Distinguishing attacks on Feistel ciphers based on linear and differential attacks

Denis Fomin

HSE University

- $q \in \mathbb{N}$, $Q \in \mathbb{N}$
- $K \in \mathbb{Z}_q^k$ — key
- $T \in \mathbb{Z}_q^t$ — tweak
- $I \in \mathbb{N}$ rounds
- Encryption function of GTFN: $E_{K,T} \colon \mathbb{Z}_q^Q \to \mathbb{Z}_q^Q$
- round function of GTFN is a key, tweak and round-dependent mapping:

$$F \colon \mathbb{Z}_q^k \times \mathbb{Z}_q^t \times \mathbb{Z}_I \times \mathbb{Z}_q^R \to \mathbb{Z}_q^L,$$

  $L, R \in \mathbb{N}$, $L + R = Q$
- Let for some $h \in \mathbb{N}$, $h < Q$, $L = \lceil Q/h \rceil$, then $R = Q - \lceil Q/h \rceil$

An internal state of i round of GTFN: $S^{(i)} = S_0^{(i)} \| S_1^{(i)}$, where $S_0^{(i)} \in \mathbb{Z}_q^L$, $S_1^{(i)} \in \mathbb{Z}_q^R$

$S^{(0)}$ is a plaintext, $S^{(I)}$ is a ciphertext.

The round function is evaluated as follows:

$$S^{(i)} = S_1^{(i-1)} \left\| \left(S_0^{(i-1)} + F\left(K, T, i, R^{(i-1)}\right) \right) \right. ,$$

where "$+$" is either

- an operation of group $\mathbb{Z}_{q^L}$, that we will denote as $\boxplus$;
- or operator of vector space $\mathbb{Z}_q^L$, that we will denote as $\oplus$.

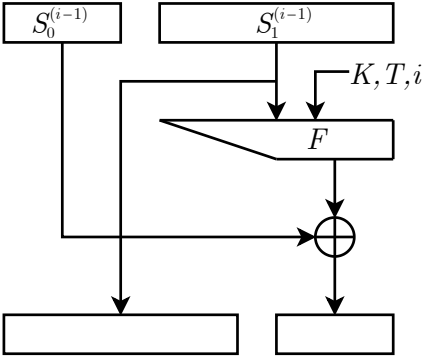


Figure: GTFN$_{\oplus}$, $q = 2$ and "+" operator in round function is $\oplus$

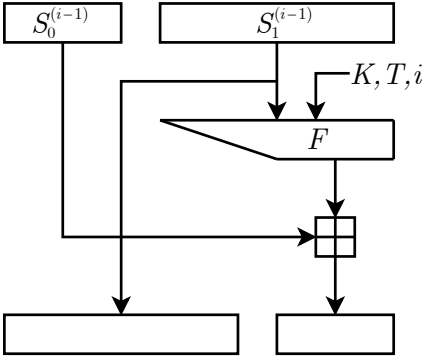


Figure: GTFN$_{\boxplus}$, $q = L$ and "+" operator in round function is $\boxplus$

F for fixed K, T is realized a random function $\mathbb{Z}_q^R \to \mathbb{Z}_q^L$ according to the discrete distribution D

- Uniform discrete distribution $\mathsf{U}(\mathbb{Z}_q^L)$
- Distribution $\mathsf{M}(q, L)$ of the following random variable:

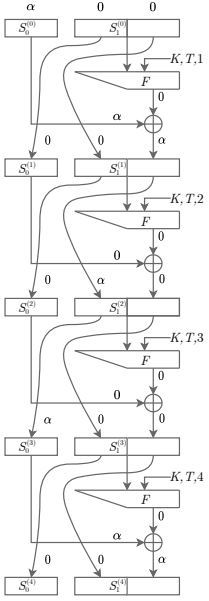
$$\zeta = \xi \pmod{(q^L)}, \text{ where } \xi \sim \mathsf{U}\left(\mathbb{Z}_2^{\lceil L \cdot \log_2(q) \rceil}\right),$$

- Let $R = (h-1) \cdot L$, internal state is a concatenation of h elements of $\mathbb{Z}_q$
- With probability 1, the following difference relationship for h rounds holds:

$$(\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{1} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha) \xrightarrow{1}$$

$$\xrightarrow{1} (\underbrace{0 \| \ldots \| 0}_{h-2} \| \alpha \| \star) \xrightarrow{1} \cdots \xrightarrow{1} (\alpha \| \underbrace{\star \| \ldots \| \star}_{h-1}),$$

