Security Proof of ORANGE-Zest

Bishwajit Chakraborty and Mridul Nandi

Indian Statistical Institute, Kolkata

Abstract. ORANGE is a Round 1 candidate in the NIST LwC competition. The authenticated encription mode in ORANGE is called ORANGE-Zest. This AEAD scheme is important in the sense that it has rate 1 which is highest among all other Sponge based submissions. In this paper, we give a security bound on ORANGE-Zest. Our security bounds show that it is secure within NIST prescribed Data complexity of 2^{50} bytes and Time complexity 2^{112}.

1 Introduction

In recent years the popularity and requirement of lightweight cryptographic designs has increased immensely. The switching of modern computing from desktop computers to small devices has created a restrain in the availability of resources as such the classical schemes are not always efficient or in some cases not even implementable. In all those cases lightweight schemes must be used to provide security. Hence, to increase the research interest of the scientific community, NIST has initiated a Light Weight Cryptography competition.

For sponge based authenticated encryption scheme the best known adversarial advantage for forgery is $\frac{DT}{2^c}$ where c is the capacity, D is the data complexity and T is the time complexity. ORANGE-Zest [3] is a round 1 candidate in the NIST-LwC competition. In many ways it is similar to the Beetle [2] type construction other than the fact that it uses an external state and it has full rate message absorption. In comparison to other submissions ORANGE-Zest has the major advantage that it is the only submission that has rate 1 i.e. at the cost of this small additional state, it has full state message absorption and as such maximum among all other NIST round 1 candidates which are based on Sponge[1].

In this paper we give a security proof of ORANGE-Zest in the Ideal Cipher Model. We note that we make a minor revision of the original submission. The details of the changes and the complete algorothm are given in Section 3. Our main result is the following.

Theorem 1. *(Main Result) For any adversary $\mathscr{A}$ making at most q_p many primitive queires, q_e many encryption queries with total σ_e many blocks and q_v decryption queries with total σ_v many blocks,*

$$\mathsf{Adv}^{aead}_{\mathsf{ORANGE\text{-}Zest}}(\mathscr{A}) \leq \frac{q_p}{2^\kappa} + \frac{5\sigma_e q_p}{2^b} + \frac{4\sigma_v q_p}{2^b} + \frac{2q_v}{2^\tau} + \frac{2rq_p\sigma_e}{2^b} + \frac{4\sigma_e\sigma_v}{2^c} + \frac{rq_p\sigma_v\sigma_e}{2^{b+c}} + \frac{\tau q_v q_p}{2^{\tau+c}} + \frac{rq_v q_p}{2^b} + \frac{b\sigma_v q_p^2}{2^{b+c}}$$

where κ is the key size, τ is the tag size and $b = r + c$ is the state size.

According to NIST requirement $q_p \leq 2^{112}$ and $\sigma_e + \sigma_v \leq 2^{45}$. ORANGE-Zest uses PHOTON-256[4] as the underlying permutation which has a state size of 256-bit. The Nonce size, Key size and capacity size of ORANGE-Zest is 128-bit. Plugging the appropriate values in Theorem 1, the dominating terms are $\frac{q_p}{2^\kappa} \leq \frac{1}{2^{16}}$ and $\frac{4\sigma_v q_p}{2^b} \leq \frac{1}{2^{36}}$, which are all well within the NIST requirements. All other terms are bounded by $\frac{1}{2^{80}}$ and hence negligible. Thus, we can conclude that ORANGE-Zest is secure with respect to the NIST requirement.

