Security of Permutation-based Compression Function $lp231$

Jooyoung Lee\textsuperscript{1⋆} and Daesung Kwon\textsuperscript{2}

\textsuperscript{1} Sejong University, Seoul, Korea, jlee05@sejong.ac.kr
\textsuperscript{2} The Attached Institute of Electronics and Telecommunications Research Institute, Daejeon, Korea, ds_kwon@ensec.re.kr

Abstract. In this paper, we study security of a certain class of permutation-based compression functions. Denoted $lp231$ in [10], they are $2n$-to-$n$-bit compression functions using three calls to a single $n$-bit random permutation. We prove that $lp231$ is asymptotically preimage resistant up to $2^{2n^3/n}$ query complexity and collision resistant up to $2^{2n^2/n^{1+\epsilon}}$ query complexity for any $\epsilon > 0$. Based on a single permutation, $lp231$ provides both efficiency and almost optimal collision security.

1 Introduction

A cryptographic hash function takes a message of arbitrary length, and returns a bit string of fixed length. The most common way of hashing variable length messages is to iterate a fixed-size compression function according to the Merkle-Damgård paradigm. The underlying compression function can either be constructed from scratch, or be built upon off-the-shelf cryptographic primitives such as blockciphers. For example, the Whirlpool hash function, adopted as ISO/IEC 10118-3 standard, is based on the Miyaguchi-Preneel construction using a modified version of AES [1]. Compression functions based on blockciphers have been widely studied [4–8, 14, 15]. Recently, researchers has begun to pay attention to building compression functions from fixed key blockciphers, where just a small number of constants are used as keys [2, 3, 9, 10, 12, 13]. Since each key of a blockcipher defines an independent random permutation in the ideal cipher model, such compression functions are often called permutation-based. Permutation-based compression functions have an obvious advantage over conventional blockcipher-based ones, since fixing the keys allows to save computational overload for key scheduling.

In earlier work, Black, Cochran and Shrimpton showed that any “highly-efficient” compression function using exactly one permutation call for each message block allows a query-efficient collision-finding attack [3]. Rogaway and Steinberger extended this result to a wide class of compression functions that map $mn$ bits to $rn$ bits using $k$ calls to $n$-bit permutations [11]. Such compression functions, denoted $m \xrightarrow{k} r$, allow collision-finding attacks with $2^{n(1-(m-r)/k)}$ query complexity, and preimage finding attacks with $2^{n(1-(m-r)/k)}$ query complexity.

In [10], the authors focused on the security of a special class of permutation-based compression functions, where the input to each permutation is given by a linear combination of the inputs to the compression function and the outputs of the previously called permutations. Such compression functions are called linearly-dependent permutation-based, and denoted by $LPmkr$ if the compression function is based on independent random permutations, and by

\* This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2013R1A1A2007488).
lpmkr if the compression function is based on a single random permutation. Taking into account the attacks presented in [11], they investigated the security of LP231, LP241, LP352, LP362 and their “lp variants”. From a practical point of view, it is obvious that lp compression functions are more efficient compared to LP ones since an lp compression function uses its basing blockcipher with only one fixed key. However, [10] gives a concrete analysis only for LP231. The analysis of the other compression functions rest on computer-aided approximation. Especially, the authors claim that analyzing lp231 by hand would require about 30 times as much paper as LP231.

In this paper, we give a concrete security analysis of lp231 in terms of preimage resistance and collision resistance. Specifically, we prove preimage resistance up to \((2^{2n}/n^3)\) query complexity and collision resistance up to \((2^{n}/n^{1+\epsilon})\) query complexity for any \(\epsilon > 0\). Our analysis is not only simpler than the authors of [10] estimated, but also elegant based on a recursive approach.

2 Preliminaries

General Notations For a positive integer \(n\), let \([1, n] = \{1, 2, \ldots, n\}\) and let \(\Pi_n\) be the set of permutations on \(\{0, 1\}^n\). We let \(\mathbb{F}_{2^n}\) denote a finite field of order \(2^n\). Throughout our work, we will identify \(\mathbb{F}_{2^n}\) and \(\{0, 1\}^n\), assuming a fixed mapping between the two sets.

For positive integers \(s\) and \(t\), let \(\mathcal{M}_{\mathbb{F}_{2^n}}^{s \times t}\) denote the set of all \(s \times t\) matrices over \(\mathbb{F}_{2^n}\). For \(A, B \in \mathcal{M}_{\mathbb{F}_{2^n}}^{2 \times 1}\), \([A, B]\) is the \(2 \times 2\) matrix obtained by the concatenation of \(A\) and \(B\). The concatenation is similarly denoted for more than two matrices. For a \(2 \times 1\) matrix

\[
A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}
\]

let

\[
A^* = \begin{bmatrix} a_2 \\ a_1 \end{bmatrix}.
\]

Note that \(A^*A = 0\). For \(A, B \in \mathcal{M}_{\mathbb{F}_{2^n}}^{2 \times 1}\), \(A^*B\) is the determinant of \([A, B]\).

We write \(u \xleftarrow{s} U\) to denote uniform random sampling from the set \(U\) and assignment to \(u\). For a multiset \(U\), \(\text{mult}(U, u)\) is the multiplicity of \(u\) in \(U\), and \(\text{mult}(U) = \max_{u \in U} \text{mult}(U, u)\).

Linearly-dependent Permutation-based Compression Functions For positive integers \(m, k\) and \(r\) with \(m > r\), let \(\mathcal{M}_{\mathbb{F}_{2^n}}(m, k, r)\) be a set of \((k + r) \times (m + k)\) matrices \(A = (a_{ij})\) over \(\mathbb{F}_{2^n}\) such that

\[
\text{a}_{ij} = 0 \text{ for } 1 \leq i \leq k \text{ and } j \geq m + i.
\]

Then each matrix \(A \in \mathcal{M}_{\mathbb{F}_{2^n}}(m, k, r)\) defines a compression function \(\text{lp}_{mkr}^A\) with oracle access to a random permutation \(\pi \in \Pi_n\) as follows.

\[
\text{lp}_{mkr}^A : (\{0, 1\}^n)^m \longrightarrow (\{0, 1\}^n)^r
\]

\[
(v_1, \ldots, v_m) \longmapsto (w_1, \ldots, w_r)
\]

where \((w_1, \ldots, w_r)\) is computed by the algorithm described in Figure 1(a). A function \(\text{lp}_{mkr}^A\) is called linearly-dependent single-permutation-based, and often simply denoted as \(\text{lp}_{mkr}\) or \(\text{lp}^A\). A compression function \(\text{lp}_{231}^A\) for \(A \in \mathcal{M}_{\mathbb{F}_{2^n}}(2, 3, 1)\) is separately described in Figure 1(b).
Collision Resistance and Preimage Resistance For simplicity of notations, we will define collision resistance and preimage resistance focusing on linearly-dependent single-permutation-based compression functions, while these security notions can be extended in an obvious way to any hash function based on public ideal primitives.