- There is an efficient algorithm to distinguish h rounds of the GTFN algorithm from a random substitution
- Difficulty and amount of material are about $O\left(q^L\right)$

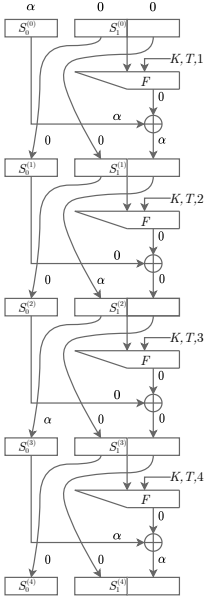


- With probability $q^{-(h-1)L}$, the following difference relationship for $h+1$ rounds holds:

$$(\alpha\|\underbrace{0\|\ldots\|0}_{h-1}) \xrightarrow{1} (\underbrace{0\|\ldots\|0}_{h-1}\|\alpha) \xrightarrow{q^{-L}}$$
$$\xrightarrow{q^{-L}} (\underbrace{0\|\ldots\|0}_{h-2}\|\alpha\|0) \xrightarrow{q^{-L}} \cdots$$
$$\cdots \xrightarrow{q^{-L}} (\alpha\|\underbrace{0\|\ldots\|0}_{h-1}) \xrightarrow{1} (\underbrace{0\|\ldots\|0}_{h-1}\|\alpha),$$

- the following difference relation holds for 4 rounds of the GTFN$_\oplus$ when $h=3$:

$$(\alpha\|0\|0) \xrightarrow{1} (0\|0\|\alpha) \xrightarrow{2^{-b_L}} (0\|\alpha\|0) \xrightarrow{2^{-b_L}} (\alpha\|0\|0) \xrightarrow{1} (0\|0\|\alpha)$$



The idea of this attack is based on the statistical problem of distinguishing between two hypotheses:

- random sample observation from Bernoulli distribution with "success" probability equals to $q^{-(h-1)L}$;
- random sample observation from Bernoulli distribution with "success" probability equals to $q^{-(h)L}$.

The difficulty of differental attack based on this test is about $O\left(q^{hL}\right)$.

Let there are $M_j \leq M/2$ pairs of plain texts encrypted using $t_j, j = 1, \ldots, T$, tweaks that have a difference $(\alpha \| 0 \| \ldots \| 0)$ for some fixed α.

Then the statistics equivalent to the likelihood ratio statistics:

$$S_j(M_j) = \sum_{i=1}^{M_j} z_{i,j},$$

where $z_{i,j}$ is–an indicator that equals 1 if and only if i-th pair of plaintexts that has an input difference $(\alpha \| 0 \| \ldots \| 0)$ is have the same difference between ciphertexts.

Simply increasing the material using different tweaks, values of α is generally speaking not correct.

However, we can consider $S_j(M_j)$ at one tweak with fixed α as a random variable that has a binomial distribution with parameters Bin (M_j, q_i).

In that case we can consider N such observations (N tweaks) and the statistic equivalent to the likelihood ratio statistic equals to:

$$K(N, M) = \sum_{j=1}^{N} S_j(M_j).$$

Considering different values of α also leads to an increase in the efficiency of the attack.

Note that if the adversary has the ability to encrypt arbitrary texts, then he can choose texts in such a way as to obtain up to M different values of α for which there will be about $M/2$ pairs of plaintexts for the chosen values of α.

Indeed, if the cryptoanalyst can encrypt $M = q^e$ plaintexts $(x_1, x_2, \ldots, x_h)$, where $x_1 \leq M$, the difference relations described above are fulfilled for any value of $\alpha \in \mathbb{Z}_q^e \backslash \{0\}$, $\alpha \leq M$.

This potentially could increase the amount of material (like in a multidimensional linear cryptanalysis)

Let's consider $\mathrm{GTFN}_{\boxplus}$ with round function that are chosen according M distribution.

