# Preimage Resistance Beyond the Birthday Barrier – The Case of Blockcipher Based Hashing

Matthias Krause[1], Frederik Armknecht[1], and Ewan Fleischmann[2]

[1] Arbeitsgruppe Theoretische Informatik und Datensicherheit, University of Mannheim, Germany
[2] Chair of Media Security, Bauhaus-University Weimar, Germany

**Abstract.** We provide the first preimage resistance bounds for block cipher based double length, double call hash functions that go *beyond* the birthday bound. More precisely, we consider hash functions using two calls to an ideal block cipher with an $r$-bit key and $n$-bit plain- and ciphertext where $r > n$. For several practical and well-known double length design principles, we introduce techniques for proving that no adversary asking less than $\Omega(2^{1.5n})$ queries can find a preimage with probability greater than 1/2. These techniques can be applied to a series of existing constructions, e.g., Hirose's FSE'06 scheme, for deriving bounds that significantly outmatch previous results of $\Omega(2^n)$. In the case that two independent block ciphers are used, we are even able to state an asymptotically optimal bound of $\Omega(2^{2n})$.

**Keywords:** Hash Function, Preimage Resistance, Block Cipher, Beyond Birthday Bound, Foundations

## 1 Introduction

**Motivation.** A cryptographic hash function is a function which maps an input of arbitrary length to an output of fixed length and is one of the most important primitives in cryptography [15]. In recent years, a rise in research for block cipher based hash functions is clearly noticeable. Reasons for this certainly include the usage of such constructions in some candidates of the SHA-3 contest or the more widespread availability of resource restricted devices. Also there currently seems to be a deeper understanding on how to design good block ciphers compared to the knowledge on how to design good hash functions. The approach of block cipher based hashing might be used to transfer this knowledge to the field of hash functions – although one has to pay attention to the details. Due to the short output length of most practical block ciphers, *e.g.*, $n = 128$-bit, one is mainly interested in sound design principles for *double length* (DL) hash functions. Such constructions usually use one block cipher with $n$-bit output as the building block by which messages are projected to a fixed $2n$-bit string. Here, the two most important criteria are efficiency and security. By efficiency, one usually concentrates on the rate of the construction, being

$$\frac{\text{message bits processed per compression function}}{(\text{number of block cipher calls in H}) \times n},$$

although it might be interesting to consider other factors as, *e.g.*, the number of key schedules or one might want to incorporate the key length of the block ciphers. Regarding the security, the common notions refer to collision attacks (finding two inputs that map to the same output), preimage attacks (given an output, find a matching input), and second-preimage attacks (given an input-output pair, find another different input that maps to

the same output). Recall that for an ideal hash function with output length $2n$, the effort for finding a collision is in $\Theta(2^n)$ and for finding a preimage resp. second-preimage is in $\Theta(2^{2n})$. Thus, the security of constructions should be measured in comparison to these bounds.

There has been a considerable amount of publications that have analyzed the preimage resistance of block cipher based constructions. For single length compression functions, near-optimal bounds are known [2]. Quite the contrary is true for double length compression functions where several authors [5, 6, 13, 17, 7, 8] have tried to analyze preimage resistance, but *all* got stuck at the birthday bound $2^n$. The result usually stated is, that the adversary has negligible advantage of finding a preimage if she asks $q < 2^n$ queries while nothing is known about the advantage if $q \geq 2^n$. Indeed, several authors, *e.g.*, [5, 6, 13], called the challenge of finding more satisfying preimage bounds as one of the interesting open problems in the field of block cipher based hash functions. The problem does not seem to be the lack of promising constructions where stronger preimage bounds *might* hold, but rather that no techniques are known for *assessing* these questions. Summing up, several concepts on how to *design* block cipher based hash function are known today, but only little is known on how to *analyze* them thoroughly.

**Our Contribution.** In this paper, we present several new techniques for deriving minimum security bounds on the preimage resistance far beyond the birthday barrier. We use these techniques to analyze the following three simple and straightforward designs

- $H_1(K,X) = \Big(E_{K||0}(X) \oplus X, E_{\overline{K}||1}(X) \oplus X\Big)$,
- $H_2(K,X) = \big(E_K(X) \oplus X, E_{\overline{K}}(X) \oplus X\big)$, and
- $H_3(K,X) = \big(E_K(X) \oplus X, E_K(\overline{X}) \oplus \overline{X}\big)$,

where $\overline{K}$ denotes the bit-by-bit complement of $K$ and '||' the concatenation of bit strings. All these constructions are in principle using a concatenation of the established Davies-Meyer principle. The differences are that in $H_1$, two different ciphers are used, in $H_2$ two different keys, and in $H_3$ two different plaintexts. In that sense, this selection seems to be quite natural and a good starting point. Observe that all these constructions have a rate equal or very close to 1/2. Although some recent publications [6, 23, 24] have put some effort in developing and using more generalized forms, it seems to result in a tradeoff between clarity of presentation and broad applicability. We chose to use specific and very simple representation since some of the proofs are – even in these cases – rather involved. However, we show how to adapt these techniques to other designs.

We achieve the following new results:

1. No adversary can find a preimage with probability greater than 1/2 for $H_1$, $H_2$, or $H_3$ with less than $\Omega(2^{1.5n})$ queries.
2. We prove that no adversary can find a preimage with probability greater than 1/2 for $H_1$ with less than $2^{2n-9}$ queries. For this, we make the reasonable assumption that an adversary makes only queries for which the success probability of finding a preimage is non-zero.
3. We give examples for concrete designs where our results are directly applicable.

| Scheme | Best Known Preimage Bounds |
|---|---|
| Cyclic-DM [6] | n/a |
| ADD/3-DM [6] | n/a |
| Abreast-DM [11] | $\Omega(2^n)$ [6, 12] |
| Tandem-DM [11] | $\Omega(2^n)$ [13] |
| Hirose [8] | $\Omega(2^n)$<br>$\mathbf{\Omega(2^{1.5n})}$, **Section 5** |
| $H_1$, Hirose [7] | $\Omega(2^n)$<br>$\mathbf{\Omega(2^{1.5n})}$, **Section 3**<br>$\mathbf{\Omega(2^{2n})}$, **Section 4** |
| $H_2$ | $\mathbf{\Omega(2^{1.5n})}$, **Section 3** |
| $H_3$ | $\mathbf{\Omega(2^{1.5n})}$, **Section 3** |

**Table 1.** Comparison of our results for $H_1, H_2$ and $H_3$ with existing related work. The bounds refer to the minimum number of queries necessary for achieving a success probability of at least 1/2. Our results are marked in bold font.

In Table 1, we put our results into relation to existing comparable work.

We remark that there are other constructions that might be worth considering, *e.g.*, double length hash function with only one block cipher call and an additional operation [24, 14]. But this operation usually takes considerably longer than a plain additional block cipher call. Our table also does not contain proposals for which strong attacks are known, *e.g.*, MDC-2 [10, 25], or that have a weak rate considerable lower than 1/2, *e.g.*, MDC-4 [19], Merkle's DES-based scheme [16], or permutation based approaches [21, 22]. Currently, there are no schemes known in literature that employ a block cipher with an $n$-bit plain-/ciphertext and $2n$-bit key with a rate $> 1/2$ that are known to be secure. For block ciphers with $n$-bit key/plain-/ciphertext not even rate 1/2 schemes with these features are known.