Let $\text{lp}_{\text{mkr}}^A$ be a compression function for $A \in \mathcal{M}_{\mathbb{F}_n}(m, k, r)$. Given an information-theoretic adversary $A$ with oracle access to $\pi$ and $\pi^{-1}$, we execute the experiment $\text{Exp}_{\text{col}}^A$ described in Figure 2(a) in order to quantify the collision resistance of $\text{lp}_{\text{mkr}}^A$. The experiment records the query-response pairs that the adversary $A$ obtains into a query history $Q$. A pair $(x, y)$ is in the query history if $A$ makes $\pi(x)$ and gets back $y$, or it makes $\pi^{-1}(y)$ and gets back $x$. Given a query history $Q$, then $\text{Map}_{\text{lp}}^A(Q) \subseteq (\{0, 1\}^n)^m \times (\{0, 1\}^n)^r$ is defined to be the set of pairs $(v, w)$ such that there exist evaluations $(x_i, y_i) \in Q$ satisfying the following equations.

$$x_i = \sum_{j=1}^{m} a_{ij}v_j + \sum_{j=1}^{i-1} a_{i(m+j)}y_j, \quad i = 1, \ldots, k,$$

$$w_i = \sum_{j=1}^{m} a_{(k+i)j}v_j + \sum_{j=1}^{k} a_{(k+i)(m+j)}y_j, \quad i = 1, \ldots, r,$$

where we write $v = (v_1, \ldots, v_m)$ and $w = (w_1, \ldots, w_r)$. Informally, $\text{Map}_{\text{lp}}^A(Q)$ is the set of the evaluations of $\text{lp}_{\text{mkr}}^A$ that are determined by the query history $Q$. Now the collision-finding advantage of $A$ is defined to be

$$\text{Adv}^\text{col}_{\text{lp}}^A(A) = \Pr \left[ \text{Exp}_{\text{col}}^A = 1 \right].$$

The probability is taken over the random permutation $\pi$, and $A$’s coins (if any). For $q > 0$, we define $\text{Adv}^\text{col}_{\text{lp}}^A(q)$ as the maximum of $\text{Adv}^\text{col}_{\text{lp}}^A(A)$ over all adversaries $A$ making at most $q$ queries.

The preimage resistance of $\text{lp}_{\text{mkr}}^A$ is quantified similarly using the experiment $\text{Exp}_{\text{pre}}^A$ described in Figure 2(b). The adversary $A$ takes as input a random $w \in (\{0, 1\}^n)^r$ before it begins making queries to $\pi$ and $\pi^{-1}$. The preimage-finding advantage of $A$ is defined to be

$$\text{Adv}^\text{pre}_{\text{lp}}^A(A) = \Pr \left[ \text{Exp}_{\text{pre}}^A = 1 \right].$$

For $q > 0$, $\text{Adv}^\text{pre}_{\text{lp}}^A(q)$ is the maximum of $\text{Adv}^\text{pre}_{\text{lp}}^A(A)$ over all adversaries $A$ making at most $q$ queries.
Fig. 2. Experiments for quantification of collision resistance and preimage resistance

3 Security of lp231

In this section, we will prove that a linearly-determined permutation-based compression function $lp_{231}^A$ achieves good collision resistance and preimage resistance as long as the matrix $A \in \mathcal{M}_{2^n} (m, k, r)$ satisfies a certain condition. Our strategy is to first define a certain “bad event” parameterized by a matrix, and then show that collision or preimage finding is hard without the occurrence of the bad event. The bad event is the union of some “auxiliary events”, defined in the following subsection. Throughout this paper, we will write $N = 2^n$ and $N' = N - q$.

3.1 Auxiliary Events

In order to analyze the security of lp231, we need to define some auxiliary events. Suppose that an adversary $\mathcal{A}$ makes $q$ adaptive queries to a random permutation $\pi$ and its inverse $\pi^{-1}$, and records a query history

$$Q = \{(x^j, y^j) \in \{0, 1\}^n \times \{0, 1\}^n : 1 \leq j \leq q\}$$

where $(x^j, y^j)$ denotes the query-response pair obtained by the $j$-th query. Throughout this work, we will assume that $x^j$’s are all distinct. This means that $y^j$’s are also all distinct since $\pi$ is a permutation. It would not affect $\mathcal{A}$’s collision finding advantage making any redundant query.

For $t \geq 1$ and, $a_i, b_i \in \mathbb{F}_{2^n}$ and $A_i, B_i \in \mathcal{M}_{\mathbb{F}_{2^n}}^{2 \times 1}$, $i = 1, \ldots, t$, we define the following multisets.

$$U^t (a_1, b_1, \ldots, a_t, b_t) = \left\{ \sum_{i=1}^{t} (a_i x^{j_i} + b_i y^{j_i}) : j_1, \ldots, j_t \in [1, q] \right\},$$

$$U^{t}_{\neq} (a_1, b_1, \ldots, a_t, b_t) = \left\{ \sum_{i=1}^{t} (a_i x^{j_i} + b_i y^{j_i}) : j_1, \ldots, j_t \in [1, q] \text{ are all distinct} \right\},$$

$$V^t (A_1, B_1, \ldots, A_t, B_t) = \left\{ \sum_{i=1}^{t} (A_i x^{j_i} + B_i y^{j_i}) : j_1, \ldots, j_t \in [1, q] \right\}.$$
For a positive integer $l$, these multisets are associated with the following events.

$$E^l(a_1, b_1, \ldots, a_t, b_t; l) \leftrightarrow \mathcal{A} \text{ sets } \mu(U^l(a_1, b_1, \ldots, a_t, b_t)) > l,$$

$$E^l_x(a_1, b_1, \ldots, a_t, b_t; l) \leftrightarrow \mathcal{A} \text{ sets } \mu(U^l(a_1, b_1, \ldots, a_t, b_t)) > l,$$

$$F^l(A_1, B_1, \ldots, A_t, B_t; l) \leftrightarrow \mathcal{A} \text{ sets } \mu(V^l(A_1, B_1, \ldots, A_t, B_t)) > l.$$

We will often write $E^l(l) = E^l(a_1, b_1, \ldots, a_t, b_t; l)$, $E^l_x(l) = E^l_x(a_1, b_1, \ldots, a_t, b_t; l)$ and $F^l(l) = F^l(A_1, B_1, \ldots, A_t, B_t; l)$ for simplicity. The rest of this section is devoted to the estimation of the probability of these auxiliary events for small $t$’s (say $t = 1, 2, 3$).

**Theorem 1.** Let $a_1, a_2, b_1, b_2$ be nonzero elements in $\mathbb{F}_{2^n}$, and let

$$f_1 = f_1(d_1) = N \left( \frac{q}{d_1 + 1} \right) \left( \frac{1}{N} \right)^{d_1 + 1},$$

$$f_2 = f_2(d_2) = N \left( \frac{q}{d_2 + 1} \right) \left( \frac{2q}{N} \right)^{d_2 + 1}$$

for positive integers $d_1$ and $d_2$. Then the following hold.

1. $\Pr[E^1(a_1, b_1; d_1)] \leq f_1$.

2. $\Pr[E^2_x(a_1, b_1, a_2, b_2; 2d_1d_2)] \leq f_2 + 2f_1$.

3. If either $a_1 + a_2 \neq 0$ or $b_1 + b_2 \neq 0$, then

$$\Pr[E^2(a_1, b_1, a_2, b_2; 2d_1d_2 + d_1)] \leq f_2 + 3f_1.$$

**Proof.** First we give a proof for the first inequality. Fix $c \in \mathbb{F}_{2^n}$. When $\mathcal{A}$ makes the $j$-th forward query $y = \pi(x^*)$, the probability that $a_1x^* + b_1y = c$, which is equivalent to $y = b_1^{-1}(c + a_1x^*)$, is not greater than $1/(2^n - (j - 1))$. Similarly, when $\mathcal{A}$ makes the $j$-th backward query $x = \pi^{-1}(y^*)$, the probability that $a_1x + b_1y^* = c$ is not greater than $1/(2^n - (j - 1))$. The event $E^1(a_1, b_1; d_1)$ occurs when there exists a set $\{j_1, \ldots, j_{d_1+1}\} \subset [1, q]$ such that

$$a_1x^{j_1} + b_1y^{j_1} = \ldots = a_1x^{j_{d_1+1}} + b_1y^{j_{d_1+1}} = c$$

for some $c \in \mathbb{F}_{2^n}$. Since $1/(2^n - (j - 1)) \leq 1/N'$, it follows that

$$\Pr[E^1(a_1, b_1; d_1)] \leq N \left( \frac{q}{d_1 + 1} \right) \left( \frac{1}{N'} \right)^{d_1 + 1} = f_1.$$