Then the probability of the difference relation $F(x) + F(x+a) = b$ is equal to:

$$\mathsf{P}\{F(x) + F(x+a) = b\} = \frac{4W_{0,0}}{(N')^2} + \frac{2W_{0,1}}{(N')^2} + \frac{W_{1,1}}{(N')^2} = p_1(b),$$

where

$$W_{0,0} = \max\{N' - N - b, 0, N' - 2N + b, 2N' - 3N\}$$

$$W_{0,1} = \min\{2b, 2(N' - N), 2(N - b), 4N - 2N'\}$$

$$W_{1,1} = N - W_{0,0} - W_{0,1},$$

$N = 2^{\lceil L \log_2(q) \rceil}$, $N' = q^L$

The graph of this probability for the case $q = 10$, $h = 3$, $L = 3$ is shown in figure:

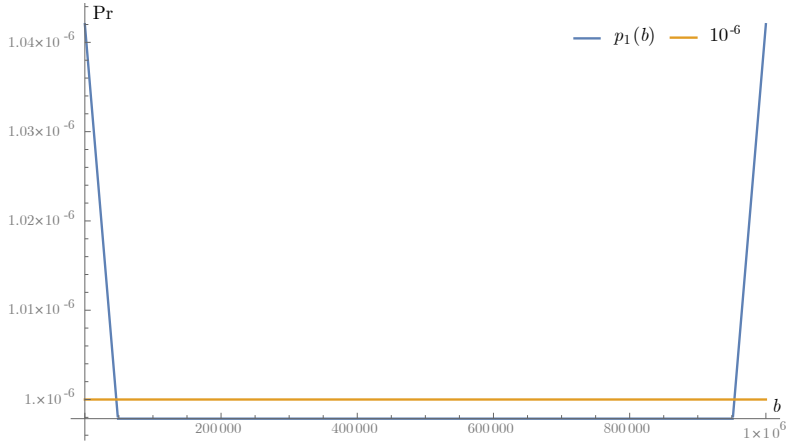


Figure: Graph of probability $p_1(b)$ in case $q = 10$, $h = 3$, $L = 3$

This property helps to reduce the amount of material needed to apply the difference attack compared to the equal-probability case ($\mathsf{U}(\mathbb{Z}_{q^L})$).

It also allows to apply a difference attack for more rounds. Without losing generality, let us consider the special case of $\mathrm{GTFN}_{\boxplus}$ with $q = 10$, $h = 3$. Let's find the probability of the following $2h + 1$-rounds differential relation:

$$(\alpha \| 0 \| 0) \xrightarrow[2h+1 \text{ rounds}]{} (0 \| 0 \| \star),$$

where $\alpha \in \mathbb{Z}_{10^\iota}$ — a fixed value, $\star$ — any value of the set $\mathbb{Z}_{10^\iota}$. The differential above can be descripted as follows:

$$(\alpha \| 0 \| 0) \to (0 \| 0 \| \alpha) \to (0 \| \alpha \| \gamma) \to (\alpha \| \gamma \| \delta) \to (\gamma \| \delta \| \beta) \to \\ \to (\delta \| \beta \| 0) \to (\beta \| 0 \| 0) \to (0 \| 0 \| \beta),$$

where $\alpha \in \mathbb{Z}_{10^\iota}$ — a fixed value, β, γ, δ — some values of the set $\mathbb{Z}_{10^\iota}$.

This probability is different from the case of an equal probability distribution:

	Distribution	
L	M	U
3	$10^{-6} + 1.61 \cdot 10^{-11}$	10^{-6}
4	$10^{-8} + 4.01 \cdot 10^{-11}$	10^{-8}
5	$10^{-10} + 7.24 \cdot 10^{-13}$	10^{-10}
6	$10^{-12} + 1.17 \cdot 10^{-16}$	10^{-12}

Let $\alpha \in \mathbb{Z}_q^L$, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_L)$, $w = L - (Q - (h-1)L) = hL - Q$,
$\alpha_1 = \ldots = \alpha_w = 0$.

Then the following difference relationship for h rounds holds:

$$(\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow[h \text{ rounds}]{1} (\alpha' \| \star)\,,$$

where $\alpha' = (\alpha_{w+1}, \alpha_{w+2}, \ldots, \alpha_L)$, $\star$ — some element of $\mathbb{Z}_q^{Q-L+w}$.

As we can see in case $L = Q/h$ the value $w = 0$ and all statements shown earlier are correct.

