Collision attacks against 22-step SHA-512

Somitra Kumar Sanadhya* and Palash Sarkar

Applied Statistics Unit,
Indian Statistical Institute,
203, B.T. Road, Kolkata,
India 700108.
somitra_r@isical.ac.in, palash@isical.ac.in

3^{rd} April 2008

Abstract. In this work, we present two attacks against 22-step SHA-512. Our first attack succeeds with probability about 2^{-8} whereas the second attack is deterministic. To construct the attack, we use a single local collision and handle conditions on the colliding pair of messages. All but one condition can be satisfied deterministically in our first attack while in the second attack all conditions can be satisfied deterministically. There are four free words in our second attack and hence we get exactly 2^{256} collisions for 22-step SHA-512.

Recently, attacks against up to 24-step SHA-256 have been reported in the literature which use a local collision given earlier by Nikolić and Biryukov at FSE'08. We provide evidence which shows that using this local collision is unlikely to produce collisions for step reduced SHA-512. Consequently, our attacks are currently the best against reduced round SHA-512. The same attacks also work against SHA-256. Since our second attack is a deterministic construction, it is also the best attack against 22-step SHA-256.

Keywords: Cryptanalysis, SHA-2 hash family, reduced round attacks

1 Introduction

At FSE '08, Nikolić and Biryukov [2] presented attacks against 20-step SHA-256 and 21-step SHA-256. Their 20-step attack succeeds with probability about 1/3 and 21-step attack with probability about 2^{-19}. Using the local collision of [2], Indesteege et al. [1] developed 23-step and 24-step attacks against SHA-256. We note that no 22-step or more attack against SHA-512 has been published in the literature till date. In this work, we present two attacks against 22-step SHA-256 and 22-step SHA-512. We use two variations of another local collision, given recently by Sanadhya and Sarkar [4], to develop our two attacks. The first one is a probabilistic attack while the second one is deterministic. We also show that the 22, 23 and 24 step SHA-256 attacks described in [1] (in various versions) are not likely to succeed against SHA-512.

2 Notation

In this paper we use the following notation:

* This author is supported by the Ministry of Information Technology, Govt. of India.

- Message words: $W_i \in \{0,1\}^n$, $W'_i \in \{0,1\}^n$; n is 32 for SHA-256 and 64 for SHA-512.
- Colliding message pair: $\{W_0, W_1, W_2, \ldots W_{15}\}$ and $\{W'_0, W'_1, W'_2, \ldots W'_{15}\}$.
- Expanded message pair: $\{W_0, W_1, W_2, \ldots W_{N-1}\}$ and $\{W'_0, W'_1, W'_2, \ldots W'_{N-1}\}$. The number of steps N is 64 for SHA-256 and 80 for SHA-512.
- The internal registers for the two messages at step i: $\text{REG}_i = \{a_i, \ldots, h_i\}$ and $\text{REG'}_i = \{a'_i, \ldots, h'_i\}$.
- $\text{ROTR}^k(x)$: Right rotation of an n-bit string x by k bits.
- $\text{SHR}^k(x)$: Right shift of an n-bit string x by k bits.
- $\oplus$: bitwise XOR.
- $+, -$: addition and subtraction modulo 2^n.
- $\delta X = X' - X$ where X is an n-bit quantity.
- $\delta\Sigma_1(e_i) = \Sigma_1(e'_i) - \Sigma_1(e_i)$.
- $\delta\Sigma_0(a_i) = \Sigma_0(a'_i) - \Sigma_0(a_i)$.
- $\delta f^i_{MAJ}(x,y,z)$: Output difference of the f_{MAJ} function in step i when its inputs differ by x, y and z. That is, $\delta f^i_{MAJ}(x,y,z) = f_{MAJ}(a_i + x, b_i + y, c_i + z) - f_{MAJ}(a_i, b_i, c_i)$.
- $\delta f^i_{IF}(x,y,z)$: Output difference of the f_{IF} function in step i when its inputs differ by x, y and z. That is, $\delta f^i_{IF}(x,y,z) = f_{IF}(e_i + x, f_i + y, g_i + z) - f_{IF}(e_i, f_i, g_i)$.

3 The SHA-2 Hash Family

The SHA-2 hash function was standardized by NIST in 2002 [5]. There are 2 differently designed functions in this standard: the SHA-256 and SHA-512. The number in the name of the hash function refers to the length of message digest produced by that function. Next we briefly describe the structure of SHA-2.

Eight registers are used in the evaluation of SHA-2. The initial value in the registers is specified by an $8 \times n$ bit IV, n=32 for SHA-256 and 64 for SHA-512. In Step i, the 8 registers are updated from (a_{i-1}, b_{i-1}, c_{i-1}, d_{i-1}, e_{i-1}, f_{i-1}, g_{i-1}, h_{i-1}) to (a_i, b_i, c_i, d_i, e_i, f_i, g_i, h_i) according to the following Equations:

$$\left.\begin{aligned}
a_i &= \Sigma_0(a_{i-1}) + f_{MAJ}(a_{i-1}, b_{i-1}, c_{i-1}) + \Sigma_1(e_{i-1}) + f_{IF}(e_{i-1}, f_{i-1}, g_{i-1}) \\
&\quad + h_{i-1} + K_i + W_i \\
b_i &= a_{i-1} \\
c_i &= b_{i-1} \\
d_i &= c_{i-1} \\
e_i &= d_{i-1} + \Sigma_1(e_{i-1}) + f_{IF}(e_{i-1}, f_{i-1}, g_{i-1}) + h_{i-1} + K_i + W_i \\
f_i &= e_{i-1} \\
g_i &= f_{i-1} \\
h_i &= g_{i-1}
\end{aligned}\right\} \quad (1)$$

The initial register values $\{a_{-1}, b_{-1}, \ldots h_{-1}\}$ are specified by the IV. The f_{IF} and the f_{MAJ} are three variable bitwise boolean functions defined as:

$$\begin{aligned}
f_{IF}(x,y,z) &= (x \wedge y) \oplus (\neg x \wedge z), \\
f_{MAJ}(x,y,z) &= (x \wedge y) \oplus (y \wedge z) \oplus (z \wedge x).
\end{aligned}$$

For SHA-256, the functions Σ_0 and Σ_1 are defined as:

$$\begin{aligned}
\Sigma_0(x) &= ROTR^2(x) \oplus ROTR^{13}(x) \oplus ROTR^{22}(x), \\
\Sigma_1(x) &= ROTR^6(x) \oplus ROTR^{11}(x) \oplus ROTR^{25}(x).
\end{aligned}$$

For SHA-512, the corresponding functions are:

$$\begin{aligned}
\Sigma_0(x) &= ROTR^{28}(x) \oplus ROTR^{34}(x) \oplus ROTR^{39}(x), \\
\Sigma_1(x) &= ROTR^{14}(x) \oplus ROTR^{18}(x) \oplus ROTR^{41}(x).
\end{aligned}$$

Round i uses a n-bit word W_i which is derived from the message and a constant word K_i. There are $N = 64$ steps in SHA-256 and $N = 80$ steps in SHA-512. The hash function operates on a 512-bit (resp. 1024-bit) block specified as 16 words of 32 (resp. 64) bits for SHA-256 (resp. SHA-512). Given the message words m_0, m_1, $\ldots m_{15}$, the W_i 's are computed using the Equation:

$$W_i = \begin{cases} m_i & \text{for } 0 \leq i \leq 15 \\ \sigma_1(W_{i-2}) + W_{i-7} + \sigma_0(W_{i-15}) + W_{i-16} & \text{for } 16 \leq i \leq (N-1) \end{cases} \tag{2}$$

For SHA-256, the functions σ_0 and σ_1 are defined as:

$$\sigma_0(x) = ROTR^7(x) \oplus ROTR^{18}(x) \oplus SHR^3(x),$$
$$\sigma_1(x) = ROTR^{17}(x) \oplus ROTR^{19}(x) \oplus SHR^{10}(x).$$

And for SHA-512, they are defined as:

$$\sigma_0(x) = ROTR^1(x) \oplus ROTR^8(x) \oplus SHR^7(x),$$
$$\sigma_1(x) = ROTR^{19}(x) \oplus ROTR^{61}(x) \oplus SHR^6(x).$$

The IV $= (a_{-1}, b_{-1}, c_{-1}, d_{-1}, e_{-1}, f_{-1}, g_{-1}, h_{-1})$ is defined by the standard to be some random looking constant words.

The output hash value of a one block (512-bit for SHA-256 and 1024-bit for SHA-512) message is obtained by chaining the IV with the register values at the end of the final round as per the Merkle-Damgård construction. A similar strategy is used for multi-block messages, where the IV for next block is taken as the hash output of the previous block. For complete details of the SHA-2 family, see [5].

An important relationship between register values From Equation 1, we get:

$$e_i = a_{i-4} + a_i - \Sigma_0(a_{i-1}) - f_{MAJ}(a_{i-1}, a_{i-2}, a_{i-3}). \tag{3}$$

We call this relationship "the cross dependence equation (CDE)".

