$$

This holds only for $(0, -v)$ and $(-u, 0)$. Thus, in this case the $\mathsf{IP}_{\mathrm{f}}(0^2) = \frac{2}{p^2}$. When $Z = (z_0, z_1) \neq (0, 0)$, $f(x, y) = Z$ implies $xy = z_0$ and $(x + u)(y + v) = z_1$. Now since $u \neq 0$,

$$(x + u)(y + v) = z_1 \implies y = u^{-1}(z_1 - z_0 - vx - uv) \; .$$

Substituting the value of y in $xy = z_0$, we obtain a quadratic equation in y. This has at most 2 solutions. So for $Z \neq (0, 0)$, $\mathsf{IP}_{\mathrm{f}}(Z) \leq \frac{2}{p^2}$. □

So, for the duplicated field multiplication, $\mathrm{MIP}_{\mathrm{f}} = \frac{2}{p^2}$, which is indeed achieved for $Z = 0^2$.

Lemma 5. *When* f *is the duplicated multiplication,* $\mathsf{DP}_{\mathrm{f}}(A, \Delta)$ *is given by:*

$$\mathsf{DP}_{\mathrm{f}}(A, \Delta) = \begin{cases} \frac{1}{p} & , \textit{ when } va + ub = \delta_2 - \delta_1 \\ 0 & , \textit{ otherwise.} \end{cases}$$

Proof. An input difference $A = (a, b)$ propagates to an output difference $\Delta = (\delta_1, \delta_2)$ under a key $K = (h, k)$ if

$$\begin{aligned} (a + h)(b + k) - hk &= \delta_1 \; , \\ (a + h + u)(b + k + v) - (h + u)(k + v) &= \delta_2 \; . \end{aligned}$$

This simplifies to

$$bh + ak + ab = \delta_1 \quad \text{and} \quad va + ub = \delta_2 - \delta_1 \; . \tag{3}$$

Thus $\mathsf{DP}_{\mathrm{f}}(A, \Delta) > 0$ iff $va + ub = \delta_2 - \delta_1$ and in that case we must have $bh + ak + ab = \delta_1$, which again describes a line in $\mathbb{F}_p \times \mathbb{F}_p$, i.e., $\mathsf{DP}_{\mathrm{f}}(A, Z) = 1/p$. □

So $\mathrm{MDP}_{\mathrm{f}} = \frac{1}{p} > \mathrm{MIP}_{\mathrm{f}} = \frac{2}{p^2}$. Thus we conclude by Theorem 1 that parallelized duplicated multiplication is $1/p$-Δuniversal.

7 The Multiply-transform-multiply Construction

We now define a construction for building a public function from finite field multiplication and two linear transformations, that we call the *multiply-transform-multiply* (MTM) construction. We show that instances of this construction with maximum-distance-separable (MDS) linear transformations provide very good uniformity in the parallelization. As such we briefly remind the reader of the branch number of a matrix[10] and the definition of MDS matrices in terms of their branch numbers.

Definition 7 (Branch number). *Given a $n \times n$ matrix* N *defined over a field $\mathbb{F}_p$, its branch number is defined as* $\min_{\mathbf{x} \in \mathbb{F}_p^n / \{0^n\}} (\mathrm{w}(\mathbf{x}) + \mathrm{w}(\mathrm{N} \cdot \mathbf{x}))$.

We have a trivial upper-bound for the branch number of a matrix given by: branch number of $\mathrm{N} \leq n + 1$.

Definition 8 (MDS matrix). *An $n \times n$ matrix* N *defined over a over a field $\mathbb{F}_p$ is said to be MDS if* N *has branch number $n + 1$.*

Before looking at the multiply-transform-multiply construction, we define the coordinate-wise product of vectors that will help us to explain the construction.

Definition 9 (Coordinate-wise product). *Given $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})^\intercal$ and $\mathbf{y} = (y_0, y_1, \ldots, y_{n-1})^\intercal$, their coordinate wise product denoted as $\mathbf{x} \odot \mathbf{y}$ is given by*

$$\mathbf{x} \odot \mathbf{y} = (x_0 y_0, x_1 y_1, \ldots, x_{n-1} y_{n-1})^\intercal \ .$$

Definition 10 ($\mathrm{MTM}(n, p)$). *For a positive integer $n > 1$ and prime p such that $\log_2 p \geq n$, $\mathrm{MTM}(n, p)$ denotes a family of functions where any function $\mathrm{f}[\alpha, \beta] \in \mathrm{MTM}(n, p)$ is given by:*

$$\mathrm{f}[\alpha, \beta] \colon \mathbb{F}_p^{2n} \mapsto \mathbb{F}_p^{2n} \colon \mathrm{f}[\alpha, \beta](X) = \mathrm{f}[\alpha, \beta](\mathbf{x}, \mathbf{y}) = (\mathbf{x} \odot \mathbf{y}, \mathrm{N}_\alpha \cdot \mathbf{x} \odot \mathrm{N}_\beta \cdot \mathbf{y}) \ ,$$

where N_α, N_β are any $n \times n$ matrices.

To simplify notations, we will denote $\mathrm{N}_\alpha \cdot \mathbf{x} = \mathbf{p}$ and $\mathrm{N}_\beta \cdot \mathbf{y} = \mathbf{q}$, which means $\mathrm{f}[\alpha, \beta](\mathbf{x}, \mathbf{y}) = (\mathbf{x} \odot \mathbf{y}, \mathbf{p} \odot \mathbf{q})$. The matrices N_α and N_β are denoted as $\mathrm{N}_\alpha = \left[\alpha_{i,j}\right]$ and $\mathrm{N}_\beta = \left[\beta_{i,j}\right]$ with $0 \leq i, j \leq n - 1$.

We now provide an intuitive design rationale for our construction. Our goal is to build a key-then-hash function that is close to $1/p^n$-Δuniversal, i.e., by Theorem 1 we must have $\mathrm{MDP}_{\mathrm{f}[\alpha,\beta]} \leq \frac{1}{p^n}$ and $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]} \leq \frac{1}{p^n}$. $\mathrm{f}[\alpha, \beta]$ computes a total of $2n$ multiplications and this requires that for any input difference A to $\mathrm{f}[\alpha, \beta]$, that difference propagates to maximal number of multiplications in $\mathrm{f}[\alpha, \beta]$.

