# Two-Permutation-Based Hashing with Binary Mixing

Atul Luykx, Bart Mennink, Bart Preneel, and Laura Winnen

Dept. Electrical Engineering, ESAT/COSIC, KU Leuven, and iMinds, Belgium
{atul.luykx,bart.mennink,bart.preneel}@esat.kuleuven.be

**Abstract.** We consider the generic design of compression functions based on two $n$-bit permutations and XOR-based mixing functions. It is known that any such function mapping $n + \alpha$ to $\alpha$ bits, with $1 \leq \alpha \leq n$, can achieve at most $\min\{2^{\alpha/2}, 2^{n/2-\alpha/4}\}$ collision security. Using techniques similar to Mennink and Preneel (CRYPTO 2012), we show that there is only one equivalence class of these functions achieving optimal collision security, and additionally $\min\{2^{\alpha}, 2^{n/2}\}$ preimage security. The equivalence class compares well with existing functions based on two or three permutations, and is well-suited for wide-pipe hashing.

**Keywords.** hash function, permutation-based, wide-pipe, collision resistance, preimage resistance.

## 1 Introduction

The classical approach to hash function design is blockcipher-based. The first appearance of this idea is the construction $F(h, m) = \mathsf{DES}_m(h)$ by Rabin [8] using the Data Encryption Standard, but many other constructions and further analysis followed; see, for instance, [4, 7] for an extensive discussion of the most basic modes. Yet these designs assume that blockciphers are close to ideal, which in particular means having a strong key schedule.

An increasingly popular approach removes the need for a key schedule by basing hash functions on a limited number of permutations, at the cost of increasing the primitive calls per message block. Concretely, Black et al. [3] demonstrated that a $2n$-to-$n$-bit compression function $F$ using one $n$-bit permutation $\pi$ cannot be secure. A generalization of this result by Rogaway and Steinberger [11] showed that for any $(n + \alpha)$-to-$\alpha$-bit compression function using $r$ $n$-bit permutations, collisions can be found in about $2^{\frac{(r-1)n-\alpha/2}{r}}$ queries and preimages in about $2^{\frac{(r-1)n}{r}}$, provided the function satisfies a "uniformity assumption."

If a $2n$-to-$n$-bit compression function is optimally collision resistant, then Rogaway and Steinberger's bounds imply that its efficiency must be limited to at least three permutation calls per message block. Many three permutation compression functions have been constructed, including those by Rogaway and Steinberger (RS) [10], Shrimpton and Stam (SS) [12], and Mennink and Preneel (MP) [6].

Yet, taking efficiency as priority, it becomes natural to ask what level of security a $2n$-to-$n$-bit compression function can achieve if it only uses *two* permutations. The bounds of Rogaway and Steinberger dictate at most $2^{n/4}$ security,

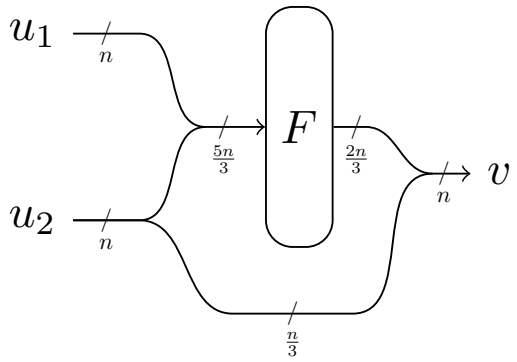


Fig. 1: Expanding a $5n/3$-to-$2n/3$-bit compression function to a $2n$-to-$n$-bit compression function with no security loss.

but Stam [13] showed how a violation of the uniformity assumption can result in a compression function secure *beyond* this bound. In more detail, he observed that increasing $\alpha$ actually *reduces* the security. For example a $5n/3$-to-$2n/3$-bit compression function using two permutations can achieve security up to $2^{n/3}$ queries, whereas a $2n$-to-$n$-bit function can only achieve security up to $2^{n/4}$ queries. In contrast, if one converts the $5n/3$-to-$2n/3$-bit compression function to a $2n$-to-$n$-bit compression function by leaving part of the input unprocessed, as depicted in Figure 1, then no security is lost.

This example suggests that one should fix the number of permutations $r$, and derive $\alpha$ to maximize the security. For $r = 2$, Figure 2 plots the Rogaway-Steinberger bound and the classical birthday bound. It shows that the former bound only improves over the birthday bound for $\alpha \geq 2n/3$. Based on this, Stam proved that if $f_1, f_2$ are uniformly random functions, the compression function

$$F(u_1, u_2) = \mathsf{msb}_{2n/3}\left(f_2(f_1(u_1) \oplus u_2 0^{n/3})\right) \oplus u_2\,, \tag{1}$$

depicted in Figure 3, achieves approximately $2^{n/3}$ security. By forwarding $n/3$ bits of $u_2'$, as in Figure 1, we can expand $F$ to a $2n$-to-$n$-bit function:

$$F'(u_1, u_2') = F(u_1, \mathsf{msb}_{2n/3}(u_2')) \parallel \mathsf{lsb}_{n/3}(u_2')\,.$$

As a matter of fact, Stam generally conjectured that collisions can typically be found in about $2^{\frac{n(r-1)}{r+1}}$ queries, a bound later proven by Steinberger et al. [14, 15]. This bound is intuitively explained by the observation that the Rogaway-Steinberger bound and the birthday bound meet for $\alpha = \frac{2n(r-1)}{r+1}$.

### Our Contribution

We consider the generic design of $(n + \alpha)$-to-$\alpha$-bit compression functions based on two permutations and XOR-based mixing functions. According to Figure 2,

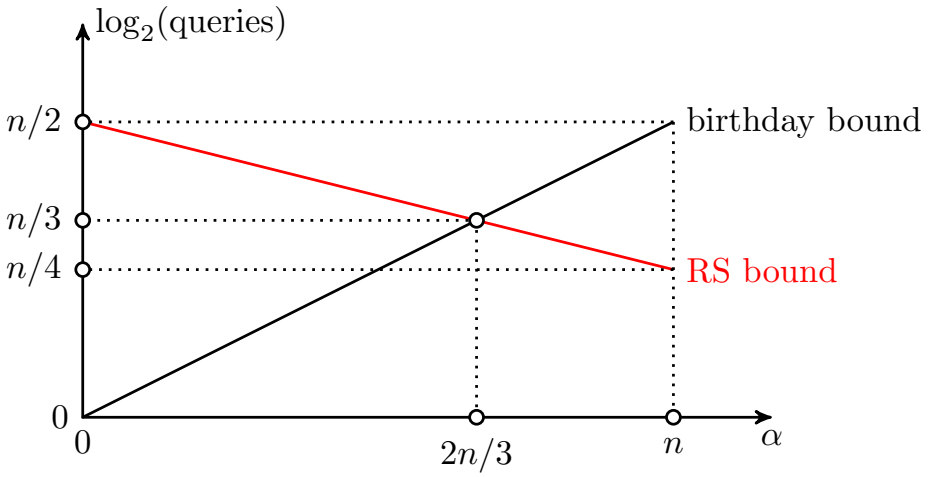


Fig. 2: The Rogaway and Steinberger bound only improves on the birthday bound when $\alpha \geq 2n/3$, for $r = 2$.