2 Priliminaries

2.1 H-coefficient Technique

Consider a computationally unbounded and deterministic adversary $\mathscr{A}$ that tries to distinguish the real oracle, say $\mathcal{O}_1$, from the ideal oracle, say $\mathcal{O}_0$. We denote the query-response tuple of $\mathscr{A}$'s interaction with its oracle by a transcript ω. Sometimes, this may also include any additional information that the oracle chooses to reveal to the distinguisher at the end of the query-response phase of the game. We will consider this extended definition of transcript. We denote by Θ_1 (res. Θ_0) the random transcript variable when $\mathscr{A}$ interacts with $\mathcal{O}_1$ (res. $\mathcal{O}_0$). The probability of realizing a given transcript ω in the security game with an oracle $\mathcal{O}$ is known as the *interpolation probability* of ω with respect to $\mathcal{O}$. Since $\mathscr{A}$ is deterministic, this probability depends only on the oracle $\mathcal{O}$ and the transcript ω. A transcript ω is said to be *attainable* if $\Pr[\Theta_0 = \omega] > 0$. In this paper, $\mathcal{O}_1 = (\mathsf{enc_K}, \mathsf{dec_K}, f^{\pm})$, $\mathcal{O}_0 = (\Gamma, \perp, f^{\pm})$, and the adversary is trying to distinguish $\mathcal{O}_1$ from $\mathcal{O}_0$ in AEAD sense. Now we state a simple yet powerful tool due to Patarin, known as the H-coefficient technique (or simply the H-technique).

Theorem 2 (H-coefficient technique [5]). *Let Ω be the set of all realizable transcripts. For some $\epsilon_{\mathsf{bad}}, \epsilon_{\mathsf{ratio}} > 0$, suppose there is a set $\Omega_{\mathsf{bad}} \subseteq \Omega$ satisfying the following:*

- $\Pr[\Theta_0 \in \Omega_{\mathsf{bad}}] \leq \epsilon_{\mathsf{bad}}$;
- *For any $\omega \notin \Omega_{\mathsf{bad}}$,*

$$\frac{\Pr[\Theta_1 = \omega]}{\Pr[\Theta_0 = \omega]} \geq 1 - \epsilon_{\mathsf{ratio}}.$$

Then for any adversary $\mathscr{A}$ we have the following bound on its AEAD distinguishing advantage:

$$\mathsf{Adv}^{aead}_{\mathcal{O}_1}(\mathscr{A}) \leq \epsilon_{\mathsf{bad}} + \epsilon_{\mathsf{ratio}}.$$

A proof of this theorem is available in multiple papers including [5].

3 The ORANGE-Zest Mode

Let $P : \{0,1\}^n \to \{0,1\}^n$ be any permutation function. Let α be any primitive polynomial. Then the ORANGE-Zest mode can be best understood from the diagrams given in 1 2.

Note, that there are two minor revisions from the original ORANGE proposal.

1. While processing the first message block state output of the last associate data block is used as the value of the additional state instead of the key. The reason behind this is that in this scenario it will be similar to the other message blocks processing and hence will require less number of multiplexers which would be efficient from the hardware designing point of view.
2. To make the above change go through for empty associated data encryption we consider empty associated data as an incomplete associated data block and do 0^*1 padding to it. Formally,

$$\mathsf{pad}(A) = \begin{cases} 0^{2n-1}1 \text{ if } |A| = 0 \\ A \text{ if } 2n \mid |A| \\ 0^{2n-r-1}1 \| A \text{ if } |A| = r \mod 2n, r \neq 0. \end{cases}$$

Due to this new consideration the empty associated data case will not be treated differently as in the case of the submitted version of ORANGE and hence will use less number of multiplexers in hardware implementation. For the sake of completeness, we give the complete algorithm following these changes.

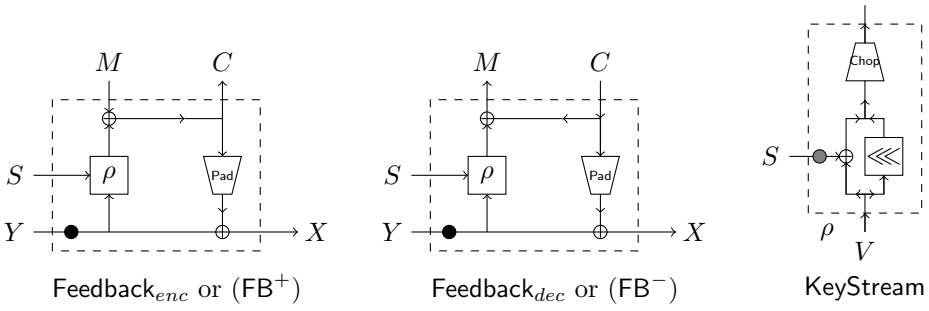

Fig. 1: Feedback process for ORANGE-Zest. the black dot represents α^{δ_M} multiplication to the MSB $\frac{n}{2}$-bits of Y where $\delta_M = 0/1/2$ depending on whether the message block is a intermediate/last and full or last and partial block. The gray dot represents α multiplication.