**Outline.** In Section 2, we give some definitions and statements that are either well known from literature or are used for our later discussion. Section 3 gives a lower bound of preimage resistance of $\Omega(2^{1.5n})$ queries. This bound holds for all three constructions $H_1$, $H_2$, and $H_3$. For $H_1$ we also give a tighter bound that asymptotically matches the optimal $\Omega(2^{2n})$ barrier. The proof can be found in Section 4. In Section 5, we demonstrate that our techniques can be used for analyzing other designs as well, *e.g.*, Hirose's construction from FSE'06. In Section 6 we discuss our results and conclude the paper.

## 2 Preliminaries

### 2.1 General Notations

An $(r, n)$-block cipher is a keyed family of permutations consisting of two paired algorithms $E : \{0,1\}^r \times \{0,1\}^n \to \{0,1\}^n$ and $E^{-1} : \{0,1\}^r \times \{0,1\}^n \to \{0,1\}^n$ both accepting a key of size $r$ bits and an input block of size $n$ bits for some $n, r > 0$. For positive $n$, $Block(r, n)$ is the set of all $(r, n)$-block ciphers. For any $E \in Block(r, n)$ and any fixed key $K \in \{0,1\}^r$, decryption $E_K^{-1} := E^{-1}(K, \cdot)$ is the inverse function of encryption $E_K := E(K, \cdot)$, so that $E_K^{-1}(E_K(X)) = X$ holds for any input $X \in \{0,1\}^n$. In the ideal cipher model [2, 4, 9] $E$ is modeled as a family of random permutations $\{E_K\}$ whereas the random permutations are chosen independently for each key $K$, *i.e.*, formally $E$ is selected randomly from $Block(r, n)$. We use the convention to write oracles, that are provided to an algorithm, as superscripts. For example $\mathcal{A}^E$ is an algorithm $\mathcal{A}$ with oracle access to $E$ to which $\mathcal{A}$ can request forward- and backward queries. For ease of presentation, we identify the sets $\{0,1\}^{a+b}$ and $\{0,1\}^a \times \{0,1\}^b$. Similarly for $A \in \{0,1\}^a$ and $B \in \{0,1\}^b$, the concatenation of these bit strings is denoted by $A||B \in \{0,1\}^{a+b} = \{0,1\}^a \times \{0,1\}^b$.

### 2.2 Block Cipher Based Compression Functions

A compression function $H$ takes an $m$-bit message and an $l$-bit chaining value and compresses them to an $l$-bit value, i.e., $H : \{0,1\}^{m+l} \to \{0,1\}^l$ for some $m, l > 0$. A block cipher based compression function is a compression function that has access to usually one block cipher $E \in Block(r, n)$. We say a block cipher based compression function $H$ is *double-length, double-call* (DL) if a $2n$-bit value $(U, V)$ is computed using two calls to $E$.

### 2.3 Preimage Resistance

Insecurity is quantified by the success probability of an optimal resource-bounded adversary. The resource is the number of queries (forward and backward) to the block cipher $E$. An adversary is a computationally unbounded but always-halting algorithm $\mathcal{A}$ with access to $E \in Block(r, n)$. The adversary may make a *forward* query $(K, X)$ to discover the corresponding value $Y = E(K, X)$, or a *backward* query $(K, Y)$, so as to learn the corresponding value $X = E^{-1}(K, Y)$ such that $E(K, X) = Y$. Either way, the result of the query is stored in a triple $(K_i, X_i, Y_i) := (K, X, Y)$ and the *query history* $\mathcal{Q}$ is the tuple $(Q_1, \ldots, Q_q)$ where $Q_i = (K_i, X_i, Y_i)$ and $q$ is the total number of queries made by the adversary. Without loss of generality, we assume that $\mathcal{A}$ asks at most once on a triplet of a key $K_i$, a plaintext $X_i$ and a ciphertext $Y_i$ obtained by a query and the corresponding reply. For a set $S$, let $z \xleftarrow{\$} S$ represent random sampling from $S$ under the uniform distribution. For a probabilistic algorithm $\mathcal{M}$, let $z \xleftarrow{\$} \mathcal{M}$ mean that $z$ is an output of $\mathcal{M}$ and its distribution is based on the random choices of $\mathcal{M}$.

A preimage finding adversary is an algorithm whose goal is to find a preimage of a specific compression function. There are several methods known on how to define this notion [20]. We opt for the strongest notion, preimage resistance (Pre), which intuitively states that a function is a one-way function. This notion does imply weaker notions as, *e.g.*, everywhere preimage resistance (ePre) and always preimage resistance (aPre).

**Definition 1.** ***(Preimage Resistance* Pre *[20])*** *Let $H$ be a block cipher based compression function, $H : \{0,1\}^{m+l} \to \{0,1\}^l$. Fix an adversary $\mathcal{A}$ with access to oracles $E, E^{-1}$. The advantage of $\mathcal{A}$ of inverting $H$ is the real number*

$$\mathbf{Adv}_H^{Pre}(\mathcal{A}) = \Pr[E \xleftarrow{\$} Block(r,n); A \xleftarrow{\$} \{0,1\}^{m+l}; B \xleftarrow{\$} H(A); A' \xleftarrow{\$} \mathcal{A}^E(B) : H(A') = B].$$

Again, for $q \geq 1$, we write

$$\mathbf{Adv}_H^{Pre}(q) = \max_{\mathcal{A}}\{\mathbf{Adv}_H^{Pre}(\mathcal{A})\}$$

where the maximum is taken over all adversaries that ask at most $q$ oracle queries.

### 2.4 Analysis Preliminaries

For our analysis, we focus on the following three double call, double length block cipher based compression functions, $H_1 : \{0,1\}^{r-1} \times \{0,1\}^n \longrightarrow \{0,1\}^{2n}$ and $H_2, H_3 : \{0,1\}^r \times \{0,1\}^n \longrightarrow \{0,1\}^{2n}$ that all have access to a block cipher $E$ and are defined as follows. For any $X$, $K$ let

$$\begin{aligned} H_1(K,X) &= \left(E_{K||0}(X) \oplus X, E_{\overline{K}||1}(X) \oplus X\right), \\ H_2(K,X) &= \left(E_K(X) \oplus X, E_{\overline{K}}(X) \oplus X\right), \quad \text{and} \\ H_3(K,X) &= \left(E_K(X) \oplus X, E_K(\overline{X}) \oplus \overline{X}\right). \end{aligned}$$

Note that – in particular $H_2$ and $H_3$ – can be defined in a similar way by replacing the bit by bit complement with an arbitrary fixed point free involution without changing any arguments used in the proofs.

We analyze the preimage resistance of $H_1$, $H_2$ and $H_3$, *i.e.*, we derive lower bounds on the number of queries to $E$ necessary for reaching at least a success probability of $1/2$ in finding a preimage for a random hash value $(U,V) \in \{0,1\}^{2n}$.