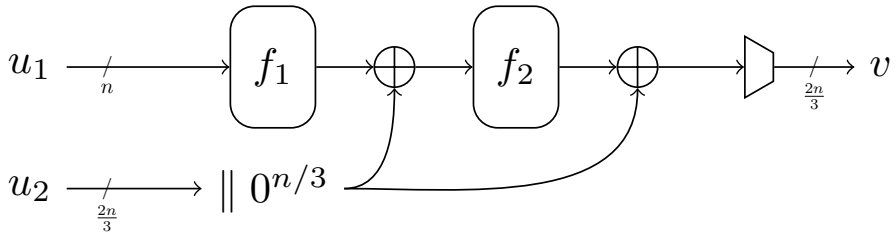


Fig. 3: Stam's scheme achieving $2^{n/3}$ security.

the optimal security for any such function is $2^{\alpha/2}$ for $1 \leq \alpha \leq 2n/3$ and $2^{n/2-\alpha/4}$ for $2n/3 \leq \alpha \leq n$, ignoring the input-forwarding trick, which shows that the bound at $\alpha = 2n/3$ can be extended to larger values of $\alpha$.[1] We consider all possible functions of above-mentioned form (512 in total), and identify the ones that achieve optimal collision security. Our approach is to analyze functions as equivalence classes, similar to Mennink and Preneel [6]. We also inherit some equivalence reductions, but the analysis in this work is more complex than the one in [6] because we do not restrict $\alpha$ to $n$.

We show that out of the 512 functions, there is exactly one class of functions that achieves optimal collision security for $1 \leq \alpha \leq 2n/3$, namely the class

[1] We remark that if the input-forwarding trick were not ignored, the optimal security bound for the region $2n/3 \leq \alpha \leq n$ would be $2^{n/3}$, as explained before.

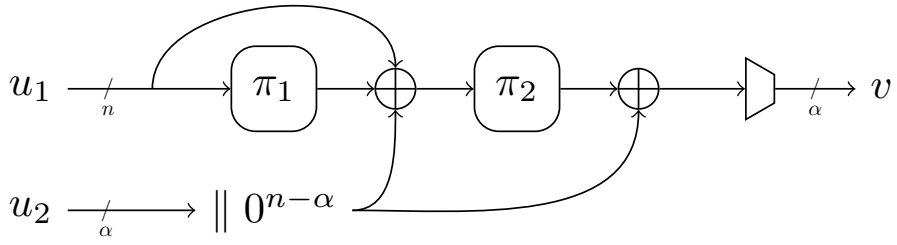


Fig. 4: The class of schemes achieving optimal collision security for $1 \leq \alpha \leq 2n/3$.

Table 1: Collision security comparison of $F_{2n/3}$ with existing permutation-based functions.

| design | input | output | #calls | security | reference |
|---|---|---|---|---|---|
| RS/SS/MP | $2n$ | $n$ | 3 | $n/2$ | [6, 10, 12] |
| Grøstl | $2n$ | $n$ | 2 | $n/4$ | [5] |
| $F_{2n/3}$ | $5n/3$ | $2n/3$ | 2 | $n/3$ | this work |

defined by:

$$F_\alpha(u_1, u_2) = \mathsf{msb}_\alpha\left(\overline{u_2} \oplus \pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right) , \tag{2}$$

where $\overline{u_2} = u_2 \| 0^{n-\alpha} \in \{0,1\}^n$, depicted in Figure 4. Any scheme not equivalent to $F_\alpha$ allows for collisions in about $2^{2(n-\alpha)/3}$ queries or less. In particular, the permutation-based equivalent of (1) — with $f_i(x) = x \oplus \pi_i(x)$ — is not equivalent to $F_\alpha$. Additionally, we show that $F_\alpha$ achieves relatively good preimage resistance, up to about $2^{\min\{\alpha/2, n/2\}}$ queries.

**Comparison**

The approaches of RS [10], SS [12], and MP [6] give $2n$-to-$n$-bit compression functions with $n/2$-bit security based on three permutations. The function $F_{2n/3}$ from equation (2) is a $5n/3$-to-$2n/3$-bit compression function based on two permutations, that achieves $2^{n/3}$ security. We note that $F_{2n/3}$ is particularly suited to be employed in a wide-pipe mode of operation. For example, to design a hash function with output size $m$ bits, one could use $F_{2n/3}$ for $n = 3m/2$ and obtain optimal security using two primitives of size $3m/2$ bits. The same approach is followed by the $2n$-to-$n$-bit Grøstl compression function, which uses two permutations and achieves $n/4$-bit security, resulting in a permutation of size $n = 2m$. These results are summarized in Table 1.

### Outline

The security model is outlined in Section 2. Our generic two-call permutation-based compression function is introduced in Section 3. Then, in Section 4, we synthetically analyze each of these functions.

## 2 Security Model

Throughout, we have $\alpha, n \in \mathbb{N}$ with $1 \leq \alpha \leq n$. By $\{0,1\}^n$ we denote the set of bit strings of length $n$. We denote by $\mathsf{P}(n)$ the set of all permutations on $n$ bits. The concatenation of two bit strings $x$ and $y$ is denoted $x\|y$, and if they are of the same size their bitwise XOR is denoted $x \oplus y$. For $x \in \{0,1\}^\alpha$, we write $\overline{x} = x\|0^{n-\alpha} \in \{0,1\}^n$. For $x \in \{0,1\}^n$, $\mathsf{msb}_\alpha(x)$ denotes the $\alpha$ most significant bits of $x$ and $\mathsf{lsb}_\alpha(x)$ the $\alpha$ least significant bits, in such a way that $x = \mathsf{msb}_\alpha(x)\|\mathsf{lsb}_{n-\alpha}(x)$. For $0 \leq i < 2^n$, by $\langle i \rangle_n$ we denote the encoding of $i$ as an $n$-bit string. For a set $\mathcal{X}$, $x \xleftarrow{\$} \mathcal{X}$ denotes the uniformly random sampling of an element from $\mathcal{X}$. For a matrix A, by $\mathbf{a}_{i,*}$ we denote the $i$th row of A, and by $\mathbf{a}_{*,j}$ its $j$th column.

Let $F : \{0,1\}^m \to \{0,1\}^n$ be a compression function based on 2 permutations. We consider the security of $F$ in the ideal permutation model, where $\pi_1, \pi_2 \xleftarrow{\$} \mathsf{P}(n)$. An adversary $\mathcal{A}$ is a probabilistic algorithm that has oracle access to $\pi_i, \pi_i^{-1}$. It is information-theoretic and its complexity is measured by the number of queries it makes to its oracles. The queries made by $\mathcal{A}$ are stored in a query history $\mathcal{Q}$ as tuples of the form $(\pm, x_k, y_k)$, where $\pi_k(x_k) = y_k$ and the query is made in forward $(+)$ or inverse $(-)$ direction. We denote by $\mathcal{Q}_q$ the history after $q \geq 0$ queries. We require that $\mathcal{Q}_q$ always contains the queries required for the attack, and, without loss of generality, we assume that $\mathcal{A}$ never makes duplicate queries.

We consider the definitions of security and preimage resistance which were also used in [6].