function ORANGE-ZEST$_{[\mathsf{P}]}$.enc(K, N, A, M)
  $(A_{a-1}, \ldots, A_0) \xleftarrow{n} A$
  $(M_{m-1}, \ldots, M_0) \xleftarrow{n} M$
  if $a = 0, m = 0$ **then**
    $(T, *) \leftarrow \mathsf{P}((K \oplus 2) \| N)$
    return (λ, T)
  $(U, S) \leftarrow \mathsf{proc\_hash}(K \| N, A, 1, 2)$
  if $m \neq 0$ **then**
    $(C, U) \leftarrow \mathsf{proc\_txt}(S, U, M, +)$
  return $(C, \mathsf{proc\_tg}(U))$

function ORANGISH(D)
  if $|D| = 0$ **then**
    $Z_1 \leftarrow \mathsf{P}(0^{2n})$
  else
    $(D_{d-1}, \ldots, D_0) \xleftarrow{n} D$
    $D_d \leftarrow (n \nmid |D_{d-1}|)?\ 0^{n-2}10 : 0^{n-1}1$
    $D_{d-1} \leftarrow \mathsf{pad}(D_{d-1})$
    $X \leftarrow (0^n || D_0)$
    for $i = 0$ to $d - 1$ **do**
      $A_i \leftarrow (D_i \| D_{i+1})$
    $A \leftarrow (A_{d-1} \| \ldots \|, A_0)$
    $(Z, \star) \leftarrow \mathsf{proc\_hash}(X, A, 0, 0)$
    $Z_1 \leftarrow \mathsf{P}(Z)$
  $Z_2 \leftarrow \mathsf{P}(Z_1)$
  return $\lfloor Z_2 \rfloor_n \| \lfloor Z_1 \rfloor_n$

function $\mathsf{proc\_txt}(S_0, U_0, D, \mathsf{dir})$
  $(D_{d-1}, \ldots, D_0) \xleftarrow{2n} D$
  for $i = 0$ **to** $d - 1$ **do**
    $V_i \leftarrow \mathsf{P}(U_i)$
    if $i = d - 1$ **then**
      $c \leftarrow (2n \mid |D_{d-1}|)?1 : 2$
      $V_i \leftarrow \mathsf{mult}(c, V_i)$
    $KS_i \leftarrow \rho(S_i, V_i)$
    $D'_i \leftarrow D_i \oplus \lfloor KS_i \rfloor_{|D_i|}$
    if $\mathsf{dir} =$ " + " **then** $D_i \leftarrow D'_i$
    $S_{i+1} \leftarrow \lceil V_i \rceil_n$
    $U_{i+1} \leftarrow V_i \oplus \mathsf{pad}(D_i)$
  return (D', U_d)

function ORANGE-ZEST$_{[\mathsf{P}]}$.dec(K, N, A, C, T)
  $(A_{a-1}, \ldots, A_0) \xleftarrow{n} A$
  $(C_{m-1}, \ldots, C_0) \xleftarrow{n} C,\ M \leftarrow \lambda$
  $(U, S) \leftarrow \mathsf{proc\_hash}(N \| K, A, 1, 2)$
  if $m \neq 0$ **then**
    $(M, U) \leftarrow \mathsf{proc\_txt}(S, U, C, -)$
  $T' \leftarrow \mathsf{proc\_tg}(U)$
  if $T \neq T'$ **then**
    return $\perp$
  else
    return $(M, \top)$