Although the adversary can pose by definition arbitrary queries to the $E$-oracle, one sees easily that certain pairs of queries can be naturally polled. For example, asking in the $H_1$ case for the value $E_{K||0}(X)$ is not sufficient for confirming a preimage without asking (or knowing) the value $E_{K||1}(X)$. Therefore, we consider meta queries instead where certain $E$-queries are given for "free". These queries are selected by the following criteria:

1. One meta query should provide all information necessary for deciding whether the asked parameters yield a preimage.
2. The set of meta queries should not restrict the set of possible $E$-queries. That is, for each possible values $X$, $Y$, and $K$, there should exist meta queries that provide the value $E_K(X)$ resp. $E_K^{-1}(Y)$.

Based on these considerations, we use the following meta queries:

**Definition 2.** *For $H_1$, the adversary always asks one of the following oracle queries.*

– *Type-I queries $Q$ with respect to $X \in \{0,1\}^n$, $K \in \{0,1\}^{r-1}$, denoted $Q = (I, X, K)$, yielding the response $R(Q,E)$ consisting of $Y = E_{K||0}(X)$ and $Y' = E_{\overline{K}||1}(X)$.)*
– *Type-II queries with respect to $Y \in \{0,1\}^n$, $K \in \{0,1\}^{r-1}$, denoted $Q = (II, Y, K)$, yielding the response $R(Q,E)$ consisting of $X = E^{-1}_{K||0}(Y)$ and $Y' = E_{\overline{K}||1}(X)$.)*
– *Type-III queries with respect to $Y' \in \{0,1\}^n$, $K \in \{0,1\}^{r-1}$, denoted $Q = (III, Y, K)$, yielding the response $R(Q,E)$ consisting of $X = E^{-1}_{\overline{K}||1}(Y')$ and $Y = E_{K||0}(X)$.)*

For $H_2$ and $H_3$ the queries are adapted accordingly. In more detail, for $H_2$, the Type-I query response is $(Y, Y') = (E_K(X), E_{\overline{K}}(X))$, the Type-II query response is $(X, Y') = (E_K^{-1}(Y), E_{\overline{K}}(X))$ and the Type-III query response is $(X, Y) = (E^{-1}_{\overline{K}}(Y), E_K(X))$. For $H_3$, the Type-I query response is $(Y, Y') = (E_K(X), E_K(\overline{X}))$ and the Type-II query response is $(X, Y') = (E_K^{-1}(Y), E_K(\overline{X}))$. It is easy to see that for $H_3$ it is not necessary to consider Type-III queries.

In any case, the 5-tupel $(T, X, Y, Y', K)$, $T \in \{I, II, III\}$, is called the query-response pair $(Q, R(Q))$. Observe that there are two equivalent versions of the operation mode of the oracle.

– In the offline version, the oracle in advance chooses the random $(r,n)$-block cipher $E$.
– In the online version, for each new query, the oracle chooses the answer at random with respect to the uniform distribution from the set of possible answers.

The $\Omega(2^{1.5n})$ bound will be derived with respect to the the offline version of the $E$-oracle, while the $\Omega(2^{2n})$ bound is based on the online version.

## 3 Lower $\Omega(2^{1.5n})$ Preimage Resistance Bounds for $H_1$, $H_2$ and $H_3$

We first analyze in Section 3.1 the preimage resistance of $H_1$ and give an $\Omega(2^{3/2 \cdot n})$ lower bound on the number of queries necessary for finding a preimage with a chance of success of at least $1/2$. Then, in Section 3.2, we discuss how we can generalize this result to $H_2$ and $H_3$.

### 3.1 A Preimage Bound for $H_1$

We start with an important observation on $H_1$.

**Lemma 1.** *For any $(U,V) \in \{0,1\}^{2n}$, the complexity of finding a preimage to the hash value $(U,V)$ w.r.t. $H_1$ is the same as the complexity of finding a preimage to the hash value $(0,0)$ w.r.t. $H_1$.*

*Proof.* Finding a preimage for $H_1$ to the image $(0,0) \in \{0,1\}^{2n}$ is equivalent to finding a key-prefix $K \in \{0,1\}^{r-1}$ and an input $X \in \{0,1\}^n$ such that $E_{K||0}(X) = E_{\overline{K}||1}(X) = X$. We call such an $X$ a *double fixed point* of the permutations $E_{K||0}$ and $E_{\overline{K}||1}$.

Finding a preimage to the image $(U,V) \in \{0,1\}^{2n}$ is equivalent to finding a key-prefix $K \in \{0,1\}^{r-1}$ and an input $X \in \{0,1\}^n$ such that $\tilde{E}_{K||0}(X) = \tilde{E}_{K||1}(X) = X$, where the $(r,n)$-block cipher $\tilde{E}$ is defined as follows: For all keys $\kappa = (\kappa_1, \cdots, \kappa_r)$ and $X \in \{0,1\}^n$ let $\tilde{E}_\kappa(X) := E_\kappa(X) \oplus V$ if $\kappa_r = 0$ and $\tilde{E}_\kappa(X) := E_\kappa(X) \oplus W$ if $\kappa_r = 1$.

Note that assigning $\tilde{E}$ to $E$ defines a bijective mapping over $Block(r,n)$. The outputs of $E$ and $\tilde{E}$ are uniformly distributed. Consequently, the effort for finding a preimage to the image $(U,V)$ is the same as of finding a preimage to the image $(0,0)$. □

This lemma naturally implies the following definition:

**Definition 3 (Successful Queries).** *Let $\mathcal{Q} = (Q_1, \cdots, Q_q)$ be an arbitrarily fixed sequence of $q$ queries,* i.e., *$Q_i = (T_i, Z_i, K_i)$, where $T_i \in \{I, II, III\}$, $Z_i \in \{0,1\}^n$, $K_i \in \{0,1\}^{r-1}$.*

*We say that the $i$-th query $Q_i = (T_i, Z_i, K_i)$ is* successful*, if the corresponding query-response pair $(Q_i, R(Q_i)) := (T_i, X_i, Y_i, Y_i', K_i)$ satisfies $X_i = Y_i = Y_i'$. We denote this event by $\texttt{Succ}(Q_i)$.*

*Analogously, we say $\mathcal{Q}$ is successful, iff at least one query in $\mathcal{Q}$ is* successful*. This event will be denoted by $\texttt{Succ}(\mathcal{Q})$.*

We now upper bound $\Pr[\texttt{Succ}(\mathcal{Q})]$, where the probability is measured over the uniform distribution over $Block(r,n)$. To this end, we define by $K(\mathcal{Q}) = \{K_1, \cdots, K_{q'}\}$ the set of keys involved in $\mathcal{Q}$. Naturally, $q' \leq q$. For any $K \in K(\mathcal{Q})$, the sequence of queries $Q_i$, $1 \leq i \leq q$, with $K_i = K$ is denoted by $\mathcal{Q}_K$. Clearly,

$$\Pr[\texttt{Succ}(\mathcal{Q})] \leq \sum_{K \in K(\mathcal{Q})} \Pr[\texttt{Succ}(\mathcal{Q}_K)]. \tag{1}$$

Next, we upper bound $\Pr[\texttt{Succ}(\mathcal{Q}_K)]$ for $K \in K(\mathcal{Q})$. Given two permutations $\pi, \pi'$ over $\{0,1\}^n$, we denote by $DFP(\pi,\pi') = \{X : \pi(X) = \pi'(X) = X\}$ the set of double fixed points of $\pi, \pi'$. Furthermore, we define

- $\mathcal{K}_0(E) = \{K \in \{0,1\}^{r-1}, DFP(E_{K||0}, E_{\overline{K}||1}) = \emptyset\}$,
- $\mathcal{K}_1(E) = \{K \in \{0,1\}^{r-1}, |DFP(E_{K||0}, E_{\overline{K}||1})| = 1\}$, and
- $\mathcal{K}_2(E) = \{K \in \{0,1\}^{r-1}, |DFP(E_{K||0}, E_{\overline{K}||1})| \geq 2\}$.

