Exponential sums in linear cryptanalysis

Tim Beyne [1,2] and Clémence Bouvier [2]

[1] COSIC, KU Leuven, Leuven, Belgium
tim.beyne@esat.kuleuven.be
[2] Ruhr University Bochum, Faculty of Computer Science, Bochum, Germany
clemence.bouvier@rub.de

Abstract. It is shown how bounds on exponential sums derived from modern algebraic geometry, and ℓ-adic cohomology specifically, can be used to upper bound the absolute correlations of linear approximations for cryptographic constructions of low algebraic degree. This is illustrated by applying results of Deligne, Denef and Loeser, and Rojas-León, to obtain correlation bounds for a generalization of the Butterfly construction, three-round Feistel ciphers, and a generalization of the Flystel construction. For each of these constructions, bounds obtained using other methods are significantly weaker. In the case of the Flystel construction, our bounds resolve a conjecture by the designers.
Correlation bounds of this type are relevant for the development of security arguments against linear cryptanalysis, especially in the weak-key setting or for primitives that do not involve a key. Since the methods used in this paper are applicable to constructions defined over arbitrary finite fields, the results are also relevant for arithmetization-oriented primitives such as Anemoi, which uses S-boxes based on the Flystel construction.

Keywords: linear cryptanalysis · algebraic exponential sums · Butterfly · Feistel · Flystel

1 Introduction

Linear cryptanalysis is one of a few general cryptanalytic techniques that define the design space of symmetric-key primitives. Although contemporary designs come with useful heuristic security arguments against linear cryptanalysis, rigorously proving that the correlations of all linear approximations are close to zero is currently out of reach.

The most straightforward security argument consists of showing that the absolute correlations of all linear trails are small. For primitives based on S-boxes and a linear layer, the wide-trail design strategy [18] is often used to achieve this goal: by carefully choosing the linear layer, one can guarantee that all trails contain a large number of active S-boxes. This leads to bounds on the correlations of trails, provided that the linearity of the S-boxes is known. From the study of Boolean functions, several systematic constructions of S-boxes with low linearity are known.

Bounding the correlations of linear trails is necessary, but not sufficient. This is because the correlation of a linear approximation is a sum of the correlations of multiple trails, and a large number of small numbers may nonetheless add up to a high correlation. In the context of block ciphers, stronger security arguments can be obtained by upper bounding key-averaged squared correlations (variance of the correlation). For example, for four rounds of the AES with independent and uniform random round keys, such bounds were obtained by Hong *et al.* [26], Park *et al.* [26] and Canteaut and Roué [16]. The exact maximum was computed by Keliher and Sui [27].

Bounds on key-averaged squared correlations are still a far cry from a rigorous security argument against linear cryptanalysis. At best one can deduce that the correlation of individual linear approximations is low for many keys, but the strength of the conclusions is inherently limited. For example, for the 16-bit superbox of the block cipher Midori-64, the maximum variance of the correlation is equal to 2^{-8} [6, §4.1.1]. Chebyshev's inequality then shows that the correlation of one linear approximation can be ± 1 for at most one in 2^8 keys. It was shown in [5] using nonlinear invariants that there indeed exists a linear approximation with this behavior. In fact, the number of weak keys is larger because three such approximations exist. All current key-averaged results also depend on using independent round keys. Moreover, in the general study of vectorial Boolean functions, constructions typically do not involve a key.

Using current techniques, upper bounding maximum absolute correlations is not feasible for any practical cipher. Even for building blocks such as the AES superbox, little is known despite the relevance of such results for cryptanalysis. In particular, such bounds could be leveraged to improve trail or variance bounds for larger constructions, and are of interest to the analysis of Boolean functions. In fact, most known results originated in the study of vectorial Boolean functions. For example, the AES S-box is based on inversion in a finite field of order 2^8. As first observed by Carlet and reported by Nyberg [29], its linearity follows from estimates of Kloosterman sums. More recently, low-degree monomial functions have been used in arithmetization-oriented primitives such as MiMC [2], Rescue [3], Poseidon [24] and Anemoi [12]. The linearity of low-degree univariate polynomials can be understood using Weil's bound for exponential sums [33]. For constructions based on low-degree multivariate polynomials, only the quadratic case is well understood. For example, the analysis of the Butterfly construction by Canteaut, Duval and Perrin [14] is based on the fact that, in characteristic two, the function $x \mapsto x^3$ is quadratic as a multivariate polynomial over the base field. The linearity of instances with non-quadratic functions is not known. Similarly, for the Flystel construction (a variant of the Butterfly structure used in Anemoi), determining the linearity is an open problem [11, 13].

Contribution. This paper shows that results from modern algebraic geometry, and ℓ-adic cohomology in particular, have applications in linear cryptanalysis. Specifically, we use bounds for algebraic exponential sums deriving from Grothendieck's trace formula for ℓ-adic sheaves to upper bound the correlations

of linear approximations for various cryptographic constructions that are not amenable to other methods.

To illustrate how these results can be applied, we prove correlation bounds for three important constructions, each defined over an arbitrary finite field $\mathbf{F}_q$: a generalization of the Butterfly structure, three-round Feistel ciphers, and a generalization of the Flystel construction. The resulting bounds are of the order $\mathcal{O}(1/q)$, which is optimal up to constants. In contrast, using linear trails and Weil's bound yields estimates of order at best $\mathcal{O}(1/\sqrt{q})$. Despite the similarities between these three constructions, each of our proofs is based on a different result about algebraic exponential sums. Section 3 provides a high-level overview of the results that we use, and their relation to ℓ-adic cohomology. Since this approach is applicable to constructions over arbitrary finite fields, it is particularly relevant for the analysis of arithmetization-oriented primitives.

In Section 4, we use a theorem of Deligne to prove correlation upper bounds for a generalization of the (open and closed) Butterfly construction over arbitrary finite fields. Deligne's theorem [19] is probably the best known extension of Weil's bound to the multivariate setting. With some technical conditions, Theorem 4 establishes a correlation bound of $(d-1)^2/q$ for the degree-d generalized Butterfly construction.

Deligne's theorem is not sufficient to obtain correlation bounds for three-round Feistel ciphers. Instead, we rely on a theorem of Denef and Loeser [22]. This result replaces a smoothness condition in Deligne's theorem by the weaker requirement of non-degeneracy with respect to a Newton polyhedron. Up to technical conditions, Theorem 5 shows that the maximum absolute correlation of linear approximations of three-round Feistel ciphers with round functions of degree d is at most $(d+1)(d-1)^2/q$. Note that we focus on three round Feistel networks because two rounds are not enough to ensure low correlations.

In Section 6, we apply a different generalization of Deligne's theorem, due to Rojas-León [32], to prove correlation bounds for a generalization of the (open and closed) Flystel construction. The result of Rojas-León avoids the smoothness condition in Deligne's theorem for special types of singularities. Theorem 6 establishes a correlation bound of $(d-1)^2/q$ for generalized Flystel constructions of degree d. This resolves the conjecture from [11, 13] on the linearity of the Flystel construction.

Prior work by the authors. Deligne's theorem was used by the first author in the ePrint note [10, §3.2] from 2020. The result on the Flystel construction first appeared in the authors' joint ePrint note [9]. Neither of these notes have been submitted elsewhere for publication. Part of the motivation of this paper was to provide a more systematic treatment of these methods.

2 Linear cryptanalysis

We assume that the reader is familiar with ordinary linear cryptanalysis, *i.e.* over groups of the form $\mathbf{F}_2^n$. Since many applications of our results are relevant

for primitives defined over finite fields of odd characteristic, we briefly review linear cryptanalysis for groups of the form $\mathbf{F}_q^n$ with q a power of a prime p.

A partial generalization of linear cryptanalysis to arbitrary finite Abelian groups was given by Baignères, Stern and Vaudenay [4]. Since our results do not depend on averaging with respect to a random key, we instead follow the description given in [7, §3.3] and [8, Chapter 3]. Concretely, a linear approximation of a function $\mathsf{F}\colon \mathbf{F}_q^n \to \mathbf{F}_q^m$ consists of a pair (ψ, χ) of characters of the groups $\mathbf{F}_q^n$ and $\mathbf{F}_q^m$ respectively. The correlation of (ψ, χ) is defined as[3]

$$C^{\mathsf{F}}_{\chi,\psi} = \frac{1}{q^n} \sum_{x \in \mathbf{F}_q^n} \chi\big(\mathsf{F}(x)\big)\, \psi(-x)\,.$$

For a uniform random function or permutation, the absolute values of the correlations are close to $1/\sqrt{q^n}$ with high probability. If there exists a linear approximation with absolute correlation c significantly larger than this, then this leads to a distinguisher with data-complexity approximately $1/c^2$.

The matrix C^{F} with coordinates $C^{\mathsf{F}}_{\chi,\psi}$ labeled by characters ψ and χ is called the *correlation matrix* of F. If $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, then $C^{\mathsf{F}} = C^{\mathsf{F}_r} \cdots C^{\mathsf{F}_2} C^{\mathsf{F}_1}$ [7, Theorem 3.2] so that

$$C^{\mathsf{F}}_{\chi_{r+1},\chi_1} = \sum_{\chi_2,\cdots,\chi_r} C^{\mathsf{F}_r}_{\chi_{r+1},\chi_r} \cdots C^{\mathsf{F}_2}_{\chi_3,\chi_2} C^{\mathsf{F}_1}_{\chi_2,\chi_1}\,,$$

where the sum is over all sequences of characters $(\chi_1, \chi_2, \ldots, \chi_{r+1})$ with the endpoints χ_1 and χ_{r+1} fixed. These sequences are called *linear trails* and the product $\prod_{i=1}^{r} C^{\mathsf{F}_i}_{\chi_{i+1},\chi_i}$ is called the correlation of the corresponding trail.

Up to an arbitrary choice of a primitive character $\omega\colon \mathbf{F}_q \to \mathbf{C}^\times$ of $\mathbf{F}_q$, such as $\omega\colon x \mapsto \zeta_p^{\mathrm{Tr}(x)}$ with ζ_p a primitive p^{th} root of unity, the characters ψ and χ can be written as $\psi(x) = \omega(u^{\mathsf{T}}x)$ and $\chi(x) = \omega(v^{\mathsf{T}}x)$ with u and v vectors in $\mathbf{F}_q^n$. Hence, the correlations of linear approximations can be expressed as

$$C^{\mathsf{F}}_{\chi,\psi} = \frac{1}{q^n} \sum_{x \in \mathbf{F}_q^n} \omega\Big(v^{\mathsf{T}}\mathsf{F}(x) - u^{\mathsf{T}}x\Big)\,.$$

This is called a complete (algebraic) exponential sum. Since sums of this type have many applications, several techniques have been developed to bound their absolute value. The next section reviews a specific class of such methods which will be used throughout this paper.