function $\mathsf{proc\_hash}(X, D, c_0, c_1)$
  $c \leftarrow (|D| \neq 0 \cap 2n \mid |D|)?c_0 : c_1$
  $(D_{d-1}, \ldots, D_0) \xleftarrow{2n} \mathsf{pad}(D)$
  $X_0 \leftarrow X$
  for $i = 0$ **to** $d - 2$ **do**
    $Y_i \leftarrow \mathsf{P}(X_i)$
    $X_{i+1} \leftarrow Y_i \oplus D_i$
  $Y_{d-1} \leftarrow \mathsf{P}(X_{d-1})$
  $S \leftarrow \lceil Y_{d-1} \rceil_n$
  $Y_{d-1} \leftarrow \mathsf{mult}(c, Y_{d-1})$
  $X_d \leftarrow Y_{d-1} \oplus D_{d-1}$
  return (X_d, S)

function $\rho(S, Y)$
  $(Y^b, Y^t) \xleftarrow{n} Y$
  $Z \leftarrow (Y^b \oplus \alpha S) \| (Y^t \lll 1)$
  return Z

function $\mathsf{mult}(c, V)$
  $(V^b, V^t) \xleftarrow{n} V$
  return $\alpha^c \cdot V^b \parallel V^t$

function $\mathsf{proc\_tg}(U)$
  $(U^b, U^t) \xleftarrow{n} U$
  return $\mathsf{P}(U^t \| U^b)$

4 Multichain Security game

Let $\mathcal{L} = ((u_1, v_1), \ldots, (u_t, v_t))$ be a list of pairs of b-bit elements such that $\forall i \neq j, u_i \neq u_j, v_i \neq v_j$. For any such list we define $\mathsf{domain}(\mathcal{L}) = \{u_1, \ldots, u_t\}$ and $\mathsf{range}(\mathcal{L}) = \{v_1, \ldots, v_t\}$.

Given a list $\mathcal{L}$ we define a directed graph $\mathcal{G}_{\mathcal{L}}$ as follows: $\mathsf{range}(\mathcal{L})$ is the set of vertices of $\mathcal{G}_{\mathcal{L}}$. There are to types of edges:

Given any $i, j \in [t]$, there exist a directed edge $v_i \xrightarrow{x} v_j$ where $x = v_i \oplus u_j$.

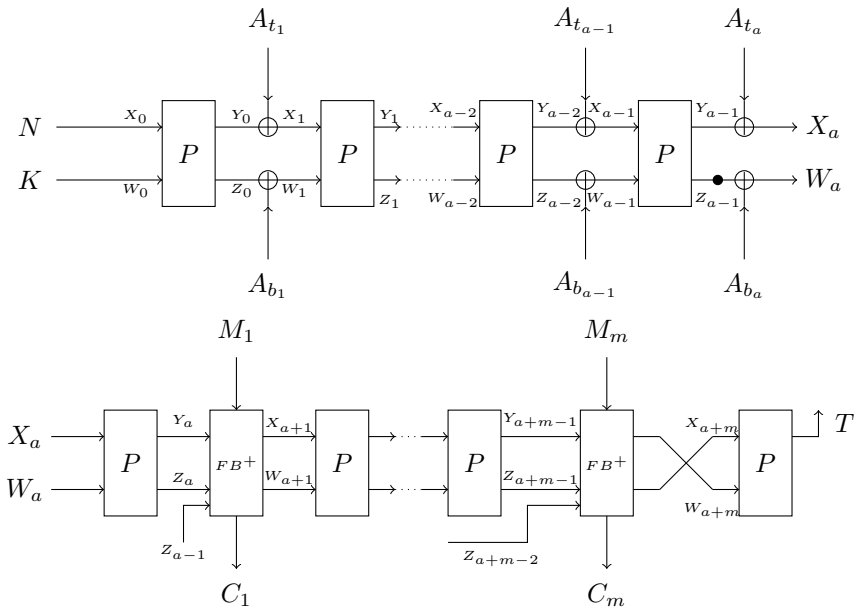

Fig. 2: The ORANGE-Zest Mode of AEAD

Given any $i, j \in [t]$ there exist a directed edge $v_i \underset{x}{\rightarrow} v_j \iff u_j = (\lfloor x \rfloor_r \oplus \lfloor v_i \rfloor_r) \| (\lceil x \rceil_c \oplus \alpha^{\delta_x} \cdot \lceil v_i \rceil_c)$ similarly we can extend this definition to define a labled walk $\mathcal{W}$ from ω_0 to ω_k by