For the proof of the second inequality, we define events

$$E^2_j(c) \leftrightarrow \mathcal{A} \text{ sets } a_1x^{j_1} + b_1y^{j_1} + a_2x^{j_2} + b_2y^{j_2} = c \text{ with } j = \max\{j_1, j_2\},$$

$$E^2(c; 2d_1d_2) \leftrightarrow \mathcal{A} \text{ sets } \mu(U^2, c) > 2d_1d_2$$

where we simply write $U^2 = U^2(a_1, b_1, a_2, b_2)$. Note that $E^2_j(c)$ occurs when the $j$-th query increases the multiplicity of $c$ in $U^2$ at least by one. In order to estimate $\Pr[E^2(2d_1d_2)]$, we decompose $E^2(2d_1d_2)$ as follows.

$$E^2(2d_1d_2) = \bigvee_{c \in \mathbb{F}_{2^n}} E^2_j(c; 2d_1d_2) \Rightarrow E^2(c) \lor \bigvee_{c \in \mathbb{F}_{2^n}} (E^2(c; 2d_1d_2) \land \neg E^2(c)) \quad (2)$$
where
\[ E_{ex} = E^1(a_1, b_1; d_1) \lor E^1(a_2, b_2; d_1). \]

By the first inequality, it follows that
\[ \Pr[E_{ex}] \leq 2f_1. \]  \hspace{1cm} (3)

We now analyze the event \( E^2(c; 2d_1d_2) \land \neg E_{ex} \) for a fixed \( c \in \mathbb{F}_{2^n} \). Suppose that a certain query, say the \( j \)-th query, completes equation
\[ a_1x^j + b_1y^j + a_2x^{j'} + b_2y^{j'} = c \]
for some \( j' < j \). For the fixed \( j \)-th query, the \( j' \)-th query should satisfy
\[ a_2x^{j'} + b_2y^{j'} = c + a_1x^{j'} + b_1y^{j'} \]
and the number of such queries is at most \( d_1 \) without the occurrence of \( E^1(a_2, b_2; d_1) \). Taking into account the other position where the last query might contributes, we see that each query increases the multiplicity \( \text{mult}(U^2, c) \) at most by \( 2d_1 \) without the occurrence of \( E_{ex} \). This implies that the number of queries that increase \( \text{mult}(U^2, c) \) should be at least \( d_2 + 1 \). Therefore we have
\[ E^2(c; 2d_1d_2) \land \neg E_{ex} \Rightarrow \bigvee_{J \subseteq [1,q]} \left( \bigwedge_{j \in J} \left( E^2_j(c) \land \neg E_{ex} \right) \right). \]  \hspace{1cm} (4)

In order to compute \( \Pr \left[ \bigwedge_{j \in J} \left( E^2_j(c) \land \neg E_{ex} \right) \right] \) for a fixed \( J \subseteq [1, q] \) such that \( |J| = d_2 + 1 \), suppose that \( A \) makes the \( j^* \)-th query \( \pi(x^*) \) for \( j^* \in J \). Then we can upper bound the number of responses \( y = \pi(x^*) \) that contribute the equation
\[ a_1x^{j_1} + b_1y^{j_1} + a_2x^{j_2} + b_2y^{j_2} = c, \]
with \( j^* = \max\{j_1, j_2\} \). If \( j_1 = j^* > j_2 \), then it should hold that
\[ y = b_2^{-1}(c + a_1x^* + a_2x^{j_2} + b_2y^{j_2}) \]
for some \( j_2 < j^* \leq q \). So the number of possible responses for this case is at most \( q \). The case \( j_2 = j^* > j_1 \) is analyzed similarly. Therefore the total number of possible responses \( y = \pi(x^*) \) is at most \( 2q \). With an analogous argument for \( \pi^{-1} \), we conclude that
\[ \Pr \left[ \bigwedge_{j \in J} \left( E^2_j(c) \land \neg E_{ex} \right) \right] \leq \left( \frac{2q}{N} \right)^{d_2 + 1}. \]  \hspace{1cm} (5)

Now by (2), (3), (4), and (5), we have
\[ \Pr[E^2(a_1, b_1, a_2, b_2; 2d_1d_2)] \leq N \left( \frac{q}{d_2 + 1} \right) \left( \frac{2q}{N} \right)^{d_2 + 1} + 2f_1 = f_2 + 2f_1. \]

In order to prove the third inequality, note that
\[ U^2(a_1, b_1, a_2, b_2) = U^2_\pi(a_1, b_1, a_2, b_2) \cup U^1(a_1 + a_2, b_1 + b_2). \]
If \( a_1 + a_2 \neq 0 \) and \( b_1 + b_2 \neq 0 \), then by the first inequality,
\[
\Pr \left[ E^1 (a_1 + a_2, b_1 + b_2; d_1) \right] \leq f_1.
\]
This inequality also holds for the case that either \( a_1 + a_2 \neq 0 \) or \( b_1 + b_2 \neq 0 \): since \( \pi \) is a permutation, this special case implies
\[
\Pr \left[ E^1 (a_1 + a_2, b_1 + b_2; 1) \right] = 0.
\]
Therefore by the first two inequalities,
\[
\Pr \left[ E^2(a_1, b_1, a_2, b_2; 2d_1d_2 + d_1) \right] \leq \Pr \left[ E^2_2(a_1, b_1, a_2, b_2; 2d_1d_2) \right] + \Pr \left[ E^1 (a_1 + a_2, b_1 + b_2; d_1) \right] \leq f_2 + 3f_1.
\]

**Definition 1.** For \( t \geq 1 \), a matrix \( M = [A_1, B_1, \ldots, A_t, B_t] \in \mathcal{M}^{2 \times 2t} \) is called column-sum independent if \( M \) satisfies the following conditions.

1. \( \left[ \sum_{i \in I_1} A_i, \sum_{i \in I_2} A_i \right] \) and \( \left[ \sum_{i \in I_1} B_i, \sum_{i \in I_2} B_i \right] \) are invertible for any pair of distinct nonempty subsets \( I_1, I_2 \subset [1, t] \).
2. \( \left[ \sum_{i \in I_1} A_i, \sum_{i \in I_2} B_i \right] \) are invertible for any pair of (not necessarily distinct) nonempty subsets \( I_1, I_2 \subset [1, t] \).

Definition 1 will be used for compact statement of the following corollary. We point out some useful properties of column-sum independent matrices.

**Property 1** If \([A_1, B_1, \ldots, A_t, B_t]\) is column-sum independent, then \([A_{i_1}, B_{i_1}, \ldots, A_{i_s}, B_{i_s}]\) is also column-sum independent for every nonempty subset \( \{i_1, \ldots, i_s\} \subset [1, t] \).

**Property 2** If \([A_1, B_1, \ldots, A_t, B_t]\) is column-sum independent, then \( \sum_{i \in I} A_i \neq 0 \) and \( \sum_{i \in I} B_i \neq 0 \) for every nonempty subset \( I \subset [1, t] \).

**Property 3** Column-sum independence of \( M \) stipulates nonsingularity of
\[
2 \left( 2^t - 1 \right) \left( 2^t - 2 \right) + (2^t - 1)^2
\]
matrices determined by \( M \).