**Definition 1.** *Let $F : \{0,1\}^{n+\alpha} \to \{0,1\}^\alpha$ be a compression function. The advantage of a collision-finding adversary $\mathcal{A}$ for $F$ is defined as*

$$\mathbf{Adv}_F^{\mathrm{col}}(\mathcal{A}) = \mathbf{Pr}\left(\pi_1, \pi_2 \xleftarrow{\$} \mathsf{P}(n),\ u, u' \leftarrow \mathcal{A}^{\pi_i, \pi_i^{-1}} \ :\ u \neq u' \ \wedge\ F(u) = F(u')\right).$$

*By $\mathbf{Adv}_F^{\mathrm{col}}(q)$ we define the maximum advantage taken over all adversaries making $q$ queries.*

For preimage resistance, we use *everywhere* preimage resistance [9], which implies preimage security for every range point.

**Definition 2.** *Let $F : \{0,1\}^{n+\alpha} \to \{0,1\}^\alpha$ be a compression function. The advantage of a preimage-finding adversary $\mathcal{A}$ for $F$ is defined as*

$$\mathbf{Adv}_F^{\mathrm{pre}}(\mathcal{A}) = \max_{v \in \{0,1\}^\alpha} \mathbf{Pr}\left(\pi_1, \pi_2 \xleftarrow{\$} \mathsf{P}(n),\ u \leftarrow \mathcal{A}^{\pi_i, \pi_i^{-1}}(v) \ :\ F(u) = v\right).$$

*By $\mathbf{Adv}_F^{\mathrm{pre}}(q)$ we define the maximum advantage taken over all adversaries making $q$ queries.*

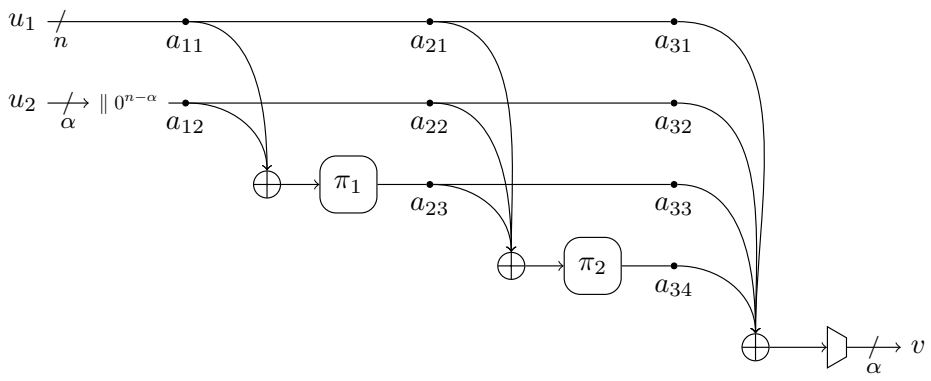


Fig. 5: The permutation-based compression function $F_{\alpha,\mathrm{A}}$ of (4).

## 3 Generic Two-Call Construction

Let $\pi_1, \pi_2 \in \mathsf{P}(n)$ be two $n$-bit permutations. Let $\mathrm{A} \in (\{0,1\})^{3\times 4}$ be a matrix of the form

$$\mathrm{A} = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} . \tag{3}$$

We define the compression function $F_{\alpha,\mathrm{A}} : \{0,1\}^{n+\alpha} \to \{0,1\}^{\alpha}$ as

$$\begin{aligned} F_{\alpha,\mathrm{A}}(u_1,u_2) = v\,, \text{ where } \; & y_1 \leftarrow \pi_1(a_{11}u_1 \oplus a_{12}\overline{u_2})\,, \\ & y_2 \leftarrow \pi_2(a_{21}u_1 \oplus a_{22}\overline{u_2} \oplus a_{23}y_1)\,, \\ & v \leftarrow \mathsf{msb}_{\alpha}\left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus a_{33}y_1 \oplus a_{34}y_2\right)\,, \end{aligned} \tag{4}$$

depicted in Figure 5. The construction is that of Mennink and Preneel [6] restricted to two permutations, but generalized to any output size with the parameter $\alpha$.

## 4 Classification of Secure Functions

We consider all possible compression functions of the form (4), for arbitrary $\alpha, \mathrm{A}$, and derive a classification based on their collision and preimage security guarantees. There are, however, many schemes that are related by simple transformations on the inputs or permutations. We therefore group compression functions together in equivalence classes via security as done by Mennink and Preneel [6].

**Definition 3.** *Two compression functions $F_{\alpha,\mathrm{A}}$ and $F_{\alpha,\mathrm{A}'}$ are* equivalent *if for both collision and preimage security there exist tight reductions between $F_{\alpha,\mathrm{A}}$ and $F_{\alpha,\mathrm{A}'}$. Formally, they are equivalent if there is a small constant $c \in \mathbb{N}$ such that $\mathbf{Adv}^{\mathrm{col/pre}}_{F_{\alpha,\mathrm{A}}}(q) \leq \mathbf{Adv}^{\mathrm{col/pre}}_{F_{\alpha,\mathrm{A}'}}(q+c)$ and $\mathbf{Adv}^{\mathrm{col/pre}}_{F_{\alpha,\mathrm{A}'}}(q) \leq \mathbf{Adv}^{\mathrm{col/pre}}_{F_{\alpha,\mathrm{A}}}(q+c)$.*

The definition is similar to random-oracle reducibility [1, 2]. Using Definition 3, we derive our main result.

**Theorem 1.** *Let $F_{\alpha,\mathrm{A}}$ be as in (4). If $F_{\alpha,\mathrm{A}}$ is equivalent to $F_{\alpha,\mathrm{B}}$ with* $\mathrm{B} = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{smallmatrix}\right)$*, then*

$$\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(q) \leq \begin{cases} \dfrac{nq^3}{2^n} + \dfrac{n^2q^2}{2^\alpha} + \left(\dfrac{4eq}{n2^{n-\alpha}}\right)^n & \text{if } \alpha \leq 2n/3\,, \\ \dfrac{nq^3}{2^{3n/2-3\alpha/4}} + \dfrac{n^2q^2}{2^{n-\alpha/2}} + \left(\dfrac{4eq}{n2^{n/2-\alpha/4}}\right)^n & \text{if } \alpha \geq 2n/3\,. \end{cases}$$

*Otherwise,*

$$\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(2q) \geq q\binom{q}{2}/2^{2(n-\alpha)}\,.$$

The theorem states that there is only *one* class of functions of the form (4) achieving optimal collision security (cf. Figure 2). Remarkably, the permutation-based equivalent of Stam's construction (1) is not equivalent to $F_{\alpha,\mathrm{B}}$, but $F_{\alpha,\mathrm{C}}$ (where C is described in the theorem's proof), and collisions can be found in roughly $2^{2(n-\alpha)/3}$ queries.

We remark that the lower bound of Theorem 1 for $F_{\alpha,\mathrm{A}}$ being in-equivalent to $F_{\alpha,\mathrm{B}}$ is not necessarily tight for all $\alpha$. Indeed, for $\alpha < 4n/7$ it is worse than the trivial $2^{\alpha/2}$ security bound. It may be possible that any function $F_{\alpha,\mathrm{A}}$ that is not equivalent to $F_{\alpha,\mathrm{B}}$ may still achieve optimal security for some values $\alpha$.

We also prove that $F_{\alpha,\mathrm{B}}$ achieves good preimage security.