4 Nonlinear Local Collision for SHA-2

We use two variations of a 9-step non-linear local collision for our attacks. This local collision was given recently by Sanadhya and Sarkar [4]. This local collision starts by introducing a perturbation message difference of 1 in the first message word. Next eight message words are chosen suitably to obtain the desired differential path. Table 1 shows the local collision used. The message word differences are different for the two variations of the local collision. Columns headed I and II under δW_i in Table 1 show the message word differences for the first and the second variations of the local collision respectively.

In the local collision, the registers (a_{i-1}, $\ldots$, h_{i-1}) and W_i are inputs to Step i of the hash evaluation and this step outputs the registers (a_i, $\ldots$, h_i).

4.1 Conditions on the Differential Path of Table 1 for the First Attack (Column I)

The message word differences, $\delta W_{i+1}, \delta W_{i+2}$, δW_{i+3} and δW_{i+7} are computed from the following equations:

$$\delta W_{i+1} = -1 - \delta f_{IF}^{i}(1, 0, 0) - \delta\Sigma_1(e_i), \tag{4}$$

$$\delta W_{i+2} = -1 - \delta f_{IF}^{i+1}(-1, 1, 0) - \delta\Sigma_1(e_{i+1}), \tag{5}$$

$$\delta W_{i+3} = -\delta f_{IF}^{i+2}(-1, -1, 1) - \delta\Sigma_1(e_{i+2}), \tag{6}$$

Table 1. The 9-step Sanadhya-Sarkar local collision [4] used in the present work. Our two attacks use unequal message word differences to achieve the same differential path.

Step	δW_i		Register differences							
	I	II	δa_i	δb_i	δc_i	δd_i	δe_i	δf_i	δg_i	δh_i
$i-1$	0	0	0	0	0	0	0	0	0	0
i	1	1	1	0	0	0	1	0	0	0
$i+1$	δW_{i+1}	δW_{i+1}	0	1	0	0	-1	1	0	0
$i+2$	δW_{i+2}	0	0	0	1	0	-1	-1	1	0
$i+3$	δW_{i+3}	δW_{i+3}	0	0	0	1	0	-1	-1	1
$i+4$	0	0	0	0	0	0	1	0	-1	-1
$i+5$	0	0	0	0	0	0	0	1	0	-1
$i+6$	0	0	0	0	0	0	0	0	1	0
$i+7$	δW_{i+7}	0	0	0	0	0	0	0	0	1
$i+8$	-1	-1	0	0	0	0	0	0	0	0

$$\delta W_{i+7} = -\delta f_{IF}^{i+6}(0,0,1). \tag{7}$$

Intermediate registers need to satisfy the following conditions:

$$\begin{aligned} &a_{i-2} = a_{i-1} = a_i = -1,\ a_{i+1} = a_{i+2} = 0, \\ &e_{i+2} = 0,\ e_{i+3} = e_{i+4} = e_{i+5} = -1. \end{aligned} \tag{8}$$

All the conditions in Equation 8 can be deterministically satisfied by choosing message words carefully. This ensures the success probability of 1 for this local collision. These conditions can be derived in the same way as in [2]. The same local collision has also been used by Sanadhya and Sarkar recently to attack 20-step SHA-512 in [4] and 21-step SHA-512 in [3]. The derivation of these conditions is similar to the derivation of conditions for the next variation of the local collision that is used in Section 4.2 ahead.

Satisfying the conditions Note that, using Equation 1, the message word W_k can be chosen to set either a_k or e_k to a desired value. Therefore the conditions on a_{i-2}, a_{i-1}, a_i, a_{i+1}, a_{i+2}, e_{i+3}, e_{i+4} and e_{i+5} can be satisfied deterministically. After this, we are left with condition on e_{i+2} only. Next we show that this condition is satisfied automatically.

From Equation 3, we get:

$$\begin{aligned} e_{i+2} &= a_{i-2} + a_{i+2} - \Sigma_0(a_{i+1}) - f_{MAJ}(a_{i+1}, a_i, a_{i-1}) \\ &= -1 + 0 - \Sigma_0(0) - f_{MAJ}(0, -1, -1) \\ &= 0. \end{aligned}$$

4.2 Conditions on the Differential Path of Table 1 for the second attack (Column II)

The message word differences, δW_{i+1} and δW_{i+3} are computed from the following equations:

$$\delta W_{i+1} = -1 - \delta f_{IF}^{i}(1,0,0) - \delta\Sigma_1(e_i), \tag{9}$$

$$\delta W_{i+3} = -\delta f_{IF}^{i+2}(-1,-1,1) - \delta\Sigma_1(e_{i+2}). \tag{10}$$

Intermediate registers need to satisfy the following conditions:

$$\begin{aligned} &a_{i-3} = -2, a_{i-2} = a_{i-1} = a_i = -1,\ a_{i+1} = a_{i+2} = 0, \\ &e_{i+1} = 0, e_{i+2} = 0, e_{i+3} = e_{i+4} = e_{i+5} = e_{i+6} = -1, \\ &e_i - e_{i-1} + 1 = 0. \end{aligned} \tag{11}$$

All the conditions in Equation 11 can be deterministically satisfied by choosing message words carefully. This ensures the success probability of 1 for this local collision. These conditions can be derived in the same way as in [2]. Detailed derivation of these conditions is provided in Section A.

Satisfying the conditions As in Section 4.1, conditions on a_{i-3}, a_{i-2}, a_{i-1}, a_i, a_{i+2}, e_{i+3}, e_{i+4}, e_{i+5} and e_{i+6} can be satisfied deterministically. After this, we are left with conditions on e_{i+1}, e_{i+2} and e_i only. Two out of these three conditions are satisfied automatically as shown next.

From Equation 3, we get:

$$\begin{aligned} e_{i+1} &= a_{i-3} + a_{i+1} - \Sigma_0(a_i) - f_{MAJ}(a_i, a_{i-1}, a_{i-2}) \\ &= -2 + 0 - \Sigma_0(-1) - f_{MAJ}(-1, -1, -1) \\ &= 0. \end{aligned}$$

Similarly, we get $e_{i+2} = 0$. Now we consider the last remaining condition $e_i - e_{i-1} + 1 = 0$. Using Equation 3, we get:

$$\begin{aligned} 0 &= 1 + e_i - e_{i-1} \\ &= 1 + (a_{i-4} + a_i - \Sigma_0(a_{i-1}) - f_{MAJ}(a_{i-1}, a_{i-2}, a_{i-3})) \\ &\quad -(a_{i-5} + a_{i-1} - \Sigma_0(a_{i-2}) - f_{MAJ}(a_{i-2}, a_{i-3}, a_{i-4})) \\ &= a_{i-4} - a_{i-5} + 2 + f_{MAJ}(-1, -2, a_{i-4}). \end{aligned}$$

This implies,

$$a_{i-5} = a_{i-4} + 2 + f_{MAJ}(-1, -2, a_{i-4}). \tag{12}$$

Equation 12 defines the register value a_{i-5} in terms of the register value a_{i-4}. But a_{i-4} will be computed only after a_{i-5} is available. To resolve this, we first choose any arbitrary value for a_{i-4} first and then compute the required value of a_{i-5}. From Equation 1, we can ensure the deterministic success of the required condition using the free words W_{i-5} and W_{i-4}.

Note Even though both local collisions hold with probability 1, it does not imply that they give rise to collisions on higher number of rounds with probability 1. The difference will become clear later when we discuss the two attacks.

5 The Probabilistic Attack

In [2], a single local collision spanning from Step 6 to Step 14 is used and a 21-step collision for SHA-256 is obtained probabilistically. We use a similar method for our attack but this time we use the local collision of Table 1 spanning from Step 8 to Step 16. Message words are given by Column (I). The SHA-2 design has freedom of message words W_0 to W_{15} only. The rest of the message words are generated by message recursion. Therefore we put restrictions on the messages chosen so that the word W_{16} has the desired differential behaviour.

First of all, note that the local collision starts from Step 8. It can be seen from the structure of the local collision that $\delta W_8 = 1$ and $\delta W_{12} = \delta W_{13} = \delta W_{14} = 0$. In addition, δW_{16} is expected to be -1. Messages outside the span of the local collision are taken to have zero differentials. Therefore $\delta W_i = 0$ for $i \in \{0, 1, 2, 3, 4, 5, 6, 7\}$. Consider the first 6 steps of message expansion for SHA-2 next.

$$\left.\begin{aligned} W_{16} &= \sigma_1(W_{14}) + \underline{W_9} + \sigma_0(W_1) + W_0, \\ W_{17} &= \underline{\sigma_1(W_{15}) + W_{10}} + \sigma_0(W_2) + W_1, \\ W_{18} &= \underline{\sigma_1(W_{16}) + W_{11}} + \sigma_0(W_3) + W_2, \\ W_{19} &= \underline{\sigma_1(W_{17})} + W_{12} + \sigma_0(W_4) + W_3, \\ W_{20} &= \underline{\sigma_1(W_{18})} + W_{13} + \sigma_0(W_5) + W_4, \\ W_{21} &= \underline{\sigma_1(W_{19})} + W_{14} + \sigma_0(W_6) + W_5. \end{aligned}\right\} \tag{13}$$

Terms which *may have* non-zero differentials in the above equations are underlined. To obtain 22-step collisions in SHA-2, it is sufficient to ensure that the following conditions are satisfied:

1. $\delta W_9 = -1$ so that $\delta W_{16} = -1$ and the local collision terminates at Step 16 as desired.
2. $\delta\{\sigma_1(W_{15}) + W_{10}\} = 0$ so that $\delta W_{17} = 0$.
3. $\delta\{\sigma_1(W_{16}) + W_{11}\} = 0$ so that $\delta W_{18} = 0$.