Definition 11 (Active multiplication). *A multiplication in* $\mathrm{f}[\alpha,\beta]$ *is said to be active corresponding to an input difference if that multiplication has a non-zero input difference in at least one of its multiplicands as a result of propagation of that difference inside* $\mathrm{f}[\alpha,\beta]$.

The minimum number of active multiplications over all possible input differences is determined by the minimum of the branch numbers of N_α and N_β. To that end, we show that when N_α and N_β are both MDS, Parallel $[\mathrm{f}[\alpha,\beta]]$ is $2/p^n$-Δuniversal.

$\varepsilon = 2/p^n$ instead of $1/p^n$ is due to the fact that $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]} = 2/p^n$ by Corollary 1. As we saw in Section 6, the maximum value of $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}$ can indeed be reduced if the entries to some of the multiplications were offset by a non-zero quantity, or in other words, if we used affine maps instead of the linear maps N_α and N_β. However such an offset will not have any bearing on the DP of the differentials and thus by using offsets ε is only improved by a factor of 2. So the gain by having an offset is insignificant and thus we decided that none of the multiplications in $\mathrm{MTM}(n,p)$ family of functions will have any offsets.

For the remainder of this section, we assume $\mathrm{f}[\alpha,\beta] \in \mathrm{MTM}(n,p)$ is chosen such that the underlying matrices N_α and N_β are both MDS.

7.1 Maximum Image Probability of $\mathrm{f}[\alpha,\beta]$

In this section we obtain the value of $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]}$ and show that the value is obtained when output $Z = 0^{2n}$.

Lemma 6. *Let* $\mathrm{f}[\alpha,\beta] \in \mathrm{MTM}(n,p)$ *be chosen such that both* N_α *and* N_β *are MDS. Then,* $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(0^{2n}) = \frac{2p^n-1}{p^{2n}}$.

Proof. We prove that $\mathrm{f}[\alpha,\beta](\mathbf{x},\mathbf{y}) = 0^{2n}$ if and only if $\mathbf{x} = 0^n$ or $\mathbf{y} = 0^n$.

If $\mathbf{x} = 0^n$, then we must have $\mathbf{p} = 0^n$ and if $\mathbf{y} = 0^n$, we must have $\mathbf{q} = 0^n$ since the matrices N_α and N_β, being MDS, are invertible. Thus if $\mathbf{x} = 0^n$ or $\mathbf{y} = 0^n$, then $\mathbf{x} \odot \mathbf{y} = 0^n$ and $\mathbf{p} \odot \mathbf{q} = 0^n$, i.e., $\mathrm{f}[\alpha,\beta](\mathbf{x},\mathbf{y}) = 0^{2n}$.

We now show that if both $\mathbf{x} \neq 0^n$, $\mathbf{y} \neq 0^n$, then $\mathrm{f}[\alpha,\beta](\mathbf{x},\mathbf{y}) \neq 0^{2n}$. Indeed,

$$\mathrm{f}[\alpha,\beta](\mathbf{x},\mathbf{y}) = (\mathbf{x} \odot \mathbf{y}, \mathbf{p} \odot \mathbf{q}) = 0^{2n} \implies \mathbf{x} \odot \mathbf{y} = 0^n \text{ and } \mathbf{p} \odot \mathbf{q} = 0^n \ .$$

Let $\mathrm{w}(\mathbf{x}) = w$ for some some $w \in \{1,\ldots,n\}$. We will argue only on the basis of $\mathrm{w}(\mathbf{x})$ and thus we assume without loss of generality that $x_0, x_1, \ldots, x_{w-1} \neq 0$ and $x_w = x_{w+1} = \ldots = x_{n-1} = 0$. Thus for $\mathbf{x} \odot \mathbf{y} = 0^n$, we must have $y_0 = y_1 = \ldots = y_{w-1} = 0$, which means $\mathrm{w}(\mathbf{y}) \leq n - w$. Now N_α and N_β are both $n \times n$ MDS matrix. Thus we have

$$\mathrm{w}(\mathbf{x}) + \mathrm{w}(\mathbf{p}) \geq n+1 \implies \mathrm{w}(\mathbf{p}) \geq n - w + 1 \ , \tag{4.1}$$

$$\mathrm{w}(\mathbf{y}) + \mathrm{w}(\mathbf{q}) \geq n+1 \text{ and } \mathrm{w}(\mathbf{y}) \leq n - w \implies \mathrm{w}(\mathbf{q}) \geq w + 1 \ . \tag{4.2}$$

But, (4.1) together with $\mathbf{p} \odot \mathbf{q} = 0^n$ implies that $\mathrm{w}(\mathbf{q}) \leq w - 1$, a contradiction with (4.2). Thus $\mathrm{f}[\alpha,\beta](\mathbf{x},\mathbf{y}) = 0^{2n}$ iff either $\mathbf{x} = 0^n$ or $\mathbf{y} = 0^n$. Hence, $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(0^{2n}) = \frac{p^n+p^n-1}{p^{2n}} = \frac{2p^n-1}{p^{2n}}$. □

Lemma 7. *Let* $\mathrm{f}[\alpha,\beta] \in \mathrm{MTM}(n,p)$ *be chosen such that both* N_α *and* N_β *are MDS. Then, for any* $Z \neq 0^{2n}$, $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(Z) \leq \frac{2(2p-1)^{n-1}+p-3}{p^{2n}}$.

Proof. Since $Z \neq 0^{2n}$, $z_i \neq 0$ for some $i \in \{0,\ldots,2n-1\}$. First let $z_i \neq 0$ for some $i \in \{0,\ldots n-1\}$. Without loss of generality we assume $z_0 \neq 0$. This implies $y_0 = z_0/x_0$. The $(n-1)$ equations $x_i y_i = z_i$ for $i = 1,\ldots,(n-1)$ can have at most $(2p-1)^{n-1}$ solutions. Let one such solution be

$$x_i = \lambda_i, \quad y_i = \mu_i \text{ for } i = 1,\ldots(n-1)\ . \tag{5}$$

Now, we see that all the variables x_i, y_i for $i \in \{1,2,\ldots,n-1\}$ have been evaluated and the only unknowns are x_0 and y_0. We now find out the number of possible values of x_0 and y_0 corresponding to each solution in (5).