**Theorem 2.** *Let $F_{\alpha,\mathrm{A}}$ be as in (4). If $F_{\alpha,\mathrm{A}}$ is equivalent to $F_{\alpha,\mathrm{B}}$ with* $\mathrm{B} = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{smallmatrix}\right)$*, then*

$$\mathbf{Adv}^{\mathrm{pre}}_{F_{\alpha,\mathrm{A}}}(q) \leq \begin{cases} \dfrac{q^2}{2^n} + \dfrac{2nq}{2^\alpha} + \left(\dfrac{4eq}{n2^{n-\alpha}}\right)^n & \text{if } \alpha \leq n/2\,, \\ \dfrac{q^2}{2^n} + \dfrac{2nq}{2^{n/2}} + \left(\dfrac{4eq}{n2^{n/2}}\right)^n & \text{if } \alpha \geq n/2\,. \end{cases}$$

The proofs start by considering any possible matrix A, leaving 512 schemes to be analyzed per $\alpha$. In Section 4.1 three reductions on permutation-based compression functions from [6] are described. Then, in Section 4.2 we derive attacks on various schemes and reduce the number of classes. This analysis leaves one scheme, $F_{\alpha,\mathrm{B}}$. In Section 4.3, this remaining scheme is proven collision and preimage secure.

### 4.1 Equivalence Reductions

Four reductions were introduced in [6]: the $x$-, $\oplus$-, $\pi$-swap-, and $\pi$-inverse-reduction. Since one input is padded with 0's in $F_{\alpha,\mathrm{A}}$, the $x$-reduction is not applicable and the other reductions are only applicable to the input $u_1$. We introduce the $\oplus$-reduction, $\pi$-swap-reduction, and $\pi$-inverse-reduction to our setting. The proofs of [6] apply to the reductions in the current setting.

**Proposition 1 ($\oplus$-reduction).** *Consider a matrix* $\mathrm{A} = \left(\mathbf{a}_{*,1}\,;\,\mathbf{a}_{*,2}\,;\,\mathbf{a}_{*,3}\,;\,\mathbf{a}_{*,4}\right)$*, and let* $k$ *be the row number of the first non-zero coefficient of* $u_1$*, that is* $k = \min\{\, i \mid a_{i1} \neq 0 \,\}$*. Let* $c_1, c_2, c_3 \in \{0,1\}$*. Consider the matrix*

$$\mathrm{A}' = \mathrm{A} \oplus \left(\mathbf{0}\,;\, c_1\mathbf{a}_{*,1}\,;\, [k \geq 2]c_2\mathbf{a}_{*,1}\,;\, [k \geq 3]c_3\mathbf{a}_{*,1}\,;\, \mathbf{0}\right),$$

*where* $[X] = 1$ *if* $X$ *holds and* $0$ *otherwise. Then, the compression functions* $F_{\alpha,\mathrm{A}}$ *and* $F_{\alpha,\mathrm{A}'}$ *are equivalent.*

The $\pi$-swap-reduction corresponds to swapping the roles of $\pi_1$ and $\pi_2$. This is only possible if $\pi_1$'s output is input to $\pi_2$.

**Proposition 2 ($\pi$-swap-reduction).** *Consider a matrix* $\mathrm{A}$ *with* $a_{23} = 0$*. Consider the matrix* $\mathrm{A}'$ *obtained from* $\mathrm{A}$ *by swapping rows* $\mathbf{a}_{1,*}$ *and* $\mathbf{a}_{2,*}$ *and swapping columns* $\mathbf{a}_{*,3}$ *and* $\mathbf{a}_{*,4}$*. Then, the compression functions* $F_{\alpha,\mathrm{A}}$ *and* $F_{\alpha,\mathrm{A}'}$ *are equivalent.*

The $\pi$-inverse-reduction corresponds to replacing $\pi_1$ by $\pi_1^{-1}$, which can only be done if the input to $\pi_1$ is independent of $u_2$.

**Proposition 3 ($\pi$-inverse-reduction).** *Consider a matrix* $\mathrm{A}$ *with* $(a_{11}, a_{12}) = (1, 0)$*. Consider the matrix* $\mathrm{A}'$ *obtained from* $\mathrm{A}$ *by swapping* $(a_{21}, a_{31})$ *and* $(a_{23}, a_{33})$*. Then, the compression functions* $F_{\alpha,\mathrm{A}}$ *and* $F_{\alpha,\mathrm{A}'}$ *are equivalent.*

The $\oplus$-reduction corresponds to replacing $u_1$ by $u_1 \oplus c_1\overline{u_2}$ for $c_1 \in \{0,1\}$. If the first permutation, $\pi_1$, does not have $u_1$ as input, i.e. $a_{11} = 0$, then the $\oplus$-reduction corresponds to replacing $u_1$ by $u_1 \oplus c_1\overline{u_2} \oplus c_2 y_1$ for $c_1, c_2 \in \{0,1\}$, where $y_1$ is $\pi_1$'s output. The case where additionally $\pi_2$ does not take $u_1$ as input is similar. Note that the $\oplus$-reduction is asymmetric, as swapping the roles of $u_1$ and $u_2$ may make the reduction incompatible with the definition of $F_{\alpha,\mathrm{A}}$.

### 4.2 Elimination of Insecure Classes

Our aim is to classify the security of $F_{\alpha,\mathrm{A}}$ for various A, but for some choices the scheme is trivially insecure. As a first step, we "rule out" matrices for which the resulting construction can be generically attacked in a constant number of queries.

**Lemma 1.** *If*

- *the matrix* $\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$ *is non-invertible, or*
- *the third or fourth column of* $\mathrm{A}$ *is zero,*

*then* $\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(4) = 1$.

*Proof.* Firstly, note that $\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ is equivalent to $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ because of an $\oplus$-reduction. Therefore, we assume A satisfies one of the following properties:

(i) $a_{11} = a_{21} = 0$ or $a_{12} = a_{22} = 0$;
(ii) $a_{11} = a_{12} = 0$ or $a_{21} = a_{22} = 0$;
(iii) $a_{23} = a_{33} = 0$ or $a_{34} = 0$.

If $a_{11} = a_{21} = 0$, then $u_1$ is not used as input to $\pi_1$ or $\pi_2$, and a collision can be found in at most 2 queries: $u_1, u_2, u_2'$ are chosen arbitrarily such that $u_2 \neq u_2'$ (if $a_{31} \neq 0$) or such that $u_2 = u_2'$ (if $a_{31} = 0$), and $u_1$ is adapted to satisfy the collision. Similar analysis holds for the case $a_{12} = a_{22} = 0$.

Assume A does not satisfy property (i), and consider (ii). If $a_{11} = a_{12} = 0$, then $\pi_1$ is evaluated on 0 and the construction is a function based on 1 permutation,

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha\left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus a_{33}\pi_1(0) \oplus a_{34}y_2\right) ,$$

where $y_2 = \pi_2(u_1 \oplus \overline{u_2} \oplus a_{23}\pi_1(0))$. By an $\oplus$-reduction, it is equivalent to

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha\left(a_{31}u_1 \oplus (a_{31} \oplus a_{32})\overline{u_2} \oplus a_{33}\pi_1(0) \oplus a_{34}y_2\right) ,$$

where $y_2 = \pi_2(u_1 \oplus a_{23}\pi_1(0))$. Since $a_{21} = 0$ and $a_{22} = 0$, the function satisfies property (i), and a collision can be found easily. For the second case of property (ii), when $a_{21} = a_{22} = 0$, if $a_{23} = 0$ we can apply similar reasoning as when $a_{11} = a_{12} = 0$. If $a_{23} = 1$ we have

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha\left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus a_{33}y_1 \oplus a_{34}\pi_2(y_1)\right) ,$$

where $y_1 = \pi_1(u_1 \oplus \overline{u_2})$. The attack is the same as before with $a_{34}y_2$ replaced by $(a_{33}id \oplus a_{34}\pi_2)(y_1)$.