3 Estimates for algebraic exponential sums

Let $f\colon \mathbf{F}_q^n \to \mathbf{F}_q$ be a function. As discussed in Section 2, bounding the absolute correlation of a linear approximation over a function is equivalent to upper

[3] Unlike [7], we include the minus sign in ψ rather than in χ. This corresponds to the group action $x \mapsto x+t$ instead of $x \mapsto x-t$. The motivation for this is that the former convention works in a more general setting that includes integral cryptanalysis.

bounding the absolute value of a certain exponential sum of the form

$$S(f) = \sum_{x \in \mathbf{F}_q^n} \omega(f(x)) \,, \tag{1}$$

with ω an additive character of $\mathbf{F}_q$. In this section, we review several algebro-geometric estimates for such sums when the algebraic degree of f is low. A first example of such an estimate is Weil's bound [33] for $n = 1$,

$$|S(f)| \leq (d-1)\sqrt{q} \,,$$

for all f of degree $d \geq 2$ coprime to the characteristic p of $\mathbf{F}_q$. For $n \geq 2$, obtaining estimates close to $q^{n/2}$ is more difficult and requires additional assumptions on f. Nevertheless, comparable results do exist. These results arise from a cohomological interpretation of (1), which we describe from a high level in Section 3.1.

3.1 Cohomological framework

In pursuit of Weil's conjectures on the number of points of algebraic varieties over finite fields, Grothendieck and his collaborators developed ℓ-adic cohomology theory [20,25]. In this framework, exponential sums such as $S(f)$ can be expressed in terms of the traces of linear maps (the action of geometric Frobenius) on certain vector spaces (ℓ-adic cohomology spaces) over an algebraic closure $\overline{\mathbf{Q}}_\ell$ of the field of ℓ-adic numbers. Throughout, ℓ is a prime different from the characteristic p of $\mathbf{F}_q$.

A complete description of this approach would lead us too far and is not necessary to apply the results, but we give a surface-level overview so that readers unfamiliar with the theory can understand its advantages and limitations. For an additive character ω of $\mathbf{F}_q$ and a function $f\colon \mathbf{F}_q^n \to \mathbf{F}_q$, one can construct a certain object called an ℓ-adic sheaf $\mathcal{L}$ on the affine space $\mathbf{A}^n$ over $\overline{\mathbf{F}}_q$. The stalk of $\mathcal{L}$ at a point x on $\mathbf{A}^n$ is a one-dimensional vector space $\mathcal{L}_x$ over $\overline{\mathbf{Q}}_\ell$ with an action of the Frobenius automorphism $x \mapsto x^q$. The action is given by the geometric Frobenius map σ and satisfies

$$\operatorname{Tr}\left(\sigma_x \mid \mathcal{L}_x\right) = \omega(f(x)) \,.$$

Since the map σ_x on the stalk $\mathcal{L}_x$ corresponds to multiplication by a constant, taking the trace is not doing a lot of work. However, the same formula holds for more general sheaves with higher-dimensional stalks. For our purposes, the importance of this fact is that $S(f)$ can be expressed as

$$S(f) = \sum_{x \in \mathbf{F}_q^n} \operatorname{Tr}\left(\sigma_x \mid \mathcal{L}_x\right) . \tag{2}$$

The sum (2) ranges over the points of $\mathbf{A}^n$ that are fixed under the Frobenius map. In analogy with the Lefschetz fixed-point theorem from algebraic topology,

Grothendieck [25, Exposé III] proved that (2) can be expressed in terms of the traces of the global geometric Frobenius automorphism F on ℓ-adic cohomology:

$$S(f) = \sum_{x \in \mathbf{F}_q^n} \operatorname{Tr}\left(\sigma_x \mid \mathcal{L}_x\right) = \sum_{i=0}^{2n} (-1)^i \operatorname{Tr}\left(F \mid H_c^i(\mathbf{A}^n, \mathcal{L})\right). \tag{3}$$

In particular, $H_c^i(\mathbf{A}^n, \mathcal{L})$ is the i^{th} ℓ-adic cohomology group with compact supports. The precise definition of ℓ-adic cohomology would lead us too far. For our discussion, it is enough to say that $H_c^i(\mathbf{A}^n, \mathcal{L})$ is a finite-dimensional vector space over $\overline{\mathbf{Q}}_\ell$ and F is a linear operator. Hence, the trace of F on $H_c^i(\mathbf{A}^n, \mathcal{L})$ is the sum of its eigenvalues $\lambda_1, \lambda_2, \ldots$. The number of eigenvalues is equal to $\dim H_c^i(\mathbf{A}^n, \mathcal{L})$. Suppose that, for all i, $|\lambda_i| \leq \kappa$ for some constant κ, then one obtains the following bound on the absolute value of $S(f)$:

$$|S(f)| \leq \kappa \sum_{i=0}^{2n} \dim H_c^i(\mathbf{A}^n, \mathcal{L}).$$

This approach allows one to bound the absolute correlations of linear approximations, provided that one has (i) a bound on the eigenvalues of F (ii) a bound on the dimensions of the cohomology spaces. In certain cases, these bounds are provided by the 'vanishing and purity' results in Sections 3.2 to 3.4.

3.2 Vanishing and purity from smoothness

Deligne proved a general result for 'pure' and 'lisse' ℓ-adic sheaves [21, Corollaire 3.3.4] that implies that the eigenvalues of F on $H_c^i(\mathbf{A}^n, \mathcal{L})$ are at most $q^{i/2}$ in absolute value. This is sometimes summarized by saying that $H_c^i(\mathbf{A}^n, \mathcal{L})$ is mixed of weight i. If equality holds, the cohomology space is called pure of weight i.

More detailed results can be obtained with additional conditions on f. In particular, for the applications in Sections 4 to 6, we will rely on results showing that all cohomology spaces $H_c^i(\mathbf{A}^n, \mathcal{L})$ with $i \neq n$ vanish and that $H_c^n(\mathbf{A}^n, \mathcal{L})$ is pure of weight n. A first example of such a result is the following theorem of Deligne, which will be used in Section 4.

Theorem 1 (Deligne [19, Théorème 8.4]). *Let f be a polynomial over $\mathbf{F}_q$ in n variables and with degree d coprime to the characteristic of $\mathbf{F}_q$. Let f_d be the degree d homogeneous component of f. If the projective hypersurface in $\mathbf{P}^{n-1}$ defined by $f_d = 0$ is smooth, then for the ℓ-adic sheaf $\mathcal{L}$ associated to (1),*

1. *For all $i \neq n$, we have vanishing cohomology $H_c^i(\mathbf{A}^n, \mathcal{L}) = 0$.*
2. *$H_c^n(\mathbf{A}^n, \mathcal{L})$ is pure of weight n with dimension at most $(d-1)^n$.*

Explicitly, the smoothness condition in Theorem 1 requires that the system of equations $f_d = \partial f_d/\partial x_1 = \cdots = \partial f_d/\partial x_n = 0$ has no nonzero solutions. As explained at the end of Section 3.1, Theorem 1 implies the estimate

$$|S(f)| \leq (d-1)^n \sqrt{q^n}.$$

Weil's bound is the special case of this inequality with $n = 1$, as the smoothness condition is then automatically satisfied.

Deligne's theorem requires that the maximum-degree homogeneous component of f defines a smooth hypersurface in $\mathbf{P}^{n-1}$, but this is often a problem for applications in linear cryptanalysis. For example, in unpublished work [10, §3.2], the first author applied Theorem 1 to two rounds of RESCUE with mixed success: it led to a good bound for most linear approximations, but for some it was necessary to resort to other techniques – leading to suboptimal bounds. This is problematic from a cryptanalytic point of view, since the worst case is particularly important.

3.3 Vanishing and purity from Newton polyhedra

Adolphson and Sperber showed that the smoothness condition in Theorem 1 can be replaced by a weaker non-degeneracy condition [1, Theorems 4.2 and 5.18]. The conditions of their result were relaxed as a consequence of subsequent work by Denef and Loeser [22, Theorem 9.2]. The results of Adolphson-Sperber and Denef-Loeser depend on the *Newton polyhedron* of f. Suppose that the algebraic normal form of f is equal to

$$f(x_1, \ldots, x_n) = \sum_{e_1,\ldots,e_n} c_{e_1,\ldots,e_n} \prod_{i=1}^{n} x_i^{e_i} ,$$

where the exponents $e_1, \ldots, e_n$ are in $\{0, \ldots, q-1\}$ and the coefficients $c_{e_1,\ldots,e_n}$ in $\mathbf{F}_q$. The Newton polyhedron $\Delta(f)$ of f at infinity is equal to the convex hull of the following set of points in $\mathbf{R}^n$:

$$\big\{(0, \ldots, 0)\big\} \cup \big\{(e_1, \ldots, e_n) \mid c_{e_1,\ldots,e_n} \neq 0\big\} \subset \mathbf{R}^n .$$

The function f is called *commode* if there exist nonzero integers $d_1, \ldots, d_n$ so that $(d_1, 0, 0, \ldots, 0), (0, d_2, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, d_n) \in \Delta(f)$. Since $\Delta(f)$ is a polyhedron, its boundary consists of faces. For a face τ of $\Delta(f)$, let

$$f_\tau(x_1, \ldots, x_n) = \sum_{(e_1,\ldots,e_n) \in \tau} c_{e_1,\ldots,e_n} \prod_{i=1}^{n} x_i^{e_i} ,$$

where the sum is over all integer points of the face τ and $c_{e_1,\ldots,e_n} = 0$ if $\prod_{i=1}^{n} x_i^{e_i}$ is not a monomial in f. The function f is called non-degenerate with respect to its Newton polyhedron if, for every face τ of $\Delta(f)$ not containing the origin,

$$\frac{\partial f_\tau}{\partial x_1} = \cdots = \frac{\partial f_\tau}{\partial x_n} = 0 ,$$

has no nonzero solutions. For a set of indices I, let $\mathrm{Vol}_I\, \Delta(f)$ be the volume of the intersection of Δ_f and the hyperplanes defined by $x_i = 0$ for all i in I. The Newton number of $\Delta(f)$ is by definition equal to

$$\nu(f) = \sum_{I \subseteq \{1,\ldots,n\}} (-1)^{|I|} (n - |I|)!\, \mathrm{Vol}_I\, \Delta(f) .$$

Denef and Loeser provide the following theorem.

Theorem 2 (Denef-Loeser [22, Theorem 9.2]). *If f is a polynomial over $\mathbf{F}_q$ in n variables that is commode and non-degenerate with respect to its Newton polyhedron, then for the ℓ-adic sheaf $\mathcal{L}$ associated to (1),*

1. *For all $i \neq n$, we have vanishing cohomology $H_c^i(\mathbf{A}^n, \mathcal{L}) = 0$.*
2. *$H_c^n(\mathbf{A}^n, \mathcal{L})$ is pure of weight n with dimension equal to $\nu(f)$.*

Theorem 2 implies the following upper bound on $|S(f)|$:

$$|S(f)| \leq \nu(f)\sqrt{q^n}\,.$$

This bound will be applied in Section 5. In some cases, such as for the Flystel construction that we consider in Section 6, non-degeneracy is not satisfied and more specialized results are necessary.

3.4 Vanishing and purity for isolated singularities

The following theorem due to Rojas-Léon [32, Theorem 2] shows that the conclusions of Theorem 1 are still true if the singularities of the hypersurface defined by $f_d = 0$ are well-behaved. In particular, the singularities should be of *isolated quasi-homogeneous* type.