Conditions 2 and 3 above ensure that next 3 steps of the message expansion will not produce any difference, and we will have a 22-step collision.

We satisfy each of the three conditions above one by one.

5.1 Ensuring $\delta\{\sigma_1(W_{15}) + W_{10}\} = 0$

We need the following condition to be satisfied:

$$\sigma_1(W_{15}) - \sigma_1(W_{15} + \delta W_{15}) = \delta W_{10}. \tag{14}$$

The local collision defines the register values $e_{12} = e_{13} = -1$ and the difference δW_{15} as follows:

$$\begin{aligned}
\delta W_{15} &= -\delta f_{IF}^{14}(0,0,1) \\
&= -f_{IF}(e_{14}, f_{14}, g_{14} + 1) + f_{IF}(e_{14}, f_{14}, g_{14}) \\
&= -f_{IF}(e_{14}, e_{13}, e_{12} + 1) + f_{IF}(e_{14}, e_{13}, e_{12}) \\
&= -f_{IF}(e_{14}, -1, -1 + 1) + f_{IF}(e_{14}, -1, -1) \\
&= -f_{IF}(e_{14}, -1, 0) + f_{IF}(e_{14}, -1, -1) \\
&= -e_{14} + (-1) \\
&= -(e_{14} + 1).
\end{aligned}$$

It can be seen from Equation 1 that e_{14} can be governed by W_{14}. The message word W_{14} is free and hence we can have any value of e_{14}, and consequently any value of δW_{15}, by appropriate choice of W_{14}.

Next, suppose $W_{15} = -\delta W_{15}$. In this case, Equation 14 gives $\sigma_1(W_{15}) - \sigma_1(0) = \delta W_{10}$. That is, $\sigma_1(W_{15}) = \delta W_{10}$.

The difference δW_{10} will be available after Step 10 of the hash evaluation for the two messages. At that point itself, we can choose $W_{15} = \sigma_1^{-1}(\delta W_{10})$ and $\delta W_{15} = -W_{15}$. This implies that we need $e_{14} = \sigma_1^{-1}(\delta W_{10}) - 1$. This allows the satisfaction of the required condition deterministically.

The particular solution suggested above will work only if we can invert the 32×32 bit map σ_1 for SHA-256 and the 64×64 bit map σ_1 for SHA-512. We note that the map σ_1 is a linear function. Therefore $\sigma_1(x)$ can be expressed as multiplication of a matrix with x. For both these hash functions, the corresponding matrix is of full rank. Therefore σ_1 is indeed invertible.

5.2 Ensuring $\delta W_9 = -1$

As already remarked, having $\delta W_9 = -1$ ensures that $\delta W_{16} = -1$. This will terminate the local collision successfully if other conditions have been fulfilled.

From the condition of the local collision, we have that:

$$\begin{aligned}
\delta W_9 &= -1 - f_{IF}^{8}(1,0,0) - \delta\Sigma_1(e_8) \\
&= -1 - f_{IF}(e_8 + 1, f_8, g_8) + f_{IF}(e_8, f_8, g_8) - \Sigma_1(e_8 + 1) + \Sigma_1(e_8) \\
&= -1 - f_{IF}(e_8 + 1, e_7, e_6) + f_{IF}(e_8, e_7, e_6) - \Sigma_1(e_8 + 1) + \Sigma_1(e_8).
\end{aligned}$$

If we can have $e_8 = -1$ then the above expression can be simplified considerably. We also have from the conditions of the local collision that $a_6 = a_7 = a_8 = -1$. Therefore, from the CDE, we have that:

$$\begin{aligned} e_8 &= a_4 + a_8 - \Sigma_1(a_7) - f_{MAJ}(a_7, a_6, a_5) \\ &= a_4 - 1 - \Sigma_1(-1) - f_{MAJ}(-1, -1, a_5) \\ &= a_4 - 1 + 1 + 1 \\ &= a_4 + 1. \end{aligned}$$

Hence, to get $e_8 = -1$, we will set $a_4 = -2$. This can be achieved since a_4 can be influenced by the message word W_4.

When $e_8 = -1$, then the expression for δW_9 simplifies to:

$$\begin{aligned} \delta W_9 &= -1 - f_{IF}(0, e_7, e_6) + f_{IF}(-1, e_7, e_6) - \Sigma_1(0) + \Sigma_1(-1) \\ &= -1 - e_6 + e_7 - 0 + (-1) \\ &= e_7 - e_6 - 2. \end{aligned} \tag{15}$$

The CDE can also be used for e_7. In this case, we get:

$$\begin{aligned} e_7 &= a_3 + a_7 - \Sigma_0(a_6) - f_{MAJ}(a_6, a_5, a_4) \\ &= a_3 - 1 - \Sigma_0(-1) - f_{MAJ}(-1, a_5, -2) \\ &= a_3 - 1 + 1 - f_{MAJ}(-1, a_5, -2) \\ &= a_3 - f_{MAJ}(-1, a_5, -2). \end{aligned}$$

Suppose we also choose $a_3 = -1$, then the CDE for e_6 gives:

$$\begin{aligned} e_6 &= a_2 + a_6 - \Sigma_0(a_5) - f_{MAJ}(a_5, a_4, a_3) \\ &= a_2 - 1 - \Sigma_0(a_5) - f_{MAJ}(a_5, -2, -1). \end{aligned}$$

Hence, we get:

$$\begin{aligned} e_7 - e_6 - 1 &= a_3 - a_2 + \Sigma_0(a_5) - f_{MAJ}(-1, a_5, -2) + f_{MAJ}(a_5, -2, -1) \\ &= -1 - a_2 + \Sigma_0(a_5). \end{aligned}$$

From Equation 15, getting $\delta W_9 = -1$ means getting $e_7 - e_6 - 1 = 0$. The above expression simplifies this condition to $-1 - a_2 + \Sigma_0(a_5) = 0$.

This is a kind of dependency which can be solved in a deterministic way. The message words W_2 and W_5 are both free. Using the word W_5 we can obtain any desired value of a_5. Let a_5 be fixed a-priori to a value α. This implies, we need $a_2 = \Sigma_0(\alpha) - 1$. The free word W_2 can be used to obtain this value which will be available to us a-priori. Therefore it is possible to satisfy the condition $\delta W_9 = \delta W_{16} = -1$ in a deterministic way.

5.3 Ensuring $\delta\{\sigma_1(W_{16}) + W_{11}\} = 0$

We need the following condition to be satisfied:

$$\sigma_1(W_{16}) - \sigma_1(W_{16} - 1) = \delta W_{11}. \tag{16}$$

In Equation 16 above, we have assumed that $\delta W_{16} = -1$. The satisfaction of this condition for δW_{16} has already been discussed.

We consider the right hand side of Equation 16 now.

$$\begin{aligned}
\delta W_{11} &= -\delta f_{IF}^{10}(-1,-1,1) - \delta\Sigma_1(e_{10}) \\
&= -f_{IF}(e_{10}-1, f_{10}-1, g_{10}+1) + f_{IF}(e_{10}, f_{10}, g_{10}) - \Sigma_1(e_{10}-1) + \Sigma_1(e_{10}) \\
&= -f_{IF}(e_{10}-1, e_9-1, e_8+1) + f_{IF}(e_{10}, e_9, e_8) - \Sigma_1(e_{10}-1) + \Sigma_1(e_{10}) \\
&= -f_{IF}(-1, e_9-1, e_8+1) + f_{IF}(0, e_9, e_8) - \Sigma_1(-1) + \Sigma_1(0) \\
&= -(e_9-1) + e_8 - (-1) + 0 \\
&= -e_9 + e_8 + 2.
\end{aligned}$$

By using the CDE for e_9 and e_8, the above expression can be simplified to:

$$\delta W_{11} = -(a_5 + 1).$$

The register value a_5 has already been chosen to be an a-priori fixed value α so that the condition $\delta W_9 = -1$ is satisfied. Note that we have not yet specified the exact value of α.

Refer to Equation 13 for W_{16}. The message word W_{16} is derived from words W_0, W_1, W_9 and W_{14}. We have generated words W_9 and W_{14} in some specific ways to satisfy conditions on the desired differential path. However, words W_0 and W_1 are allowed to be random. This suggests that the word W_{16} should be random. *It is interesting to note that despite the apparent randomness of* W_{16}*, the term* $\sigma_1(W_{16}) - \sigma_1(W_{16}-1)$ *is highly non-random.*

We observe that some values of $\sigma_1(W_{16}) - \sigma_1(W_{16}-1)$ occur very frequently. Some such values of this term are listed in Table 2.

Table 2. Some frequently occurring values of $\sigma_1(W_{16}) - \sigma_1(W_{16}-1)$ for SHA-256 and SHA-512.