The scalar product of two functions f_1, f_2 with values in $\mathbb{C}^\times$ is defined as follows:

$$\langle f_1, f_2 \rangle = \sum_{x \in X} f_1(x)\overline{f_2(x)}.$$

The Fourier coefficients of function $f \in \mathbb{C}^X$ is a function $C_\alpha^f \in \mathbb{C}^{\widehat{X}}$:

$$C_\alpha^f = \langle f, \overline{\chi_\alpha} \rangle = \sum_{x \in X} f(x)\overline{\chi_\alpha(x)}, \ \alpha \in X.$$

These coefficients are defined the Fourier transform of f:

$$f = \frac{1}{|X|} \sum_{\alpha \in X} C_\alpha^f \chi_\alpha.$$

Let D is a distribution of values of finite Abelian group X:

$$\Pr_{\mathsf{D}}\{x\} = p(x).$$

The function $p(x)$ can be represented using the Fourier transform as function of $\mathbb{C}^X$:

$$p(x) = \frac{1}{|X|}\sum_{\alpha\in X} C_\alpha^P \chi_\alpha(x).$$

Then C_α^p is the expected number of $\overline{\chi_\alpha}$:

$$C_\alpha^p = \sum_{x\in X} p(x)\overline{\chi_\alpha(x)} = \mathbf{E}\overline{\chi_\alpha}.$$

Statement

Let $f \in Y^X$ be a function with arguments in finite Abelian group X and with values in finite Abelian group Y. Then

$$\mathbf{E}\psi_\beta(f(x)) = \frac{1}{|X|}\sum_{\alpha\in X} C^f_{\beta,\alpha}\cdot \mathbf{E}\chi_\alpha.$$

Statement

Under the conditions of the previous statement:

$$\Pr\{f(x)=b\} = \frac{1}{|Y|}\sum_{\beta\in Y}\mathbf{E}\psi_\beta(f)\overline{\psi_\beta(b)} = \frac{1}{|Y|}\sum_{\beta\in Y}\mathbf{E}\overline{\psi_\beta(f)}\psi_\beta(b).$$

Let's consider the function $F(x)$ of the form $F(x) = (f(x), -x)$. In that case $F(x) \in (Y \dot{+} X)^X$. If X and Y are finite Abelian groups then $Z = Y \dot{+} X$ also a finite Abelian group and

$$Z = Y \dot{+} X = H_1 \dot{+} \ldots \dot{+} H_t \dot{+} G_1 \dot{+} \ldots \dot{+} G_k.$$

Let ϕ_γ, $\gamma \in Z$, $\gamma = \beta \| \alpha$ — are characters of group Z. Then for function F:

$$\Pr\{F(x) = b\} = \frac{1}{|Z|} \sum_{\gamma \in Z} \mathbf{E} \phi_\gamma(F) \overline{\phi_\gamma(b)} = \sum_{\gamma \in Z} \mathbf{E} \left(\psi_\beta(f(x)) \overline{\chi_\alpha(x)} \right) \overline{\psi_\beta(f(x))} \chi_\alpha(x).$$

We can see that $\mathbf{E}\left(\psi_\beta(f(x))\overline{\chi_\alpha(x)}\right)$ is a Fourier coefficient of function F when $\mathsf{D} = \mathsf{U}$. In this work we call correlation coefficient of the linear approximation $(\chi_\alpha, \phi_\beta)$ of function f the value

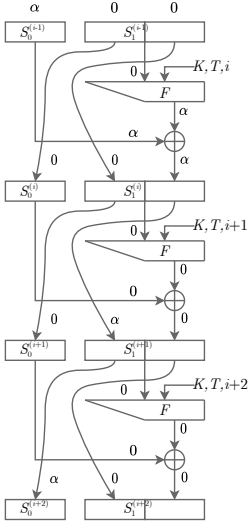
$$\mathbf{L}^F_{\beta,\alpha} = \mathbf{E}\left(\psi_\beta(f(x))\overline{\chi_\alpha(x)}\right).$$

If Y and X are the same groups the equation above can be rewritten as follows:

$$\mathbf{L}^F_{\beta,\alpha} = \mathbf{E}\left(\chi_\beta(f(x))\overline{\chi_\alpha(x)}\right).$$