Substituting the values of x_i, y_i for $i \in \{1,\ldots,n-1\}$ from (5) and $y_0 = z_0/x_0$ in $p_i q_i = z_{n+i}$ for $i \in \{0,1,\ldots,n-2\}$, we have a system of $n-1$ equations, where the i-th equation is given by:

$$(\alpha_{i,0}x_0 + \alpha_{i,1}\lambda_1 + \ldots + \alpha_{i,n-1}\lambda_{n-1})(\beta_{i,0}\frac{z_0}{x_0} + \beta_{i,1}\mu_1 + \ldots + \beta_{i,n-1}\mu_{n-1}) = z_{n+i}\ . \tag{6}$$

Now, for $i \in \{0,1,\ldots,n-2\}$ let us denote by $\alpha_i = \alpha_{i,1}\lambda_1 + \ldots + \alpha_{i,n-1}\lambda_{n-1}$ and $\beta_i = \beta_{i,1}\mu_1 + \ldots + \beta_{i,n-1}\mu_{n-1}$. Then the set of $(n-1)$ linear equations in (6) converts to:

$$(\alpha_{i,0}x_0 + \alpha_i)(\beta_{i,0}\frac{z_0}{x_0} + \beta_i) = z_{n+i} \text{ for } i \in \{0,1,\ldots,n-2\}\ . \tag{7}$$

Now, whenever one of $\alpha_i \neq 0$ or $\beta_i \neq 0$, $(\alpha_{i,0}x_0 + \alpha_i)(\beta_{i,0}\frac{z_0}{x_0} + \beta_i) = z_{n+i}$ describes a quadratic equation. But whenever both $\alpha_i = 0$ and $\beta_i = 0$, this becomes independent of x_0. This leads to the following two possibilities:

1. $\alpha_i = \beta_i = 0\ \forall\ i \in \{0,\ldots n-2\}$.
2. $\exists 0 \leq i \leq n-2$ such that $\alpha_i \neq 0$ or $\beta_i \neq 0$.

We first look at the case $\alpha_i = \beta_i = 0$ for each $i \in \{0,\ldots n-2\}$. Let $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_{n-1})^\intercal \in \mathbb{F}_p^{n-1}, \mu = (\mu_1,\mu_2,\ldots\mu_{n-1})^\intercal \in \mathbb{F}_p^{n-1}$. Furthermore let N'_α and N'_β denote the $(n-1)\times(n-1)$ submatrices of N_α and N_β respectively, where N'_α and N'_β are obtained by removing the 0-th column and $(n-1)$-th row from N_α and N_β respectively. So, $\mathrm{N}'_\alpha = \left[\alpha_{i,j}\right]$ and $\mathrm{N}'_\beta = \left[\beta_{i,j}\right]$ with $0 \leq i \leq n-2$, $1 \leq j \leq n-1$. Now the set of $(n-1)$ linear equations $\alpha_i = \beta_i = 0$ for $i \in \{0,\ldots n-2\}$ can be written as:

$$\mathrm{N}'_\alpha \cdot \lambda = 0^n \text{ and } \mathrm{N}'_\beta \cdot \mu = 0^n\ . \tag{8}$$

N'_α and N'_β are submatrices of the MDS matrices N_α and N_β. Thus from (8), N'_α and N'_β both must be invertible. Thus $\lambda = \mu = 0^n$. In this case (7) converts to $\alpha_{i,0}\beta_{i,0}z_0 = z_{n+i}$ for $i \in \{0,\ldots,n-2\}$, i.e., the equations are independent of x_0. Thus, the number of solutions, if it exists, is bounded by the number of

solutions to $x_0y_0 = z_0$, i.e., $p-1$. In other words when each of $x_i = y_i = 0$ for $i \in \{1, \ldots, n-1\}$, there can be at most $(p-1)$ values of the pair (x_0, y_0)

When there exists $0 \leq i \leq n-2$ such that $\alpha_i \neq 0$ or $\beta_i \neq 0$, for that value of i, (9) is a quadratic equation.

$$(\alpha_{i,0}x_0 + \alpha_i)(\beta_{i,0}\frac{z_0}{x_0} + \beta_i) = z_{n+i} \; . \tag{9}$$

Thus (9) has at most 2 solutions and this holds true for each solution from (5) except when each $x_i = y_i = 0$ for $i \in \{1, \ldots, n-1\}$.

Now, each $x_i = y_i = 0$ for $i \in \{1, \ldots, n-1\}$ implies that $z_i = 0$ for each $i \in \{1, \ldots, n-1\}$. Thus we can conclude that for a $Z \neq 0^{2n}$,

$$\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(Z) \leq \begin{cases} \frac{2(2p-1)^{n-1}+p-3}{p^{2n}} & , \text{ if } z_i \neq 0 \text{ for exactly one } i \in \{0, \ldots, n-1\} \\ \frac{2(2p-1)^{n-1}}{p^{2n}} & , \text{ if } z_i \neq 0 \text{ for more than one } i \in \{0, \ldots, n-1\}. \end{cases}$$

Now let $z_i \neq 0$ for some $i \in \{n, \ldots, 2n-1\}$. Since inverse of an MDS matrix is necessarily MDS and invertible, we can apply the same arguments as when $z_i \neq 0$ for some $i \in \{0, \ldots, n-1\}$ to obtain a solution in p_i, q_i and the number of such solutions is again at most $\frac{2(2p-1)^{n-1}+p-3}{p^{2n}}$. Since N_α^{-1} and N_β^{-1} are both invertible, each solution in p_i, q_i corresponds to a unique solution in x_i, y_i and hence for such a Z as well $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(Z) \leq \frac{2(2p-1)^{n-1}+p-3}{p^{2n}}$. □

Corollary 1. *For any* $\mathrm{f}[\alpha, \beta] \in \mathrm{MTM}(n, p)$*, where* N_α *and* N_β *are both MDS,* $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]} = \frac{2p^n-1}{p^{2n}} \leq \frac{2}{p^n}$

Proof. By Lemmas 6 and 7, we see that $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(0^{2n}) = \frac{2p^n-1}{p^{2n}}$ and for any $Z \neq 0^{2n}$, $\mathrm{IP}_{\mathrm{f}[\alpha,\beta]}(Z) \leq \frac{2(2p-1)^{n-1}+p-3}{p^{2n}}$. Now, since p is chosen such that $\log_2 p > n$, we have

$$2p^n - 2(2p-1)^{n-1} > 2p^{n-1}(p - 2^{n-1}) > 2^n p^{n-1} > p - 2 \; .$$

So, $2p^n - 1 > 2(2p-1)^{n-1} + p - 3$ and thus $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]} = \frac{2p^n-1}{p^{2n}} \leq \frac{2}{p^n}$. □