SHA-256				SHA-512			
1	0001a000	2	fffe2000	1	ffffdfffffffffe8	2	fffe00000000008
3	00602000	4	005a3000	3	0000dfffffffffff8	4	00001fffffffff8
5	fffca000	6	00006000	5	0006200000000098	6	0000200000000028
7	ffffa000	8	00002000	7	ffffdfffffffffa8	8	ffff200000000038

We want to have a value of $\sigma_1(W_{16}) - \sigma_1(W_{16}-1) = \beta$ (say) such that it is equal to δW_{11}. The term δW_{11} is equal to $-(a_5+1)$. Thus we want $a_5 = -\beta - 1$. We had chosen a_5 to be an arbitrary value α earlier. Now we choose a specific value of α which is equal to $-\beta - 1$, where β is one of the values from Table 2. If a run of the attack produces the correct value of $\delta\sigma_1(W_{16})$ then this step has been satisfied, otherwise we repeat the process. We observe that this step succeeds with very high probability and within few runs of the attack, we get the desired value of $\delta\sigma_1(W_{16})$. This is the only step in the attack which requires probabilistic satisfaction. All the other conditions can be fulfilled deterministically.

5.4 Summary of conditions required for the first attack

The local collision is started from Step 8 and ends at Step 16. The register values and some of the message words/ differences must be as listed below. First two sets of conditions below are imposed by the local collision used. Refer to Equation 8 where the starting step of the local collision is $i = 8$. Next four sets of conditions are required to have the message expansion steps from Step 16 to Step 21 to behave as desired.

1. $a_6 = a_7 = a_8 = -1$, $a_9 = a_{10} = 0$.

2. $e_{10} = 0$, $e_{11} = e_{12} = e_{13} = -1$.
3. $W_{15} = \sigma_1^{-1}(\delta W_{10})$, $e_{14} = \sigma_1^{-1}(\delta W_{10}) - 1$.
4. $e_8 = -1$, $a_4 = -2$, $a_3 = -1$.
5. $a_5 = \alpha$, $a_2 = \Sigma_0(\alpha) - 1$, where $\alpha = -\beta - 1$ and β is one of the values from Table 2.
6. Hope to get: $\sigma_1(W_{16}) - \sigma_1(W_{16} - 1) = \beta$.

5.5 Algorithm to obtain 22-step collisions

Recall that Equation 1 is used at Step i of the hash evaluation. Registers $(a_{i-1}, b_{i-1}, \ldots h_{i-1})$ are available at this step and the output register a_i or e_i can be controlled by selecting W_i suitably. For instance, if we wish to make a_i to be zero, then we can calculate the suitable value of W_i from Equation 1 which will make this happen. We define two functions which return the required message word W_i to set the register value a_i or e_i to desired values, say `desired_a` and `desired_e`, at Step i. Equation 1 provides the definitions of these two functions.

1. `W_to_set_register_A(Step` i`, desired_a, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\})$:
   $=$ (`desired_a` $-\Sigma_0(a_{i-1}) - f_{MAJ}(a_{i-1}, b_{i-1}, c_{i-1}) - \Sigma_1(e_{i-1}) - f_{IF}(e_{i-1}, f_{i-1}, g_{i-1}) - h_{i-1} - K_i$)
2. `W_to_set_register_E(Step` i`, desired_e, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\})$:
   $=$ (`desired_e` $-d_{i-1} - \Sigma_1(e_{i-1}) \; - f_{IF}(e_{i-1}, f_{i-1}, g_{i-1}) - h_{i-1} - K_i$)

The algorithm to obtain message pairs leading to 22-step collisions for SHA-2 family in described in Table 3.

5.6 Success probability of the attack

All but one condition of the 22-step attack described above can be fulfilled deterministically. The only step left is to satisfy $\delta\sigma_1(W_{16}) = \delta W_{11}$. This step is also likely to succeed if we choose a suitable β value in the attack. We used $\beta =$ `0001a000` for the 22-step SHA-256 attack and $\beta =$ `ffffdfffffffffe8` for the 22-step SHA-512 attack. For these chosen values, our 22-step SHA-256 attack succeeds with average probability of $2^{-4.86}$ and our 22-step SHA-512 attack succeeds with average probability of $2^{-5.46}$. The worst case probability for these attacks were $2^{-7.05}$ and $2^{-7.92}$ respectively. All the probability estimates are experimental values averaged over 2^{15} trials. It appears possible to use some other suitable value of β and improve on the success probability. We have not yet experimented with a large number of values of β.

6 The Deterministic Attack

In Section 5, a single local collision spanning from Step 8 to 16 was used. In contrast, this time we use the second variation (i.e., the message words are given by Column (II)) of the same local collision spanning from Step 7 to Step 15. The SHA-2 design has freedom of message words W_0 to W_{15}. Since the local collision spans this range only, we can deterministically satisfy all the conditions from Equation 11. The message words after Step 16 are generated by message expansion. The local collision is chosen in such a way that the message expansion produces no difference in words W_i and W_i' for $i \in \{16, 17, \ldots 21\}$. This results in a deterministic 22-step attack. We explain this fact below.

First of all, note that the local collision starts from Step 7 and ends at Step 15. It can be seen from the structure of the local collision that $\delta W_7 = 1$, $\delta W_{15} = -1$ and $\delta W_9 = \delta W_{11} = \delta W_{12} = \delta W_{13} = \delta W_{14} = 0$. Messages outside the span of the local collision are taken to have zero differentials. Therefore $\delta W_i = 0$ for $i \in \{0, 1, 2, 3, 4, 5, 6\}$. Next consider the first 6 steps of message expansion for SHA-2 from (13). To obtain 22-step collisions in SHA-2, it is sufficient to ensure that $\delta\{\sigma_1(W_{15}) + W_{10}\} = 0$ so that $\delta W_{17} = 0$. This also ensures that next 4 steps of the message expansion do not produce any difference, and we have a 22-step collision.

Table 3. Probabilistic algorithm to obtain message pairs leading to collisions for 22-step SHA-2. This corresponds to our first attack.

external `W_to_set_register_A(Step` i`, desired_a, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\}$`)` :
Returns the required message W_i to be used in step i so that a_i is set to the given value.

external `W_to_set_register_E(Step` i`, desired_e, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\}$`)` :
Returns the required message W_i to be used in step i so that e_i is set to the given value.

First Message words:

1. Select W_0 and W_1 randomly.
2. Choose $\alpha = -\beta - 1$ where β is one of the values from Table 2.
3. Run Steps 0 and 1 of hash evaluation to define $\{a_1, b_1, \ldots h_1\}$.
4. Choose W_2 = `W_to_set_register_A(2,` $\Sigma_0(\alpha) - 1$`,` $\{a_1, b_1, \ldots h_1\}$`)`.
5. Run Step 2 of hash evaluation to define $\{a_2, b_2, \ldots h_2\}$.
6. Choose W_3 = `W_to_set_register_A(3,` -1`,` $\{a_2, b_2, \ldots h_2\}$`)`.
7. Run Step 3 of hash evaluation to define $\{a_3, b_3, \ldots h_3\}$.
8. Choose W_4 = `W_to_set_register_A(4,` -2`,` $\{a_3, b_3, \ldots h_3\}$`)`.
9. Run Step 4 of hash evaluation to define $\{a_4, b_4, \ldots h_4\}$.
10. Choose W_5 = `W_to_set_register_A(5,` α`,` $\{a_4, b_4, \ldots h_4\}$`)`.
11. Run Step 5 of hash evaluation to define $\{a_5, b_5, \ldots h_5\}$.
12. Choose W_6 = `W_to_set_register_A(6,` -1`,` $\{a_5, b_5, \ldots h_5\}$`)`.
13. Run Step 6 of hash evaluation to define $\{a_6, b_6, \ldots h_6\}$.
14. Choose W_7 = `W_to_set_register_A(7,` -1`,` $\{a_6, b_6, \ldots h_6\}$`)`.
15. Run Step 7 of hash evaluation to define $\{a_7, b_7, \ldots h_7\}$.
16. Choose W_8 = `W_to_set_register_A(8,` -1`,` $\{a_7, b_7, \ldots h_7\}$`)`.
17. Run Step 8 of hash evaluation to define $\{a_8, b_8, \ldots h_8\}$.
18. Choose W_9 = `W_to_set_register_A(9,` 0`,` $\{a_8, b_8, \ldots h_8\}$`)`.
19. Run Step 9 of hash evaluation to define $\{a_9, b_9, \ldots h_9\}$.
20. Compute $\delta W_{10} = -1 - f_{IF}(e_9 - 1, f_9 + 1, g_9) + f_{IF}(e_9, f_9, g_9) - \Sigma_1(e_9 - 1) + \Sigma_1(e_9)$. (Refer Equation 5)
21. Choose W_{10} = `W_to_set_register_A(10,` 0`,` $\{a_9, b_9, \ldots h_9\}$`)`.
22. Run Step 10 of hash evaluation to define $\{a_{10}, b_{10}, \ldots h_{10}\}$.
23. Choose W_{11} = `W_to_set_register_E(11,` -1`,` $\{a_{10}, b_{10}, \ldots h_{10}\}$`)`.
24. Run Step 11 of hash evaluation to define $\{a_{11}, b_{11}, \ldots h_{11}\}$.
25. Choose W_{12} = `W_to_set_register_E(12,` -1`,` $\{a_{11}, b_{11}, \ldots h_{11}\}$`)`.
26. Run Step 12 of hash evaluation to define $\{a_{12}, b_{12}, \ldots h_{12}\}$.
27. Choose W_{13} = `W_to_set_register_E(13,` -1`,` $\{a_{12}, b_{12}, \ldots h_{12}\}$`)`.
28. Run Step 13 of hash evaluation to define $\{a_{13}, b_{13}, \ldots h_{13}\}$.
29. Choose W_{14} = `W_to_set_register_E(14,` $\sigma_1^{-1}(\delta W_{10}) - 1$`,` $\{a_{12}, b_{12}, \ldots h_{12}\}$`)`.
30. Run Step 14 of hash evaluation to define $\{a_{14}, b_{14}, \ldots h_{14}\}$.
31. Choose $W_{15} = \sigma_1^{-1}(\delta W_{10})$.
32. Run Step 15 of hash evaluation to define $\{a_{15}, b_{15}, \ldots h_{15}\}$.
33. Compute $W_{16} = \sigma_1(W_{14}) + W_9 + \sigma_0(W_1) + W_0$.
34. If $\sigma_1(W_{16}) - \sigma_1(W_{16} - 1) = \beta$ then the attack has succeeded, hence proceed to compute the second message. Otherwise go back to Step 1 above.