**Lemma 2.** *For any $K \in \{0,1\}^{r-1}$ and $N \geq 3$ it holds that $\Pr[K \in \mathcal{K}_1(E)] < 1/N$ and $\Pr[K \in \mathcal{K}_2(E)] < 1/N^2$.*

*Proof.* The probability for any $X \in \{0,1\}^n$ being a double fixed point is upper bounded by $((N-1)!/N!)^2$, *i.e.* dividing the number of possible permutations of $N$ elements with one element fixed by the total number of permutations of $N$ elements. Consequently,

$$\Pr[K \in \mathcal{K}_1(E)] < N \cdot \left(\frac{(N-1)!}{N!}\right)^2 = 1/N.$$

With a similar argument one can show that

$$\Pr[K \in \mathcal{K}_2(E)] < \binom{N}{2} \left(\frac{(N-2)!}{N!}\right)^2 = \frac{1}{2N(N-1)} < 1/N^2 \tag{2}$$

for $N \geq 3$. Note that these two bounds are in fact upper bounds since they essentially upper bound the cases 'one or more double fixed points' and 'two or more double fixed points'. □

Since $\Pr[\mathtt{Succ}(\mathcal{Q}_K)|K \in \mathcal{K}_0(E)] = 0$, we obtain

$$\begin{aligned}\Pr[\mathtt{Succ}(\mathcal{Q})] &\leq \sum_{K\in K(\mathcal{Q})} \sum_{j=0}^{2} \Pr[\mathtt{Succ}(\mathcal{Q}_K) \,|\, K \in \mathcal{K}_j(E)] \cdot \Pr[K \in \mathcal{K}_j(E)] \\ &< \sum_{K\in K(\mathcal{Q})} \left(1/N \cdot \Pr[\mathtt{Succ}(\mathcal{Q}_K) \,|\, K \in \mathcal{K}_1(E)] + 1/N^2\right) \\ &\leq 1/N^2 \cdot |K(\mathcal{Q})| + \sum_{K\in K(\mathcal{Q})} 1/N \cdot \Pr[\mathtt{Succ}(\mathcal{Q}_K) \,|\, K \in \mathcal{K}_1(E)]. \qquad (3)\end{aligned}$$

We now derive an upper bound for $\Pr[\mathtt{Succ}(\mathcal{Q}_K) \,|\, K \in \mathcal{K}_1(E)]$. This probability equals the probability $\Pr_1[\mathtt{Succ}(\mathcal{Q}_K)]$, where the probability measure $\Pr_1$ is defined w.r.t. the uniform distribution over the set of all keys $K \in \mathcal{K}_1(E)$.

Let $\mathcal{Q}_K = (S_1, \cdots, S_t)$, $S_i = (T_i, Z_i, K)$ for $i = 1, \cdots, t$, where $Z_i \in \{0,1\}^n$ and $T_i \in \{I, II, III\}$. The event that some query of $\mathcal{Q}_K$ is successful implies the existence of some $i$, $0 \leq i \leq t-1$, such that $S_{i+1}$ is successful but $S_1, \cdots, S_i$ have not been successful before. Consequently,

$$\begin{aligned}\Pr_1[\mathtt{Succ}(\mathcal{Q}_k)] &\leq \sum_{i=0}^{t-1} \Pr_1\left[\mathtt{Succ}(S_{i+1}) \wedge \bigwedge_{j=1}^{i} \mathtt{Fail}(S_j)\right] \\ &= \sum_{i=0}^{t-1} \Pr_1\left[\mathtt{Succ}(S_{i+1}) \,\middle|\, \bigwedge_{j=1}^{i} \mathtt{Fail}(S_j)\right] \cdot \Pr_1\left[\bigwedge_{j=0}^{i} \mathtt{Fail}(S_j)\right] \\ &\leq \sum_{i=0}^{t-1} \Pr_1\left[\mathtt{Succ}(S_{i+1}) \,\middle|\, \bigwedge_{j=1}^{i} \mathtt{Fail}(S_j)\right].\end{aligned}$$

Hereby, $\mathtt{Succ}(S_i)$ denotes the event that the query $S_i$ is successful and $\mathtt{Fail}(S_j)$ denotes the event that query $S_j$ is is not successful.

**Lemma 3.** *It holds that* $\Pr_1\left[\mathtt{Succ}(S_{i+1}) \,\middle|\, \bigwedge_{j=1}^{i} \mathtt{Fail}(S_j)\right] \leq \frac{1}{N-3i}$.

*Proof.* Let us fix a key $K \in \mathcal{K}_1(E)$ and let $X^*$ denote the unique double fixed point, i.e., $E_{K||0}(X) = E_{\overline{K}||1}(X) = X$. We suppose that the queries $S_1, \cdots, S_i$ have not been successful.

Let $(S_1, R(S_1)), \cdots, (S_i, R(S_i))$ denote the corresponding sequence of query-response pairs, where $(S_j, R(S_j)) = (T_j, X_j, Y_j, Y'_j, K)$ for $1 \leq j \leq i$.

Let $\mathcal{A} = \{X_1, \cdots, X_i\} \cup \{Y_1, \cdots, Y_i\} \cup \{Y'_1, \cdots, Y'_i\}$. It can be easily checked that $X^* \notin \mathcal{A}$, since otherwise one of the queries $S_j$, $1 \leq j \leq i$, would have been successful. Note that any $X^{**} \in \{0,1\}^n \setminus \mathcal{A}$ is equally likely to be $X^*$. The success probability of query $S_{i+1}$ is zero if $X_{i+1} \in \mathcal{A}$ and $\leq 1/(N - |\mathcal{A}|)$ if not. As $|\mathcal{A}| \leq 3i$ we obtain the desired result. □

We fix a parameter $\alpha$, $1 \leq \alpha \leq N$, which will be determined later, and define

$$\mathcal{A}_1 = \{K \in K(\mathcal{Q}), |\mathcal{Q}_K| \leq \alpha\} \text{ and } \mathcal{A}_2 = \{K \in K(\mathcal{Q}), |\mathcal{Q}_K| > \alpha\}.$$

According to (3) we obtain that

$$\begin{aligned}\Pr[\mathtt{Succ}(\mathcal{Q})] &\leq 1/N^2 \cdot |K(\mathcal{Q})| \ + 1/N \cdot \sum_{K \in \mathcal{A}_1} \Pr_1[\mathtt{Succ}(\mathcal{Q}_K)] + 1/N \cdot \sum_{K \in \mathcal{A}_2} \Pr_1[\mathtt{Succ}(\mathcal{Q}_K)] \\ &\leq 1/N^2 \cdot |K(\mathcal{Q})| + 1/N \cdot |\mathcal{A}_1| \cdot \frac{\alpha}{N - 3\alpha} + 1/N \cdot |\mathcal{A}_2|. \end{aligned} \tag{4}$$