7.2 Maximum Differential Probability of f[α, β]

Before looking into the differential probability of $\mathrm{f}[\alpha, \beta]$, we introduce some new notation. For an input difference $A = (\mathbf{a}, \mathbf{b})$, we denote $\mathrm{N}_\alpha \cdot \mathbf{a} = \mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$ and $\mathrm{N}_\beta \cdot \mathbf{b} = \mathbf{d} = (d_0, d_1, \ldots, d_{n-1})$ respectively. Due to $\mathrm{N}_\alpha, \mathrm{N}_\beta$ being MDS, for any choice of $\mathbf{a}$ and $\mathbf{b}$, we must have $\mathrm{w}(\mathbf{a}) + \mathrm{w}(\mathbf{c}) \geq (n+1)$ and $\mathrm{w}(\mathbf{b}) + \mathrm{w}(\mathbf{d}) \geq (n+1)$. For any vector $\mathbf{x} \in \mathbb{F}_p^n$, $\mathrm{D}_\mathbf{x}$ denotes the $n \times n$ diagonal matrix with i-th diagonal entry being x_i.

Lemma 8. *An input difference* $A = (\mathbf{a}, \mathbf{b})$ *propagates to the output difference* $\Delta = (\delta_0, \delta_1, \ldots, \delta_{2n-1})^\intercal$ *under* $\mathrm{f}[\alpha, \beta]$ *for a key* $K = (\mathbf{h}, \mathbf{k})$ *if*

$$\mathrm{N}_A \cdot K + \mathrm{O}_A = \Delta \; . \tag{10}$$

Here N_A *and* O_A *are given by:*

$$\mathrm{N}_A = \begin{bmatrix} \mathrm{D}_{\mathbf{b}} & \mathrm{D}_{\mathbf{a}} \\ \mathrm{N}_\alpha \cdot \mathrm{D}_{\mathbf{d}} & \mathrm{N}_\beta \cdot \mathrm{D}_{\mathbf{c}} \end{bmatrix} ,$$
$$\mathrm{O}_A = \begin{bmatrix} a_0 b_0 \ \ldots \ a_{n-1} b_{n-1} \ c_0 d_0 \ \ldots \ c_{n-1} d_{n-1} \end{bmatrix}^{\mathsf{T}} .$$

Proof. An input difference $A = (\mathbf{a}, \mathbf{b})$ propagates to an output difference $\Delta = (\delta_0, \delta_1, \ldots, \delta_{2n-1})^{\mathsf{T}}$ under a key $K = (\mathbf{h}, \mathbf{k})$ if

$$\mathrm{f}[\alpha, \beta](\mathbf{h} + \mathbf{a}, \mathbf{k} + \mathbf{b}) - \mathrm{f}[\alpha, \beta](\mathbf{h}, \mathbf{k}) = \Delta .$$

From the definition of $\mathrm{f}[\alpha, \beta]$ we have,

$$((\mathbf{h} + \mathbf{a}) \odot (\mathbf{k} + \mathbf{b}), \mathrm{N}_\alpha \cdot (\mathbf{h} + \mathbf{a}) \odot \mathrm{N}_\beta \cdot (\mathbf{k} + \mathbf{b})) - (\mathbf{h} \odot \mathbf{k}, \mathrm{N}_\alpha \cdot \mathbf{h} \odot \mathrm{N}_\beta \cdot \mathbf{k}) = \Delta .$$

This leads to two sets of n equations, where for $i \in \{0, \ldots, n-1\}$ the i-th such equations are given by:

$$(h_i + a_i) \cdot (k_i + b_i) - h_i \cdot k_i = \delta_i ,$$
$$\left(\sum_{j=0}^{n-1} \alpha_{i,j} (h_j + a_j) \right) \cdot \left(\sum_{j=0}^{n-1} \beta_{i,j} (k_j + b_j) \right) - \left(\sum_{j=0}^{n-1} \alpha_{i,j} h_j \right) \cdot \left(\sum_{j=0}^{n-1} \beta_{i,j} k_j \right) = \delta_{n+i} .$$

For each $0 \leq i \leq n-1$, these sets of equations respectively simplify to:

$$b_i h_i + a_i k_i + a_i b_i = \delta_i ,$$
$$d_i \sum_{j=0}^{n-1} \alpha_{i,j} h_i + c_i \sum_{j=0}^{n-1} \beta_{i,j} k_j + c_i d_i = \delta_{n+i} .$$

Thus, we now have a set of $2n$ linear equations in $2n$ variables h_i, k_i. Writing this set of equations in terms of matrices, we arrive at our desired lemma. □

We call the matrix N_A the *difference matrix* corresponding to the input difference A. Given any A we denote by $\mathcal{C}_A$ the column space generated by N_A.

Corollary 2. *Given a differential* (A, Δ) *to* $\mathrm{f}[\alpha, \beta]$, $\mathsf{DP}_{\mathrm{f}[\alpha,\beta]}(A, \Delta)$ *is given by:*

$$\mathsf{DP}_{\mathrm{f}[\alpha,\beta]}(A, \Delta) = \begin{cases} 0 & , \ \textit{when } (\Delta - \mathrm{O}_A) \notin \mathcal{C}_A \\ \frac{1}{p^r} & , \ \textit{otherwise.} \end{cases}$$

Here r *denotes the rank of* N_A*.*

Proof. When $(\Delta - \mathrm{O}_A) \notin \mathcal{C}_A$, (10) is inconsistent and thus $\mathsf{DP}_{\mathrm{f}[\alpha,\beta]}(A, \Delta) = 0$. Otherwise (10) has p^{2n-r} solutions and hence $\mathsf{DP}_{\mathrm{f}[\alpha,\beta]}(A, \Delta) = \frac{p^{2n-r}}{p^{2n}} = \frac{1}{p^r}$. □

Lemma 9. *Given any input difference A to $\mathrm{f}[\alpha, \beta]$ where the underlying matrices are MDS, the rank of N_A is at least n.*