Finally, the analysis of property (iii) is the same as for (ii). □

Using this lemma, we apply reductions to show that it suffices to consider matrices with first row $(1, 0, 0, 0)$.

**Lemma 2.** *Any compression function $F_{\alpha,\mathrm{A}}$ can be reduced to a compression function $F_{\alpha,\mathrm{A}'}$ where the matrix $\mathrm{A}'$ satisfies one of the properties of Lemma 1, or has $a'_{11}a'_{12} = 10$.*

*Proof.* Assume $a_{11}a_{12} \neq 10$. If $a_{11}a_{12} = 00$, Lemma 1 applies, and the claim holds. If $a_{11}a_{12} = 11$, we perform an $\oplus$-reduction to find $\mathrm{A}'$ with $a'_{11}a'_{12} = 10$. Finally, the case $a_{11}a_{12} = 01$ is a bit harder to analyze. Note that by the conditions of Lemma 1, $a_{21} = 1$. The second row thus satisfies $a_{21}a_{22}a_{23} \in \{100, 101, 110, 111\}$, all of which are equivalent to the first element in the set by the $\oplus$-reduction. In other words, A is equivalent to $\mathrm{A}'$ with $a'_{11}a'_{12} = 01$ and $a'_{21}a'_{22}a'_{23} = 100$. A $\pi$-swap-reduction gives the required result. □

Lemmas 1 and 2 imply that it suffices to consider matrices of the form

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 1 \end{pmatrix} ,$$

where $a_{23} + a_{33} \geq 1$. We continue with the second row.

**Lemma 3.** *If $a_{21} = 0$ or $a_{23} = 0$, then $\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(2q) \geq \binom{q}{2}^2 / 2^\alpha$.*

*Proof.* Both cases are equivalent by the $\pi$-inverse-reduction, and we focus on the case $a_{23} = 0$. By Lemma 1, we can assume that $a_{33} = 1$. In other words,

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha \left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus \pi_1(u_1) \oplus \pi_2(a_{21}u_1 \oplus \overline{u_2})\right) .$$

Collision-finding adversary $\mathcal{A}$ proceeds as follows. For $i = 1, \ldots, q$, it queries $x_1^{(i)} = \langle i \rangle_\alpha 0^{n-\alpha}$ to $\pi_1$ to obtain $y_1^{(i)}$. Similarly, for $j = 1, \ldots, q$, it queries $x_2^{(j)} = \langle j \rangle_\alpha 0^{n-\alpha}$ to $\pi_2$ to obtain $y_2^{(j)}$. A collision for $F_{\alpha,\mathrm{A}}$ is found if for some $i, i', j, j'$,

$$\begin{aligned} &\mathsf{msb}_\alpha \left(a_{31}x_1^{(i)} \oplus a_{32}(x_2^{(j)} \oplus a_{21}x_1^{(i)}) \oplus y_1^{(i)} \oplus y_2^{(j)}\right) \\ = &\mathsf{msb}_\alpha \left(a_{31}x_1^{(i')} \oplus a_{32}(x_2^{(j')} \oplus a_{21}x_1^{(i')}) \oplus y_1^{(i')} \oplus y_2^{(j')}\right) . \end{aligned}$$

This is a generalized birthday problem, and any such collision happens with probability at least $\binom{q}{2}^2/2^\alpha$. □

By Lemma 3, it suffices to consider

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 1 \end{pmatrix} .$$

These are eight matrices (or in fact six due to the $\pi$-inverse-swap, but we refrain from using that one right now).

**Lemma 4.** *If* $a_{31}a_{32}a_{33} \in \{000, 100, 110, 001, 011, 111\}$*, then* $\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(q) \geq \binom{q}{2}/2^{n-\alpha}$.

*Proof.* Note that

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha \left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus a_{33}\pi_1(u_1) \oplus \pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right) .$$

We consider a collision-finding adversary $\mathcal{A}$ that aims at finding $(u_1, u_2)$ and $(u_1', u_2')$ that collide on the input to $\pi_2$ *as well as* in the remaining term, or in other words such that

$$\begin{pmatrix} u_1 \oplus \overline{u_2} \oplus \pi_1(u_1) \\ \mathsf{msb}_\alpha \left(a_{31}u_1 \oplus a_{32}\overline{u_2} \oplus a_{33}\pi_1(u_1)\right) \end{pmatrix} = \begin{pmatrix} u_1' \oplus \overline{u_2'} \oplus \pi_1(u_1') \\ \mathsf{msb}_\alpha \left(a_{31}u_1' \oplus a_{32}\overline{u_2'} \oplus a_{33}\pi_1(u_1')\right) \end{pmatrix} .$$

Adding the first to the second row and/or using the $\pi$-inverse-reduction, it suffices to consider the cases $a_{31}a_{32}a_{33} \in \{000, 100\}$, for which

$$F_{\alpha,\mathrm{A}}(u_1, u_2) = \mathsf{msb}_\alpha \left(a_{31}u_1 \oplus \pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right) .$$

Adversary $\mathcal{A}$ proceeds as follows. For $i = 1, \ldots, q$, it queries $x_1^{(i)} = 0^\alpha \langle i \rangle_{n-\alpha}$ to $\pi_1$ to obtain $y_1^{(i)}$. A collision for $F_{\alpha,\mathrm{A}}$ is found if for some $i, i'$,

$$\mathsf{lsb}_{n-\alpha} \left(x_1^{(i)} \oplus y_1^{(i)}\right) = \mathsf{lsb}_{n-\alpha} \left(x_1^{(i')} \oplus y_1^{(i')}\right) ,$$

as in this case, we put $u_1 = x_1^{(i)}$, $u_2 = \mathsf{msb}_\alpha \left(x_1^{(i)} \oplus y_1^{(i)}\right)$, and similarly for $u_1', u_2'$. Any such collision happens with probability at least $\binom{q}{2}/2^{n-\alpha}$. □

We note that for $\alpha < n/2$, the trivial birthday bound of $\binom{q}{2}/2^\alpha$ is tighter than the one of Lemma 4. We are eventually left with two matrices:

$$\mathrm{B} = \begin{pmatrix} 1\,0\,0\,0 \\ 1\,1\,1\,0 \\ 0\,1\,0\,1 \end{pmatrix}, \qquad \mathrm{C} = \begin{pmatrix} 1\,0\,0\,0 \\ 1\,1\,1\,0 \\ 1\,0\,1\,1 \end{pmatrix}.$$

Notice that the permutation-based equivalent of Stam's construction, or in more detail (1) with $f_1(x) = x \oplus \pi_1(x)$ and $f_2(x) = x \oplus \pi_2(x)$, is equivalent to C. It turns out that in the permutation-based setting, this scheme does not achieve optimal security.

**Lemma 5.** *We have* $\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{C}}}(2q) \geq q\binom{q}{2}/2^{2(n-\alpha)}$.