$$\mathcal{W} : \omega_0 \xrightarrow{x_1} \omega_1 \xrightarrow{x_2} \omega_2 \cdots \omega_{k-1} \xrightarrow{x_k} \omega_k$$

We simple denote this by $\omega_0 \xrightarrow{x} \omega_k$ where $x = (x_1, \ldots, x_k)$. k is the length of the walk. Similary by $w_0 \underset{y}{\xrightarrow{x}} w_{k+1}$ we denote the walk $\omega_0 \xrightarrow{x} \omega_k \underset{y}{\rightarrow} \omega_{k+1}$.

4.1 Multichain Structure

Definition 1 (multi-chain Structure). *Let $r, \tau \leq b$ be some parameters. We say that a set of labled walks $\{\mathcal{W}_1, \ldots, \mathcal{W}_p\}$ forms a multi-chain of a given lable $x = (x_1, x_2, \ldots, x_k)$ in the graph $\mathcal{G}_{\mathcal{L}}$ if $\forall 1 \leq i \leq p$ We have $\mathcal{W}_i : v_0^i \xrightarrow[x_k]{(x_1,\ldots,x_{k-1})} v_k^i$ such that $\forall 1 \leq i, j \leq p; \lfloor u_0^i \rfloor_r = \lfloor u_0^j \rfloor_r; \lfloor v_k^i \rfloor_\tau = \lfloor v_k^j \rfloor_\tau$. We call it a multi-chain of length k.*

Let W_k denote the maximum size of a multi-chain of length k (of a given lable x) induced by $\mathcal{L}$. Now consider an adversary $\mathscr{A}$ interacting at most t times with $f^{\pm}$. Let (x_i, dir_i) denote ith query where $x_i \in \{0,1\}^b$ and dir_i is either $+$ or $-$ (representing forward or inverse query). If $dir_i = +$, it gets response y_i as $f(x_i)$, else the response y_i is set as $f^{-1}(x_i)$. After t many interactions,

we define a list $\mathcal{L}$ of pairs $(u_i, v_i)_i$ where $(u_i, v_i) = (x_i, y_i)$ if $dir_i = +$, and $(u_i, v_i) = (y_i, x_i)$ otherwise. So we have $f(u_i) = v_i$ for all i. We call the tuple of triples $\theta := ((u_1, v_1, dir_1), \ldots, (u_t, v_t, dir_t))$ the transcript of the adversary $\mathscr{A}$ interacting with $f^{\pm}$. We also write $\theta' = ((u_1, v_1), \ldots, (u_t, v_t))$ which only stores the information about the random permutation. We write

$$\mu_{k,\mathscr{A}} := \mathsf{Ex}\left[W_k\right].$$

Here W_k is defined for the labeled graph induced by the list θ' as defined above and expectation is defined over the randomness of the random permutation f and the random coin of the adversary $\mathscr{A}$. Finally, we define $\mu_{k,t} = \max_{\mathscr{A}} \mu_{k,\mathscr{A}}$ where maximum is taken over all adversaries making at most t queries.

5 Security Analysis of ORANGE-Zest

We fix a deterministic non-repeating query making distinguisher $\mathcal{A}$ that interacts with either (1) the real oracle $(\mathcal{O}^f, f)$ or (2) the ideal oracle $(\$^f, f)$ making at most,

1. q_e encryption queries $(N^i, A^i, M^i)_{i\in(q_e]}$ with an aggregate of total σ_e many blocks.
2. q_f offline or direct forward queries $(U^i, V^i, +)_{i\in(q_f]}$ to f.
3. q_b direct backward queries $(U^i, V^i, -)_{i\in(q_b]}$ to f.
4. attempts to forge with q_v many queries $(N^{\star i}, A^{\star i}, C^{\star i}, T^{\star i})_{i\in(q_v]}$ having a total of σ_v many blocks.