A singular point of a hypersurface will be called isolated if there exists a Zariski neighborhood of the point that contains no other singular points. For concreteness, we define 'quasi-homogeneous singularity' in the case of an affine hypersurface, with a singularity that can be assumed to be in the origin by translation. The projective case follows by choosing an open affine set that contains the singularity. If the affine hypersurface in $\mathbf{A}^n$ defined by $h(x_1, \ldots, x_n) = 0$ has an isolated singularity at the origin, then we say that it is quasi-homogeneous if there exists a quasi-homogeneous polynomial g and an $\overline{\mathbf{F}}_q$-algebra isomorphism

$$\frac{\overline{\mathbf{F}}_q[[x_1, \ldots, x_n]]}{(h)} \cong \frac{\overline{\mathbf{F}}_q[[x_1, \ldots, x_n]]}{(g)}\,,$$

with $\overline{\mathbf{F}}_q[[x_1, \ldots, x_n]]$ the ring of formal power series in $x_1, \ldots, x_n$ over $\overline{\mathbf{F}}_q$. A polynomial g is called quasi-homogeneous of degree δ if there exist 'weights' $w_1, \ldots, w_n$ such that

$$g(\lambda^{w_1} x_1, \ldots, \lambda^{w_n} x_n) = \lambda^{\delta} g(x_1, \ldots, x_n)\,,$$

for all λ in $\overline{\mathbf{F}}_q$. The *Milnor number* of the singularity is equal to $\prod_{i=1}^{n}(\delta/w_i - 1)$ and it can be shown that this does not depend on g. Rojas-Léon has proven the following refinement of Theorem 1.

Theorem 3 (Rojas-Léon [32, Theorem 2]). *Let f be a degree-d polynomial over $\mathbf{F}_q$ in n variables with $f = f_d + f_{d'} + \cdots$, where f_d is the degree-d homogeneous component of f and $f_{d'}$ is its homogeneous component of degree $d' = \deg f - f_d$. Suppose that d and d' are coprime to the characteristic p of $\mathbf{F}_q$ and $d'/d > p/(p + (p-1)^2)$. If the projective hypersurface in $\mathbf{P}^{n-1}$ defined*

by $f_d = 0$ has at worst quasi-homogeneous isolated hypersurface singularities of degrees prime to p with Milnor numbers $\mu_1, \ldots, \mu_s$, and if the projective hypersurface in $\mathbf{P}^{n-1}$ defined by $f_{d'} = 0$ contains none of these singularities, then for the ℓ-adic sheaf $\mathcal{L}$ associated to (1),

1. *For all $i \neq n$, we have vanishing cohomology $H_c^i(\mathbf{A}^n, \mathcal{L}) = 0$.*
2. *$H_c^n(\mathbf{A}^n, \mathcal{L})$ is pure of weight n with dimension $(d-1)^n - (d-d')\sum_{i=1}^s \mu_i$.*

Theorem 3 implies the following estimate for the exponential sum (1):

$$|S(f)| \leq \left((d-1)^n - (d-d')\textstyle\sum_{i=1}^s \mu_i\right)\sqrt{q^n}\,.$$

This result will be used in Section 6.

4 Generalized Butterfly construction

The Butterfly construction was introduced by Perrin, Udovenko and Biryukov [31], and was originally defined over $\mathbf{F}_{2^n}^2$ with n odd. Several generalizations were investigated in [11, 14, 15, 28]. In what follows we study a generalization of the Butterfly construction that does not require the internal functions to be monomials.

4.1 Definition

Let $\mathsf{G} : \mathbf{F}_q \to \mathbf{F}_q$ be a permutation, $\mathsf{H} : \mathbf{F}_q \to \mathbf{F}_q$ a function, and α in $\mathbf{F}_q$. The generalized closed Butterfly construction is depicted in Figure 1b. In terms of polynomials, $\mathsf{F}\colon \mathbf{F}_q^2 \to \mathbf{F}_q^2$ is given by $\mathsf{F}(x_1, x_2) = (y_1, y_2)$, where

$$\begin{aligned} y_1 &= \mathsf{G}(x_1 + \alpha x_2) + \mathsf{H}(x_2) \\ y_2 &= \mathsf{G}(x_2 + \alpha x_1) + \mathsf{H}(x_1)\,. \end{aligned}$$

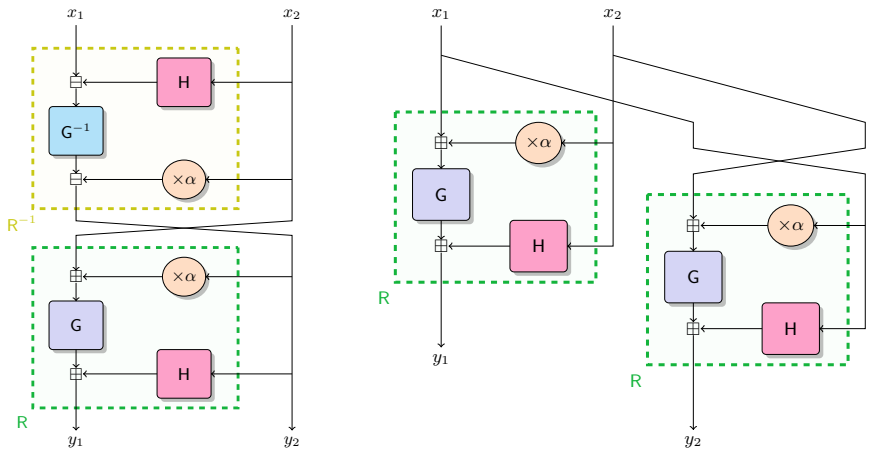
For brevity, we introduce the notation $\mathsf{F} = \textsc{Butterfly}[\mathsf{G}, \mathsf{H}, \alpha]$.

There is a variant of the closed Butterfly construction, called the open Butterfly construction, that resembles a Feistel construction. It is shown in Figure 1a. In terms of polynomials, the open Butterfly construction is defined by the map $(x_1, x_2) \mapsto (y_1, y_2)$, where

$$\begin{aligned} y_1 &= \mathsf{G}(x_2 + \alpha y_2) + \mathsf{H}(y_2) \\ y_2 &= \mathsf{G}^{-1}(x_1 - \mathsf{H}(x_2)) - \alpha x_2\,. \end{aligned}$$

The open and closed variants of the Butterfly construction are CCZ-equivalent [17], meaning that there is an affine bijection between their graphs.

The proof of CCZ-equivalence for the general case follows from the observation that the proof for the original Butterfly construction [31, Lemma 2] only uses the bijectivity of G. One consequence of the CCZ-equivalence between the closed and open Butterfly constructions is that their correlation matrices are the same up to sign and up to a permutation of their entries. Since the algebraic degree of the closed Butterfly is guaranteed to be low if G and H have low degree, in what follows we focus on the closed variant.



(a) Open variant.

(b) Closed variant.

Fig. 1: The generalized Butterfly construction.

4.2 Related work

The original Butterfly construction corresponds to $\mathsf{G}(x) = x^3$ and $\mathsf{H}(x) = \beta x^3$, with β a non-zero constant in $\mathbf{F}_q$, defined on a field $\mathbf{F}_q$ of characteristic two. Hence, to ensure that G is invertible, $q = 2^n$ with n odd. Canteaut, Duval and Perrin proved in [14, Main Theorem] that for $n > 3$, the maximum absolute correlation is $1/2^{n-1}$ unless $\beta = (\alpha + 1)^3$, in which case it is $1/2^{(n-1)/2}$. These results have been generalized to $\mathsf{G}(x) = x^d$ and $\mathsf{H}(x) = \beta x^d$ for other exponents d of Hamming weight two [23, 28].

The reason that previous work has been concerned only with exponents of Hamming weight two is that, in this case, the Butterfly construction is quadratic as a multivariate polynomial over the base field $\mathbf{F}_2$. Up to a linear change of variables, quadratic forms can be brought into a 'block diagonal' normal form. From there, one can see that absolute correlations can be expressed in terms of the rank of the form. Although the details may be technical, this approach is generally workable. The same approach works for fields of odd characteristic.

The results below are applicable to instances of the generalized Butterfly construction that are not quadratic over the base field $\mathbf{F}_p$. This is worth mentioning, because this case would be out of reach for the methods that have previously been used in the literature.

4.3 Correlation bound

In this section, we upper bound the absolute correlations of linear approximations over the generalized Butterfly construction using Theorem 1. This requires some conditions on G, H and α. In particular, we assume that $\deg \mathsf{G} \neq \deg \mathsf{H}$ and

if $\deg \mathsf{G} > \deg \mathsf{H}$, then $\alpha \neq \pm 1$. As discussed below, some of these conditions can be removed but this requires more sophisticated results such as Theorem 3.

Lemma 1. *Let x_1 and x_2 be elements of $\mathbf{F}_q$, and v in $\mathbf{F}_q^\times$. If $d \geq 2$ is an integer indivisible by the characteristic p of $\mathbf{F}_q$, then the projective hypersurface in $\mathbf{P}^1$ defined by $x_1^d + v x_2^d = 0$ is smooth.*

Proof. Let $f(x_1, x_2) = x_1^d + v x_2^d$. The locus of singular points is defined by the equations $f = 0$, $\partial f / \partial x_1 = 0$ and $\partial f / \partial x_2 = 0$. Since $d \not\equiv 0 \pmod{p}$ and $v \neq 0$, the latter two equations are equivalent to $x_1^{d-1} = 0$ and $x_2^{d-1} = 0$. Hence, the projective hypersurface defined by $f = 0$ contains no singular points.

Theorem 4. *Let $\mathsf{F} = \textsc{Butterfly}[\mathsf{G}, \mathsf{H}, \alpha]$ be the generalized closed Butterfly construction with $\mathsf{G} \colon \mathbf{F}_q \to \mathbf{F}_q$ a permutation, $\mathsf{H} \colon \mathbf{F}_q \to \mathbf{F}_q$ a function and α in $\mathbf{F}_q$. If either $\deg \mathsf{G} > \deg \mathsf{H} > 1$ with $\deg \mathsf{G}$ indivisible by p and $\alpha \neq \pm 1$ or $\deg \mathsf{H} > \deg \mathsf{G} > 1$ with $\deg \mathsf{H}$ indivisible by p, then for every linear approximation (ψ, χ) of F with $\chi = (\chi_1, \chi_2) \neq (1, 1)$,*

$$\left| C^{\mathsf{F}}_{\chi,\psi} \right| \leq \frac{1}{q} \begin{cases} (\deg \mathsf{G} - 1)(\deg \mathsf{H} - 1) & \text{if } \chi_1 = 1 \text{ or } \chi_2 = 1\,, \\ (\max\{\deg \mathsf{G}, \deg \mathsf{H}\} - 1)^2 & \text{else}\,. \end{cases}$$

Proof. If $\chi_1 = 1$ or $\chi_2 = 1$, then there is at most one linear trail through the generalized Butterfly construction. In particular, if $\psi = (\psi_1, \psi_2)$ and $\chi_1 = 1$, then the correlation is equal to

$$C^{\mathsf{F}}_{\chi,\psi} = C^{\mathsf{G}}_{\chi_2,\psi_2} C^{\mathsf{H}}_{\chi_2,\psi_1/\psi_2^\alpha}\,.$$

Since G and H correspond to univariate polynomials, Weil's bound implies

$$\left| C^{\mathsf{F}}_{\chi,\psi} \right| \leq (\deg \mathsf{G} - 1)/\sqrt{q} \times (\deg \mathsf{H} - 1)/\sqrt{q}\,.$$

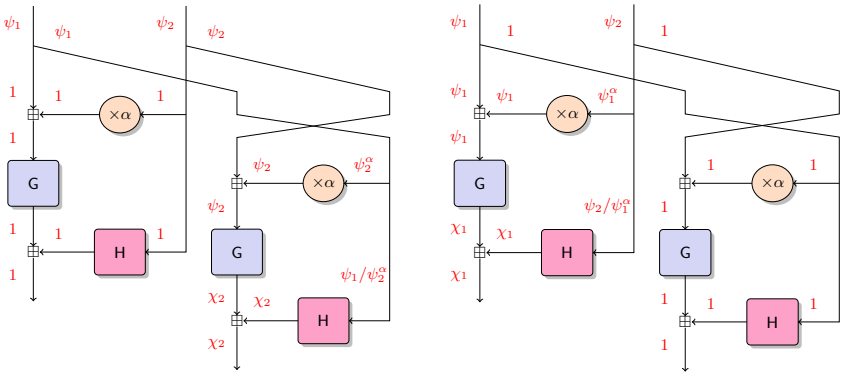
By symmetry, the case $\chi_2 = 1$ is analogous.