*Proof.* Note that

$$F_{\alpha,\mathrm{C}}(u_1, u_2) = \mathsf{msb}_\alpha\left(u_1 \oplus \pi_1(u_1) \oplus \pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right).$$

We consider a collision-finding adversary $\mathcal{A}$ that restricts itself to $u_1 = u_1'$. In other words, the goal is to find $u_1, u_2, u_2'$ such that

$$\mathsf{msb}_\alpha\left(\pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right) = \mathsf{msb}_\alpha\left(\pi_2(u_1 \oplus \overline{u_2'} \oplus \pi_1(u_1))\right).$$

For $i = 1, \ldots, q$, it queries $x_1^{(i)} = \langle i \rangle_n$ to $\pi_1$ to obtain $y_1^{(i)}$. For $j = 1, \ldots, q$, it queries $y_2^{(j)} = 0^\alpha \langle j \rangle_{n-\alpha}$ to $\pi_2^{-1}$ to obtain $x_2^{(j)}$. A collision for $F_{\alpha,\mathrm{A}}$ is found if for some $i, j, j'$,

$$\mathsf{lsb}_{n-\alpha}\left(x_1^{(i)} \oplus y_1^{(i)}\right) = \mathsf{lsb}_{n-\alpha}\left(x_2^{(j)}\right) = \mathsf{lsb}_{n-\alpha}\left(x_2^{(j')}\right),$$

as in this case, we put $u_1 = u_1' = x_1^{(i)}$, $u_2 = \mathsf{msb}_\alpha\left(x_1^{(i)} \oplus y_1^{(i)} \oplus x_2^{(j)}\right)$, and similarly for $u_2'$. Any such collision happens with probability at least $q\binom{q}{2}/2^{2(n-\alpha)}$. □

Again, for $\alpha < 4n/7$, the trivial birthday bound of $\binom{q}{2}/2^\alpha$ is tighter than the one of Lemma 5.

### 4.3 Collision and Preimage Security of $F_{\alpha,\mathrm{B}}$

The only surviving scheme is (see also Figure 4):

$$F_{\alpha,\mathrm{B}}(u_1, u_2) = \mathsf{msb}_\alpha\left(\overline{u_2} \oplus \pi_2(u_1 \oplus \overline{u_2} \oplus \pi_1(u_1))\right).$$

We prove that it indeed achieves optimal security.

**Collision Resistance (Theorem 1)**

Finding a collision for $F_{\alpha,\mathrm{B}}$ corresponds to finding queries $(\pm, x_1, y_1)$, $(\pm, x_1', y_1')$, $(\pm, x_2, y_2)$, and $(\pm, x_2', y_2')$ such that

$$(x_1, x_2) \neq (x_1', x_2')\,, \tag{5a}$$
$$\mathsf{lsb}_{n-\alpha}\left(x_1 \oplus y_1 \oplus x_2\right) = 0\,, \tag{5b}$$
$$\mathsf{lsb}_{n-\alpha}\left(x_1' \oplus y_1' \oplus x_2'\right) = 0\,, \tag{5c}$$
$$\mathsf{msb}_{\alpha}\left(x_1 \oplus y_1 \oplus x_2 \oplus y_2\right) = \mathsf{msb}_{\alpha}\left(x_1' \oplus y_1' \oplus x_2' \oplus y_2'\right)\,. \tag{5d}$$

Indeed, in this case we have

$$F_{\alpha,\mathrm{B}}(x_1, \mathsf{msb}_\alpha(x_1 \oplus y_1 \oplus x_2)) = F_{\alpha,\mathrm{B}}(x_1', \mathsf{msb}_\alpha(x_1' \oplus y_1' \oplus x_2'))\,.$$

We define the event that $\mathcal{A}$ finds a solution to (5) by $\mathsf{col}(\mathcal{Q}_q)$. We make a further distinction depending on whether $x_1$ equals $x_1'$ or not, and whether $x_2$ equals $x_2'$ or not. Clearly, the case $(x_1, x_2) = (x_1', x_2')$ violates (5a) and can be omitted. In other words,

$$\begin{aligned}\mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q)\right) \leq\ & \mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 = x_1' \wedge x_2 \neq x_2'\right) + \\ & \mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 \neq x_1' \wedge x_2 = x_2'\right) + \\ & \mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 \neq x_1' \wedge x_2 \neq x_2'\right)\,,\end{aligned}$$

and these probabilities are analyzed in Lemmas 6-8.

**Lemma 6.** *We have* $\mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 = x_1' \wedge x_2 \neq x_2'\right) \leq \binom{q}{2}\frac{2^{n-\alpha}}{2^n - q}$.

*Proof.* Finding a solution to (5) is at least as hard as finding a solution to

$$\mathsf{msb}_{\alpha}\left(x_2 \oplus y_2\right) = \mathsf{msb}_{\alpha}\left(x_2' \oplus y_2'\right)\,.$$

Note that for any query $\mathcal{A}$ makes to $\pi_2$, the response is randomly drawn from a set of size at least $2^n - q$. Hence, after $q$ queries to $\pi_2$, it finds a solution to this equation with probability at most $\binom{q}{2}\frac{2^{n-\alpha}}{2^n - q}$. □

**Lemma 7.** *We have* $\mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 \neq x_1' \wedge x_2 = x_2'\right) \leq \binom{q}{2}\frac{2^{n-\alpha}}{2^n - q}$.

*Proof.* The proof is identical to the one of Lemma 6, and henceforth omitted. □

**Lemma 8.** *Let* $\tau \geq n$ *be an integral threshold. We have*

$$\mathbf{Pr}\left(\mathsf{col}(\mathcal{Q}_q) \wedge x_1 \neq x_1' \wedge x_2 \neq x_2'\right) \leq \frac{(\tau-1)q^3}{2(2^n - q)} + \frac{2^{n-\alpha}(\tau-1)^2 q^2}{2(2^n - q)} + \left(\frac{2e2^\alpha q}{\tau(2^n - q)}\right)^\tau\,.$$

*Proof.* Write the event as $\mathsf{col}_{\neq,\neq}(\mathcal{Q}_q)$. Denote by $\mathsf{X}(\mathcal{Q}_q)$ the event

$$\mathsf{X}(\mathcal{Q}_q):\ \max_{v \in \{0,1\}^{n-\alpha}} \left|\left\{(\pm, x_k, y_k) \in \mathcal{Q}_q \;\middle|\; \mathsf{lsb}_{n-\alpha}(x_k \oplus y_k) = v\right\}\right| \geq \tau\,.$$

Then, by basic probability theory,

$$\mathbf{Pr}\left(\mathsf{col}_{\neq,\neq}(\mathcal{Q}_q)\right) \leq \mathbf{Pr}\left(\mathsf{col}_{\neq,\neq}(\mathcal{Q}_q) \mid \neg\mathsf{X}(\mathcal{Q}_q)\right) + \mathbf{Pr}\left(\mathsf{X}(\mathcal{Q}_q)\right) . \tag{6}$$

We start with $\mathbf{Pr}\left(\mathsf{X}(\mathcal{Q}_q)\right)$. Let $v \in \{0,1\}^{n-\alpha}$. Any query $(\pm, x_k, y_k)$ to $\pi_k$ satisfies $\mathsf{lsb}_{n-\alpha}(x_k \oplus y_k) = v$ with probability at most $\frac{2^\alpha}{2^n-q}$. At least $\tau$ solutions to this equation are found with probability at most $\binom{q}{\tau}\left(\frac{2^\alpha}{2^n-q}\right)^\tau$. This term is at most $\left(\frac{e2^\alpha q}{\tau(2^n-q)}\right)^\tau$ by Stirling's approximation. Considering any choice of $v$, we find

$$\mathbf{Pr}\left(\mathsf{X}(\mathcal{Q}_q)\right) \leq 2^{n-\alpha}\left(\frac{e2^\alpha q}{\tau(2^n-q)}\right)^\tau \leq \left(\frac{2e2^\alpha q}{\tau(2^n-q)}\right)^\tau ,$$

as $\tau \geq n$.