Second message words:

35. Define $\delta W_i = 0$ for $i \in \{0, 1, 2, 3, 4, 5, 6, 7, 12, 13, 14\}$.
36. Define $\delta W_8 = 1$ and $\delta W_9 = \delta W_{16} = -1$. Step 20 above has already defined δW_{10}
37. Define $\delta W_{11} = -f_{IF}(e_{10} - 1, f_{10} - 1, g_{10} + 1) + f_{IF}(e_{10}, f_{10}, g_{10}) - \Sigma_1(e_{10} - 1) + \Sigma_1(e_{10})$. (Refer Equation 6)
38. Define $\delta W_{15} = -W_{15}$.
39. Compute $W_i' = W_i + \delta W_i$ for $0 \leq i \leq 15$.

6.1 Ensuring $\delta\{\sigma_1(W_{15}) + W_{10}\} = 0$

Since $\delta W_{15} = -1$, we need the following condition to be satisfied:

$$\sigma_1(W_{15}) - \sigma_1(W_{15} - 1) = \delta W_{10}. \tag{17}$$

The local collision defines the register values as per Equation 11. The message word difference δW_{10} is defined by Equation 10. Simplifying this expression (where the starting step $i = 7$), we get:

$$\begin{aligned}
\delta W_{10} &= -\delta f^9_{IF}(-1,-1,1) - \Sigma_1(e_9) \\
&= -f_{IF}(e_9 - 1, f_9 - 1, g_9 + 1) + f_{IF}(e_9, f_9, g_9) - \Sigma_1(e_9 - 1) + \Sigma_1(e_9) \\
&= -f_{IF}(e_9 - 1, e_8 - 1, e_7 + 1) + f_{IF}(e_9, e_8, e_7) - \Sigma_1(e_9 - 1) + \Sigma_1(e_9) \\
&= -f_{IF}(-1, -1, e_7 + 1) + f_{IF}(0, 0, e_7) - \Sigma_1(-1) + \Sigma_1(0) \\
&= -(-1) + e_7 - (-1) + 0 \\
&= e_7 + 2.
\end{aligned}$$

Using the CDE, we can express the right hand side of the above expression as:

$$\begin{aligned}
e_7 + 2 &= a_3 + a_7 - \Sigma_0(a_6) - f_{MAJ}(a_6, a_5, a_4) + 2 \\
&= a_3 + (-1) - \Sigma_0(-1) - f_{MAJ}(-1, -1, -2) + 2 \\
&= a_3 - 1 - (-1) - (-1) + 2 \\
&= a_3 + 3.
\end{aligned}$$

In satisfying Equation 12, we needed to choose any arbitrary value of $a_{i-4} = a_3$ first. The above analysis implies that we can deterministically satisfy $\delta(\sigma_1(W_{15}) + W_{10})$ as follows:

1. Choose any arbitrary value for W_{15}. This defines the difference $\sigma_1(W_{15}) - \sigma_1(W_{15} - 1)$. Let it be called "`DELTA`".
2. From Equation 17, δW_{10} must take the value `DELTA`. This can be obtained by setting $a_3 =$ `DELTA` $- 3$.

Using the functions defined in Section 5.5, the algorithm for obtaining colliding message pairs is described in Table 4.

7 Infeasibility of the Nikolić and Biryukov Local Collision for SHA-512

In the previous section, we described a deterministic attack against 22-step SHA-2 using a local collision different from the one described by Nikolić and Biryukov [2]. This local collision has recently been used to attack 22, 23 and 24-step SHA-256 in different versions of [1]. The authors of [1] remark that their attacks should also succeed against SHA-512. We show this to be unlikely, i.e., we show that it is unlikely that the Nikolić-Biryukov local collision can be used to obtain 22, 23 or 24-step SHA-512 collisions.

The Nikolić-Biryukov local collision is shown in Table 5.

7.1 Conditions on the Diff. Path of Table 5

The message word differences $\delta W_{i+1}, \delta W_{i+2}$ and δW_{i+3} are computed from the following equations:

$$\delta W_{i+1} = -1 - \delta f^i_{IF}(1,0,0) - \delta\Sigma_1(e_i), \tag{18}$$

$$\delta W_{i+2} = -\delta f^{i+1}_{IF}(-1,1,0) - \delta\Sigma_1(e_{i+1}), \tag{19}$$

Table 4. Deterministic algorithm to obtain message pairs leading to collisions for 22-step SHA-2. This corresponds to our second attack.

external `W_to_set_register_A(Step` i`, desired_a, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\}$`)` :
Returns the required message W_i to be used in step i so that a_i is set to the given value.

external `W_to_set_register_E(Step` i`, desired_e, Current State` $\{a_{i-1}, b_{i-1}, \ldots h_{i-1}\}$`)` :
Returns the required message W_i to be used in step i so that e_i is set to the given value.

First Message words:

1. Select W_0, W_1, W_{14} and W_{15} randomly.
2. Set `DELTA` $= \sigma_1(W_{15}) - \sigma_1(W_{15} - 1)$.
3. Run Steps 0 and 1 of hash evaluation to define $\{a_1, b_1, \ldots h_1\}$.
4. Choose W_2 = `W_to_set_register_A(2, DELTA` $- 1 + f_{MAJ}(-1, -2,$ `DELTA` $- 3)$, $\{a_1, b_1, \ldots h_1\})$.
5. Run Step 2 of hash evaluation to define $\{a_2, b_2, \ldots h_2\}$.
6. Choose W_3 = `W_to_set_register_A(3, DELTA` $- 3$, $\{a_2, b_2, \ldots h_2\})$.
7. Run Step 3 of hash evaluation to define $\{a_3, b_3, \ldots h_3\}$.
8. Choose W_4 = `W_to_set_register_A(4,` -2, $\{a_3, b_3, \ldots h_3\})$.
9. Run Step 4 of hash evaluation to define $\{a_4, b_4, \ldots h_4\}$.
10. Choose W_5 = `W_to_set_register_A(5,` -1, $\{a_4, b_4, \ldots h_4\})$.
11. Run Step 5 of hash evaluation to define $\{a_5, b_5, \ldots h_5\}$.
12. Choose W_6 = `W_to_set_register_A(6,` -1, $\{a_5, b_5, \ldots h_5\})$.
13. Run Step 6 of hash evaluation to define $\{a_6, b_6, \ldots h_6\}$.
14. Choose W_7 = `W_to_set_register_A(7,` -1, $\{a_6, b_6, \ldots h_6\})$.
15. Run Step 7 of hash evaluation to define $\{a_7, b_7, \ldots h_7\}$.
16. Choose W_8 = `W_to_set_register_A(8,` 0, $\{a_7, b_7, \ldots h_7\})$.
17. Run Step 8 of hash evaluation to define $\{a_8, b_8, \ldots h_8\}$.
18. Choose W_9 = `W_to_set_register_A(9,` 0, $\{a_8, b_8, \ldots h_8\})$.
19. Run Step 9 of hash evaluation to define $\{a_9, b_9, \ldots h_9\}$.
20. Choose W_{10} = `W_to_set_register_E(10,` -1, $\{a_9, b_9, \ldots h_9\})$.
21. Run Step 10 of hash evaluation to define $\{a_{10}, b_{10}, \ldots h_{10}\}$.
22. Choose W_{11} = `W_to_set_register_E(11,` -1, $\{a_{10}, b_{10}, \ldots h_{10}\})$.
23. Run Step 11 of hash evaluation to define $\{a_{11}, b_{11}, \ldots h_{11}\}$.
24. Choose W_{12} = `W_to_set_register_E(12,` -1, $\{a_{11}, b_{11}, \ldots h_{11}\})$.
25. Run Step 12 of hash evaluation to define $\{a_{12}, b_{12}, \ldots h_{12}\}$.
26. Choose W_{13} = `W_to_set_register_E(13,` -1, $\{a_{12}, b_{12}, \ldots h_{12}\})$.