**Theorem 2.** Let \( f_1 = f_1(d_1) \) and \( f_2 = f_2(d_2) \) be defined as in Theorem 1, and let
\[
g_2 = g_2(d_1, d_2) = \frac{d_2 (4d_1q + 2d_1 + 1) q}{N'}
\]
\[
g_3 = g_3(d_1, d_2, d_3) = N^2 q \left( \frac{6d_1d_2 + 6d_1 + 1}{N'} \right)^{d_3+1}
\]
for positive integers \( d_1, d_2 \) and \( d_3 \). If a matrix \([A_1, B_1, A_2, B_2, A_3, B_3] \in \mathcal{M}^{2 \times 6} \) is column-sum independent, then the following hold.

1. \( \Pr \left[ F^1(A_1, B_1; 1) \right] = 0. \)
2. \( \Pr \left[ F^2(A_1, B_1, A_2, B_2; 1) \right] \leq g_2 + 6f_2 + 14f_1. \)
3. \( \Pr \left[ F^3(A_1, B_1, A_2, B_2, A_3, B_3; 7d_3) \right] \leq g_3 + 3g_2 + 24f_2 + 66f_1. \)

**Proof.** The proof of the first equality is straightforward since \([A_1, B_1]\) is invertible. In order to prove the second inequality, we define the following events.

\[ F_{\text{col}}^2(j) \iff \mathcal{A} \text{ sets } A_1x^{j_1} + B_1y^{j_1} + A_2x^{j_2} + B_2y^{j_2} = A_1x^{j_3} + B_1y^{j_3} + A_2x^{j_4} + B_2y^{j_4} \]

where \( j_3 < j_1 \leq j, \ j_4 < j_2 \leq j, \) and \( j = \max\{j_1, j_2\} \)

and

\[
E_{\text{ex}} = E^2 ((B_1 + B_2)^* A_1, (B_1 + B_2)^* B_1, (B_1 + B_2)^* A_2, (B_1 + B_2)^* B_2; 2d_1d_2 + d_1) \\
\vee E^2 ((A_1 + A_2)^* A_1, (A_1 + A_2)^* B_1, (A_1 + A_2)^* A_2, (A_1 + A_2)^* B_2; 2d_1d_2 + d_1) \\
\vee E_{\neq}^2 (B_1^* A_2, B_1^* B_2, B_1^* A_2, B_1^* B_2; 2d_1d_2) \\
\vee E_{\neq}^2 (A_1^* A_2, A_1^* B_2, A_1^* A_2, A_1^* B_2; 2d_1d_2) \\
\vee E_{\neq}^2 (B_2^* A_1, B_2^* B_1, B_2^* A_1, B_2^* B_1; 2d_1d_2) \\
\vee E_{\neq}^2 (B_2^* A_2, B_2^* B_2, B_2^* A_2, B_2^* B_2; 2d_1d_2) .
\]

Then it follows that

\[
F^2(A_1, B_1, A_2, B_2; 1) \Rightarrow \bigvee_{1 \leq j \leq q} F_{\text{col}}^2(j) \Rightarrow E_{\text{ex}} \vee \bigvee_{1 \leq j \leq q} (F_{\text{col}}^2(j) \land \neg E_{\text{ex}})
\]

and

\[
\Pr [E_{\text{ex}}] \leq 2 (f_2 + 3f_1) + 4 (f_2 + 2f_1) = 6f_2 + 14f_1
\]

by Theorem 1.

We now estimate the probability \( \Pr \left[ F_{\text{col}}^2(j) \land \neg E_{\text{ex}} \right] \). Suppose that \( \mathcal{A} \) makes the \( j \)-th query \( \pi(x^*) \), and consider the following three cases where \( y = \pi(x^*) \) contributes the equation

\[
A_1x^{j_3} + B_1x^{j_1} + A_2x^{j_2} + B_2x^{j_2} = A_1x^{j_3} + B_1y^{j_3} + A_2x^{j_4} + B_2y^{j_4}.
\]

**Case 1:** \( j_1 = j_2 = j \). The equality (8) is reduced to

\[
(A_1 + A_2)x^* + (B_1 + B_2)y = A_1x^{j_3} + B_1y^{j_3} + A_2x^{j_4} + B_2y^{j_4}.
\]

Any response \( y \) satisfying (9) corresponds to a pair \((j_3, j_4) \in [1, j - 1]^2\) such that

\[
(B_1 + B_2)^* A_1x^{j_3} + (B_1 + B_2)^* B_1y^{j_3} + (B_1 + B_2)^* A_2x^{j_4} + (B_1 + B_2)^* B_2y^{j_4}
\]

\[
= (B_1 + B_2)^* (A_1 + A_2)x^* .
\]

The number of such pairs is at most \( 2d_1d_2 + d_1 \) without the occurrence of \( E_{\text{ex}} \).

**Case 2:** \( j_1 = j \) and \( j_2 \neq j \). The equality (8) is reduced to

\[
A_1x^* + B_1y = A_2x^{j_2} + B_2y^{j_2} + A_1x^{j_3} + B_1y^{j_3} + A_2x^{j_4} + B_2y^{j_4}.
\]

Any response \( y \) satisfying (11) corresponds to a triple \((j_2, j_3, j_4) \in [1, j - 1]^3\) such that \( j_2 \neq j_4 \) and

\[
B_1^* A_1x^{j_3} + (B_1^* A_2x^{j_2} + B_1^* B_2y^{j_2} + B_1^* A_2x^{j_4} + B_1^* B_2y^{j_4}) = B_1^* A_1x^* .
\]

For each \( j_3 \in [1, j - 1] \), the number of pairs \((j_2, j_4) \in [1, j - 1]^2\) satisfying (12) and \( j_2 \neq j_4 \) is at most \( 2d_1d_2 \) without the occurrence of \( E_{\text{ex}} \). Therefore the number of the triples satisfying (11) is at most \( 2d_1d_2q \) without the occurrence of \( E_{\text{ex}} \).
By Theorem 1 and the first two inequalities, it follows that

\[
\Pr \left[ F_{\text{col}}^2(j) \land \neg e_{\text{ex}} \right] \leq \frac{d_2 (4d_1 q + 2d_1 + 1)}{N^t}.
\]  

(13)

We can apply a similar argument for \( \pi^{-1} \). So using a union bound, the second inequality is followed from (6), (7) and (13).

In order to prove the third inequality, we define events

\[
F_j^3(C) \iff \mathcal{A} \text{ sets } \sum_{i=1}^{3} (A_i x^{j_i} + B_i y^{j_i}) = C \text{ with } j = \max\{j_1, j_2, j_3\},
\]  

(14)

for \( C \in \mathcal{M}_{F_{\text{ex}}}^{2 \times 1} \) and \( j \in [1, q] \). Here we simply write \( V^3 = V^3(A_1, B_1, A_2, B_2, A_3, B_3) \). The event \( F_j^3(C) \) occurs when the \( j \)-th query increases the multiplicity of \( C \) in \( V^3 \) at least by one.