Proof. Since $A \neq 0^{2n}$, at least one of $\mathbf{a}$ and $\mathbf{b}$ must be non-zero. Let us assume without loss of generality that $\mathbf{b} \neq 0$. For $i \in \{0, 1, \ldots, n-1\}$, let C_i denote the i-th column of N_A, i.e.,

$$C_i = \left[0 \ldots b_i \ldots 0 \ \alpha_{0,i} d_0 \ldots \alpha_{i,i} d_i \ldots \alpha_{n-1,i} d_{n-1}\right]^{\mathsf{T}} .$$

We show that $\{C_0, C_1, \ldots, C_{n-1}\}$ is a set of n linearly independent column vectors and thus $\mathrm{rank}(\mathrm{N}_A) \geq n$. So, we show that for scalars $\lambda_0, \lambda_1, \ldots, \lambda_{n-1} \in \mathbb{F}_p$

$$\lambda_0 C_0 + \lambda_1 C_1 + \ldots + \lambda_{n-1} C_{n-1} = 0^n \quad \Longrightarrow \quad \lambda_0 = \lambda_1 = \ldots = \lambda_{n-1} = 0 .$$

Now, $\sum_{i=0}^{n-1} \lambda_i C_i = 0^n$ simplifies to:

$$\begin{bmatrix} \lambda_0 b_0 \\ \lambda_1 b_1 \\ \vdots \\ \lambda_{n-1} b_{n-1} \\ (\lambda_0 \alpha_{0,0} + \lambda_1 \alpha_{0,1} + \ldots \lambda_{n-1} \alpha_{0,n-1}) d_0 \\ (\lambda_0 \alpha_{1,0} + \lambda_1 \alpha_{1,1} + \ldots \lambda_{n-1} \alpha_{1,n-1}) d_1 \\ \vdots \\ (\lambda_0 \alpha_{n-1,0} + \lambda_1 \alpha_{n-1,1} + \ldots \lambda_{n-1} \alpha_{n-1,n-1}) d_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{11}$$

Let $\mathrm{w}(\mathbf{b}) = w$ for some $w \in \{1, \ldots, n\}$. Without loss of generality we assume that $b_0, b_1, \ldots, b_{w-1} \neq 0$. Then clearly from (11) we see that $\lambda_0 = \lambda_1 = \ldots = \lambda_{w-1} = 0$. Thus (11) now reduces to the following set of n linear equations in $n - w$ variables $\lambda_w, \lambda_{w+1}, \ldots, \lambda_{n-1}$.

$$\begin{bmatrix} (\lambda_w \alpha_{0,w} + \lambda_{w+1} \alpha_{0,w+1} + \ldots \lambda_{n-1} \alpha_{0,n-1}) d_0 \\ (\lambda_w \alpha_{1,w} + \lambda_{w+1} \alpha_{1,w+1} + \ldots \lambda_{n-1} \alpha_{1,n-1}) d_1 \\ \vdots \\ (\lambda_w \alpha_{n-1,w} + \lambda_{w+1} \alpha_{n-1,w+1} + \ldots \lambda_{n-1} \alpha_{n-1,n-1}) d_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{12}$$

Since, $\mathrm{w}(\mathbf{b}) = w$ and $\mathrm{w}(\mathbf{b}) + \mathrm{w}(\mathbf{d}) \geq n + 1$, at least $n - w + 1$ components of $\mathbf{d}$ must be non-zero. Again for simplicity we assume $d_0, d_1, \ldots, d_{n-w} \neq 0$. But this would imply

$$\begin{bmatrix} \alpha_{0,w} & \alpha_{0,w+1} & \ldots & \alpha_{0,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n-w,w} & \alpha_{n-w,w+1} & \ldots & \alpha_{n-w,n-1} \end{bmatrix} \cdot \begin{bmatrix} \lambda_w \\ \vdots \\ \lambda_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{13}$$

But, the $(n-w+1)\times(n-w)$ matrix in the left hand side of (13) is a submatrix of N_α and thus has full rank, i.e., rank of this matrix is $n-w$. Thus this system of equations has a unique solution given by: $\lambda_w = \ldots = \lambda_{n-1} = 0$.

Thus, we have shown that $\{C_0, \ldots, C_{n-1}\}$ is a set of n linearly independent column vectors. Hence $\mathrm{rank}(\mathrm{N}_A) \geq n$. □

Corollary 3. *For any* $\mathrm{f}[\alpha,\beta]$ *where the underlying matrices are MDS,* $\mathrm{MDP}_{\mathrm{f}[\alpha,\beta]} = \frac{1}{p^n}$.

Proof. It follows from Corollary 2 that for any differential (A, Δ) to $\mathrm{f}[\alpha,\beta]$, $\mathrm{DP}_{\mathrm{f}[\alpha,\beta]}(A,\Delta) \leq \frac{1}{p^r}$ where $\mathrm{rank}(\mathrm{N}_A) = r$. By Lemma 9 it follows that $r \geq n$ and consequently $\mathrm{DP}_{\mathrm{f}[\alpha,\beta]}(A,\Delta) \leq \frac{1}{p^n}$.

This bound is also attained for a well chosen differential. Indeed it can be seen that if A is chosen such that $\mathbf{a} = 0^n$, then $\mathrm{rank}(\mathrm{N}_A) = n$. In this case if Δ is chosen such that $(\Delta - \mathrm{O}_A) \in \mathcal{C}_A$ (in particular one can choose $\Delta = 0^{2n}$), then $\mathrm{DP}_{\mathrm{f}[\alpha,\beta]}(A,\Delta) = \frac{1}{p^n}$. □

7.3 ε-Δuniversality of Parallel $[\mathrm{f}[\alpha,\beta]]$

Theorem 2. *Let* $\mathrm{f}[\alpha,\beta] \in \mathrm{MTM}(n,p)$ *be such that its underlying matrices are MDS. Then Parallel* $[\mathrm{f}[\alpha,\beta]]$ *is* $\frac{2}{p^n}$-Δ*universal.*

Proof. From Corollaries 1 and 3 it follows that for such a $\mathrm{f}[\alpha,\beta]$, $\mathrm{MIP}_{\mathrm{f}[\alpha,\beta]} \leq \frac{2}{p^n}$ and $\mathrm{MDP}_{\mathrm{f}[\alpha,\beta]} \leq \frac{1}{p^n}$. Thus it follows from Theorem 1 that Parallel $[\mathrm{f}[\alpha,\beta]]$ is $\frac{2}{p^n}$-Δuniversal. □

Thus, whenever N_α and N_β are MDS, Parallel $[\mathrm{f}[\alpha,\beta]]$ is $\frac{2}{p^n}$-Δuniversal. However while the matrices being MDS is sufficient to obtain this universality, it is not strictly necessary. It is possible to choose a member of $\mathrm{MTM}(n,p)$, whose parallelization is $\frac{2}{p^n}$-Δuniversal, but the underlying matrices N_α and N_β are not MDS. However this requires more detailed security analysis. We look at such a public function in the next section.