We assume $q_p = q_f + q_b$ to be the total number of offline or direct queries to f. Also for simplicity assume that, $\forall i, M^i$ and A^i have m_i and 0 many blocks respectively and $C^{\star i}$ and $A^{\star i}$ have m_i and 0 many blocks respectively. Let $X^\star, Y^\star, Z^\star, W^\star$ corresponds to the imtermidiate variables of the forging queries. Let $\mathcal{E}, \mathcal{D}$ denotes the sets of indices of the encryption and decryption queries.

Theorem 3. *Let us assume that $r \geq 16$. Then for any $(q_p, q_e, q_v, \sigma_e, \sigma_v)-$ adversary $\mathscr{A}$ we have*

$$\mathsf{Adv}^{aead}_{\mathsf{ORANGE\text{-}Zest}}(\mathscr{A}) \leq \frac{q_p}{2^\kappa} + \frac{5\sigma_e q_p}{2^b} + \frac{4\sigma_v q_p}{2^b} + \frac{2q_v}{2^\tau} + \frac{2rq_p\sigma_e}{2^b} + \frac{4\sigma_e\sigma_v}{2^c} + \sum_{i\in\mathcal{D}} \frac{\mu_{m_i,q_p}}{2^c} + \frac{rq_p\sigma_v\sigma_e}{2^{b+c}}.$$

5.1 Attack transcript

The Ideal World In the ideal world there are three types of oracle queries, namely primitive query, encryption query and decryption query.

Primitive Queries The ideal world simulates $Q_{\pm}$ queries honestly and maintain a list ω_p of the query responce of Q as a partial injective list. More precisely

$$\omega_p = ((U^1, V^1, dir_1), (U^2, V^2, dir_2), \dots)$$

where $dir_i = +1$ for a direct forward query and -1 for a direct backward query. We keep ω_p as a list of direct forward queries. i.e. $f(U^i) = V^i$ for all i. Let $\omega_{p'} = (((U^1, V^1), (U^2, V^2), \dots)$ i.e. ω_p without considering the sign of the query.

Encryption Queries When the i-th query is an encryption query (N^i, M^i) where and $M^i = M^i_{m_i} \| \cdots \| M^i_2 \| M^i_1$ it first defines

$$\delta^i_j = \begin{cases} 0 \text{ for } j < m_i \\ 1 \text{ for } j = m_i, |M^i_{m_i}| = b \\ 2 \text{ otherwise} \end{cases}$$

Then it samples $(Y^i_{-1}, Y^i_0, \dots, Y^i_{m_i}) \xleftarrow{\$} \{0,1\}^r$ and $(Z^i_{-1}, Z^i_0, \dots, Z^i_{m_i}) \xleftarrow{\$} \{0,1\}^c$, and returns $T = \lfloor Z^i_{m_i} \| Y^i_{m_i} \rfloor_\tau$ and $C^i = C^i_{m_i} \| \cdots \| C^i_2 \| C^i_1$ where for $1 \le j \le m$

$$C^i_{r_j} = M^i_{r_j} \oplus lRot(Y^i_{j-1});$$
$$C^i_{c_j} = M^i_{c_j} \oplus \alpha^{\delta^i_j} \cdot Z^i_{j-1} \oplus \alpha \cdot Z^i_{j-2}$$
$$C^i_j = C^i_{c_j} \| C^i_{r_j}$$