In order to estimate \( \Pr \left[ F_j^3(7d_3) \right] \), we decompose \( F_j^3(7d_3) \) as follows.

\[
F_j^3(7d_3) = \bigvee_{C \in \mathcal{M}_{F_{\text{ex}}}^{2 \times 1}} F_j^3(C; 7d_3) \Rightarrow F_{\text{ex}} \lor \bigvee_{C \in \mathcal{M}_{F_{\text{ex}}}^{2 \times 1}} (F_j^3(C; 7d_3) \land \neg F_{\text{ex}})
\]  

(15)

where

\[
F_{\text{ex}} = F^1(A_1, B_1; 1) \lor F^1(A_2, B_2; 1) \lor F^1(A_3, B_3; 1) \\
\lor F^2(A_1, B_1, A_2, B_2; 1) \lor F^2(A_2, B_2, A_3, B_3; 1) \lor F^2(A_3, B_3, A_1, B_1; 1) \\
\lor E^1 ((A_2 + A_3)^* A_1, (A_2 + A_3)^* B_1; d_1) \lor E^1 ((A_3 + A_1)^* A_2, (A_3 + A_1)^* B_2; d_1) \\
\lor E^1 ((A_1 + A_2)^* A_3, (A_1 + A_2)^* B_3; d_1) \lor E^1 ((B_2 + B_3)^* A_1, (B_2 + B_3)^* B_1; d_1) \\
\lor E^1 ((B_3 + B_1)^* A_2, (B_3 + B_1)^* B_2; d_1) \lor E^1 ((B_1 + B_2)^* A_3, (B_1 + B_2)^* B_3; d_1) \\
\lor E^2 (A_2^* A_1, A_3^* B_1, A_3^* A_2, A_3^* B_2; 2d_1 d_2 + d_1) \\
\lor E^2 (A_1^* A_2, A_1^* B_2, A_1^* A_3, A_1^* B_3; 2d_1 d_2 + d_1) \\
\lor E^2 (A_2^* A_3, A_2^* B_3, A_2^* A_1, A_2^* B_1; 2d_1 d_2 + d_1) \\
\lor E^2 (B_2^* A_1, B_2^* B_1, B_3^* A_2, B_3^* B_2; 2d_1 d_2 + d_1) \\
\lor E^2 (B_1^* A_2, B_1^* B_2, B_1^* A_3, B_1^* B_3; 2d_1 d_2 + d_1) \\
\lor E^2 (B_2^* A_3, B_2^* B_3, B_2^* A_1, B_2^* B_1; 2d_1 d_2 + d_1).
\]  

(16)

By Theorem 1 and the first two inequalities, it follows that

\[
\Pr \left[ F_{\text{ex}} \right] \leq 3(g_2 + 6f_2 + 14f_1) + 6f_1 + 6(f_2 + 3f_1) = 3g_2 + 24f_2 + 66f_1.
\]  

(17)

We now analyze the event \( F_j^3(C; 7d_3) \land \neg F_{\text{ex}} \) for a fixed \( C \in \mathcal{M}_{F_{\text{ex}}}^{2 \times 1} \). Since each query increases the multiplicity \( \text{mult}(V^3, C) \) at most by 7 without the occurrence of \( F_{\text{ex}} \) (according to the positions where the query contributes), the number of queries that increase \( \text{mult}(V^3, C) \) should be at least \( d_3 + 1 \). Therefore we obtain

\[
F_j^3(C; 7d_3) \land \neg F_{\text{ex}} \Rightarrow \bigvee_{J \subseteq [1, q]} \left( \bigwedge_{|J| = d_3 + 1} (F_j^3(C) \land \neg F_{\text{ex}}) \right).
\]  

(18)
suppose that \( F \) of responses \( y \)

For \( 3.2 \) Concrete Security Bounds for \( \text{lp} \)

Using a union bound, the proof is complete from (15), (17), (18) and (22).

\[
\sum_{i=1}^{s} (A_i x^{j_i} + B_i y^{j_i}) + \bar{A}x^* + \bar{B}y = C
\]

with \( j^* = \max\{j_1, j_2, j_3\} \). For \( s \in \{0, 1, 2\} \), consider the case where the \( j^* \)-th query contributes \( 3 - s \) terms in equation (19). If \( j_{s+1} = \cdots = j_3 = j^* \), then the equation (19) is reduced to

Without the occurrence of \( F_{ex} \), the number of solutions \( (j_1, \ldots, j_s) \) to (21) is at most 1 if \( s = 0, d_1 \) if \( s = 1 \) and \( 2d_1d_2 + d_1 \) if \( s = 2 \). By symmetry in the positions of the \( j^* \)-th query and with an analogous argument for \( \pi^{-1} \), we obtain

\[
\Pr \left[ \bigwedge_{j \in J} \left( F_j^3(C) \land \neg F_{ex} \right) \right] \leq \left( \frac{6d_1d_2 + 6d_1 + 1}{N^s} \right)^{d_3+1}.
\]

Using a union bound, the proof is complete from (15), (17), (18) and (22).

### 3.2 Concrete Security Bounds for \( \text{lp231} \)

For \( A \in \mathcal{M}_{\mathbb{F}_2^n}(2, 3, 1) \), the system of equations (1) is rewritten as follows.

\[
x_1 = a_{11}v_1 + a_{12}v_2 \quad (23)
\]
\[
x_2 = a_{21}v_1 + a_{22}v_2 + a_{23}y_1 \quad (24)
\]
\[
x_3 = a_{31}v_1 + a_{32}v_2 + a_{33}y_1 + a_{34}y_2 \quad (25)
\]
\[
w = a_{41}v_1 + a_{42}v_2 + a_{43}y_1 + a_{44}y_2 + a_{45}y_3. \quad (26)
\]

From equations (23) and (24), we obtain the following system of equations in variables \( v_1 \) and \( v_2 \).

\[
\begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
  v_1 \\
  v_2
\end{bmatrix}
= 
\begin{bmatrix}
  x_1 \\
  x_2 + a_{23}y_1
\end{bmatrix}. \quad (27)
\]

If \( a_{11}a_{22} + a_{12}a_{21} \neq 0 \), then we can solve the system of equations (27). By substituting its solution into equations (25) and (26), we obtain an equation of the following form.

\[
A_1x_1 + B_1y_1 + A_2x_2 + B_2y_2 + A_3x_3 + B_3y_3 = Cw, \quad (28)
\]

where \( A_i \), \( B_i \) and \( C \) are matrices in \( \mathcal{M}_{\mathbb{F}_2^n}^{2 \times 1} \). We write \( M(A) = [A_1, B_1, A_2, B_2, A_3, B_3] \) and \( C(A) = C \), indicating these matrices are determined by the matrix \( A \). Note that

\[
A_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 \\ a_{45} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\]
Preimage Resistance

**Theorem 3.** Let $A = (a_{ij})$ be a matrix in $M_{F_{2^n}}(2,3,1)$ such that $a_{11}a_{22} + a_{12}a_{21} \neq 0$. If $M(A)$ is column-sum independent, then for positive integers $d_1$ and $d_2$,

$$\text{Adv}_{\text{lp}}(q) \leq \frac{(6d_1d_2 + 6d_1 + 1)q}{N'} + 6f_2 + 24f_1$$

where $f_1 = f_1(d_1)$ and $f_2 = f_2(d_2)$ are defined as in Theorem 1.

**Proof.** We will upper bound $\text{Adv}_{\text{lp}}(q)$ for an adversary $A$ makes $q$ queries to a random permutation $\pi$ and its inverse $\pi^{-1}$. Let

$$G_{\text{ex}} = E^1((A_2 + A_3)A_1, (A_2 + A_3)B_1; d_1) \lor E^1((A_3 + A_1)A_2, (A_3 + A_1)B_2; d_1)$$

$$\lor E^1((A_1 + A_2)A_3, (A_1 + A_2)B_3; d_1) \lor E^1((B_2 + B_3)A_1, (B_2 + B_3)B_1; d_1)$$

$$\lor E^1((B_3 + B_1)A_2, (B_3 + B_1)B_2; d_1) \lor E^1((B_1 + B_2)A_3, (B_1 + B_2)B_3; d_1)$$

$$\lor E^2(A_3A_1, A_3B_1, A_3A_2, A_3B_2; 2d_1d_2 + d_1)$$

$$\lor E^2(A_1A_2, A_1B_2, A_1A_3, A_1B_3; 2d_1d_2 + d_1)$$

$$\lor E^2(A_2A_3, A_2B_3, A_2A_1, A_2B_1; 2d_1d_2 + d_1)$$

$$\lor E^2(B_3A_1, B_3B_1, B_3A_2, B_3B_2; 2d_1d_2 + d_1)$$

$$\lor E^2(B_1A_2, B_1B_2, B_1A_3, B_1B_3; 2d_1d_2 + d_1)$$

$$\lor E^2(B_2A_3, B_2B_3, B_2A_1, B_2B_1; 2d_1d_2 + d_1).$$

Then by Theorem 1,

$$\Pr[G_{\text{ex}}] \leq 6f_2 + 24f_1. \tag{29}$$

For $w \in F_{2^n}$ and $j \in [1, q]$, we define an event

$$P_j(w) \iff A \text{ sets } A_1x^{j_1} + B_1y^{j_1} + A_2x^{j_2} + B_2y^{j_2} + A_3x^{j_3} + B_3y^{j_3} = C \text{ with } j = \max\{j_1, j_2, j_3\}. \tag{30}$$