8 Multi-265

We now introduce the key-then-hash function multi-265 . This is the parallelization of a public function that we denote as f-265 belonging to $\mathrm{MTM}(6, 2^{26}-5)$, i.e., multi-265 =Parallel [f-265]. The specifications of f-265 is as follows.

Definition 12. *The public function of multi-265 denoted as* f-265 *is defined as:*

$$\text{f-265}\colon \mathbb{F}_p^{12} \mapsto \mathbb{F}_p^{12}\colon \text{f-265}(X) = \text{f-265}(\mathbf{x},\mathbf{y}) = (\mathbf{x}\odot\mathbf{y}, \mathrm{N}\cdot\mathbf{x}\odot\mathrm{N}\cdot\mathbf{y}) \ ,$$

where N *is the* 6×6 *circulant matrix whose first row is given by* $\begin{pmatrix}1&1&3&1&3&0\end{pmatrix}$.

We will now motivate the design choices we made. We want to use fast 32-bit integer multiplication instructions and apply lazy modular reductions. Therefore we take a prime significantly smaller than 2^{32}. Still, we want this modular reduction to be efficient and for that purpose we take the prime of our field to be a pseudo-mersenne prime[13], i.e., of the form $2^l - \lambda$, for a small λ and we chose for $p = 2^{26} - 5$. Now, for the dimension n, we wanted $\varepsilon = 2/p^n$ to be smaller than 2^{-128}, i.e., $n \geq 5$ for our chosen prime p. Thus a 5×5 MDS matrix would guarantee $\varepsilon = 2p^{-5} \approx 2^{-129}$. However, these dimensions are not well suited in the SIMD architecture and therefore we chose for 6×6 matrices. Now, a 6×6 MDS matrix is quite expensive and thus we instead looked for a 6×6 matrix with branch number 6. Any 6×6 matrix with branch number 6 however does not assure us of $\varepsilon = 2p^{-6}$. Thus another added restriction, as we will soon see, is that matrices should have branch number at least 7 when restricted to inputs with weight at least 3.

So, we finally chose N to be a circulant matrix and chose both the matrices in f-265 to be the same to simplify our security analysis. We limited the entries in N to the set $\{-1, 0, 1, 2, 3\}$ so matrix multiplication can be efficiently implemented with only addition and subtraction. We found N by exhaustive search over all candidates checking the branch number is 6 and then selecting the ones that have branch number 7 when restricted to inputs with weight 3 or more. We did this by finding all full rank matrices having at most one entry 0 such that all 3×3, 4×4 and 5×5 submatrices have full rank. Sage code to find all matrices that satisfy these requirements can be found at https://github.com/KoustabhGhosh/Multi-265.

In this section, we use similar notations as in Section 7. So we have $\mathrm{N} \cdot \mathbf{x} = \mathbf{p}$, $\mathrm{N} \cdot \mathbf{y} = \mathbf{q}$ and for an input difference $A = (\mathbf{a}, \mathbf{b})$, $\mathrm{N} \cdot \mathbf{a} = \mathbf{c}$ and $\mathrm{N} \cdot \mathbf{b} = \mathbf{d}$.

8.1 Maximum Image Probability of f-265

In this section we show that similar to f$[\alpha, \beta]$, $\mathrm{MIP}_{\text{f-265}} \leq \frac{2}{p^6}$ and is obtained for the output 0^{12}.

Lemma 10. *For the public function* f-265, $\mathsf{IP}_{\text{f-265}}(0^{12}) = \frac{2p^6-1}{p^{12}}$.

Proof. We show that f-265$(\mathbf{x}, \mathbf{y}) = 0^{12}$ if and only if $\mathbf{x} = 0^6$ or $\mathbf{y} = 0^6$. Clearly if $\mathbf{x} = 0^6$ or $\mathbf{y} = 0^6$, f-265$(X) = 0^{12}$

We now show that when both $\mathbf{x} \neq 0^6$ and $\mathbf{y} \neq 0^6$, f-265$(\mathbf{x}, \mathbf{y}) \neq 0^{12}$. Let $\mathrm{w}(\mathbf{x}) = w$ for some some $w \in \{1, \ldots, 6\}$.

When $w \geq 3$, since N has branch number 7 for all inputs with weight at least 3, we have

$$\mathrm{w}(\mathbf{x}) + \mathrm{w}(\mathbf{p}) \geq 7 \qquad \Longrightarrow \qquad \mathrm{w}(\mathbf{p}) \geq 7 - w \; . \tag{14}$$

From (14), $\mathbf{x} \odot \mathbf{y} = 0^6$ and $\mathbf{p} \odot \mathbf{q} = 0^6$ we see that

$$\mathrm{w}(\mathbf{y}) \leq 6 - w \; , \tag{15.1}$$

$$\mathrm{w}(\mathbf{q}) \leq 6 - (7 - w) = w - 1 \; . \tag{15.2}$$

But, (15.1) together with the fact that N has branch number 6 implies that $\mathrm{w}(\mathbf{q}) \geq 6 - (6 - w) = w$, a contradiction with (15.2).

When $w < 3$, we have

$$\mathrm{w}(\mathbf{x}) + \mathrm{w}(\mathbf{p}) \geq 6 \qquad \Longrightarrow \qquad \mathrm{w}(\mathbf{p}) \geq 6 - w \; .$$

From $\mathbf{p} \odot \mathbf{q} = 0^6$ it follows that

$$\mathrm{w}(\mathbf{q}) \leq 6 - (6 - w) = w.$$

But, $\mathrm{w}(\mathbf{q}) \leq w < 3$ implies that $\mathrm{w}(\mathbf{y})$ must also be greater than 3 due to N having branch number 6. But since $\mathrm{w}(\mathbf{y}) > 3$, we must have

$$\mathrm{w}(\mathbf{y}) + \mathrm{w}(\mathbf{q}) \geq 7 \qquad \Longrightarrow \qquad \mathrm{w}(\mathbf{y}) \geq 7 - w \; . \tag{16}$$

From $\mathbf{x} \odot \mathbf{y} = 0^6$ we have $\mathrm{w}(\mathbf{y}) \leq 6 - w$, a contradiction with (16).