- $R = (h-1) \cdot L$
- Consider the following linear relation on three rounds of the GTFN algorithm:

$$(\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{c_1} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha) \xrightarrow{1} \xrightarrow{1} (\underbrace{0 \| \ldots \| 0}_{h-2} \| \alpha \| 0) \xrightarrow{1} \cdots \xrightarrow{1} (\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}),$$



Let's describe this relationship in more detail. The correlation coefficient c_1 in the first round

$$(\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{c_1} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha)$$

equals to:

$$\mathbf{E}\left(\chi_\alpha\left(K, T, 1, S_1^{(0)}\right)\overline{\chi_0\left(S_1^{(0)}\right)}\right) = \mathbf{E}\left(\chi_\alpha\left(K, T, 1, S_1^{(0)}\right)\right),$$

where $F\left(K, T, 1, S_1^{(0)}\right)$ — is F-function of the first round. In case of $\mathrm{GTFN}_\oplus$ algorithm this coefficient equals to:

$$2 \cdot \mathsf{P}\left\{\left\langle 0, S_1^{(0)}\right\rangle = \left\langle \beta, F\left(b, 1, T, S_1^{(0)}\right)\right\rangle\right\} - 1 = c_1.$$

Similarly we can consider the others correlation coefficients for the following relations:

$$(\underbrace{0\| \ldots \|0}_{h-1}\|\alpha) \xrightarrow{1} (\underbrace{0\| \ldots \|0}_{h-2}\|\alpha\|0), \ldots, (0\|\alpha\|\underbrace{0\| \ldots \|0}_{h-2}) \xrightarrow{1} (\alpha\|\underbrace{0\| \ldots \|0}_{h-1}).$$

It's easy to show, that

$$\mathbf{E}\left(\chi_0\left(F\left(K,T,i+1,S_1^{(i)}\right)\right)\overline{\chi_0\left(S_1^{(i)}\right)}\right) = 1.$$

In case of GTFN$_\oplus$ algorithm this coefficient equals to:

$$2 \cdot \mathsf{P}\left\{\left\langle 0, S_1^{(i)}\right\rangle = \left\langle 0, F\left(K,T,i+1,S_1^{(i)}\right)\right\rangle\right\} - 1 = 1.$$

As in[1] , we can use the following approach. Let the set of plaintexts have the following form: $P = \{(x_1, x_2, \ldots, x_h)\}$, where $x_2, x_3, \ldots, x_h$ are fixed by some constants from the set $\mathbb{Z}_q^L$. Then for the first three rounds of the algorithm GTFN the absolute value of correlation coefficient is equal to 1. Indeed, on the first round, the values $F\left(K, T, 1, S_1^{(0)}\right)$ will be the same and equal to some $y \in \mathbb{Z}_q^L$, from which it follows that

$$\left|\mathbf{E}\left(\chi_\alpha\left(F\left(K, T, 1, S_1^{(0)}\right)\right)\overline{\chi_0\left(S_1^{(0)}\right)}\right)\right| = \left|\mathbf{E}\left(\chi_\alpha\left(y\right)\right)\right| = \left|\left(\chi_\alpha\left(y\right)\right)\right| = 1.$$

In case of $\mathrm{GTFN}_\oplus$ algorithm this coefficient equals to:

$$2 \cdot \mathsf{P}\left\{\left\langle 0, S_1^{(0)}\right\rangle = \langle \alpha, y\rangle\right\} - 1 = \pm 1.$$

[1]Tim Beyne., "Linear Cryptanalysis of FF3-1 and FEA. Cryptology ePrint Archive, Report 2021/815, 2021. https://ia.cr/2021/815.".

- For a random vectorial Boolean function $S\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^m$ as n increases, the value $\mathbf{L}^S_{\beta,\alpha}$ will have a normal distribution with parameters $\mathcal{N}(0, 2^{-n})$.
- If X and Y are finite Abelian groups and S is a random function $S \in Y^X$ the the distribution of $\sqrt{|X|}\mathbf{L}^S_{\beta,\alpha}$ converges to the standard complex normal distribution $\mathcal{CN}(0,1)$.
- If $\mathsf{D} \neq \mathsf{U}$ then the distribution of the value

$$\mathbf{L}^S_{\alpha,0} = \mathbf{E}\left(\chi_\alpha\left(K, T, 1, S_1^{(0)}\right)\right)$$

should be estimated.