If $\chi_1 \neq 1$ and $\chi_2 \neq 1$, then let $\psi_i(x_i) = \omega(u_i x_i)$ and $\chi_i(x_i) = \omega(v_i x_i)$ with v_1 and v_2 nonzero. Up to a factor $1/q^2$, the correlation of the linear approximation (ψ, χ) is equal to an exponential sum of the form (1) with

$$f(x_1, x_2) = v_1 \mathsf{G}(x_1 + \alpha x_2) + v_2 \mathsf{G}(x_2 + \alpha x_1) + v_1 \mathsf{H}(x_2) + v_2 \mathsf{H}(x_1) - u_1 x_1 - u_2 x_2\,.$$

If $\deg \mathsf{H} > \deg \mathsf{G}$, then up to scaling the maximal degree homogeneous component of f is equal to $x_1^d + (v_1/v_2)\, x_2^d$ with $d = \deg \mathsf{H}$. By Lemma 1 with $v = v_1/v_2 \neq 0$, this defines a smooth projective hypersurface. The bound then follows from Theorem 1. If $\deg \mathsf{G} > \deg \mathsf{H}$, then the maximal degree homogeneous component of f is equal to $(x_1 + \alpha x_2)^d + (v_2/v_1)\,(x_2 + \alpha x_1)^d$ with $d = \deg \mathsf{G}$. Up to the linear transformation $(x_1, x_2) \mapsto (x_1 + \alpha x_2, x_2 + \alpha x_1)$, which is invertible if and only if $\alpha \neq \pm 1$, this defines the same hypersurface as $x_1^d + (v_2/v_1)\, x_2^d = 0$. Hence, the bound again follows from Lemma 1 and Theorem 1.

The condition $\alpha \neq \pm 1$ in Theorem 4 can be relaxed using Theorem 3. The analysis of this case is similar to that of the Flystel construction in Section 6. Avoiding the assumption $\deg \mathsf{G} \neq \deg \mathsf{H}$ is more technical, and nontrivial conditions on G and H must be imposed to obtain a bound of order $\mathcal{O}(1/q)$. In principle, however, Theorem 3 is in many cases sufficient.



(a) $\chi_1 = 1$. (b) $\chi_2 = 1$.

Fig. 2: Linear trails for a Butterfly construction.

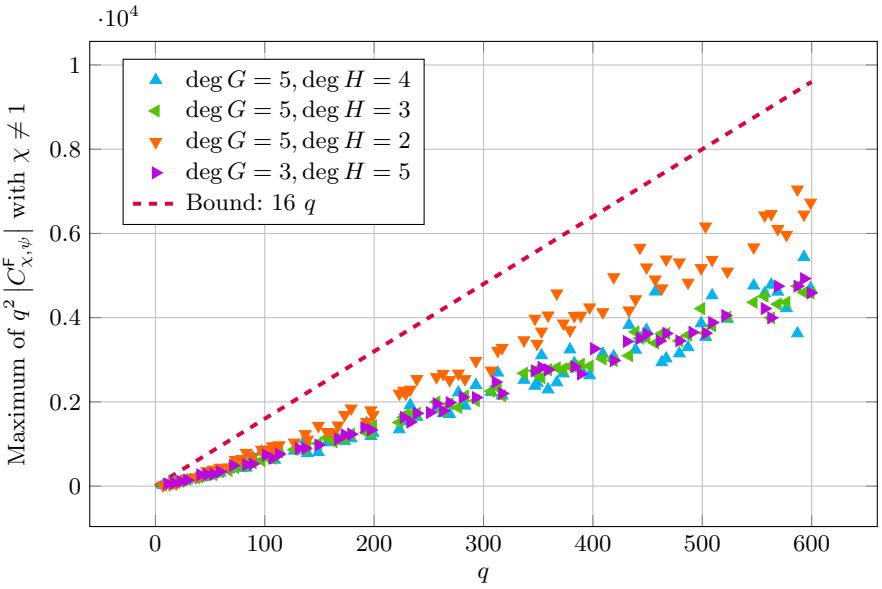
4.4 Experimental results

An experimental verification of the bound in Theorem 4 is presented in Figure 3, for q a prime and G and H monomial functions. The gap between the theoretical bound and the experimental observations may be due to a lack of tightness, or due to the fact that the experiment is necessarily limited to a small number of choices for G, H, α and q. For example, although the bound only depends on the maximum of the degrees of G and H, Figure 3a suggests that the minimum of the degrees also has some influence on the maximum absolute correlation – at least for monomial functions. In particular, for $\deg \mathsf{G} = 5$, the linearity is highest for $\deg \mathsf{H} = 2$. Similarly, Figure 3b shows that the value of α may affect the result.

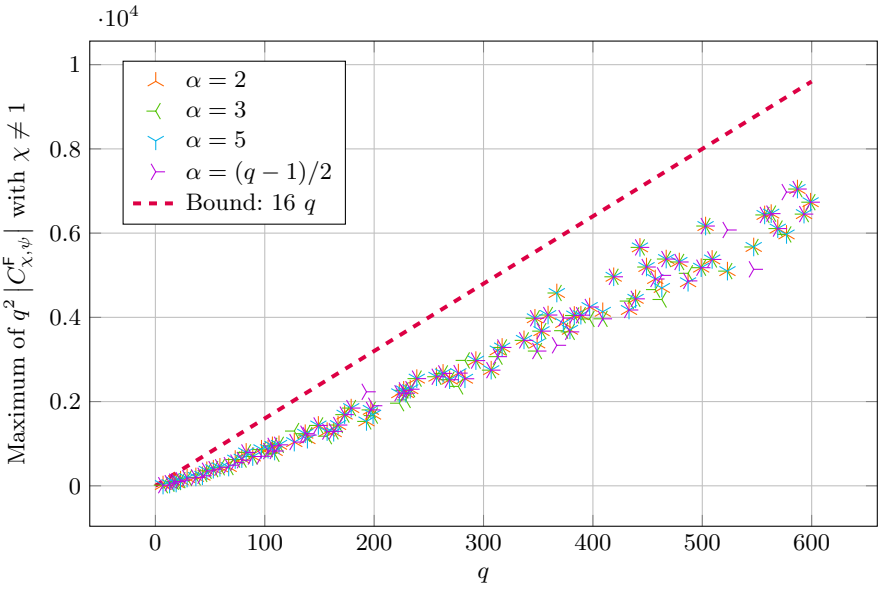
In any case, even if the bound is not tight, we expect it would be quite technical to improve over it using the same methods. Indeed, using Theorem 2, one can show that the dimension of the cohomology space $H^1_c(\mathbf{A}^2, \mathcal{L})$ matches the bound provided by Deligne's theorem. Hence, the only error introduced in the estimate is due to the relative signs of the eigenvalues of the action of Frobenius on $H^1_c(\mathbf{A}^2, \mathcal{L})$.

5 Feistel construction

In this section we analyze three-round Feistel ciphers with low-degree round functions. Although the generalized open Butterfly construction from Section 4 is closely related to the two-round Feistel construction, its analysis is different because the round functions are not of low degree. In addition, as recalled in Section 5.2, at least three rounds are necessary for the maximum absolute correlation to be of the order of $1/q$ for traditional Feistel ciphers over $\mathbf{F}_q$.



(a) Low-degree functions ($\max\{\deg \mathsf{G}, \deg \mathsf{H}\} = 5$ and $\alpha = 2$).



(b) Influence of α ($\deg \mathsf{G} = 5$ and $\deg \mathsf{H} = 2$).

Fig. 3: Experimental verification of correlation bounds from Theorem 4 for the generalized Butterfly construction $\mathsf{F} = \textsc{Butterfly}[\mathsf{G}, \mathsf{H}, \alpha]$ over a finite field $\mathbf{F}_q$ of prime order, and with G and H monomial functions.

5.1 Definition

Figure 4 depicts three rounds of a Feistel cipher with round functions F_1, F_2 and F_3, with respective degrees d_1, d_2 and d_3. This construction will be denoted by $\textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$. For the purpose of later calculations, we introduce intermediate variables z_1 and z_2, as shown in Figure 4. This leads to the following implicit equations for $\textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$:

$$\begin{aligned} x_1 &= z_1 - \mathsf{F}_1(z_2) \\ x_2 &= z_2 \\ y_1 &= z_1 + \mathsf{F}_3(z_2 + \mathsf{F}_2(z_1)) \\ y_2 &= z_2 + \mathsf{F}_2(z_1)\,. \end{aligned}$$

In the following, z_1 and z_2 will be considered to be the inputs of a function with outputs x_1, x_2, y_1 and y_2. As we will discuss in Section 5.3, the correlations of linear approximations over $\textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_2]$ are equal to the correlations of (some) linear approximations of the function $(z_1, z_2) \mapsto (x_1, x_2, y_1, y_2)$.

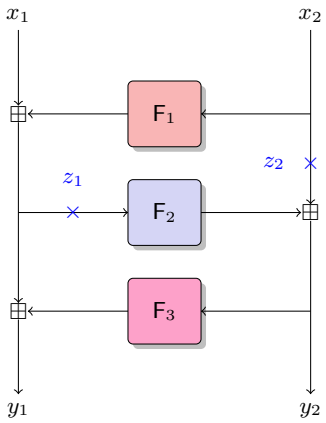


Fig. 4: A three-round Feistel construction.

5.2 Related work

For two-round Feistel ciphers, an upper bound on the absolute correlation follows from Weil's bound. This is because, as illustrated in Figure 5, there is at most one linear trail with nonzero correlation for every linear approximation. In particular, the correlation of (ψ, χ) is equal to

$$C^{\mathsf{F}_1}_{\psi_1, \psi_2/\chi_2} C^{\mathsf{F}_2}_{\chi_2, \psi_1/\chi_1}\,,$$

with $\psi = (\psi_1, \psi_2)$ and $\chi = (\chi_1, \chi_2)$. If neither $\psi_1 = 1$ and $\psi_2 = \chi_2$ nor $\chi_2 = 1$ and $\psi_1 = \chi_1$, then Weil's bound yields the upper bound $(d_1 - 1)(d_2 - 1)/q$. Otherwise, only one round function is active and Weil's bound gives the weak bound $(\max\{d_1, d_2\} - 1)/\sqrt{q}$. In fact, linear approximations with absolute correlation of the order of $1/\sqrt{q}$ cannot be avoided for a two-round Feistel cipher. At least three rounds are necessary to obtain a bound of the order of $1/q$.