Now suppose that the success probability of the adversary in this case is greater or equal than 1/2, *i.e.*, $\Pr[\mathtt{Succ}(\mathcal{Q})] \geq 1/2$ where $|\mathcal{Q}| = q$. Let $q < \frac{1}{6}N^2$. Then $\frac{1}{6}N^2 > q \geq |K(\mathcal{Q})|$ and (4) imply that

$$1/N \cdot |\mathcal{A}_1| \cdot \frac{\alpha}{N - 3\alpha} > \ 1/6 \tag{5}$$

or

$$1/N \cdot |\mathcal{A}_2| > 1/6. \tag{6}$$

Now, (5) implies (as $q \geq |\mathcal{A}_1|$) that

$$q \cdot 1/N \cdot \frac{\alpha}{N - 3\alpha} > 1/6, \qquad \textit{i.e.} \qquad q \geq B_1(N, \alpha),$$

where $B_1(N, \alpha) = 1/6 \cdot N \cdot \frac{N-3\alpha}{\alpha} = 1/6 \cdot N \cdot \left(\frac{N}{\alpha} - 3\right)$.

As $q \geq \alpha \cdot |\mathcal{A}_2|$, (6) implies $q \geq B_2(N, \alpha)$, where $B_2(N, \alpha) = 1/6 \cdot \alpha \cdot N$.

For deriving a lower bound on $q$, we fix $\alpha$ in such a way that $B_1(N, \alpha) = B_2(N, \alpha)$. We obtain $\alpha$ as the solution of $\alpha = \frac{N}{\alpha} - 3$, and, thus, as the positive solution of the quadratic equation $\alpha^2 + 3\alpha - N = 0$. This implies $\alpha = \sqrt{N + 9/4} - 3/2$. Using $q \geq B_1(N, \alpha)$, we have shown that if $\Pr[\mathtt{Succ}(\mathcal{Q})] \geq 1/2$ implies that $q \geq 2^n/6 \cdot \left(\sqrt{2^n + 9/4} - 3/2\right)$ and therefore the following Theorem 1.

**Theorem 1.** $\Pr[\mathtt{Succ}(\mathcal{Q})] \geq 1/2$ *implies that* $q \geq 2^n/6 \cdot \left(\sqrt{2^n + 9/4} - 3/2\right)$.

### 3.2 Preimage bounds for $H_2$ and $H_3$

An appealing property of the approach discussed in the previous section is that it can be easily translated to the case of $H_2$ and $H_3$ as well. In principle, the arguments are equal, but one has to pay attention to the details.

Let $(U, V)$ denote the considered image. Recall that one of the main ideas is to split the set of possible keys into the following subsets: (i) keys for which no preimage exists, (ii) keys for which one preimage exists, and (iii) keys for which more than one preimage exists. For $H_2$ and $H_3$, one can follow the same strategy. The difference is that also keys for which a preimage of $(V, U)$ exists (instead of $(U, V)$) are helpful.

Let us illustrate this for the example of $H_2$. Given an input $(K, X)$ such that $H_2(K, X) = (V, U)$, one sees easily that $H_2(\overline{K}, X) = (U, V)$. Consequently, such keys are valuable as well although they might not give *directly* a preimage of $(U, V)$. For a formalization of this issue, we define for a given key $K$ the following set of "good" inputs

$$\begin{aligned}\mathsf{Good}(K) := \{X : (E_K(X) \oplus X = U \wedge E_{\overline{K}}(X) \oplus X = V) \vee \\ (E_K(X) \oplus X = V \wedge E_{\overline{K}}(X) \oplus X = U)\}\end{aligned}$$

and consider for $i = 0, 1, 2$ the sets of keys $\mathcal{K}_i(E) := \{K \in \{0,1\}^r, |\mathtt{Good}(K)| = i\}$. One checks easily that the probabilities $Pr[K \in \mathcal{K}_i]$ are equal to the probabilities derived in Section 3.1 up to some constants. Using the same arguments as given there, one easily gets the following Theorem

**Theorem 2.** *In the case of $H_2$ and $H_3$, a probability of success $\geq 1/2$, i.e., $\Pr[\mathtt{Succ}(\mathcal{Q})] \geq 1/2$, implies that $q \in \Omega(2^{1.5n})$.*

## 4 An $\Omega(2^{2n})$ Lower Bound for $H_1$

### 4.1 Preliminaries

Now we consider another approach for estimating a lower bound on the number of queries needed to find an preimage of $(0,0)$ w.r.t. $H_1$ with probability of at least $1/2$. For this approach, we have assume an adversary that is greedy in the sense that she does not pose queries for which it is known in advance that the success probability of finding the preimage is zero. We make this formal in the following Definition:

**Definition 4 (Disjoint Query).** *Let $\mathcal{Q} = (Q_1, \cdots, Q_q)$ be a sequence of queries. Let $Q_i = (T_i, Z_i, K_i)$ for $i = 1, \cdots, q$ and $(Q_i, R(Q_i)) = (T_i, X_i, Y_i, Y_i', K_i)$ be the according query-response pair. For a fixed key $K$, we consider the set $A(\mathcal{Q}, K) \subseteq \{0,1\}^n$ of all inputs and outputs that occurred so far with respect to the same key, i.e.,*

$$A(\mathcal{Q}, K) := \bigcup_{1 \leq i \leq q : K_i = K} \{X_i, Y_i, Y_i'\}. \tag{7}$$

*We call a new query $Q = (T, Z, K)$ to be* disjoint *to $\mathcal{Q}$, if $Z \notin A(\mathcal{Q}, K)$.*

The reason for considering disjoint queries are made clear by the following Lemma:

**Lemma 4.** *Let $\mathtt{Fail}(\mathcal{Q})$ denote the event that making the queries $\mathcal{Q}$ was not successful. For a single query $Q = (T, Z, K)$, let $Pr[Q|\mathtt{Fail}(\mathcal{Q})]$ denote the probability that $Q$ is successful w.r.t. $E$ under the condition that all queries in $\mathcal{Q}$ have not been successful before w.r.t. $E$. We abbreviate $N = 2^n$. It holds*

$$Pr[Q|\mathtt{Fail}(\mathcal{Q})] \leq \begin{cases} 1/(N-q)^2 \text{ , } Q \text{ disjoint to } \mathcal{Q} \\ 0 \qquad\quad \text{, else.} \end{cases} \tag{8}$$

**Proof:** We assume the online mode for the $E$-oracle. Note that the fact that $\mathcal{Q}$ ist not successful w.r.t. $E$ implies that $X_i \neq Y_i$ or $X_i \neq Y_i'$ for all $i = 1, \cdots, q$. Furthermore, if $Z \in A(\mathcal{Q}, K)$, i.e., the query is not disjoint, then there exists an index $i \in \{1, \ldots, q\}$ such that $X_i = Z$, $Y_i = Z$, or $Y_i' = Z$. Taking both together immediately shows that $Z$ cannot be a double fixed point.