Second message words:

27. Define $\delta W_i = 0$ for $i \in \{0, 1, 2, 3, 4, 5, 6, 9, 11, 12, 13, 14\}$.
28. Define $\delta W_7 = 1$ and $\delta W_{15} = -1$.
29. Define $\delta W_8 = -1 - f_{IF}(e_7 + 1, f_7, g_7) + f_{IF}(e_7, f_7, g_7) - \Sigma_1(e_7 + 1) + \Sigma_1(e_7)$. (Refer Equation 9)
30. Define $\delta W_{10} = -f_{IF}(e_9 - 1, f_9 - 1, g_9 + 1) + f_{IF}(e_9, f_9, g_9) - \Sigma_1(e_9 - 1) + \Sigma_1(e_9)$. (Refer Equation 10)
31. Compute $W_i' = W_i + \delta W_i$ for $0 \leq i \leq 15$.

Table 5. The 9-step local collision due to Nikolić and Biryukov.

Step i	δW_i	δa_i	δb_i	δc_i	δd_i	δe_i	δf_i	δg_i	δh_i
$i-1$	0	0	0	0	0	0	0	0	0
i	1	1	0	0	0	1	0	0	0
$i+1$	δW_{i+1}	0	1	0	0	-1	1	0	0
$i+2$	δW_{i+2}	0	0	1	0	0	-1	1	0
$i+3$	δW_{i+3}	0	0	0	1	0	0	-1	1
$i+4$	0	0	0	0	0	1	0	0	-1
$i+5$	0	0	0	0	0	0	1	0	0
$i+6$	0	0	0	0	0	0	0	1	0
$i+7$	0	0	0	0	0	0	0	0	1
$i+8$	-1	0	0	0	0	0	0	0	0

$$\delta W_{i+3} = -\delta f_{IF}^{i+2}(0, -1, 1). \tag{20}$$

Intermediate registers need to satisfy the following conditions:

$$\begin{aligned} &a_{i-2} = a_{i-1} = a_{i+1} = a_{i+2},\ a_i = -1, \\ &e_{i+2} = e_{i+3},\ e_{i+4} = -1,\ e_{i+5} = 0,\ e_{i+6} = -1. \end{aligned} \tag{21}$$

In addition, an extra condition needs to be satisfied:

$$\delta f_{IF}^{i+3}(0, 0, -1) = -1. \tag{22}$$

All the conditions in Equation 21 can be deterministically satisfied by choosing message words carefully, but the condition in Equation 22 needs to be satisfied probabilistically. This causes the success probability of 1/3 for this local collision. For details refer to [2]. Note that our notation and indexing of the steps is different from [2].

7.2 Analysis of this local collision

An earlier version of [1] described a 22-step attack on SHA-256 by using the Nikolić-Biryukov local collision spanning from Steps 8 to 16. The 23 and 24 step attacks against SHA-256 avilable in the current version of [1] are obtained by shifting the local collision by one step and two steps respectively. To summarize, by spanning the Nikolić-Biryukov local collision between Step i and Step $i+8$, SHA-256 has been attacked upto $i+14$ steps.

Since the local collison ends at Step $i+8$, from the differential path of the local collision, we require the difference in the message word δW_{i+8} to be -1. The basic idea, due to Nikolić and Biryukov [2], is to ensure that the message word differences are all zero after the local collision ends. This will ensure that the two messages will not introduce any difference in the registers. Therefore all the attacks in [1] require that $\delta W_{i+9} = \delta W_{i+10} = \ldots = \delta W_{i+14} = 0$.

Now consider the second step after the local collision has ended. It can be seen from Equation 13 that we require $\delta(\sigma_1(W_{i+8}) + W_{i+3}) = 0$ to ensure that $\delta W_{i+10} = 0$. Recall that the local collision is started from Step i and δW_{i+3} is given by Equation 20.

Note that the attacks described in [1] first construct a pseudo-collision and then extend it to a collision for reduced round SHA-256. Regardless of this different attack strategy, the satisfaction of the condition $\delta W_{i+10} = 0$ will be required for the success of their $(i+14)$-step attack, either for SHA-256 or for SHA-512.

We now show the difficulty of finding values of δW_{i+3} and $\delta\sigma_1(W_{i+8})$ which are of the same order of magnitude. The values of δW_{i+3} are biased towards small magnitudes. In contrast, the values of $\sigma_1(W_{i+8}) - \sigma_1(W_{i+8} - 1)$ for SHA-512 are biased towards large magnitudes. This makes it difficult to achieve equality of the two terms as required to ensure $\delta W_{i+10} = 0$. Now we provide concrete proofs for these facts.

Magnitude of δW_{i+3} values: We first state two results which help in understanding the bias of δW_{i+3} in this case. In the discussion that follows, we use X_i to denote the i^{th} bit of a 64-bit quantity X. We also use the convention that the index of the least significant bit is 0.

Proposition 1 *$Pr[P_j \neq (P+1)_j] = 1/2^j$, where the probability is taken over random P.*

Proposition 2 *If two numbers X and Y are such that $X_i \neq Y_i$ and $X_{i-1} = Y_{i-1}$, then $|X - Y| \geq 2^{i-1} + 1$.*

Next we prove that the probability that the absolute value of δW_{i+3}, when using Nikolić-Biryukov local collision, is larger than 2^j is bounded above by $1/2^{j-1}$.

Lemma 1 *If the Nikolić-Biryukov local collision is started at Step i, then* $Pr[|\delta W_{i+3}| \geq 2^j] < 1/2^{j-1}$.

Proof. Since the local collision is started from step i, the message difference δW_{i+3} is given by Equation 20. This equation gives:

$$\begin{aligned}\delta W_{i+3} &= -\delta f_{IF}^{i+2}(0,-1,1),\\ &= -f_{IF}(e_{i+2}, f_{i+2}-1, g_{i+2}+1) + f_{IF}(e_{i+2}, f_{i+2}, g_{i+2}),\\ &= -f_{IF}(e_{i+2}, e_{i+1}-1, e_i+1) + f_{IF}(e_{i+2}, e_{i+1}, e_i).\end{aligned}$$

The two f_{IF} terms in the computation above have the same first argument e_{i+2}. The second and the third arguments have a modular difference of ± 1. If the j^{th} bit of e_{i+2} is 1 then the two f_{IF} functions will select the corresponding bit from the middle argument, else from the third argument.

Let $A = f_{IF}(e_{i+2}, e_{i+1}-1, e_i+1)$ and $B = f_{IF}(e_{i+2}, e_{i+1}, e_i)$. Further, let P_n be the event that $A_n \neq B_n$. The event $\delta W_{i+3} \geq 2^j$ can happen if and only if at least one of the bits $j, j+1, \ldots 63$ of δW_{i+3} is 1, i.e., if and only if at least one of the events $P_j, P_{j+1}, \ldots P_{63}$ holds.

Now we are ready to bound the probability of the required event. In the fourth step below, we use the fact that $f_{IF}(a,b,c) = b$ if $a = 1$ and $= c$ if $a = 0$.

$$\begin{aligned}Pr[\delta W_{i+3} \geq 2^j] &= Pr[\bigcup_{i \geq j} P_i]\\ &\leq \sum_{i \geq j} Pr[P_i]\\ &= \sum_{i \geq j} (Pr[(e_{i+2})_i = 0] \cdot Pr[P_i | ((e_{i+2})_i = 0)] + Pr[(e_{i+2})_i = 1] \cdot Pr[P_i | ((e_{i+2})_i = 1)])\\ &= \sum_{i \geq j} \left(\frac{1}{2} \cdot Pr[(e_i+1)_i \neq e_i] + \frac{1}{2} \cdot Pr[(e_{i+1}-1)_i \neq e_{i+1}] \right)\\ &= \frac{1}{2} \cdot \sum_{i \geq j} \left(\frac{1}{2^i} + \frac{1}{2^i} \right) \quad \text{(Using Proposition 1)}\\ &< \frac{1}{2^{j-1}}.\end{aligned}$$

This proves the Lemma. □

Magnitude of $\sigma_1(W) - \sigma_1(W-1)$ values: We now look at the distribution of values of $\sigma_1(W) - \sigma_1(W-1)$ for random choices of W. The function σ_1 is defined for SHA-512 as:

$$\sigma_1(W) = ROTR^{19}(W) \oplus ROTR^{61}(W) \oplus SHR^6(W). \tag{23}$$

Let the 64-bit word W be specified as $(w_{63}, w_{62}, \ldots, w_1, w_0)$ where w_0 is the least significant bit of W. Then $\sigma_1(W)$ can be expressed as bit-wise XOR of three quantities. By using the combinatorial structure of the function σ_1 and partial search using a computer program, it is possible to prove the following lemma.

Lemma 2 *For the function* σ_1 *used in SHA-512,*

$$|\sigma_1(W) - \sigma_1(W-1)| \geq (2^{42} + 2^{39} + 2^{38} + 2^{36} - 2^3),$$

where W *is any 64-bit word.*

For detailed proof of this lemma, refer to [3].

7.3 Infeasibility of the attacks described in [1] for SHA-512

From Lemma 1, we get that the probability that a value of δW_{i+3} produced when using this local collision is larger than 2^{42} is less than $1/2^{41}$. That is, on average one will require 2^{41} or more attempts with the differential path to get a value of δW_{i+3} which is larger than 2^{42}. On the other hand, Lemma 2 shows that all the values of $\sigma_1(W_{i+8}) - \sigma_1(W_{i+8} - 1)$ will be larger than 2^{42}.