We proceed with the first probability of (6). We run over all queries $i = 1, \ldots, q$, and consider the probability the $i$th query $(\pm, x_k, y_k)$ causes $\mathsf{col}_{\neq,\neq}(\mathcal{Q}_i)$ given that it does not set $\mathsf{X}(\mathcal{Q}_i)$ (and no earlier query set $\mathsf{X}(\mathcal{Q}_j)$ for $j < i$).

**Case: $i$th query is a forward or inverse query to $\pi_1$.** The cases are equivalent by symmetry, consider a forward query $(+, x_1, y_1)$. There are at most $i-1$ choices for both $(\pm, x_2, y_2)$ and $(\pm, x_2', y_2')$. By $\neg\mathsf{X}(\mathcal{Q}_i)$, there are at most $\tau - 1$ choices for $(\pm, x_1', y_1')$ to satisfy (5c). For any such combination of choices, the new query makes (5) satisfied if

$$x_1 \oplus y_1 = \mathsf{msb}_\alpha\left(x_2 \oplus y_2 \oplus x_1' \oplus y_1' \oplus x_2' \oplus y_2'\right) \parallel \mathsf{lsb}_{n-\alpha}\left(x_2\right) .$$

This happens with probability at most $\frac{(\tau-1)(i-1)^2}{2^n-q}$.

**Case: $i$th query is a forward query to $\pi_2$.** Consider a forward query $(+, x_2, y_2)$. There are at most $i-1$ choices for $(\pm, x_2', y_2')$. By $\neg\mathsf{X}(\mathcal{Q}_i)$, there are at most $\tau - 1$ choices for $(\pm, x_1, y_1)$ to satisfy (5b) and for $(\pm, x_1', y_1')$ to satisfy (5c). For any such combination of choices, the new query makes (5) satisfied if

$$\mathsf{msb}_\alpha(y_2) = \mathsf{msb}_\alpha\left(x_1 \oplus y_1 \oplus x_2 \oplus x_1' \oplus y_1' \oplus x_2' \oplus y_2'\right) .$$

This happens with probability at most $\frac{2^{n-\alpha}(\tau-1)^2(i-1)}{2^n-q}$.

**Case: $i$th query is an inverse query to $\pi_2$.** Consider an inverse query $(-, x_2, y_2)$. There are at most $i-1$ choices for both $(\pm, x_1, y_1)$ and $(\pm, x_2', y_2')$. By $\neg\mathsf{X}(\mathcal{Q}_i)$, there are at most $\tau - 1$ choices for $(\pm, x_1', y_1')$ to satisfy (5c).[2] For any such combination of choices, the new query makes (5) satisfied if

$$x_2 = \mathsf{msb}_\alpha\left(x_1 \oplus y_1 \oplus y_2 \oplus x_1' \oplus y_1' \oplus x_2' \oplus y_2'\right) \parallel \mathsf{lsb}_{n-\alpha}\left(x_1 \oplus y_1\right) .$$

This happens with probability at most $\frac{(\tau-1)(i-1)^2}{2^n-q}$.

[2] As the value $x_2$ (cf. (5b)) is not yet fixed, we cannot rely on $\neg\mathsf{X}(\mathcal{Q}_i)$ to claim that there are at most $\tau - 1$ choices for $(\pm, x_1, y_1)$.

Taking the maximum over all possible queries, we obtain

$$\begin{aligned}\mathbf{Pr}\left(\mathsf{col}_{\neq,\neq}(\mathcal{Q}_q) \mid \neg\mathsf{X}(\mathcal{Q}_q)\right) &\leq \sum_{i=1}^{q}\left[\frac{(\tau-1)(i-1)^2}{2^n-q} + \frac{2^{n-\alpha}(\tau-1)^2(i-1)}{2^n-q}\right] \\ &\leq \frac{(\tau-1)q^3}{2(2^n-q)} + \frac{2^{n-\alpha}(\tau-1)^2q^2}{2(2^n-q)}\,.\end{aligned}$$

The lemma is completed via (6). □

Combining Lemmas 6-8, we see that $\mathcal{A}$ finds a collision for $F_{\alpha,\mathrm{B}}$ with probability at most

$$\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(q) \leq 2\binom{q}{2}\frac{2^{n-\alpha}}{2^n-q} + \frac{(\tau-1)q^3}{2(2^n-q)} + \frac{2^{n-\alpha}(\tau-1)^2q^2}{2(2^n-q)} + \left(\frac{2e2^\alpha q}{\tau(2^n-q)}\right)^\tau\,.$$

Recall that we require $\tau \geq n$. Observing that $2^n - q \geq 2^{n-1}$ for $q \leq 2^{n-1}$, we find

$$\mathbf{Adv}^{\mathrm{col}}_{F_{\alpha,\mathrm{A}}}(q) \leq \frac{\tau q^3}{2^n} + \frac{\tau^2 q^2}{2^\alpha} + \left(\frac{4eq}{\tau 2^{n-\alpha}}\right)^\tau\,.$$

The proof of Theorem 1 is completed by putting $\tau = n \cdot \max\{1, 2^{3\alpha/4-n/2}\}$.

**Preimage Resistance (Theorem 2)**

Let $v \in \{0,1\}^\alpha$. Finding a preimage for $F_{\alpha,\mathrm{B}}$ corresponds to finding queries $(\pm, x_1, y_1)$, $(\pm, x_2, y_2)$ such that

$$\mathsf{lsb}_{n-\alpha}\left(x_1 \oplus y_1 \oplus x_2\right) = 0\,, \tag{7a}$$

$$\mathsf{msb}_{\alpha}\left(x_1 \oplus y_1 \oplus x_2 \oplus y_2\right) = v\,. \tag{7b}$$

Indeed, in this case we have

$$F_{\alpha,\mathrm{B}}(x_1, \mathsf{msb}_\alpha(x_1 \oplus y_1 \oplus x_2)) = v\,.$$

We define the event that $\mathcal{A}$ finds a solution to (7) by $\mathsf{pre}(\mathcal{Q}_q)$.

**Lemma 9.** *Let $\tau \geq n$ be an integral threshold. We have* $\mathbf{Pr}\left(\mathsf{pre}(\mathcal{Q}_q)\right) \leq \frac{q^2}{2(2^n-q)} + \frac{2^{n-\alpha}(\tau-1)q}{2^n-q} + \left(\frac{2e2^\alpha q}{\tau(2^n-q)}\right)^\tau$.

*Proof.* We inherit $\mathsf{X}(\mathcal{Q}_q)$ from Lemma 8. By basic probability theory,

$$\mathbf{Pr}\left(\mathsf{pre}(\mathcal{Q}_q)\right) \leq \mathbf{Pr}\left(\mathsf{pre}(\mathcal{Q}_q) \mid \neg\mathsf{X}(\mathcal{Q}_q)\right) + \left(\frac{2e2^\alpha q}{\tau(2^n-q)}\right)^\tau\,. \tag{8}$$

We proceed with the remaining probability, the same way as in Lemma 8.