Let us now suppose that $Q$ is disjoint to $\mathcal{Q}$. We estimate $Pr[Q|\mathtt{Fail}(\mathcal{Q})]$ under the condition that $T = I$. For the other two cases $T \in \{II, III\}$, the proof can be done in a similar way. The fact that $Q$ is disjoint to $\mathcal{Q}$ implies that $Z \in \{0,1\}^n \backslash X$ and $Z \in \{0,1\}^n \backslash Y$ and $Z \in \{0,1\}^n \setminus Y'$. Consequently, the probability that $E_{K||0}(Z) = Z$ and the probability that $E_{\overline{K}||1}(Z) = Z$ are both $\leq 1/(N-q)$. As both events are independent, the success probability of $Q$ is $\leq 1/(N-q)^2$. $\square$

**Definition 5 (Sequence of Disjoint Queries).** *We call $\mathcal{Q}$ a* sequence of disjoint queries *w.r.t. $E$ if for all $i$, $1 \leq i \leq q$, query $Q_{i+1}$ is disjoint to $\mathcal{Q}_{\leq i} := \{Q_1, \ldots, Q_i\}$.*

By Lemma 4, we know that this is the only kind of queries that have a non-zero success probability. Although it seems to be plausible that this strategy is the optimum one, we do not have a proof for this assumption. In other words, we cannot exclude that strategies might exist where asking some queries with zero success probability can yield globally a better success probability.

**Definition 6 (Accepting Computation).** *A sequence of queries $\mathcal{Q} = (Q_1, \cdots, Q_q)$ is called an* accepting computation *(or, for short, $\mathcal{Q}$ is* accepting*) iff*

*(1) $\mathcal{Q}$ is a sequence of disjoint queries.*
*(2) For all $i$, $1 \leq i \leq q-1$, query $Q_i = (T_i, Z_i, K_i)$ is not successful, i.e., $Z_i$ is not a double fixed point, and*
*(3) query $Q_q = (T_q, Z_q, K_q)$ is successful, i.e., $E_{K_q||0}(Z_q) = E_{\overline{K_q}||1}(Z_q) = Z_q$.*

### 4.2 Main Result

The main technical result of this section is the following estimation of the probability $Pr[\mathcal{Q}$ accepting$]$ that $\mathcal{Q}$ is an accepting computation.

**Theorem 3.** *Consider a sequence of queries $\mathcal{Q} = (Q_1, \cdots, Q_q)$ and set $N := 2^n$.*

*(i) It holds $Pr[\mathcal{Q}$ accepting$] \leq \frac{1}{(N-q)^2}$.*
*(ii) If $q \geq 15/16 \cdot N$ then $Pr[\mathcal{Q}$ accepting$] \leq e^{-1/32 \cdot N}$.*

**Proof:** The proof of part (i) is an straightforward consequence of Lemma 4. The proof of part (ii) is postponed to subsection 4.3. □

We show now how Theorem 3 can be used to derive a nearly maximal lower bound on the preimage resistance of $H_1$. Let $q \leq N^2$ and $\mathcal{Q} = (Q_1, \cdots, Q_q)$ denote an arbitrarily fixed sequence of disjoint queries asked by the adversary with $Q_i = (T_i, Z_i, K_i)$ for $1 \leq i \leq q$. We call $\mathcal{Q}$ to be successful if at least one of the queries $Q_i$ in $\mathcal{Q}$ is successful, i.e., $E_{K_i||0}(Z_i) = E_{\overline{K_i}||1}(Z_i) = Z_i$. This implies that for at least one query $Q_i \in \mathcal{Q}$ it holds that $\mathcal{Q}_{\leq i}$ is an accepting computation. Consequently,

$$Pr\left[\mathtt{Succ}(\mathcal{Q})\right] \leq \sum_{i=1}^{q} Pr\left[\mathcal{Q}_{\leq i} \text{ accepting}\right]. \tag{9}$$

Observe that the first claim of Theorem 3 does not make any useful statements beyond the birthday bound $\geq 2^n$. Indeed, the idea is now to split the set of queries into two sets, according to the statements given in Theorem 3. For all $i$, $1 \leq i \leq q$, let $rank_{\mathcal{Q}}(i)$ denote the number of queries $Q_j = (T_j, Z_j, K_j)$ with $1 \leq j < i$ and key $K_j = K_i$. Let $r_1 := \{i, rank_{\mathcal{Q}}(i) > \frac{15}{16}N\}$ and $r_2 := \{i, rank_{\mathcal{Q}}(i) \leq \frac{15}{16}N\}$. Theorem 3 and Relation (9) yield

$$Pr\left[\mathtt{Succ}(\mathcal{Q})\right] \leq \sum_{i \in r_1} Pr\left[\mathcal{Q}_{\leq i} \text{ accepting}\right] + \sum_{i \in r_2} Pr\left[\mathcal{Q}_{\leq i} \text{ accepting}\right] \tag{10}$$

$$\leq |r_1| \cdot e^{-1/32 \cdot N} + |r_2| \cdot \frac{1}{(N - 15/16 \cdot N)^2} \tag{11}$$

$$\leq q \cdot 256 \cdot N^{-2} \tag{12}$$

if $N \geq 256$. We have proved

**Theorem 4.** *For achieving a success probability of* $1/2$ *in finding a preimage of* $(0,0)$ *w.r.t.* $H_1$*, a greedy adversary has to ask at least* $1/512 \cdot 2^{2n} = 2^{2n-9}$ *queries.* □

### 4.3 The Proof of Part (ii) of Theorem 3

Let $q = 15/16 \cdot N$ and $\mathcal{Q} = (Q_1, \cdots, Q_{q+1})$ be an arbitrarily fixed sequence of $q+1$ queries w.r.t. the same key $K \in \{0,1\}^{r-1}$. Let $Q_i = (T_i, Z_i, K)$ for $i = 1, \cdots, q+1$. We derive an upper bound for the probability $Pr[\mathcal{Q} \text{ accepting}]$.

While asking $Q_1, \cdots, Q_{q+1}$, the adversary generates sets $\mathcal{X}_i = \{X_1, \cdots, X_i\}$, $\mathcal{Y}_i = \{Y_1, \cdots, Y_i\}$, and $\mathcal{Y}'_i = \{Y'_1, \cdots, Y'_i\}$ of size $i$. Let $\mathcal{A}_i := \mathcal{X}_i \cup Y_i \cup \mathcal{Y}'_i$. One sees easily that $|\mathcal{A}_i| + 1 \leq |\mathcal{A}_{i+1}| \leq |\mathcal{A}_i| + 3$ for $i = 0, \cdots, q$. (Let $\mathcal{A}_0 = \emptyset$). As $\mathcal{Q}$ is a sequence of disjoint queries, it must hold that the input $Z_{q+1}$ is outside of $\mathcal{A}_q$ and in particulkar $|\mathcal{A}_q| < N$. This implies that

$$Pr[\mathcal{Q} \text{ accepting}] \leq Pr[|\mathcal{A}_q| < N]. \tag{13}$$

We show that the latter event is rather unlikely by taking a closer look on the size of $\mathcal{A}_\ell$ for some smaller index $\ell < q$. Fix $\ell = N/8$. Because of $|\mathcal{A}_{i+1}| \geq |\mathcal{A}_i| + 1$, one has $|\mathcal{A}_q| - |\mathcal{A}_\ell| \geq q - \ell = 15/16N - 2/16N = 13/16N$. This implies that

$$|\mathcal{A}_\ell| \leq |\mathcal{A}_q| - 13/16N < N - 13/16N = 3/16N. \tag{14}$$

It follows that

$$Pr[|\mathcal{A}_q| < N] \leq Pr[|\mathcal{A}_\ell| \leq 3/16N] =: p^*. \tag{15}$$

We show that $p^* \leq e^{-1/32 \cdot N}$ which yields the initial claim by Eqs. (13) and (15).