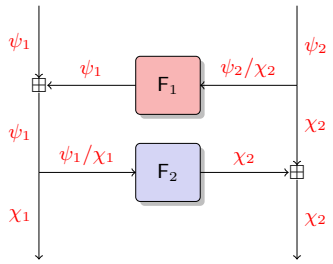


Fig. 5: Linear trail through a two-round Feistel cipher.

For three-round Feistel ciphers, little can be proven about the correlations of linear approximations with techniques from the literature. Even the case with F_1, F_2 and F_3 quadratic as multivariate polynomials over the base field $\mathbf{F}_p$ appears to be out of reach. At most, one can compute the variance of the correlations of linear approximations with respect to independent and uniform random round keys. This is comparable to the result of Nyberg and Knudsen [30] for the key-averaged probability of differentials in Feistel ciphers. However, as discussed in the introduction, such bounds are much weaker.

5.3 Correlation bound

Let $\mathsf{F} = \textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$. As discussed in Section 2, the correlation of a linear approximation (ψ, χ) for F with $\psi(x) = \omega(v^\mathsf{T} x)$ and $\chi(x) = \omega(u^\mathsf{T} x)$ is equal to

$$C^\mathsf{F}_{\chi,\psi} = \frac{1}{q^2} \sum_{x \in \mathbf{F}_q^2} \omega\big(v^\mathsf{T} \mathsf{F}(x_1, x_2) - u_1 x_1 - u_2 x_2\big)\,.$$

The substitution $x_1 = z_1 - \mathsf{F}_1(x_2)$ shows that $C^\mathsf{F}_{\chi,\psi}$ is equal to

$$C^\mathsf{F}_{\chi,\psi} = \frac{1}{q^2} \sum_{z \in \mathbf{F}_q^2} \omega\big(v^\mathsf{T} \mathsf{F}(z_1 - \mathsf{F}_1(z_2), z_2) + u_1 \mathsf{F}_1(z_2) - u_1 z_1 - u_2 z_2\big)\,.$$

As pointed out in Section 5.1, the outputs $(y_1, y_2) = \mathsf{F}(z_1 - \mathsf{F}_1(z_2), z_2)$ satisfy $y_1 = z_1 + \mathsf{F}_3(z_2 + \mathsf{F}_2(z_1))$ and $y_2 = z_2 + \mathsf{F}_2(z_1)$. It follows that the correlations can be rewritten as follows:

$$C^\mathsf{F}_{\chi,\psi} = \frac{1}{q^2} \sum_{x \in \mathbf{F}_q^2} \omega\big(v_1 z_1 + v_1 \mathsf{F}_3(z_2 + \mathsf{F}_2(z_1)) + v_2 z_2 + v_2 \mathsf{F}_2(z_1) + u_1 \mathsf{F}_1(z_2) - u_1 z_1 - u_2 z_2\big)\,.$$

Up to a factor $1/q^2$, this is an exponential sum of the form (1) with f given by

$$f(z_1, z_2) = v_1\mathsf{F}_3(z_2+\mathsf{F}_2(z_1))+v_2\mathsf{F}_2(z_1)+u_1\mathsf{F}_1(z_2)+(v_1-u_1)z_1+(v_2-u_2)z_2\,. \quad (4)$$

The proof of our upper bound on the absolute correlations of linear approximations for three-round Feistel ciphers is based on Theorem 2. This depends on the non-degeneracy of f with respect to its Newton polyhedron. Lemma 2 verifies this condition for most masks u_1 and v_1. For the remaining masks, an argument based on linear trails will be used to complete the bound.

Lemma 2. *Let F_1, F_2 and F_3 be functions on $\mathbf{F}_q$ with degrees d_1, d_2 and $d_3 \leq d_1$ respectively, all at least two and indivisible by the characteristic of $\mathbf{F}_q$. If $d_1 = d_3$, then suppose F_1 and F_3 have the same leading term. For all $u_1 \neq 0$, u_2, $v_1 \neq 0$ and v_2 such that $u_1 + v_1 \neq 0$, the Newton polyhedron of f defined by (4) is the triangle $\Delta(f)$ with vertices $(0,0)$, $(d_2d_3, 0)$ and $(0, d_1)$ shown in Figure 6. Furthermore, f is commode and non-degenerate with respect to $\Delta(f)$.*

Proof. Let a_1, a_2 and a_3 denote the leading coefficients of F_1, F_2 and F_3 respectively. If $d_1 = d_3$, then $a_1 = a_3$ by assumption. The Newton polyhedron of f is determined by its extremal points. It is sufficient to determine the highest degree terms in z_1 and z_2, and to show that all other exponents lie within the convex hull of the corresponding two points and the origin.

Since $v_1 \neq 0$, the term of f with highest degree in z_1 is equal to $v_1a_2a_3z_1^{d_2d_3}$. The term with highest degree in z_2 depends on the relative size of d_1 and d_3. If $d_1 > d_3$, then since $u_1 \neq 0$, it is $u_1a_1z_2^{d_1}$. If $d_1 = d_3$, then the term is $(u_1a_1 + v_1a_3)z_2^{d_1}$ and this is nonzero because $a_1 = a_3$ and $u_1 + v_1 \neq 0$. It follows that the Newton polyhedron of f contains the points $(d_2d_3, 0)$ and $(0, d_1)$. The convex hull of these points, together with zero, is the triangle shown in Figure 6. It suffices to verify that every monomial in f corresponds to a point within this triangle. By the binomial formula, the monomials in f are contained in the set

$$\left\{z_1^i \;\middle|\; 0 \leq i \leq d_2\right\} \cup \left\{z_2^i \;\middle|\; 0 \leq i \leq d_1\right\} \cup \left\{z_1^{i(d_3-j)}z_2^j \;\middle|\; 0 \leq i \leq d_2,\, 0 \leq j \leq d_3\right\}.$$

These points, represented in blue in Figure 6, are indeed within the triangle.

The triangle in Figure 6 has only one face τ not containing the origin, and the points on this face have coordinates equal to $((1-e)d_2d_3, ed_1)$ with e in $[0,1]$. If $d_1 > d_3$, then the only exponents of monomials that can occur in f and that are on this face are equal to $(d_2d_3, 0)$ and $(0, d_1)$. If $d_1 = d_3$, then the face τ coincides with the blue line in Figure 6. The points on this line with integer coordinates correspond to the exponents of monomials in $(z_2 + a_2z_1^{d_2})^{d_3}$. Hence,

$$f_\tau(z_1, z_2) = u_1a_1z_2^{d_1} + \begin{cases} v_1\big(z_2 + a_2\, z_1^{d_2}\big)^{d_3} & \text{if } d_1 = d_3\,, \\ v_1a_2a_3z_1^{d_2d_3} & \text{else}\,. \end{cases}$$

The non-degeneracy condition is quite different for the cases $d_1 > d_3$ and $d_1 = d_3$. If $d_1 > d_3$, then the partial derivatives are given by

$$\frac{\partial f_\tau}{\partial z_1} = v_1 a_2 a_3 d_2 d_3 z_1^{d_2 d_3 - 1}\,,$$
$$\frac{\partial f_\tau}{\partial z_2} = u_1 a_1 d_1 z_2^{d_1 - 1}\,.$$

It follows that $z_1 = z_2 = 0$ is the only solution of $\partial f_\tau / \partial z_1 = 0 = \partial f_\tau / \partial z_2$, so f_τ is non-degenerate with respect to its Newton polyhedron. If $d_1 = d_3$, then the partial derivatives are

$$\frac{\partial f_\tau}{\partial z_1} = v_1 a_2 d_2 d_3 (z_2 + a_2 z_1^{d_2})^{d_3 - 1} z_1^{d_2 - 1}$$
$$\frac{\partial f_\tau}{\partial z_2} = v_1 d_3 (z_2 + a_2 z_1^{d_2})^{d_3 - 1} + u_1 a_1 d_1 z_2^{d_1 - 1}\,.$$

Solutions to $\partial f_\tau / \partial z_1 = 0$ satisfy either $z_2 + a_2 z_1^{d_2} = 0$ or $z_1 = 0$. In the former case, $\partial f_\tau / \partial z_2 = 0$ implies $z_2 = 0$ whence $z_1 = 0$. In the latter case, we get $v_1 d_3 z_2^{d_3 - 1} = 0$ so that $z_2 = 0$. It follows that f is non-degenerate.

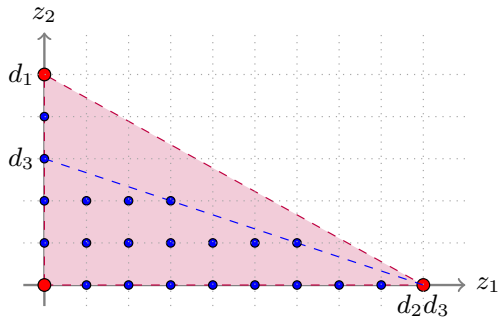


Fig. 6: Newton polyhedron of f.

The condition $d_1 \geq d_3$ in Lemma 2 is necessary to ensure non-degeneracy. Nevertheless, and although this is not stated, the bound in Theorem 5 is also applicable to Feistel ciphers with $d_1 < d_3$. Indeed, one can apply Theorem 5 to the inverse, which is the same but with F_1 and F_3 swapped. Since correlation matrices are unitary, any bound for F^{-1} transfers to a bound for F.