Since the occurrence of $P_j(w)$ means that the $j$-th query determines a preimage of $w$, it follows that

$$\text{Adv}_{\text{lp}}(q) \leq \max_{w \in F_{2^n}} \Pr \left[ \bigvee_{1 \leq j \leq q} P_j(w) \right]. \tag{31}$$

For a fixed $w \in F_{2^n}$, we use the following decomposition.

$$\Pr \left[ \bigvee_{1 \leq j \leq q} P_j(w) \right] \leq \Pr \left[ \bigvee_{1 \leq j \leq q} (P_j(w) \land \neg G_{\text{ex}}) \right] + \Pr[G_{\text{ex}}]. \tag{32}$$

Note that the event $P_j(w)$ is identical with $F_j^3(Cw)$ as defined in (14). So with the same analysis as $F_j^3(Cw)$, the number of responses that determine a preimage of $w$ is at most $6d_1d_2 + 6d_1 + 1$ for each query without the occurrence of $G_{\text{ex}}$. Therefore we obtain

$$\Pr \left[ \bigvee_{1 \leq j \leq q} (P_j(w) \land \neg G_{\text{ex}}) \right] \leq \frac{6d_1d_2 + 6d_1 + 1)q}{N'}. \tag{33}$$

The proof is complete from (29), (31), (32) and (33). \qed
Corollary 1. Let $A = (a_{ij})$ be a matrix in $\mathcal{M}_{\mathbb{F}_2^n}(2, 3, 1)$ such that $a_{11}a_{22} + a_{12}a_{21} \neq 0$. If $M(A)$ is column-sum independent, then

$$\lim_{n \to \infty} \text{Adv}^{\text{pre}}_{\text{lp}^4}(2^{\frac{2n}{N}}/n) = 0.$$  

Proof. Let $q = 2^{\frac{2n}{N}}/n$, $d_1 = 2$ and $d_2 = 5 \cdot 2^q - 1$. Since

$$\frac{(6d_1d_2 + 6d_1 + 1)q}{N'} = \frac{(6 \cdot 2^q + 1)q}{N'} \leq \frac{122}{n},$$

$$f_1 = N \left( \frac{q}{d_1 + 1} \right) \left( \frac{1}{N'} \right)^{d_1+1} \leq N \left( \frac{2e}{(d_1 + 1)N} \right)^{d_1+1} = N \left( \frac{2eq}{3N} \right)^3 = \frac{8e^3}{27n^3}$$

and

$$f_2 = N \left( \frac{q}{d_2 + 1} \right) \left( \frac{2q}{N'} \right)^{d_2+1} \leq N \left( \frac{2eq^2}{(d_2 + 1)N'} \right)^{d_2+1} \leq N \left( \frac{12e}{5n^2} \right)$$

using $N' \geq N/2$, it follows that

$$\lim_{n \to \infty} \frac{(6d_1d_2 + 6d_1 + 1)q}{N'} = \lim_{n \to \infty} f_2(d_1, d_2) = \lim_{n \to \infty} f_1(d_1) = 0.$$  

The proof is complete from Theorem 3.  

Collision Resistance

Theorem 4. Let $A = (a_{ij})$ be a matrix in $\mathcal{M}_{\mathbb{F}_2^n}(2, 3, 1)$ such that $a_{11}a_{22} + a_{12}a_{21} \neq 0$, and let $M(A) = [A_1, B_1, A_2, B_2, A_3, B_3]$ and $C(A) = C$ satisfy the following conditions.

1. $M(A)$ is column-sum independent.

2. $[A_1, C], [A_2, C], [A_3, C], [B_1, C]$ and $[B_2, C]$ are invertible.

3. The following $2 \times 6$ matrices are column-sum independent.

$$D^1 = [[B_2, C]^{-1}[A_1, B_1], [B_1, C]^{-1}[A_2, B_2], [B_1, C]^{-1}[A_3, B_3]],$$

$$D^2 = [[A_2, C]^{-1}[A_1, B_1], [A_1, C]^{-1}[A_2, B_2], [A_1, C]^{-1}[A_3, B_3]],$$

$$D^3 = [[A_3, C]^{-1}[A_1, B_1], [A_3, C]^{-1}[A_2, B_2], [A_1, C]^{-1}[A_3, B_3]],$$

$$D^4 = [[A_3, C]^{-1}[A_1, B_1], [A_3, C]^{-1}[A_2, B_2], [A_2, C]^{-1}[A_3, B_3]].$$

Then for positive integers $d_1, d_2$ and $d_3$,

$$\text{Adv}^{\text{col}}_{\text{lp}^4}(q) \leq Nq^2 \left( \frac{6d_1d_2 + 6d_1 + 1}{N'} \right)^2 + \frac{3(2d_1d_2 + d_1)q + 3q \max\{7d_3q, 2d_1d_2 + d_1\}}{N'}$$

$$+ 4g_3 + 15g_2 + 121f_2 + 333f_1,$$

where $f_1 = f_1(d_1)$, $f_2 = f_2(d_2)$, $g_2 = g_2(d_1, d_2)$ and $g_3 = g_3(d_1, d_2, d_3)$ are defined as in Theorem 1 and 2.
Proof. We will upper bound \( \text{Adv}_{lp_A}^{\text{col}}(A) \) for an adversary \( A \) makes \( q \) queries to a random permutation \( \pi \) and its inverse \( \pi^{-1} \). Let

\[
H_{ex} = F_{ex} \lor E^2 (B^*A_1, B^*B_1, B^*A_2, B^*B_2; 2d_1d_2 + d_1) \lor F^3 (D^1; 7d_3) \lor F^3 (D^2; 7d_3) \lor F^3 (D^3; 7d_3) \lor F^3 (D^4; 7d_3)
\]

where \( \bar{B} = B_1 + B_2 + B_3 \) and \( F_{ex} \) is defined as in (16). Then it follows that

\[
\Pr [H_{ex}] = 3g_2 + 24f_2 + 66f_1 + (f_2 + 3f_1) + 4(g_3 + 3g_2 + 24f_2 + 66f_1) = 4g_3 + 15g_2 + 121f_2 + 333f_1. \tag{34}
\]

For \( j \in [1, q] \) and \( \rho^1, \rho^2 \in \{0, 1\}^3 \setminus \{(0, 0, 0)\} \), we define the following event.

\[
C_j^2(\rho^1, \rho^2) \leftrightarrow A \text{ sets } A_1x_j^1 + B_1y_j^1 + A_2x_j^2 + B_2y_j^2 + A_3x_j^3 + B_3y_j^3 = Cw \text{ and } A_1x_j^1 + B_1y_j^1 + A_2x_j^2 + B_2y_j^2 + A_3x_j^3 + B_3y_j^3 = Cw,
\]

where \( w \in \mathbb{F}_{2^n} \), \( (j_1, j_2, j_3) \neq (j_1^*, j_2^*, j_3^*) \), \( j = \max \{j_i^*\} \) and \( s = 1, 2 \).