For this purpose, we introduce a set of independent random Bernoulli variables and make use of Chernov's Inequality [1, Appendix A, pp. 233-240]. We recall it here shortly: Let $\nu_1, \cdots, \nu_n$ be independent random Bernoulli variables. Let $\sigma = 1/n \cdot \sum_{i=1}^{n} \nu_i$ be the (normed) sum of these variables and $\mathbf{E}(\sigma) = 1/n \cdot \sum_{i=1}^{n} Pr[\nu_i = 1]$ its expectation value. Then, for all $\delta > 0$ it holds that $Pr[\mathbf{E}(\sigma) - \sigma > \delta] < e^{-2\delta^2 n}$.

For defining the Bernoulli variables, we take a closer look on what happens during asking a query $Q_i$. Each query $Q_i$ is composed of two separate queries $Q_i^0$ and $Q_i^1$ to the $E$-oracle. In sub-query $Q_i^0$, the adversary asks an input $Z_i^0$ and gets a response $R_i^0 = E_{K_i}(Z_i)$ if a forward query has been made or $R_i^0 = E_{K_i}^{-1}(Z_i)$ in the case of a backward query. Likewise, for the other sub-query $Q_i^1$ she requests an input $Z_i^1$ and gets an answer $R_i^1$.

For $b \in \{0,1\}$ denote by $\mathcal{R}_i^b$ the set of possible answers for query $Q_i^b$. Note that $|\mathcal{R}_i^b| = N - (i-1)$[3]. Note further that for $i \leq \ell = N/8$ it holds that

$$|\mathcal{A}_i| \leq 3i < 1/2 \cdot (N - (i-1)) = 1/2 \cdot |\mathcal{R}_i^b|. \tag{16}$$

We now introduce subsets of $\mathcal{R}_i^0$ and $\mathcal{R}_i^1$ and consider the probability that $R_i^0$ and $R_i^1$ fall into these sets, respectively. The reasons are twofold: first, it allows for deriving a lower bound on $|\mathcal{A}_\ell|$, and second do they imply Bernoulli variables $\nu_i^0$ and $\nu_i^1$ as explained above which allow for using Chernov's Inequality. These variables are defined as follows. Suppose that for $i = 1, \cdots, \ell$, in addition to asking $Q_i^0$ and $Q_i^1$, the adversary does the following.

- Before asking $Q_i^0$ she fixes a set $\tilde{\mathcal{R}}_i^0 \subseteq \mathcal{R}_i^0 \setminus (\mathcal{A}_{i-1} \cup \{Z_i\})$ of size $\lceil |\mathcal{R}_i^0|/2 \rceil$, and
- before asking $Q_i^1$ she fixes a set $\tilde{\mathcal{R}}_i^1 \subseteq \mathcal{R}_i^1 \setminus (\mathcal{A}_{i-1} \cup \{Z_i, Z_i'\})$ of size $\lceil |\mathcal{R}_i^1|/2 \rceil$.

Inequality (16) guarantees that this is always possible. For $i = 1, \cdots, s$ let $\nu_i^0 \in \{0,1\}$ denote the random Bernoulli variable taking 1 iff $R_i^0 \in \tilde{\mathcal{R}}_i^0$, and analogously let $\nu_i^1 \in \{0,1\}$ denote the random Bernoulli variables taking value 1 iff $R_i^1 \in \tilde{\mathcal{R}}_i^1$.

As we are considering the ideal cipher model, $\nu_i^0$ and $\nu_i^1$ are independent random Bernouilli variables. Let $\sigma = \frac{1}{2\ell} \cdot \sum_{i=1}^{\ell} (\nu_i^0 + \nu_i^1)$ the normed sum. We can apply Chernov's Inequality which tells that $Pr[\mathbf{E}(\sigma) - \sigma > \delta] < e^{-2\delta^2 n}$.

As each variable takes 1 with a probability $\geq 1/2$, one sees easily that that $\mathbf{E}(\sigma) \geq 1/2$ and in particular $1/4 \leq \mathbf{E}(\sigma) - 1/4$. Furthermore, let $\sigma^* := 2\ell \cdot \sigma = \sum_{i=1}^{\ell} (\nu_i^0 + \nu_i^1)$. Observe that $|\mathcal{A}_{i+1}| \geq |\mathcal{A}_i| + 1 + \nu_i^0 + \nu_i^1$ and hence $|\mathcal{A}_\ell| \geq \ell + \sigma^*$. Thus, $|\mathcal{A}_\ell| \leq 3/16N$ implies $\ell + \sigma^* \leq 3/16N \Leftrightarrow \sigma^* \leq 1/16N$ as $\ell = 1/8N$ by definition.

Putting everything together gives

$$\begin{aligned} p^* &\leq Pr[\sigma^* < 1/16 \cdot N] = Pr[2\ell \cdot \sigma < 1/16 \cdot N] && (17)\\ &= Pr[\sigma < 1/4] \leq Pr[\sigma < \mathbf{E}(\sigma) - 1/4] = Pr[\mathbf{E}(\sigma) - \sigma > 1/4] && (18)\\ &< e^{-2/16 \cdot 2s} = e^{-1/32 \cdot N}. && (19) \end{aligned}$$

*Remark 1.* Observe that one key ingredient of the proof was to show that with a high probability, the number of disjoint queries cannot grow above $15/16N$. Intuitively, one might expect that this bound is highly overrated. Indeed, computer simulations indicated that on average, only about $N/2$ disjoint queries are possible. If this bound holds in general (which is currently an open question), this would yield better concrete bounds with respect to the preimage resistance.

## 5 Applications

We now discuss how our results can be applied to derive better bounds for existing double call, double length hash functions. To this end, we consider two well-known constructions given by Hirose [8, 7]. These are depicted in Figure 1. One sees easily that the design on

[3] Note that $\mathcal{R}_i^0 = \{0,1\}^n \setminus \mathcal{X}_{i-1}$ for $T_i \in \{I, II\}$ and $\mathcal{R}_i^0 = \{0,1\}^n \setminus \mathcal{Y}_{i-1}$ if $T_i = I$. Note further that $\mathcal{R}_i^1 = \{0,1\}^n \setminus \mathcal{Y}'_{i-1}$ for $T_i \in \{I, II\}$ and $\mathcal{R}_i^1 = \{0,1\}^n \setminus \mathcal{X}_{i-1}$ for $T_i = III$.

the left side is comprised by the $H_3$ construction since the xor-operation with a non-zero value is clearly a fixed point free involution. Regarding the design on the right side, it is in principle the $H_1$ construction. Thus, the according bounds derived in Section 3 hold immediately.