Theorem 5. *Let* $\mathsf{F} = \textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$ *with round functions* F_1, F_2 *and* F_3 *of degrees* d_1, d_2 *and* $d_3 \leq d_1$ *all greater than two and indivisible by the characteristic of* $\mathbf{F}_q$. *If* $d_1 = d_3$, *then suppose that* F_1 *and* F_3 *have the same leading*

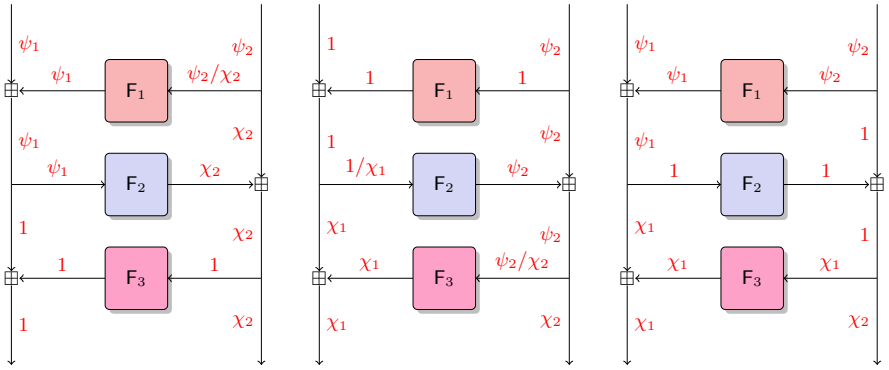
coefficient. If F_2 *is a permutation, then for all characters* $\psi = (\psi_1, \psi_2)$ *and* $\chi = (\chi_1, \chi_2) \neq (1, 1)$,

$$\left|C^{\mathsf{F}}_{\chi,\psi}\right| \leq \frac{1}{q} \begin{cases} (d_1 - 1)(d_2 - 1) & \text{if } \psi_1 \neq 1 \text{ and } \chi_1 = 1\,, \\ (d_3 - 1)(d_2 - 1) & \text{if } \psi_1 = 1 \text{ and } \chi_1 \neq 1\,, \\ (d_1 - 1)(d_3 - 1) & \text{if } \psi_1\chi_1 = 1\,, \\ (d_1 - 1)(d_2 d_3 - 1) & \text{else}\,. \end{cases}$$

Proof. Lemma 2 is applicable when $\psi_1 \neq 1$, $\chi_1 \neq 1$ and $\psi_1\chi_1 \neq 1$, which corresponds to the 'else' case in the theorem statement. Let f be as in (4). The lemma shows that f is commode and non-degenerate, so Theorem 2 is applicable. The dimension of $H^n_c(\mathbf{A}^n, \mathcal{L})$ is equal to the Newton number $\nu(f)$ of $\Delta(f)$:

$$\begin{aligned} \nu(f) &= \sum_{I \subseteq \{1,2\}} (-1)^{|I|}(2 - |I|)!\ \mathrm{Vol}_I\, \Delta(f) \\ &= 2 \cdot (d_1 d_2 d_3/2) - d_2 d_3 - d_1 + 1 \\ &= (d_2 d_3 - 1)(d_1 - 1)\,. \end{aligned}$$

This proves the bound for the 'else' case. In the first three cases, there is at most one trail with nonzero correlation (see Figure 7). Hence, Weil's bound can be used. For the first two cases, this yields bounds $(d_1 - 1)/\sqrt{q} \times (d_2 - 1)/\sqrt{q}$ and $(d_2 - 1)/\sqrt{q} \times (d_3 - 1)/\sqrt{q}$. For the third case, the fact that there is at most one trail with nonzero correlation depends on the fact that F_2 is a permutation. Weil's bound gives $(d_1 - 1)/\sqrt{q} \times (d_3 - 1)/\sqrt{q}$.



(a) $\psi_1 \neq 1$ and $\chi_1 = 1$. (b) $\psi_1 = 1$ and $\chi_1 \neq 1$. (c) $\psi_1\chi_1 = 1$.

Fig. 7: Linear trails for a three-round Feistel construction.

In Theorem 5, the condition that F_2 is a permutation cannot be dropped. Indeed, as shown in Figure 8, if F_2 is not a permutation then there exists a linear

trail that only activates F_2. Generically, this trail has absolute correlation on the order of $1/\sqrt{q}$. It follows that F_2 must be a permutation to achieve a bound of order $1/q$.

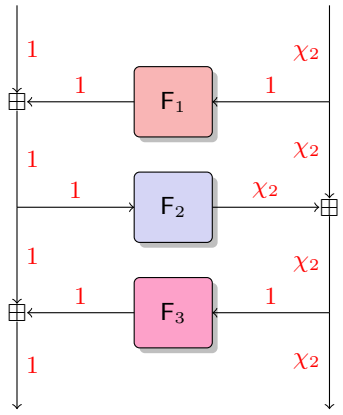


Fig. 8: Linear trail through a three-round Feistel cipher with non-bijective F_2.

5.4 Experimental results

Experimental results for two Feistel constructions with round functions of degrees that lead to the same bound are presented in Figure 9. It is worth noting that if F_2 is not a permutation, then the linearity is indeed very far above the bound from Theorem 5.

For $\textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$ with functions F_1, F_2 and F_3 of degree similar to the degree of the functions functions G and H in $\textsc{Butterfly}[\mathsf{G}, \mathsf{H}, \alpha]$, from Figures 3 and 9 the bound appears to be less tight in the case of the three-round Feistel construction. The reason for this is that the overall degree of the construction is higher, leading to a higher dimension of the cohomology space $H^1_c(\mathbf{A}^2, \mathcal{L})$ and hence more opportunities for eigenvalues to cancel.

6 Generalized Flystel construction

The Flystel construction was first introduced in [12] as the non-linear component in the family of hash functions Anemoi. In this section we analyze a generalization of the Flystel construction. Despite the similarity with the generalized Butterfly and Feistel constructions from Sections 4 and 5, a specific analysis is necessary because Theorems 1 and 2 are not applicable.

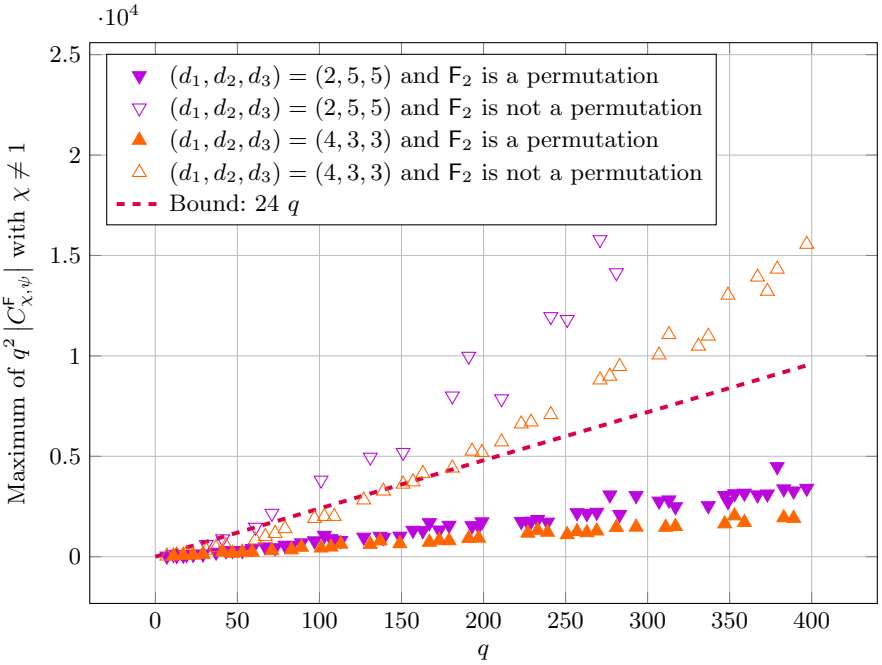


Fig. 9: Experimental verification of correlation bounds from Theorem 5 for the three-round Feistel construction $\mathsf{F} = \textsc{Feistel}[\mathsf{F}_1, \mathsf{F}_2, \mathsf{F}_3]$ over a field $\mathbf{F}_q$ of prime order and with F_1, F_2 and F_3 monomials of degrees d_1, d_2 and d_3 respectively.

6.1 Definition

Let $\mathsf{H}_1 : \mathbf{F}_q \to \mathbf{F}_q$ and $\mathsf{H}_2 : \mathbf{F}_q \to \mathbf{F}_q$ be two functions, and let $\mathsf{G} : \mathbf{F}_q \to \mathbf{F}_q$ be a permutation. The closed generalized Flystel construction (see Figure 10b) will be denoted by $\mathsf{F} = \textsc{Flystel}[\mathsf{H}_1, \mathsf{G}, \mathsf{H}_2]$. Algebraically, the function $\mathsf{F}\colon \mathbf{F}_q^2 \to \mathbf{F}_q^2$ is given by $\mathsf{F}(x_1, x_2) = (y_1, y_2)$, where

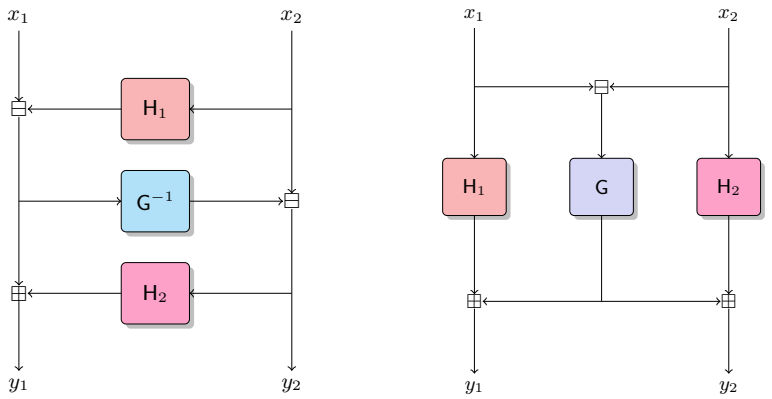
$$\begin{aligned} y_1 &= \mathsf{G}(x_1 - x_2) + \mathsf{H}_1(x_1) \\ y_2 &= \mathsf{G}(x_1 - x_2) + \mathsf{H}_2(x_2)\,. \end{aligned}$$

The open generalized Flystel construction is shown in Figure 10a. It corresponds to the map $(x_1, x_2) \mapsto (y_1, y_2)$, with

$$\begin{aligned} y_1 &= x_1 - \mathsf{H}_1(x_2) + \mathsf{H}_2(x_2 - \mathsf{G}^{-1}(x_1 - \mathsf{H}_1(x_2))) \\ y_2 &= x_2 - \mathsf{G}^{-1}(x_1 - \mathsf{H}_1(x_2))\,. \end{aligned}$$

For the original Flystel construction, G, H_1 and H_2 are given by $\mathsf{G}(x) = x^d$, $\mathsf{H}_1(x) = \beta x^2 + \gamma$ and $\mathsf{H}_2(x) = \beta x^2 + \delta$ with β a non-zero constant in $\mathbf{F}_q$.

Like for the Flystel construction, for a given tuple $(\mathsf{H}_1, \mathsf{G}, \mathsf{H}_2)$, the corresponding closed and open generalized Flystel constructions are CCZ-equivalent.



(a) Open variant. (b) Closed variant.

Fig. 10: Generalized Flystel construction.

Indeed in [12, Proposition 1] the proof of CCZ-equivalence does not require G to be a monomial permutation or H_1 and H_2 to be quadratic functions, hence the arguments carry over to the generalized Flystel construction. Therefore, as mentioned in Section 4.1, to study the linear properties of the Flystel construction, it is sufficient to consider the closed variant. The closed variant has the advantage that it is of low degree if H_1, G and H_2 are of low degree.

6.2 Related work

The original Flystel construction over prime-order fields was first investigated in [13], and it was conjectured that the absolute correlation of any nontrivial linear approximation is at most $(\log q)/q$. The same question was left as an open problem in [11]. The conjecture has been verified experimentally for different values of q and d. For small values of d, experiments have suggested even more precise bounds such as $2/q$ for $d = 3$ and $3.5/q$ for $d = 5$.

In Section 6.3, we prove the bound $(d-1)/q$. This proves the conjecture from [11, 13] for $d \leq \log q$, and matches the experimental results for low values of d. For applications such as Anemoi, these are the most relevant cases.

6.3 Correlation bound

As mentioned above, although the Flystel construction is derived from the Butterfly and three-round Feistel constructions, none of the results used in Section 4.3 and Section 5.3 apply here. Instead, we will use Theorem 3. To apply this result to the Flystel construction, the following lemma will be useful.