Here \( \rho^1 \) and \( \rho^2 \) specify the positions where the \( j \)-th query contributes within the two different evaluations. Since \( \rho^1, \rho^2 \neq (0, 0, 0) \), the occurrence of \( C_j^2(\rho^1, \rho^2) \) means that the \( \text{single} \ j \)-th query completes two colliding evaluations of \( lp_A^{231} \) at the same time. Let

\[
C^1 = \bigvee_{1 \leq j_1 < j_2 \leq q} \Pr [P_{j_1}(w) \land P_{j_2}(w)]
\]

where events \( P_{j_1}(w) \) and \( P_{j_2}(w) \) are defined as in (30) and let

\[
C^2 = \bigvee_{(\rho^1, \rho^2) \in P} C_j^2(\rho^1, \rho^2)
\]

where \( P = P_1 \cup P_2 \cup P_3 \) and

\[
\begin{align*}
P_1 &= \{(\rho^1, \rho^2) : \exists i \in \{1, 2, 3\} \text{ such that } \rho^1_i = \rho^2_i = 1\}, \\
P_2 &= \{((1, 1, 0), (0, 0, 1)), ((1, 0, 1), (0, 1, 0)), ((0, 1, 1), (1, 0, 0))\}, \\
P_3 &= \{((1, 0, 0), (0, 1, 0)), ((1, 0, 0), (0, 0, 1)), ((0, 1, 0), (0, 1, 0))\}.
\end{align*}
\]

Then by the symmetry of \( \rho^1 \) and \( \rho^2 \), it follows that

\[
\text{Adv}_{lp_A}^{\text{col}}(A) = \Pr [C^1 \lor C^2] \leq \Pr [H_{ex}] + \Pr [C^1 \land \neg H_{ex}] + \Pr [C^2 \land \neg H_{ex}]. \tag{35}
\]

**Estimation of \( \Pr [C^1 \land \neg H_{ex}] \).** If events \( P \) are defined as (30), then

\[
\Pr [C^1 \land \neg H_{ex}] = \Pr \left[ \bigvee_{1 \leq j_1 < j_2 \leq q} \Pr [P_{j_1}(w) \land P_{j_2}(w) \land \neg H_{ex}] \right].
\]
With a similar argument as the analysis of preimage resistance, we obtain

\[
\Pr \left[ P_{j_1^*}(w^*) \land P_{j_2^*}(w^*) \land \neg H_{ex} \right] \leq \left( \frac{6d_1d_2 + 6d_1 + 1}{N} \right)^2
\]

for fixed \( w^* \in \mathbb{F}_{2^n} \) and \( 1 \leq j_1^* < j_2^* \leq q \). Therefore we have

\[
\Pr \left[ C^1 \land \neg H_{ex} \right] = \Pr \left[ \bigvee_{1 \leq j_1 < j_2 \leq q} \left( P_{j_1}(w) \land P_{j_2}(w) \land \neg H_{ex} \right) \right] \\
\leq Nq^2 \left( \frac{6d_1d_2 + 6d_1 + 1}{N} \right)^2. \tag{36}
\]

**Estimation of** \( \Pr \left[ C^2 \land \neg H_{ex} \right] \). In order to use

\[
\Pr \left[ C^2 \land \neg H_{ex} \right] \leq \sum_{1 \leq j \leq q} \Pr \left[ C_j^2(\rho^1, \rho^2) \land \neg H_{ex} \right] \tag{37}
\]

we focus on the estimation of \( \Pr \left[ C_j^2(\rho^1, \rho^2) \land \neg H_{ex} \right] \) for each \((\rho^1, \rho^2) \in P \). We consider three cases as follows.

**Case 1:** We estimate the probability \( \Pr \left[ C_j^2(\rho^1, \rho^2) \land \neg H_{ex} \right] \) for \((\rho^1, \rho^2) \in P_1 \). Suppose that \( \rho^1 = \rho^2 = (1,0,0) \). If the event \( C_j^2((1,0,0), (1,0,0)) \) occurs, then it holds that

\[
A_1x^j + B_1y^j + A_2x^{j_2} + B_2y^{j_2} + A_3x^{j_3} + B_3y^{j_3} = Cw, \tag{38}
\]

\[
A_1x^j + B_1y^j + A_2x^{j_2} + B_2y^{j_2} + A_3x^{j_3} + B_3y^{j_3} = Cw \tag{39}
\]

for some \( j_2^1, j_3^1, j_2^2, j_3^2 < j \) and \( w \in \mathbb{F}_{2^n} \). The equations (38) and (39) imply that

\[
A_2x^{j_2^1} + B_2y^{j_2^1} + A_3x^{j_3^1} + B_3y^{j_3^1} = A_2x^{j_2^2} + B_2y^{j_2^2} + A_3x^{j_3^2} + B_3y^{j_3^2}.
\]

Therefore it follows that

\[
C_j^2((1,0,0), (1,0,0)) \Rightarrow F^2(A_2, B_2, A_3, B_3; 1) \Rightarrow H_{ex},
\]

and hence,

\[
\Pr \left[ C_j^2((1,0,0), (1,0,0)) \land \neg H_{ex} \right] = 0.
\]

The same argument applies to any event \( C_j^2(\rho^1, \rho^2) \) such that \( \rho^1 \land \rho^2 \neq (0,0,0) \).

**Case 2:** We estimate the probability \( \Pr \left[ C_j^2(\rho^1, \rho^2) \land \neg H_{ex} \right] \) for \((\rho^1, \rho^2) \in P_2 \). Say \( \rho^1 = (1,1,0) \) and \( \rho^2 = (0,0,1) \). Suppose that \( A \) makes the \( j \)-th query \( \pi(x^*) \). Among \((N - (j - 1)) \) possible responses for \( y = \pi(x^*) \), we need to upper bound the number of \( y = \pi(x^*) \) satisfying

\[
(A_1 + A_2)x^* + (B_1 + B_2)y + A_3x^{j_3^1} + B_3y^{j_3^1} = Cw, \tag{40}
\]

\[
A_1x^{j_1^2} + B_1y^{j_1^2} + A_2x^{j_2^2} + B_2y^{j_2^2} + A_3x^* + B_3y = Cw \tag{41}
\]
for some $j_3^1, j_1^2, j_2^2 < j$ and $w \in \mathbb{F}_{2^n}$. Adding (40) and (41), we obtain

$$Ax^* + By + A_1x^{j_3^1} + B_1y^{j_3^1} + A_2x^{j_1^2} + B_2y^{j_1^2} + A_3x^{j_2^2} + B_3y^{j_2^2} = 0$$

which implies the following equation.

$$B^* A_1x^{j_3^1} + B^* B_1y^{j_3^1} + B^* A_2x^{j_1^2} + B^* B_2y^{j_1^2} + \left( B^* A_3x^{j_2^2} + B^* B_3y^{j_2^2} \right) = B^* Ax^*.$$  \hspace{1cm} (42)

The number of solutions ($j_3^1, j_1^2, j_2^2$) to (42) is at most $(2d_1d_2 + d_1)q$ without the occurrence of $H_{ex}$. We can first fix the $j_3^1$-th query, and then count the number of pairs $(j_1^2, j_2^2)$ satisfying equation (42). For events $C_j^2((0,1,1),(1,0,0))$ and $C_j^2((1,0,1),(0,1,0))$, we have the same equation as (42). Therefore we have

$$\Pr \left[ C_j^2 (\rho^1, \rho^2) \land \neg H_{ex} \right] \leq \frac{(2d_1d_2 + d_1)q}{N'}.$$  \hspace{1cm} (43)