In addition, in our analysis we observe that the term $\sigma_1(W_{i+8}) - \sigma_1(W_{i+8} - 1)$ for any value of W_{i+8} has a peculiar and patterned structure which is far from random. Our experiments support this view further. For instance, we experimentally observed that this difference of σ_1 terms has a large trail of zero bits or a large trail of one bits in the middle. The number of ones or zeros in the continuous sequence are almost always between 20 to 35. Further, there are many 64-bit words which occur repeatedly as the value of this term for different choices of W_{i+8}. Also, some values are never achieved. It is not clear whether it is possible to achieve such a strongly structured pattern in δW_{i+3} and ensure $\delta W_{i+10} = 0$ with the use of the Nikolić-Biryukov local collision for SHA-512.

Note that this local collision succeeds for the SHA-256 case because the choice of the two rotation values used in the σ_1 function for SHA-256 are not far apart. This causes most of the bits to overlap over nearby bits and the bias of the term $\sigma_1(W_{i+8}) - \sigma_1(W_{i+8} - 1)$ is not as skewed as in the case of SHA-512.

8 Some Concluding Remarks

In this work we have presented two attacks against 22-step SHA-2. Our first attack is probabilistic while the second attack is deterministic. Both these attacks can be successfully used to find collisions for 22-step SHA-512. To the best of our knowledge, these are the currently best attacks against SHA-512. The previous approach, due to Nikolić and Biryukov [2] have been used to find 23 and 24-round collisions for SHA-256 in [1]. In contrast, we show that the Nikolić-Biryukov approach is unlikely to succeed in obtaining 22, 23 or 24-round SHA-512 collisions.

References

1. Sebastiaan Indesteege, Florian Mendel, Bart Preneel, and Christian Rechberger. Collisions and other Non-Random Properties for Step-Reduced SHA-256. *Cryptology eprint Archive*, April 2008. Available at `http://eprint.iacr.org/2008/131`.
2. Ivica Nikolić and Alex Biryukov. Collisions for Step-Reduced SHA-256. In Kaisa Nyberg, editor, *Fast Software Encryption, 15th International Workshop, FSE 2008, Lausanne, Switzerland, March 26-28, 2008*, volume Pre-proceedings version of *Lecture Notes in Computer Science*, pages 1–16. Springer, 2008.
3. Somitra Kumar Sanadhya and Palash Sarkar. Deterministic Constructions of 21-Step Collisions for the SHA-2 Hash Family. In Tzong-Chen Wu and Chin-Laung Lei, editors, *Information Security, 11th International Conference, ISC 2008, Taipei, Taiwan, September 15-18, 2008, Proceedings*, volume To appear of *Lecture Notes in Computer Science*. Springer, 2008.
4. Somitra Kumar Sanadhya and Palash Sarkar. Non-Linear Reduced Round Attacks Against SHA-2 Hash family. In Yi Mu and Willy Susilo, editors, *Information Security and Privacy - ACISP 2008, The 13th Australasian Conference, Wollongong, Australia, 7-9 July 2008, Proceedings*, volume To appear of *Lecture Notes in Computer Science*. Springer, 2008.
5. Secure Hash Standard. *Federal Information Processing Standard Publication 180-2*. U.S. Department of Commerce, National Institute of Standards and Technology(NIST), 2002. Available at `http://csrc.nist.gov/publications/fips/fips180-2/fips180-2withchangenotice.pdf`.

A Message Word Differences for Table 1, Column (II)

In Step i of SHA-2, only the registers a_i and e_i are computed. Rest of the registers are copies of the old ones. Therefore we focus on these two register evaluations only. From Equation 1, we get:

$$\delta e_i = \delta W_i + \delta\Sigma_1(e_{i-1}) + \delta f_{IF}(\delta e_{i-1}, \delta f_{i-1}, \delta g_{i-1}) + \delta d_{i-1} + \delta h_{i-1}, \tag{24}$$

$$\begin{aligned}\delta a_i &= \delta W_i + \delta\Sigma_0(a_{i-1}) + \delta f_{MAJ}(\delta a_{i-1}, \delta b_{i-1}, \delta c_{i-1}) + \delta\Sigma_1(e_{i-1}) + \\ &\quad \delta f_{IF}(\delta e_{i-1}, \delta f_{i-1}, \delta g_{i-1}) + \delta h_{i-1}, \\ &= \delta\Sigma_0(a_{i-1}) + \delta f_{MAJ}(\delta a_{i-1}, \delta b_{i-1}, \delta c_{i-1}) + \delta e_i - \delta d_{i-1}. \end{aligned} \tag{25}$$

We now try to satisfy the restriction imposed by the differential path of Table 1 by defining suitable difference of the message words in various steps.

Step i : If $\delta W_i = 1$, then this difference will propagate to both the registers a_i and e_i.

Step (i+1) : At this step $a'_i - a_i = e'_i - e_i = 1$. We want $\delta a_{i+1} = 0$ and $\delta e_{i+1} = -1$. From Equations 25 and 24, we get:

$$\begin{aligned}\delta a_{i+1} = \ 0 &= \delta\Sigma_0(a_i) + \delta f^i_{MAJ}(1,0,0) + \delta\Sigma_1(e_i) + \delta f^i_{IF}(1,0,0) + \delta W_{i+1}, \\ \delta e_{i+1} = -1 &= \delta\Sigma_1(e_i) + \delta f^i_{IF}(1,0,0) + \delta W_{i+1}.\end{aligned}$$

The conditions above translate to:

$$-1 = -\delta\Sigma_0(a_i) - \delta f^i_{MAJ}(1,0,0), \tag{26}$$

$$\delta W_{i+1} = -1 - \delta f^i_{IF}(1,0,0) - \delta\Sigma_1(e_i). \tag{27}$$

Step (i+2) : At this step $\delta W_{i+2} = 0$, $b'_{i+1} - b_{i+1} = f'_{i+1} - f_{i+1} = 1$ and $e'_{i+1} - e_{i+1} = -1$. We want $\delta a_{i+2} = 0$ and $\delta e_{i+2} = -1$. From Equations 25 and 24, we get:

$$\begin{aligned}\delta a_{i+2} = 0 &= \delta f^{i+1}_{MAJ}(0,1,0) + \delta\Sigma_1(e_{i+1}) + \delta f^{i+1}_{IF}(-1,1,0) + 0, \\ \delta e_{i+2} = -1 &= \delta\Sigma_1(e_{i+1}) + \delta f^{i+1}_{IF}(-1,1,0) + 0.\end{aligned}$$

The conditions above translate to:

$$-1 = -\delta f^{i+1}_{MAJ}(0,1,0), \tag{28}$$

$$0 = -1 - \delta f^{i+1}_{IF}(-1,1,0) - \delta\Sigma_1(e_{i+1}). \tag{29}$$

Step (i+3) : At this step $c'_{i+2} - c_{i+2} = g'_{i+2} - g_{i+2} = 1$, $e'_{i+2} - e_{i+2} = -1$ and $f'_{i+2} - f_{i+2} = -1$. We want $\delta a_{i+3} = 0$ and $\delta e_{i+3} = 0$. From Equations 25 and 24, we get:

$$\begin{aligned}\delta a_{i+3} = 0 &= \delta f^{i+2}_{MAJ}(0,0,1) + \delta\Sigma_1(e_{i+2}) + \delta f^{i+2}_{IF}(-1,-1,1) + \delta W_{i+3}, \\ \delta e_{i+3} = 0 &= \delta\Sigma_1(e_{i+2}) + \delta f^{i+2}_{IF}(-1,-1,1) + \delta W_{i+3}.\end{aligned}$$

The conditions above translate to:

$$\delta f^{i+2}_{MAJ}(0,0,1) = 0, \tag{30}$$

$$\delta W_{i+3} = -\delta f^{i+2}_{IF}(-1,-1,1) - \delta\Sigma_1(e_{i+2}). \tag{31}$$

Step (i+4) : At this step $\delta W_{i+4} = 0$, $d'_{i+3} - d_{i+3} = h'_{i+3} - h_{i+3} = 1$, $f'_{i+3} - f_{i+3} = -1$ and $g'_{i+3} - g_{i+3} = -1$. We want $\delta a_{i+4} = 0$ and $\delta e_{i+4} = 1$. From Equations 25 and 24, we get:

$$\begin{aligned}\delta a_{i+4} = 0 &= \delta f^{i+3}_{IF}(0,-1,-1) + 1 + 0, \\ \delta e_{i+4} = 1 &= \delta f^{i+3}_{IF}(0,-1,-1) + 1 + 1 + 0.\end{aligned}$$

The conditions above translate to:

$$0 = -1 - \delta f^{i+3}_{IF}(0,-1,-1). \tag{32}$$

Step (i+5) : At this step $\delta W_{i+5} = 0$, $e'_{i+4} - e_{i+4} = 1$, $g'_{i+4} - g_{i+4} = -1$ and $h'_{i+4} - h_{i+4} = -1$. We want $\delta a_{i+5} = \delta e_{i+5} = 0$. From Equations 25 and 24, we get:

$$\delta a_{i+5} = 0 = \delta\Sigma_1(e_{i+4}) + \delta f_{IF}^{i+4}(1, 0, -1) - 1 + 0,$$
$$\delta e_{i+5} = 0 = \delta\Sigma_1(e_{i+4}) + \delta f_{IF}^{i+4}(1, 0, -1) - 1 + 0.$$