Lemma 3. *Let $d \geq 2$ be an integer indivisible by the characteristic of $\mathbf{F}_q$. The projective hypersurface in $\mathbf{P}^1$ defined by $(x_1 - x_2)^d = 0$ contains a unique singular point $[1:1]$. It is of quasi-homogeneous type with Milnor number equal to $d-1$.*

Proof. Every singular point on the hypersurface defined by $(x_1 - x_2)^d = 0$ satisfies $x_1 = x_2$. Hence, projectively, $[1 : 1]$ is the only such point. To compute the Milnor number of this isolated singularity, we restrict to the open affine set defined by $x_2 = 1$. The affine hypersurface in $\mathbf{A}^1$ defined by $h(x - 1) = 0$ with $h(x) = x^d$ has $x = 1$ as its only singular point. Up to translation, we may consider the affine hypersurface defined by $h = 0$ with singularity in the origin. Since h is homogeneous, and therefore quasi-homogeneous, the Milnor number of the singularity is equal to $d - 1$ (see Section 3.4).

Using Theorem 3, together with Lemma 3 and an analysis of linear trails, we now prove the following bound for the correlations of linear approximations of the generalized Flystel construction. In Theorem 6, we focus on the case $\deg \mathsf{G} \geq \max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\}$ because this corresponds to the Flystel construction as used in Anemoi.

Theorem 6. *Let* $\mathsf{F} = \textsc{Flystel}[\mathsf{H}_1, \mathsf{G}, \mathsf{H}_2]$ *with* H_1, H_2 *and* G *functions on* $\mathbf{F}_q$ *of degree at least two and coprime to the characteristic* p *of* $\mathbf{F}_q$. *Suppose that* $\deg \mathsf{G} \geq \max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\}$, *and that the degree of the second-highest-degree term of* G *is strictly less than* $\max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\}$. *Furthermore, suppose that if* $\deg \mathsf{H}_1 = \deg \mathsf{H}_2$, *then* H_1 *and* H_2 *have the same leading coefficient. If* $\max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\}/\deg \mathsf{G} > p/(p + (p-1)^2)$, *then for all linear approximations* (ψ, χ) *of* F *with* $\chi = (\chi_1, \chi_2) \neq (1, 1)$,

$$\left|C^{\mathsf{F}}_{\chi,\psi}\right| \leq \frac{1}{q} \begin{cases} (\deg \mathsf{G} - 1)(\deg \mathsf{H}_2 - 1) & \text{if } \chi_1 = 1\,, \\ (\deg \mathsf{G} - 1)(\deg \mathsf{H}_1 - 1) & \text{if } \chi_2 = 1\,, \\ (\deg \mathsf{H}_1 - 1)(\deg \mathsf{H}_2 - 1) & \text{if } \chi_1\chi_2 = 1\,, \\ (\deg \mathsf{G} - 1)(\max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\} - 1) & \text{otherwise}\,. \end{cases}$$

Proof. Let $\psi = (\psi_1, \psi_2)$ and $\chi = (\chi_1, \chi_2)$. If $\chi_1 = 1$, $\chi_2 = 1$ or $\chi_1\chi_2 = 1$, then there is at most one linear trail with nonzero correlation. These trails are shown in Figure 11. Hence, for the first three cases, the result follows from Weil's bound. For example, for the third case, we have

$$\left|C^{\mathsf{F}}_{\chi,\psi}\right| = \left|C^{\mathsf{H}_1}_{\chi_1,\psi_1}\right| \left|C^{\mathsf{H}_2}_{\chi_2,\psi_2}\right| \leq (\deg \mathsf{H}_1 - 1)/\sqrt{q} \times (\deg \mathsf{H}_2 - 1)/\sqrt{q}\,.$$

The first two cases are analogous. For the remaining case, we rely on Theorem 3. The correlation $C^{\mathsf{F}}_{\chi,\psi}$ is equal to $S(f)/q^2$, with $S(f)$ as in (1) and

$$f(x_1, x_2) = (v_1 + v_2)\,\mathsf{G}(x_1 - x_2) + v_1\mathsf{H}_1(x_1) + v_2\mathsf{H}_2(x_2) - u_1x_1 - u_2x_2\,,$$

assuming $\psi_i(x_i) = \omega(v_ix_i)$ and $\chi_i(x_i) = \omega(u_ix_i)$. The function f is of degree $d = \deg \mathsf{G}$ and, up to a nonzero multiple, its degree-d homogeneous component is equal to $f_d(x_1, x_2) = (v_1 + v_2)(x_1 - x_2)^d$ with $v_1 + v_2 \neq 0$. By Lemma 3, the projective hypersurface defined by $(x_1 - x_2)^d = 0$ has a unique isolated quasi-homogeneous singularity at $[1 : 1]$ with Milnor number $d - 1$.

Furthermore, the projective hypersurface defined by $f_{d'}(x_1, x_2) = 0$ with $d' = \deg f - f_d$ does not contain the point $[1 : 1]$. Indeed, if $\deg \mathsf{H}_1 > \deg \mathsf{H}_2$ or

$\deg \mathsf{H}_2 > \deg \mathsf{H}_1$ then $f_{d'}$ is given by (up to scaling)

$$f_{d'}(x_1, x_2) = v_i x_i^{\deg \mathsf{H}_i} .$$

This depends on the fact that the second-highest degree term in G is of degree strictly less than $\deg \mathsf{H}_i$. Hence, $f_{d'}(1,1) \neq 0$. Otherwise, if $\deg \mathsf{H}_1 = \deg \mathsf{H}_2$, then up to scaling (since H_1 and H_2 have the same leading coefficient)

$$f_{d'}(x_1, x_2) = v_1 x_1^{\deg \mathsf{H}_1} + v_2 x_2^{\deg \mathsf{H}_2} .$$

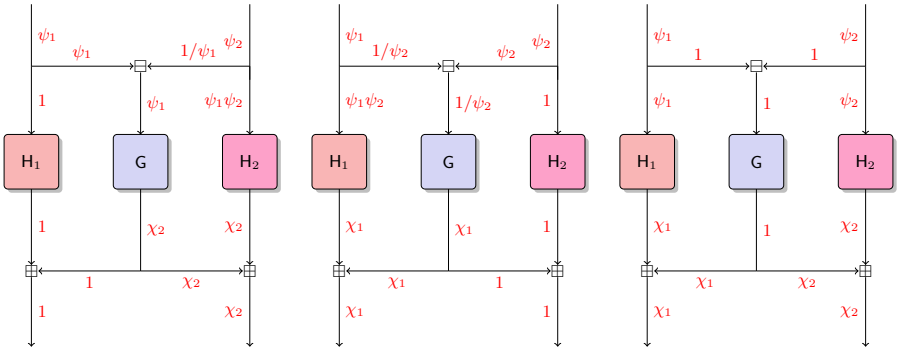
Hence, $f_{d'}(1,1) = v_1 + v_2 \neq 0$. Theorem 3 implies that $H_c^2(\mathbf{A}^2, \mathcal{L})$ is pure of weight two and of dimension

$$(d-1)^2 - (d - \max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\})\,(d-1) = (d-1)(\max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\} - 1) .$$

It follows that the exponential sum $S(f)$ satisfies the bound

$$|S(f)| \leq (\deg \mathsf{G} - 1)(\max\{\deg \mathsf{H}_1, \deg \mathsf{H}_2\} - 1)\, q ,$$

and this implies the result.



(a) $\chi_1 = 1$. (b) $\chi_2 = 1$. (c) $\chi_1\chi_2 = 1$.

Fig. 11: Linear trails for the Flystel construction.

In particular, this means that for the Flystel construction used in Anemoi, we solve Conjecture 1 in [13] and open problem 7.1 in [11]. Explicitly, we have the following corollary.

Corollary 1. *Let* $\mathsf{F} = \textsc{Flystel}[\mathsf{H}_1, \mathsf{G}, \mathsf{H}_2]$, *where* $\mathsf{G}(x) = x^d$, $\mathsf{H}_1(x) = \beta x^2 + \gamma$ *and* $\mathsf{H}_2(x) = \beta x^2 + \delta$ *with* $d \geq 2$ *indivisible by the characteristic of* $\mathbf{F}_q$, γ *and* δ *in* $\mathbf{F}_q$, *and* β *in* $\mathbf{F}_q^\times$. *For all linear approximations* (ψ, χ) *with* $\chi = (\chi_1, \chi_2) \neq (1,1)$,

$$\left|C^{\mathsf{F}}_{\chi,\psi}\right| \leq \frac{1}{q} \begin{cases} 1 & \text{if } \chi_1\chi_2 = 1 \text{ or } \chi_1 = 1 \text{ or } \chi_2 = 1 , \\ d-1 & \text{otherwise} . \end{cases}$$

6.4 Experimental results

An experimental verification of Theorem 6 is shown in Figure 12. For simplicity, only results for monomial functions H_1, G and H_2 are shown. Like in Sections 4 and 5, Figure 12a suggests that for functions H_1, G and H_1 of low degree our bound is tighter than if the degree is higher.

Note that Theorem 6 is valid regardless of whether or not G is a permutation. This is confirmed by Figure 12b, but the experimental results also suggest that the bound may be refined – at least for the case of monomial functions – if G is a permutation, which is the case for the generalized Flystel construction.

7 Conclusions

The main message of this work is that bounds on exponential sums derived from purity and vanishing theorems for ℓ-adic cohomology have direct applications to linear cryptanalysis. To demonstrate the potential of this approach, we have applied three different results (Theorem 1 due to Deligne, Theorem 2 due to Denef and Loeser, and Theorem 3 due to Rojas-León) to obtain correlation bounds for several important constructions that could not be dealt with using other methods. We expect that the same results will be useful to analyze many other constructions beyond those considered here.

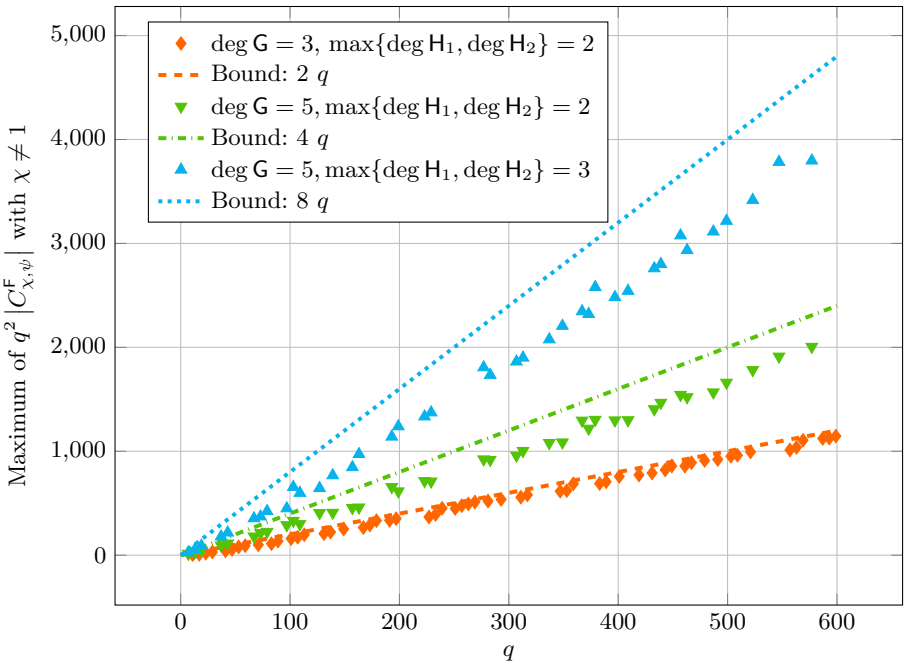
Although Theorems 1 to 3 all impose some conditions that usually do not hold for all choices of the masks, in our examples these edge cases generally coincided with linear approximations for which there is at most one linear trail with nonzero correlation. It would be of interest to understand this, and other aspects of linear cryptanalysis, from the point of view of ℓ-adic cohomology. In addition, detailed calculations of the cohomology spaces might allow refining our bounds.