Thus f-265$(X) = 0^{12}$ iff either $\mathbf{x} = 0^6$ or $\mathbf{y} = 0^6$. □

Lemma 11. *Let $Z \neq 0^{12}$. Then for the public function* f-265, $\mathsf{IP}_{\text{f-265}}(Z) \leq \frac{2(2p-1)^5+p-3}{p^{12}}$

Proof. When $Z \neq 0^{12}$, $\mathsf{IP}_{\text{f-265}}(Z) \leq \frac{2(2p-1)^5+p-3}{p^{12}}$ follows directly from Lemma 7. The only property of $n \times n$ MDS matrices that we used in proof to Lemma 11 was that the MDS matrix itself and all its $(n-1) \times (n-1)$ are invertible. N follows these properties since N itself and all its 5×5 submatrices are invertible. □

Corollary 4. *For the public function* f-265, $\mathrm{MIP}_{\text{f-265}} = \frac{2p^6-1}{p^{12}} \leq \frac{2}{p^6}$.

Proof. The proof follows directly from Lemmas 10 and 11. □

8.2 Maximum Differential Probability of f-265

By applying Lemma 8, we see that an input difference $A = (\mathbf{a}, \mathbf{b})$ propagates to the output difference $\Delta = (\delta_0, \delta_1, \ldots, \delta_{11})$ under f-265 for a key $K = (\mathbf{h}, \mathbf{k})$ if $\mathrm{N}_A \cdot K + \mathrm{O}_A = \Delta$. For f-265, N_A is given by

$$\mathrm{N}_A = \begin{bmatrix}
b_0 & 0 & 0 & 0 & 0 & 0 & a_0 & 0 & 0 & 0 & 0 & 0 \\
0 & b_1 & 0 & 0 & 0 & 0 & 0 & a_1 & 0 & 0 & 0 & 0 \\
0 & 0 & b_2 & 0 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 \\
0 & 0 & 0 & b_3 & 0 & 0 & 0 & 0 & 0 & a_3 & 0 & 0 \\
0 & 0 & 0 & 0 & b_4 & 0 & 0 & 0 & 0 & 0 & a_4 & 0 \\
0 & 0 & 0 & 0 & 0 & b_5 & 0 & 0 & 0 & 0 & 0 & a_5 \\
d_0 & d_0 & 3d_0 & d_0 & 3d_0 & 0 & c_0 & c_0 & 3c_0 & c_0 & 3c_0 & 0 \\
0 & d_1 & d_1 & 3d_1 & d_1 & 3d_1 & 0 & c_1 & c_1 & 3c_1 & c_1 & 3c_1 \\
3d_2 & 0 & d_2 & d_2 & 3d_2 & d_2 & 3c_2 & 0 & c_2 & c_2 & 3c_2 & c_2 \\
d_3 & 3d_3 & 0 & d_3 & d_3 & 3d_3 & c_3 & 3c_3 & 0 & c_3 & c_3 & 3c_3 \\
3d_4 & d_4 & 3d_4 & 0 & d_4 & d_4 & 3c_4 & c_4 & 3c_4 & 0 & c_4 & c_4 \\
d_5 & 3d_5 & d_5 & 3d_5 & 0 & d_5 & c_5 & 3c_5 & c_5 & 3c_5 & 0 & c_5
\end{bmatrix} .$$

We know by Corollary 2 that for a differential (A, Δ), its DP is upper bounded by $1/p^r$, where r is the rank of N_A. Before we can obtain a lower bound for N_A, we first look at an important property of N.

Lemma 12. *For $r \in \{1, 2, 3, 4, 5\}$, all the $(r+1) \times r$ submatrices of* N *have rank r.*

Proof. For $r = 1$, all 2×1 submatrices of N has rank 1 since $\begin{bmatrix} 0 & 0 \end{bmatrix}^\intercal$ is not a submatrix of N. For $r = 2$, all 3×2 submatrices must have rank 2 since otherwise there must exist a 3×3 submatrix with rank 2, a contradiction. For $r \geq 3$, the proof is trivial since all $r \times r$ submatrices of N have rank r for $r \geq 3$. □

Lemma 13. *Given any input difference A, rank* $(\mathrm{N}_A) \geq 6$.

Proof. Since $A = (\mathbf{a}, \mathbf{b}) \neq 0^{12}$, we assume without loss of generality that $\mathbf{b} \neq 0^6$.

Let $\mathrm{w}(\mathbf{b}) = w$. In the proof to Lemma 9 we used the following facts about the underlying $n \times n$ MDS matrix N_α of $\mathrm{f}[\alpha, \beta]$:

1. For any $\mathbf{b}$ with $\mathrm{w}(\mathbf{b}) = w$, we must have $\mathrm{w}(\mathbf{d}) \geq n - w + 1$.
2. Every $(n - w + 1) \times (n - w)$ submatrix of N_α has rank $(n - w)$.

In this case, N however is not MDS and has branch number 6. But, we can use the fact that for $w \geq 3$, its branch number is 7.

Indeed when $w < 3$, we must have $\mathrm{w}(\mathbf{d}) \geq 6 - w > 3$. Consequently arguing similarly to the proof of Lemma 9, we see that instead of requiring that every $(7-w) \times (6-w)$ submatrix of N have rank $(6-w)$, we instead require that each $(6 - w) \times (6 - w)$ submatrix of N must have rank $6 - w$. This is true since for $w < 3$, all $(6 - w) \times (6 - w)$ submatrices of N have full rank by design.

For $w \geq 3$, since N has branch number 7, we only require that each $(7-w) \times (6 - w)$ have rank $6 - w$, which is indeed true by Lemma 12. □

8.3 ε-Δuniversality of Multi-265

Theorem 3. *multi-265 is 2^{-154}-Δuniversal.*

Proof. From Lemmas 10 and 13, it follows that $\mathrm{MIP}_{\text{f-265}} \leq \frac{2}{p^6}$ and $\mathrm{MDP}_{\text{f-265}} \leq \frac{1}{p^6}$. Now, since $\frac{2}{p^6} = \frac{2}{(2^{26}-5)^6} \leq 2^{-154}$, it follows from Theorem 1 that multi-265 is 2^{-154}-Δuniversal. □

8.4 Implementation Aspects

Multi-265 can be implemented on any platform with SIMD architecture. These instructions process vectors of the same type elements that are packed together in parallel. As a result, operations like addition, multiplication etc. can be performed on multiple entries at the same time and this increases the performance of the implementation. Newer SIMD architecture has 128-bit vector registers that can be seen as 16, 8, 4, or 2 elements with size 8, 16, 32, and 64 bits respectively

by defining the arrangement-specifier accordingly. This specifier determines the packing unit of data.