The conditions above translate to:

$$0 = 1 - \delta f_{IF}^{i+4}(1, 0, -1) - \delta\Sigma_1(e_{i+4}). \tag{33}$$

Step (i+6) : At this step $\delta W_{i+6} = 0$, $f'_{i+5} - f_{i+5} = 1$ and $h'_{i+5} - h_{i+5} = -1$. We want $\delta a_{i+6} = \delta e_{i+6} = 0$. From Equations 25 and 24, we get:

$$\delta a_{i+6} = 0 = \delta f_{IF}^{i+5}(0, 1, 0) - 1 + 0,$$
$$\delta e_{i+6} = 0 = \delta f_{IF}^{i+5}(0, 1, 0) - 1 + 0.$$

The conditions above translate to:

$$0 = 1 - \delta f_{IF}^{i+5}(0, 1, 0). \tag{34}$$

Step (i+7) : At this step $\delta W_{i+7} = 0$, $g'_{i+6} - g_{i+6} = x$. We want $\delta a_{i+7} = \delta e_{i+7} = 0$. From Equations 25 and 24, we get:

$$\delta a_{i+7} = 0 = \delta f_{IF}^{i+6}(0, 0, 1) + 0,$$
$$\delta e_{i+7} = 0 = \delta f_{IF}^{i+6}(0, 0, 1) + 0.$$

The conditions above translate to:

$$0 = -\delta f_{IF}^{i+6}(0, 0, 1). \tag{35}$$

Step (i+8) : At this step $h'_{i+7} - h_{i+7} = 1$. We want $\delta a_{i+8} = \delta e_{i+8} = 0$. This will happen as desired if we have:

$$\delta W_{i+8} = -1. \tag{36}$$

A.1 Solution of Equations

To find the local collision, we need message pairs which will satisfy Equations 26 to 36. We use techniques similar to [2] to derive sufficient conditions to ensure these.

- Equation 26 is satisfied if we have $a_i = -1$ (and hence from the differential path $a'_i = 0$) and $a_{i-1} = a_{i-2}$. This ensures that the $\delta\Sigma_0$ term propagates a difference of -1 and $\delta f_{MAJ}^{i} = 0$.
- Equation 28 is satisfied by ensuring that $a_{i+1} = 0$, $a_i = a_{i-1} = -1$ (and hence from the differential path $a'_i = 0$).
- Equation 30 is satisfied by ensuring $a_{i+2} = a_{i+1}$.
- Equation 32 is satisfied if $\delta f_{IF}^{i+3}(0, -1, -1) = -1$. This happens as desired if we have $e_{i+3} = -1$ so that the middle argument of f_{IF}^{i+3} is selected which has a difference of -1.
- Equation 33 is satisfied if $\delta\Sigma_1(e_{i+4}) = 1$ and $\delta f_{IF}^{i+4}(1, 0, -1) = 0$. The condition on $\Sigma_1(e_{i+4})$ is satisfied if we can have $e_{i+4} = -1$ and $e'_{i+4} = 0$. For the δf_{IF}^{i+4} condition, we need:

$$\begin{aligned} & f_{IF}^{i+4}(e_{i+4} + 1, e_{i+3}, e_{i+2} - 1) - f_{IF}^{i+4}(e_{i+4}, e_{i+3}, e_{i+2}) = 0 \\ \Rightarrow\ & f_{IF}^{i+4}(0, e_{i+3}, e_{i+2} - 1) - f_{IF}^{i+4}(-1, e_{i+3}, e_{i+2}) = 0 \\ \Rightarrow\ & e_{i+2} - 1 = e_{i+3}. \end{aligned}$$

Since e_{i+3} has already been taken to be -1, we need $e_{i+2} = 0$.

– Equation 34 is satisfied if $\delta f_{IF}^{i+5}(0,1,0) = 1$. This condition implies:

$$f_{IF}^{i+5}(e_{i+5}, e_{i+4}+1, e_{i+3}) - f_{IF}^{i+5}(e_{i+5}, e_{i+4}, e_{i+3}) = 1.$$

Since the middle term of f_{IF}^{i+5} has a difference of 1 which we want to be propagated, we need to ensure $e_{i+5} = -1$. This causes the middle term to be selected by the f_{IF} function.

– Equation 35 is satisfied if the f_{IF} always selects its middle argument which has zero difference. This will happen if we have $e_{i+6} = -1$.

– Equation 29 is satisfied if

$$\begin{aligned} 0 &= -1 - \delta f_{IF}^{i+1}(-1,1,0) - \delta \Sigma_1(e_{i+1}) \\ &= -1 - f_{IF}(e_{i+1}-1, e_i+1, e_{i-1}) + f_{IF}(e_{i+1}, e_i, e_{i-1}) - \Sigma_1(e_{i+1}-1) + \Sigma_1(e_{i+1}). \end{aligned}$$

From Section 4.2, we have that $e_{i+1} = 0$. Therefore the above condition simplifies to:

$$\begin{aligned} 0 &= -1 - f_{IF}(-1, e_i+1, e_{i-1}) + f_{IF}(0, e_i, e_{i-1}) - \Sigma_1(-1) + \Sigma_1(0) \\ &= -1 - (e_i+1) + e_{i-1} - (-1) + 0 \\ &= -1 - e_i + e_{i-1}. \end{aligned}$$

– Equations 27, 31 and 36 merely define message word differences. No condition is imposed by them.

The conditions derived above correspond to Equations 9, 10 and 11.

B Colliding message pairs

Colliding message pairs for 22-step SHA-512 and 22-step SHA-256 generated by the algorithm of Tables 3 and 4 are provided in Tables 6, 7, 8 and 9 respectively.

Table 6. Colliding message pair for 22-step SHA-512 with standard IV. These messages have been generated using the algorithm of Table 3.

W_1	0-3	0000000000000000	550b57fb514c9b79	cff1882089fc9d67	2810605c1bd7ed0c
	4-7	75ad3c8a1c93c5ed	f20d9ad5246bd372	24bfb9a1eb7aceff	f15320a1acd4b2f0
	8-11	92aaca629e0027df	fe30a1bcb92fedda	db6c1a412c9b4d4d	aaf3823c2a004b1f
	12-15	8d41a28b0d847693	7f212e01c4e96937	7eeeca5c84ba3bda	1acad103aa814e0e
W_2	0-3	0000000000000000	550b57fb514c9b79	cff1882089fc9d67	2810605c1bd7ed0c
	4-7	75ad3c8a1c93c5ed	f20d9ad5246bd372	24bfb9a1eb7aceff	f15320a1acd4b2f0
	8-11	92aaca629e0027e0	fe30a1bcb92fedd9	db687a412d1b4d65	aaf3623c2a004b07
	12-15	8d41a28b0d847693	7f212e01c4e96937	7eeeca5c84ba3bda	0000000000000000

Table 7. Colliding message pair for 22-step SHA-256 with standard IV. These messages have been generated using the algorithm of Table 3.

W_1	0-7	00000000	6b1525df	ba8df484	6dd804eb	d048b076	762152d2	ca960cd9	01e340a9
	8-15	03fa80ac	c787b892	e2e01390	aaf3823e	8d41a28e	7f22ee02	7c625999	183e603f
W_2	0-7	00000000	6b1525df	ba8df484	6dd804eb	d048b076	762152d2	ca960cd9	01e340a9
	8-15	03fa80ad	c787b891	defe7410	aaf5223e	8d41a28e	7f22ee02	7c625999	00000000

Table 8. Colliding message pair for 22-step SHA-512 with standard IV. These messages have been generated using the algorithm of Table 4.

W_1	0-3	0000000000000000	0000000000000000	c2bc8e9a85e2eb5a	6d623c5d5a2a1442
	4-7	cd38e6dee1458de7	acb73305cddb1207	148f31a512bbade5	ecd66ba86d4ab7e9
	8-11	92aafb1e9cfa1fcb	533c19b80a7c8968	e3ce7a41b11b4d75	aef3823c2a004b20
	12-15	8d41a28b0d847692	7f214e01c4e96950	0000000000000000	0000000000000000
W_2	0-3	0000000000000000	0000000000000000	c2bc8e9a85e2eb5a	6d623c5d5a2a1442
	4-7	cd38e6dee1458de7	acb73305cddb1207	148f31a512bbade5	ecd66ba86d4ab7ea
	8-11	90668fd7ec6718ee	533c19b80a7c8968	dfce7a41b11b4d76	aef3823c2a004b20
	12-15	8d41a28b0d847692	7f214e01c4e96950	0000000000000000	ffffffffffffffff

Table 9. Colliding message pair for 22-step SHA-256 with standard IV. These messages have been generated using the algorithm of Table 4.

W_1	0-7	00000000	00000000	0be293bf	99c539c9	1c672194	99b6a58a	5bf1d0ae	0a9a18d3
	8-15	0c18cf1c	329b3e6e	dc4e7a43	ab33823f	8d41a28d	7f214e03	00000000	00000000
W_2	0-7	00000000	00000000	0be293bf	99c539c9	1c672194	99b6a58a	5bf1d0ae	0a9a18d4
	8-15	07d56809	329b3e6e	dc0e7a44	ab33823f	8d41a28d	7f214e03	00000000	ffffffff