Acknowledgements

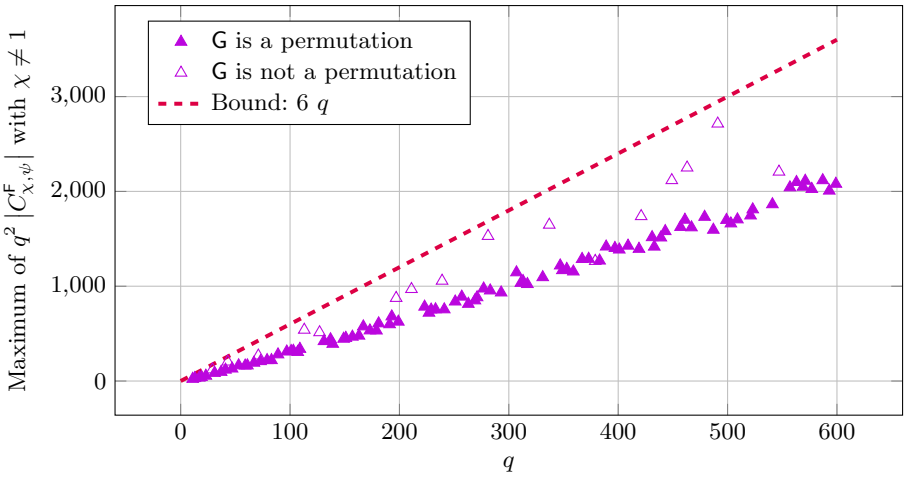
Tim Beyne is supported by a junior postdoctoral fellowship from the Research Foundation – Flanders (FWO) with reference number 1274724N. He would like to thank Wouter Castryck for helpful discussions about exponential sums, and Raf Cluckers for pointing out the result of Denef and Loeser [22] in June of 2020. Clémence Bouvier is supported by the European Research Council (ERC, grant number 101097056 "SymTrust").

References

1. Alan Adolphson and Steven Sperber. Exponential sums and Newton polyhedra: cohomology and estimates. *Annals of Mathematics*, 130(2):367–406, 1989.



(a) Low-degree permutations G, H_1 and H_2.



(b) $\deg \mathsf{G} = 7$ and $\deg \mathsf{H}_1 = \deg \mathsf{H}_2 = 2$.

Fig. 12: Experimental verification of correlation bounds from Theorem 6 for the generalized Flystel construction $\mathsf{F} = \textsc{Flystel}[\mathsf{H}_1, \mathsf{G}, \mathsf{H}_2]$ over a field $\mathbf{F}_q$ of prime order and with H_1, G and H_2 monomials.

2. Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *Advances in Cryptology – ASIACRYPT 2016, Part I*, volume 10031 of *Lecture Notes in Computer Science*, pages 191–219, Hanoi, Vietnam, December 4–8, 2016. Springer, Berlin, Heidelberg, Germany.
3. Abdelrahaman Aly, Tomer Ashur, Eli Ben-Sasson, Siemen Dhooghe, and Alan Szepieniec. Design of symmetric-key primitives for advanced cryptographic protocols. *IACR Transactions on Symmetric Cryptology*, 2020(3):1–45, 2020.
4. Thomas Baignères, Jacques Stern, and Serge Vaudenay. Linear cryptanalysis of non binary ciphers. In Carlisle M. Adams, Ali Miri, and Michael J. Wiener, editors, *SAC 2007: 14th Annual International Workshop on Selected Areas in Cryptography*, volume 4876 of *Lecture Notes in Computer Science*, pages 184–211, Ottawa, Canada, August 16–17, 2007. Springer, Berlin, Heidelberg, Germany.
5. Tim Beyne. Block cipher invariants as eigenvectors of correlation matrices. In Thomas Peyrin and Steven Galbraith, editors, *Advances in Cryptology – ASIACRYPT 2018, Part I*, volume 11272 of *Lecture Notes in Computer Science*, pages 3–31, Brisbane, Queensland, Australia, December 2–6, 2018. Springer, Cham, Switzerland.
6. Tim Beyne. *Linear cryptanalysis in the weak key model*. Master's thesis, KU Leuven, 2019.
7. Tim Beyne. A geometric approach to linear cryptanalysis. In Mehdi Tibouchi and Huaxiong Wang, editors, *Advances in Cryptology – ASIACRYPT 2021, Part I*, volume 13090 of *Lecture Notes in Computer Science*, pages 36–66, Singapore, December 6–10, 2021. Springer, Cham, Switzerland.
8. Tim Beyne. *A geometric approach to symmetric-key cryptanalysis*. PhD thesis, KU Leuven, June 2023.
9. Tim Beyne and Clémence Bouvier. Linear approximations of the Flystel construction. Cryptology ePrint Archive, Paper 2024/1465, 2024.
10. Tim Beyne, Anne Canteaut, Gregor Leander, María Naya-Plasencia, Léo Perrin, and Friedrich Wiemer. On the security of the rescue hash function. Cryptology ePrint Archive, Report 2020/820, 2020.
11. Clémence Bouvier. *Cryptanalysis and design of symmetric primitives defined over large finite fields*. PhD thesis, Sorbonne Université, November 2023.
12. Clémence Bouvier, Pierre Briaud, Pyrros Chaidos, Léo Perrin, Robin Salen, Vesselin Velichkov, and Danny Willems. New design techniques for efficient arithmetization-oriented hash functions: Anemoi permutations and Jive compression mode. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023, Part III*, volume 14083 of *Lecture Notes in Computer Science*, pages 507–539, Santa Barbara, CA, USA, August 20–24, 2023. Springer, Cham, Switzerland.
13. Clémence Bouvier, Pierre Briaud, Pyrros Chaidos, Léo Perrin, Robin Salen, Vesselin Velichkov, and Danny Willems. New design techniques for efficient arithmetization-oriented hash functions: Anemoi permutations and Jive compression mode. Cryptology ePrint Archive, Paper 2022/840, 2022.
14. Anne Canteaut, Sébastien Duval, and Léo Perrin. A generalisation of Dillon's APN permutation with the best known differential and nonlinear properties for all fields of size 2^{4k+2}. *IEEE Trans. Inf. Theory*, Vol. 63(11):7575–7591, Nov 2017.
15. Anne Canteaut, Léo Perrin, and Shizhu Tian. If a generalised butterfly is APN then it operates on 6 bits. *Cryptography and Communications*, Vol. 11(6):1147–1164, 2019.

16. Anne Canteaut and Joëlle Roué. On the behaviors of affine equivalent sboxes regarding differential and linear attacks. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology – EUROCRYPT 2015, Part I*, volume 9056 of *Lecture Notes in Computer Science*, pages 45–74, Sofia, Bulgaria, April 26–30, 2015. Springer, Berlin, Heidelberg, Germany.
17. Claude Carlet, Pascale Charpin, and Victor Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, Vol. 15(2):125–156, 1998.
18. Joan Daemen and Vincent Rijmen. The wide trail design strategy. In Bahram Honary, editor, *8th IMA International Conference on Cryptography and Coding*, volume 2260 of *Lecture Notes in Computer Science*, pages 222–238, Cirencester, UK, December 17–19, 2001. Springer, Berlin, Heidelberg, Germany.
19. Pierre Deligne. La conjecture de Weil: I. *Publication Mathématiques de l'IHÉS*, 43:273–307, 1974.
20. Pierre Deligne. Séminaire de géométrie algébrique du Bois Marie – cohomologie étale (SGA $4\frac{1}{2}$). *Lecture Notes in Mathematics*, 1977.
21. Pierre Deligne. La conjecture de Weil: II. *Publications Mathématiques de l'IHÉS*, 52:137–252, 1980.
22. Jan Denef and François Loeser. Weights of exponential sums, intersection cohomology, and newton polyhedra. *Inventiones mathematicae*, 106(1):275–294, 1991.
23. Shihui Fu, Xiutao Feng, and Baofeng Wu. Differentially 4-uniform permutations with the best known nonlinearity from butterflies. *IACR Transactions on Symmetric Cryptology*, 2017(2):228–249, 2017.
24. Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A new hash function for zero-knowledge proof systems. In Michael Bailey and Rachel Greenstadt, editors, *USENIX Security 2021: 30th USENIX Security Symposium*, pages 519–535. USENIX Association, August 11–13, 2021.
25. Alexander Grothendieck. Séminaire de géométrie algébrique du Bois Marie – cohomologie ℓ-adique et fonctions L. *Lecture Notes in Mathematics*, 1977.
26. Seokhie Hong, Sangjin Lee, Jongin Lim, Jaechul Sung, Dong Hyeon Cheon, and Inho Cho. Provable security against differential and linear cryptanalysis for the SPN structure. In Bruce Schneier, editor, *Fast Software Encryption – FSE 2000*, volume 1978 of *Lecture Notes in Computer Science*, pages 273–283, New York, NY, USA, April 10–12, 2001. Springer, Berlin, Heidelberg, Germany.
27. Liam Keliher and Jiayuan Sui. Exact maximum expected differential and linear probability for two-round advanced encryption standard. *IET Information Security*, 1(2):53–57, 2007.
28. Yongqiang Li, Shizhu Tian, Yuyin Yu, and Mingsheng Wang. On the generalization of butterfly structure. *IACR Transactions on Symmetric Cryptology*, 2018(1):160–179, 2018.
29. Kaisa Nyberg. Differentially uniform mappings for cryptography. In Tor Helleseth, editor, *Advances in Cryptology – EUROCRYPT'93*, volume 765 of *Lecture Notes in Computer Science*, pages 55–64, Lofthus, Norway, May 23–27, 1994. Springer, Berlin, Heidelberg, Germany.
30. Kaisa Nyberg and Lars R. Knudsen. Provable security against a differential attack. *Journal of Cryptology*, 8(1):27–37, December 1995.
31. Léo Perrin, Aleksei Udovenko, and Alex Biryukov. Cryptanalysis of a theorem: Decomposing the only known solution to the big APN problem. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016,*

Part II, volume 9815 of *Lecture Notes in Computer Science*, pages 93–122, Santa Barbara, CA, USA, August 14–18, 2016. Springer, Berlin, Heidelberg, Germany.
32. Antonio Rojas-León. Purity of exponential sums on $\mathbb{A}^n$. *Compositio Mathematica*, 142(2):295–306, 2006.
33. André Weil. On some exponential sums. *Proceedings of the National Academy of Sciences*, 34(5):204–207, 1948.