A message block being an element of $\mathbb{F}_p^{12}$ can be treated as 288-bits since $p = 2^{26} - 5$. Now, each message block can be stored in 3 128-bit vector registers. Thus, a rearranging procedure is done after loading each block of message from memory. The first byte of each 32-bit word in the Neon vectors is set to zero and the remaining 3-bytes are filled with the corresponding 3-bytes of the message block.

Each 32-bit word contains 24-bits of data initially. So we can defer modular reductions for all the linear operations to the output of each call to f-265 and we are only required to do 12 modular reductions at the end of each round. Moreover our choice of prime $p = 2^{26} - 5$, being a pseudo mersenne prime, means that the reduction can be done very efficiently[13].

The reference code for multi-265 in a keyless setting is available at `https://github.com/KoustabhGhosh/Multi-265`.

Acknowledgements. Koustabh Ghosh is supported by the Netherlands Organisation for Scientific Research (NWO) under TOP grant TOP1.18.002 SCALAR, Joan Daemen and Jonathan Fuchs are supported by the European Research Council under the ERC advanced grant agreement under grant ERC-2017-ADG Nr. 788980 ESCADA and Parisa Amiri Eliasi is supported by the Cryptography Research Center of the Technology Innovation Institute (TII), Abu Dhabi (UAE), under the TII-Radboud project with title Evaluation and Implementation of Lightweight Cryptographic Primitives and Protocols.

References

1. Albrecht, M.R., Grassi, L., Rechberger, C., Roy, A., Tiessen, T.: Mimc: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In: Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I. Lecture Notes in Computer Science, vol. 10031, pp. 191–219 (2016)
2. Bellare, M., Canetti, R., Krawczyk, H.: Keying Hash Functions for Message Authentication. In: Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings. LNCS, vol. 1109, pp. 1–15. Springer (1996)
3. Bellare, M., Kilian, J., Rogaway, P.: The Security of Cipher Block Chaining. In: Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings. Lecture Notes in Computer Science, vol. 839, pp. 341–358. Springer (1994)
4. Bernstein, D.J.: The Poly1305-AES Message-Authentication Code. In: Gilbert, H., Handschuh, H. (eds.) Fast Software Encryption: 12th International Workshop, FSE 2005, Paris, France, February 21-23, 2005, Revised Selected Papers. LNCS, vol. 3557, pp. 32–49. Springer (2005)

5. Bertoni, G., Daemen, J., Hoffert, S., Peeters, M., Assche, G.V., Keer, R.V.: The authenticated encryption schemes Kravatte-SANE and Kravatte-SANSE. IACR Cryptol. ePrint Arch. p. 1012 (2018)
6. Black, J., Rogaway, P.: CBC MACs for Arbitrary-Length Messages: The Three-Key Constructions. In: Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings. LNCS, vol. 1880, pp. 197–215. Springer (2000)
7. Black, J., Rogaway, P.: A Block-Cipher Mode of Operation for Parallelizable Message Authentication. In: Advances in Cryptology - EUROCRYPT 2002, International Conference on the Theory and Applications of Cryptographic Techniques, Amsterdam, The Netherlands, April 28 - May 2, 2002, Proceedings. LNCS, vol. 2332, pp. 384–397. Springer (2002)
8. Daemen, J., Hoffert, S., Assche, G.V., Keer, R.V.: The design of Xoodoo and Xoofff. IACR Trans. Symmetric Cryptol. **2018**(4), 1–38 (2018)
9. Daemen, J., Rijmen, V.: The Pelican MAC Function. IACR Cryptol. ePrint Arch. p. 88 (2005)
10. Daemen, J., Rijmen, V.: The Design of Rijndael - The Advanced Encryption Standard (AES), Second Edition. Information Security and Cryptography, Springer (2020)
11. Fuchs, J., Rotella, Y., Daemen, J.: On the security of keyed hashing based on an unkeyed block function. IACR Cryptol. ePrint Arch. p. 1172 (2022)
12. Grassi, L., Rechberger, C., Rotaru, D., Scholl, P., Smart, N.P.: Mpc-friendly symmetric key primitives. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016. pp. 430–443. ACM (2016)
13. Greuet, A., Montoya, S., Vermeersch, C.: Quotient Approximation Modular Reduction. Cryptology ePrint Archive, Paper 2022/411 (2022)
14. Ishai, Y., Sahai, A., Wagner, D.A.: Private circuits: Securing hardware against probing attacks. In: Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings. Lecture Notes in Computer Science, vol. 2729, pp. 463–481. Springer (2003)
15. Kocher, P.C., Jaffe, J., Jun, B.: Differential power analysis. In: Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings. Lecture Notes in Computer Science, vol. 1666, pp. 388–397. Springer (1999)
16. McGrew, D.A., Viega, J.: The Security and Performance of the Galois/Counter Mode (GCM) of Operation. In: Progress in Cryptology - INDOCRYPT 2004, 5th International Conference on Cryptology in India, Chennai, India, December 20-22, 2004, Proceedings. LNCS, vol. 3348, pp. 343–355. Springer (2004)
17. Nikova, S., Rijmen, V., Schläffer, M.: Secure hardware implementation of nonlinear functions in the presence of glitches. J. Cryptology **24**(2), 292–321 (2011)
18. Shoup, V.: On Fast and Provably Secure Message Authentication Based on Universal Hashing. In: Koblitz, N. (ed.) Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings. Lecture Notes in Computer Science, vol. 1109, pp. 313–328. Springer (1996)
19. Stinson, D.R.: On the Connections Between Universal Hashing, Combinatorial Designs and Error-Correcting Codes. Electron. Colloquium Comput. Complex. **TR95-052** (1995)

20. Wegman, M.N., Carter, J.: New hash functions and their use in authentication and set equality. Journal of Computer and System Sciences **22**(3), 265–279 (1981)