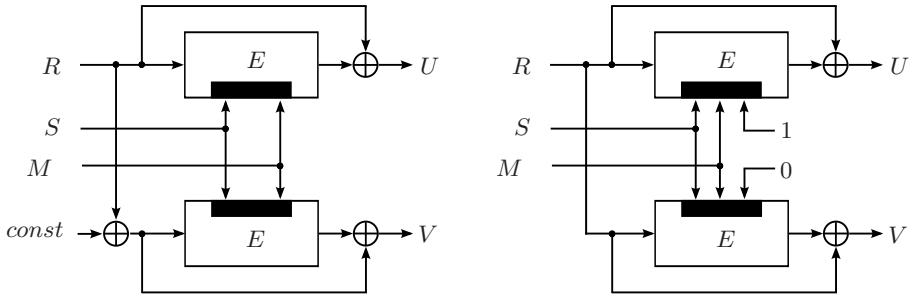


**Fig. 1.** Double call, double length examples from literature.

## 6 Discussion and Conclusion

In this work, we developed and applied new techniques for determining lower bounds with respect to preimage resistance. For the considered constructions, the given results outmatch significantly the best known bounds.

Despite this landmark result, there are still a lot of challenges open in the field of block cipher based hashing. For example, is it possible to show that the $\Omega(2^{2n})$ bound on the preimage resistance of $H_1$ does hold for any attacker and/or that the conjecture stated in Remark 1 is true? Likewise, can this bound be extended to the other constructions considered in this work, i.e., $H_2$ and $H_3$? More general, can our techniques be adapted for assessing other known constructions like Abreast-DM or Tandem-DM? Going one step further, one may ask whether other interesting generalizations are possible, as, *e.g.*, the alleviation of the necessity of two keys. Closely related is the question of how these techniques can be applied to single call double length hash functions.

## References


[1] Noga Alon and Joel Spencer. *The Probabilistic Method*. John Wiley, 1992.

[2] John Black, Phillip Rogaway, and Thomas Shrimpton. Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV. In Moti Yung, editor, *CRYPTO*, volume 2442 of *Lecture Notes in Computer Science*, pages 320–335. Springer, 2002.

[3] Orr Dunkelman, editor. *Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers*, volume 5665 of *Lecture Notes in Computer Science*. Springer, 2009.

[4] Shimon Even and Yishay Mansour. A Construction of a Cipher From a Single Pseudorandom Permutation. In Hideki Imai, Ronald L. Rivest, and Tsutomu Matsumoto, editors, *ASIACRYPT*, volume 739 of *Lecture Notes in Computer Science*, pages 210–224. Springer, 1991.

[5] Ewan Fleischmann, Michael Gorski, and Stefan Lucks. On the Security of Tandem-DM. In Dunkelman [3], pages 84–103.
[6] Ewan Fleischmann, Michael Gorski, and Stefan Lucks. Security of cyclic double block length hash functions. In Parker [18], pages 153–175.
[7] Shoichi Hirose. Provably Secure Double-Block-Length Hash Functions in a Black-Box Model. In Choonsik Park and Seongtaek Chee, editors, *ICISC*, volume 3506 of *Lecture Notes in Computer Science*, pages 330–342. Springer, 2004.
[8] Shoichi Hirose. Some Plausible Constructions of Double-Block-Length Hash Functions. In Matthew J. B. Robshaw, editor, *FSE*, volume 4047 of *Lecture Notes in Computer Science*, pages 210–225. Springer, 2006.
[9] Joe Kilian and Phillip Rogaway. How to Protect DES Against Exhaustive Key Search. In Neal Koblitz, editor, *CRYPTO*, volume 1109 of *Lecture Notes in Computer Science*, pages 252–267. Springer, 1996.
[10] Lars R. Knudsen, Florian Mendel, Christian Rechberger, and Søren S. Thomsen. Cryptanalysis of mdc-2. In Antoine Joux, editor, *EUROCRYPT*, volume 5479 of *Lecture Notes in Computer Science*, pages 106–120. Springer, 2009.
[11] Xuejia Lai and James L. Massey. Hash Function Based on Block Ciphers. In *EUROCRYPT*, pages 55–70, 1992.
[12] Jooyoung Lee and Daesung Kwon. The security of abreast-dm in the ideal cipher model. Cryptology ePrint Archive, Report 2009/225, 2009. `http://eprint.iacr.org/`.
[13] Jooyoung Lee, Martijn Stam, and John Steinberger. The collision security of tandem-dm in the ideal cipher model. Cryptology ePrint Archive, Report 2010/409, 2010. `http://eprint.iacr.org/`.
[14] Stefan Lucks. A collision-resistant rate-1 double-block-length hash function. In Eli Biham, Helena Handschuh, Stefan Lucks, and Vincent Rijmen, editors, *Symmetric Cryptography*, volume 07021 of *Dagstuhl Seminar Proceedings*. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2007.
[15] Alfred Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.
[16] Ralph C. Merkle. One Way Hash Functions and DES. In Gilles Brassard, editor, *CRYPTO*, volume 435 of *Lecture Notes in Computer Science*, pages 428–446. Springer, 1989.
[17] Onur Özen and Martijn Stam. Another glance at double-length hashing. In Parker [18], pages 176–201.
[18] Matthew G. Parker, editor. *Cryptography and Coding, 12th IMA International Conference, Cryptography and Coding 2009, Cirencester, UK, December 15-17, 2009. Proceedings*, volume 5921 of *Lecture Notes in Computer Science*. Springer, 2009.
[19] Bart Preneel. Mdc-2 and mdc-4. In Henk C. A. van Tilborg, editor, *Encyclopedia of Cryptography and Security*. Springer, 2005.
[20] Phillip Rogaway and Thomas Shrimpton. Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance. In Bimal K. Roy and Willi Meier, editors, *FSE*, volume 3017 of *Lecture Notes in Computer Science*, pages 371–388. Springer, 2004.
[21] Phillip Rogaway and John P. Steinberger. Constructing Cryptographic Hash Functions from Fixed-Key Blockciphers. In David Wagner, editor, *CRYPTO*, volume 5157 of *Lecture Notes in Computer Science*, pages 433–450. Springer, 2008.
[22] Phillip Rogaway and John P. Steinberger. Security/Efficiency Tradeoffs for Permutation-Based Hashing. In Nigel P. Smart, editor, *EUROCRYPT*, volume 4965 of *Lecture Notes in Computer Science*, pages 220–236. Springer, 2008.
[23] Martijn Stam. Blockcipher based hashing revisited. Cryptology ePrint Archive, Report 2008/071, 2008. `http://eprint.iacr.org/`.
[24] Martijn Stam. Blockcipher-based hashing revisited. In Dunkelman [3], pages 67–83.
[25] John P. Steinberger. The Collision Intractability of MDC-2 in the Ideal-Cipher Model. In Moni Naor, editor, *EUROCRYPT*, volume 4515 of *Lecture Notes in Computer Science*, pages 34–51. Springer, 2007.