**Case: $i$th query is a forward or inverse query to $\pi_1$.** The cases are equivalent by symmetry, consider a forward query $(+, x_1, y_1)$. There are at most $i-1$ choices for $(\pm, x_2, y_2)$. For any such choice, the new query makes (7) satisfied if

$$x_1 \oplus y_1 = \left(\mathsf{msb}_\alpha\left(x_2 \oplus y_2\right) \oplus v\right) \parallel \mathsf{lsb}_{n-\alpha}\left(x_2\right)\,.$$

This happens with probability at most $\frac{i-1}{2^n-q}$.

**Case: $i$th query is a forward query to $\pi_2$.** Consider a forward query $(+, x_2, y_2)$. By $\neg\mathsf{X}(\mathcal{Q}_i)$, there are at most $\tau - 1$ choices for $(\pm, x_1, y_1)$ to satisfy (7a). For any such choice, the new query makes (7) satisfied if

$$\mathsf{msb}_\alpha(y_2) = \mathsf{msb}_\alpha\left(x_1 \oplus y_1 \oplus x_2\right) \oplus v\,.$$

This happens with probability at most $\frac{2^{n-\alpha}(\tau-1)}{2^n-q}$.

**Case: $i$th query is an inverse query to $\pi_2$.** Consider an inverse query $(-, x_2, y_2)$. There are at most $i - 1$ choices for $(\pm, x_1, y_1)$. For any such choice, the new query makes (7) satisfied if

$$x_2 = \left(\mathsf{msb}_\alpha\left(x_1 \oplus y_1 \oplus y_2\right) \oplus v\right) \parallel \mathsf{lsb}_{n-\alpha}\left(x_1 \oplus y_1\right)\,.$$

This happens with probability at most $\frac{i-1}{2^n-q}$.

Taking the maximum over all possible queries, we obtain

$$\begin{aligned}\mathbf{Pr}\left(\mathsf{pre}(\mathcal{Q}_q) \mid \neg\mathsf{X}(\mathcal{Q}_q)\right) &\leq \sum_{i=1}^{q}\left[\frac{i-1}{2^n-q} + \frac{2^{n-\alpha}(\tau-1)}{2^n-q}\right] \\ &\leq \frac{q^2}{2(2^n-q)} + \frac{2^{n-\alpha}(\tau-1)q}{2^n-q}\,.\end{aligned}$$

The lemma is completed via (8). □

Observing that $2^n - q \geq 2^{n-1}$ for $q \leq 2^{n-1}$, we find

$$\mathbf{Adv}^{\mathrm{pre}}_{F_{\alpha,\mathrm{A}}}(q) \leq \frac{q^2}{2^n} + \frac{2\tau q}{2^\alpha} + \left(\frac{4eq}{\tau 2^{n-\alpha}}\right)^\tau\,.$$

Recall that we require $\tau \geq n$. The proof of Theorem 2 is completed by putting $\tau = n \cdot \max\{1, 2^{\alpha-n/2}\}$.

Acknowledgments. This work was supported in part by the Research Council KU Leuven: GOA TENSE (GOA/11/007). Atul Luykx is supported by a Ph.D. Fellowship from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). Bart Mennink is a Postdoctoral Fellow of the Research Foundation – Flanders (FWO). The authors would like to thank the anonymous reviewers of the Journal of Mathematical Cryptology for their comments and suggestions.

## References

1. Baecher, P., Farshim, P., Fischlin, M., Stam, M.: Ideal-cipher (ir)reducibility for blockcipher-based hash functions. In: Advances in Cryptology - EUROCRYPT 2013. Lecture Notes in Computer Science, vol. 7881, pp. 426–443. Springer, Heidelberg (2013)

2. Baecher, P., Fischlin, M.: Random oracle reducibility. In: Advances in Cryptology - CRYPTO 2011. Lecture Notes in Computer Science, vol. 6841, pp. 21–38. Springer, Heidelberg (2011)
3. Black, J., Cochran, M., Shrimpton, T.: On the impossibility of highly-efficient blockcipher-based hash functions. In: Advances in Cryptology - EUROCRYPT 2005. Lecture Notes in Computer Science, vol. 3494, pp. 526–541. Springer-Verlag, Berlin (2005)
4. Black, J., Rogaway, P., Shrimpton, T., Stam, M.: An analysis of the blockcipher-based hash functions from PGV. Journal of Cryptology 23(4), 519–545 (2010)
5. Gauravaram, P., Knudsen, L.R., Matusiewicz, K., Mendel, F., Rechberger, C., Schläffer, M., Thomsen, S.: Grøstl – a SHA-3 candidate (2011), submission to NIST's SHA-3 competition
6. Mennink, B., Preneel, B.: Hash functions based on three permutations: A generic security analysis. In: Advances in Cryptology – CRYPTO 2012. Lecture Notes in Computer Science, vol. 7417, pp. 330–347. Springer, Heidelberg (2012)
7. Preneel, B., Govaerts, R., Vandewalle, J.: Hash functions based on block ciphers: A synthetic approach. In: Advances in Cryptology - CRYPTO '93. Lecture Notes in Computer Science, vol. 773, pp. 368–378. Springer-Verlag, Berlin (1993)
8. Rabin, M.: Digitalized signatures. In: Foundations of Secure Computation '78. pp. 155–166. Academic Press, New York (1978)
9. Rogaway, P., Shrimpton, T.: Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance. In: Fast Software Encryption 2004. Lecture Notes in Computer Science, vol. 3017, pp. 371–388. Springer-Verlag, Berlin (2004)
10. Rogaway, P., Steinberger, J.: Constructing cryptographic hash functions from fixed-key blockciphers. In: Advances in Cryptology - CRYPTO 2008. Lecture Notes in Computer Science, vol. 5157, pp. 433–450. Springer-Verlag, Berlin (2008)
11. Rogaway, P., Steinberger, J.: Security/efficiency tradeoffs for permutation-based hashing. In: Advances in Cryptology - EUROCRYPT 2008. Lecture Notes in Computer Science, vol. 4965, pp. 220–236. Springer-Verlag, Berlin (2008)
12. Shrimpton, T., Stam, M.: Building a collision-resistant compression function from non-compressing primitives. In: International Colloquium on Automata, Languages and Programming - ICALP (2) 2008. Lecture Notes in Computer Science, vol. 5126, pp. 643–654. Springer-Verlag, Berlin (2008)
13. Stam, M.: Beyond uniformity: Better security/efficiency tradeoffs for compression functions. In: Advances in Cryptology - CRYPTO 2008. Lecture Notes in Computer Science, vol. 5157, pp. 397–412. Springer-Verlag, Berlin (2008)
14. Steinberger, J.: Stam's collision resistance conjecture. In: Advances in Cryptology - EUROCRYPT 2010. Lecture Notes in Computer Science, vol. 6110, pp. 597–615. Springer-Verlag, Berlin (2010)
15. Steinberger, J.P., Sun, X., Yang, Z.: Stam's conjecture and threshold phenomena in collision resistance. In: Advances in Cryptology - CRYPTO 2012. Lecture Notes in Computer Science, vol. 7417, pp. 384–405. Springer, Heidelberg (2012)