Consider the following linear relation on $h \cdot r + h$ rounds of the GTFN algorithm, similar to those considered in[2]:

$$(\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{1} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha) \xrightarrow{1} (\underbrace{0 \| \ldots \| 0}_{h-2} \| \alpha \| 0) \xrightarrow{1} \cdots$$

$$\cdots \xrightarrow{1} (\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{c_{h+1}} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha) \xrightarrow{1} \ldots \xrightarrow{1} (\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}) \xrightarrow{c_{2h+1}}$$

$$\xrightarrow{c_{2h+1}} \ldots \xrightarrow{c_{r-h+1}} (\underbrace{0 \| \ldots \| 0}_{h-1} \| \alpha) \xrightarrow{1} \ldots \xrightarrow{1} (\alpha \| \underbrace{0 \| \ldots \| 0}_{h-1}).$$

Using the pilling-up lemma the correlation coefficient $\mathcal{C}_1 = \mathbf{L}^{\text{GTFN}}_{(\alpha\|0\|\ldots\|0),(\alpha\|0\|\ldots\|0)}$ can be estimated as follows:

$$\mathcal{C}_1 = \prod_{i=1}^{r/h-h} c_{1+h\cdot i},$$

where $c_{1+h\cdot i} = \mathbf{L}^{F}_{\alpha,0}$.

- A random permutation will have a correlation coefficient equals to the value $\mathcal{C}_0$, which is a realization of a random variable with the uniform distribution.
- The distribution of $\mathcal{C}_0$ is well known and we also suppose that the distribution of $\mathcal{C}_1$ is also known to a cryptanalyst.

The statistics based on logarithm of likelihood function is asymptotically equivalent to:

$$\sum_{\alpha',\beta' \in X\setminus 0} \mathbf{L}^S_{\beta',\alpha'} \sum_{i=1}^{M} \overline{\chi_{\beta'}(y_i)} \chi_{\alpha'}(x_i).$$

With $M \to \infty$ the sum $\sum_{i=1}^{M} \overline{\chi_{\beta'}(y_i)} \chi_{\alpha'}(x_i)$ converges to $\overline{\mathbf{L}^S_{\beta',\alpha'}}$, then

$$\sum_{\alpha',\beta' \in X\setminus 0} \mathbf{L}^S_{\beta',\alpha'} \sum_{i=1}^{M} \overline{\chi_{\beta'}(y_i)} \chi_{\alpha'}(x_i) \to M \sum_{\alpha',\beta' \in X\setminus 0} \left|\mathbf{L}^S_{\beta',\alpha'}\right|^2.$$

As we consider plaintexts of the form $(x\|a_1\|a_2\|\ldots\|a_{h-1})$, where $a_0, a_1, \ldots, a_{h-1}$ — some fixed elements of $\mathbb{Z}_q^L$ and $\alpha' = \beta'$ of the form $(\alpha\|0\|\ldots\|0)$ then the equation above is equal to:

$$M \sum_{\alpha \in \mathbb{Z}_q^L \setminus 0} \left| \mathbf{L}^S_{(\alpha\|0\|\ldots\|0),(\alpha\|0\|\ldots\|0)} \right|^2.$$

Let $\mathbf{D}\mathcal{C}_0$ is the variance of correlation coefficient of a random function and $\mathbf{D}\mathcal{C}_1$ is the variance of a correlation coefficient

$$\mathbf{D}\mathcal{C}_1 = \mathbf{L}^{\text{GTFN}}_{(\alpha\|0\|\ldots\|0),(\alpha\|0\|\ldots\|0)} \approx \left(\mathbf{DL}^F_{\alpha,0}\right)^{r/h-h}.$$

Then for a successful attack the ratio between M, N (tweak and other plaitexts quantity) and $|X| = q^L$ should be:

$$M \cdot N \cdot |X| \approx O\left(\left(\mathbf{D}\mathcal{C}_1 - \mathbf{D}\mathcal{C}_0\right)^{-1}\right).$$

[2]Tim Beyne., "Linear Cryptanalysis of FF3-1 and FEA. Cryptology ePrint Archive, Report 2021/815, 2021. https://ia.cr/2021/815.".