**Case 3:** We estimate the probability $\Pr \left[ C_j^2 (\rho^1, \rho^2) \land \neg H_{ex} \right]$ for $(\rho^1, \rho^2) \in P_3$. Say $\rho^1 = (1,0,0)$ and $\rho^2 = (0,1,0)$. Suppose that $A$ makes the $j$-th query $\pi(x^*)$. Among $(N - (j - 1))$ possible responses for $y = \pi(x^*)$, we need to upper bound the number of $y = \pi(x^*)$ satisfying

$$A_1x^* + B_1y + A_2x^{j_1^2} + B_2y^{j_1^2} + A_3x^{j_2^2} + B_3y^{j_2^2} = Cw,$$

$$A_1x^{j_3^1} + B_1y^{j_3^1} + A_2x^{j_1^2} + B_2y^{j_1^2} + A_3x^{j_2^2} + B_3y^{j_2^2} = Cw,$$

for some $j_3^1, j_1^2, j_2^2 < j$ and $w \in \mathbb{F}_{2^n}$. Removing variables $y$ and $w$ from this system of equations, we obtain the following equation.

$$[B_1, C]^{-1} A_2x^{j_1^2} + [B_1, C]^{-1} B_2y^{j_1^2} + [B_1, C]^{-1} A_3x^{j_2^2} + [B_1, C]^{-1} B_3y^{j_2^2}$$

$$+ [B_2, C]^{-1} A_1x^{j_3^1} + [B_2, C]^{-1} B_1y^{j_3^1} + \left( [B_2, C]^{-1} A_3x^{j_2^2} + [B_2, C]^{-1} B_3y^{j_2^2} \right)$$

$$= ([B_1, C]^{-1} A_1 + [B_2, C]^{-1} A_2) x^*.$$  \hspace{1cm} (44)

For each $j_3^2$, the number of solutions $(j_1^2, j_3^1, j_2^2)$ to (44) is at most $7d_3$ without the occurrence of $F^3(D^1, 7d_3)$. Therefore the number of solutions $(j_1^2, j_3^1, j_2^2)$ to (44) is at most $7d_3q$.

One special case is when $\rho^1$ or $\rho^2$ is $(0,0,1)$ and $A$ makes a forward query $y = \pi(x^*)$. Since $[B_3, C]$ is not invertible, the above argument does not apply to this case. Say $\rho^1 = (0,0,1)$. Then the response $y = \pi(x^*)$ should satisfy

$$A_1x^{j_3^1} + B_1y^{j_3^1} + A_2x^{j_1^2} + B_2y^{j_1^2} + A_3x^* + B_3y = Cw,$$  \hspace{1cm} (45)

for some $j_1^2$ and $w \in \mathbb{F}_{2^n}$. Multiplying $B_3^* C$ on both sides of (44) and using $B_3^* C = 0$, we see that each $y$ satisfying (44) is associated with a solution $(j_1^2, j_2^2)$ to the following equation.

$$B_3^* A_1x^{j_3^1} + B_3^* B_1y^{j_3^1} + B_3^* A_2x^{j_1^2} + B_3^* B_2y^{j_1^2} = B_3^* A_3x^*.$$  \hspace{1cm} (45)

The number of solutions $(j_1^2, j_2^2)$ to (45) is at most $2d_1d_2 + d_1$ without the occurrence of $E^2(B_3^* A_1, B_3^* B_1, B_3^* A_2, B_3^* B_2; 2d_1d_2 + d_1)$. Therefore we have

$$\Pr \left[ C_j^2 (\rho^1, \rho^2) \land \neg H_{ex} \right] \leq \frac{\max\{7d_3q, 2d_1d_2 + d_1\}}{N'}.$$
To summarize the analysis for the three cases, we conclude that
\[
\Pr \left[ C^2 \land \neg H_{ex} \right] \leq \sum_{1 \leq j \leq q \atop (\rho^1, \rho^2) \in P} \Pr \left[ C_j^2(\rho^1, \rho^2) \land \neg H_{ex} \right] \\
\leq \frac{3(2d_1d_2 + d_1)q + 3q \max\{7d_3q, 2d_1d_2 + d_1\}}{N'}.
\]
\tag{46}

Now the proof is complete from (34), (35), (36) and (46).

\[\square\]

**Corollary 2.** Let \(A = (a_{ij})\) be a matrix in \(\mathcal{M}_{F_2n}(2,3,1)\) such that \(a_{11}a_{22} + a_{12}a_{21} \neq 0\). If \(M(A)\) and \(C(A)\) satisfy the conditions described in Theorem 4, then
\[
\lim_{n \to \infty} \text{Adv}^{\text{col}}_{lp^A} \left(2^{\frac{3}{2}/n^{1+\epsilon}}\right) = 0,
\]
for any \(\epsilon > 0\).

**Proof.** Let \(q = 2^{\frac{3}{2}}/n^{1+\epsilon}\) and \((d_1, d_2, d_3) = (1, n, 3)\). Then it is easy to check that
\[
\max\{7d_3q, 2d_1d_2 + d_1\} = 7d_3q
\]
and
\[
Nq^2 \left( \frac{6d_1d_2 + 6d_1 + 1}{N'} \right)^2 = O \left( \frac{1}{n^{2\epsilon}} \right), \quad 3(2d_1d_2 + d_1)q + 21d_3q^2 = O \left( \frac{1}{n^{2+2\epsilon}} \right),
\]
\[
f_1(d_1) = O \left( \frac{1}{n^{2+2\epsilon}} \right), \quad f_2(d_2) = O \left( \frac{1}{n^{(2+2\epsilon)n}} \right),
\]
\[
g_2(d_1, d_2) = O \left( \frac{1}{n^{1+2\epsilon}} \right), \quad g_3(d_1, d_2, d_3) = O \left( \frac{1}{n^{4\epsilon}} \right).
\]
Since all the terms converge to 0 as \(n\) goes to infinity, and by Theorem 4, we obtain the corollary. \(\square\)

**Example 1.** Let \(n = 128\) and let \(F_{2^{128}} = F[\zeta]/(\zeta^{128} + \zeta^7 + \zeta^2 + \zeta + 1)\) be a finite field of order \(2^{128}\), where \(f(\zeta) = \zeta^{128} + \zeta^7 + \zeta^2 + \zeta + 1\) is an irreducible polynomial over \(F_2\). For simplicity of computation, assume that \(a_{23} = 0, a_{45} = 1, \) and
\[
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]
Then we have
\[
M(A) = \begin{bmatrix} a_{31} & a_{33} & a_{32} & a_{34} & 1 & 0 \\ a_{41} & a_{43} & a_{42} & a_{44} & 0 & 1 \end{bmatrix} \quad \text{and} \quad C(A) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\]
If we set \(a_{31} = a_{44} = \zeta, a_{33} = a_{42} = \zeta^3, a_{32} = a_{43} = \zeta^2 + \zeta, \) and \(a_{34} = a_{41} = 1\), then
\[
A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \zeta & \zeta^2 + \zeta & \zeta^3 & 1 & 0 \\ 1 & \zeta^3 & \zeta^2 + \zeta & \zeta & 1 \end{bmatrix}
\]
and the corresponding matrices \(M(A)\) and \(C(A)\) satisfy all the conditions described in Theorem 4. With respect to the threshold distinguishing advantage \(1/2\), the resulting compression function \(lp^A\) is preimage resistant up to \(2^{81.5}\) queries (with \((d_1, d_2) = (2, 2^{37.5})\)) and collision resistant up to \(2^{56.6}\) queries (with \((d_1, d_2, d_3) = (2, 8, 4)\)).
References