The intermidiate values X^i_j, W^i_j are calculated as follows:

$$X^i_j = \begin{cases} \lfloor K \| N \rfloor_r \text{ for } j = -1 \\ Y^i_{-1} \oplus 1 \text{ for } j = 0 \\ Y^i_{j-1} \oplus C^i_{r_j} \text{ for } 1 < j < m_i \\ \alpha^{\delta^i_{m_i}} Z^i_{m_i - 1} \oplus C^i_{c_{m_i}} \text{ for } j = m_i \end{cases}$$

$$W^i_j = \begin{cases} \lceil K \| N \rceil_c \text{ for } j = -1 \\ \alpha^2 . Z^i_{-1} \text{ for } j = 0 \\ Z^i_{j-1} \oplus C^i_{c_j} \text{ for } j < m_i \\ Y^i_{m_i - 1} \oplus C^i_{r_{m_i}} \text{ for } j = m_i \end{cases}$$

Decryption Queries When the i-th query is a decryption query of the form (N^{*i}, C^{*i}, T^{*i}) it always returns $M^{*i} = \perp$. The decryption transcript $\omega_d = (M^{*i})_{i \in \mathcal{D}}$ where $M^{*i} = \perp$ for all $i \in \mathcal{D}$

Offline Queries After all the above queries, finally the oracle returns all the X, Y, Z, W values defined above. Let $\omega_e := (X^i_j, Y^i_j, Z^i_j, W^i_j)_{i \in \mathcal{E}, j \in [m_i]}$. The transcript of the ideal oracle is $(\omega_p, \omega_e, \omega_d)$.

Intermediate Values of the decryption queries Given the i-th decryption query $(N^{*i}, C^{*i}, T^{*i}), i \in \mathcal{D}$ we define p_i as follows.

$$p_i = \begin{cases} -1 \text{ if } N^{*i} \neq N^{i'} \forall i' \in \mathcal{E} \\ l_i \text{ if } \exists i' \in \mathcal{E} \ni N^{*i} = N^{i'}; C^{*i}_j = C^{i'}_j \forall 1 \le j \le l_i < m_i; C^{*i}_{l_i+1} \neq C^{i'}_{l_i+1} \\ l_i - 1 \text{ otherwise} \end{cases}$$

Given a statement P let

$$\chi(P) = \begin{cases} 1 \text{ if } P \text{ is true} \\ 0 \text{ otherwise} \end{cases}$$

Let

$$x_j^i = \chi(j \neq m_i) Y_{j-1}^{*i} \oplus \chi(j = m_i) \alpha^{\delta_{m_i}^i} Z_{j-1}^{*i}.$$

$$w_j^i = \chi(j = m_i) Y_{j-1}^{*i} \oplus \chi(j \neq m_i) Z_{j-1}^{*i}.$$

For any $i \in \mathcal{D}$ we define, $\forall 0 \leq j \leq p_i$

$$X_j^{*i} = X_j^{i'}; Y_j^{*i} = Y_j^{i'}; Z_j^{*i} = Z_j^{i'}; W_j^{*i} = W_j^{i'}$$

Now we further extend X, Y, Z, W values using primitive transript wherever possible. For notational simplicity let $c_j^i := \chi(j \neq m_i) C_j^{*i} \oplus \chi(j = m_i) \lfloor C_j^{*i} \rfloor_r \| \lceil C_j^{*i} \rceil_c$, $\forall p_i < j \leq m_i$. If there exist a labeled walk in the labled directed graph induced by ω_p from $Z_{p_i}^{*i} \| Y_{p_i}^{*i}$ with lable $(c_{p_i+1}^i, \ldots, c_j^i), j < m_i$, then we denote the end node as $Z_j^{*i} \| Y_j^{*i}$.

$$Z_{p_i}^{*i} \| Y_{p_i}^{*i} \xrightarrow{(c_{p_i+1}^i, \ldots, c_j^i)} Z_j^{*i} \| Y_j^{*i}$$

given $i \in \mathcal{D}$ let $p_i' < m_i$ be the maximum possible value of such j.
For all such $i \in \mathcal{D}$ and $p_i < j \leq p_i' + 1$ define

$$X_j^{*i} = x_j^i \oplus \lfloor c_j^i \rfloor_r$$

$$W_j^{*i} = w_j^i \oplus \lceil c_j^i \rceil_c$$

5.2 Identifying bad events

We say that an ideal world transcript $\omega = (\omega_p, \omega_e, \omega_d)$ is bad if any one of the following conditions holds:

Bad events due to encryption and primitive transcript:

- B1: For some $(U, V) \in \omega_p$, $K = \lceil U \rceil_\kappa$.
- B2: For some $i \in \mathcal{E}$, $j \in [m_i]$, $Z_j^i \| Y_j^i \in \mathsf{range}(\omega_p)$, (in other words, $\mathsf{range}(\omega_e) \cap \mathsf{range}(\omega_p) \neq \emptyset$)
- B3: For some $i \in \mathcal{E}, j \in [m_i]$, $W_j^i \| X_j^i \in \mathsf{domain}(\omega_p)$, (in other words, $\mathsf{domain}(\omega_e) \cap \mathsf{domain}(\omega_p) \neq \emptyset$)
- B4: For some $(i \in \mathcal{E}, j \in [m_i]) \neq (i' \in \mathcal{E}, j' \in [m_{i'}])$, $Z_j^i \| Y_j^i = Z_{j'}^{i'} \| Y_{j'}^{i'}$,
- B5: For some $(i \in \mathcal{E}, j \in [m_i]) \neq (i' \in \mathcal{E}, j' \in [m_{i'}])$, $W_j^i \| X_j^i = W_{j'}^{i'} \| X_{j'}^{i'}$,

Bad events due to decryption transcript:

- B6: For some $i \in \mathcal{D} \ni p_i \leq m_i - 1$, $(i' \in \mathcal{E}, j' \in [m_{i'}])$, $W_{p_i+1}^{*i} \| X_{p_i+1}^{*i} = W_{j'}^{i'} \| X_{j'}^{i'}$,
- B7: For some $i \in \mathcal{D}$ with $p_i \geq 0$, $p_i' = m_i - 1$ and $(W_{m_i}^{*i} \| X_{m_i}^{*i}, * \| T^{*i}) \in \omega_p$,
- B8: For some $i \in \mathcal{D}$ with $p_i \geq 0$ and $p_i' \geq p_i + 1$, $W_{p_i'+1}^{*i} \| X_{p_i'+1}^{*i} \in \mathsf{domain}(\omega_e)$.

We write BAD to denote the event that the ideal world transcript Θ_0 is bad. Then, with a slight abuse of notations, we have

$$\mathsf{BAD} = \cup_{i=1}^{8} \mathtt{Bi}$$

Lemma 1.

$$\Pr[\mathsf{BAD}] \leq \frac{q_p}{2^\kappa} + \frac{5\sigma_e q_p}{2^b} + \frac{2rq_p\sigma_e}{2^b} + \frac{4\sigma_e\sigma_v}{2^c} + \sum_{i\in\mathcal{D}} \frac{\mu_{m_i,q_p}}{2^c} + \frac{rq_p\sigma_v\sigma_e}{2^{b+c}}$$

Proof. Due to constrain of space the proof of Lemma 1 has been moved to Appendix B.

The real World The real world has the oracle $f^{\pm}$. The AE encryption and decryption queries and direct primitive queries are faithfully responded based on $f\pm$. Like the ideal, after completion of interaction, the ideal oracle returns all Y, Z -values corresponding to the encryption queries only. Note that a decryption query may return M^i which is not $\perp$.

Now consider a good transcript $\omega = (\omega_p, \omega_e, \omega_d)$. The understanding of the bad events will become clear from understanding of the good transcript. Suppose for all $1 \leq j \leq p'_i$, Y_j^{*i}, Z_j^{*i} and $X_{j+1}^{*i}, W_{j+1}^{*i}$ have been defined as described above. Then observe the following:

1. The tuple ω_e is permutation compatible and disjoint from ω_p. So union of tuples $\omega_e \cup \omega_p$ is also permutation compatible.
2. For all $i \in \mathcal{D}$ we have either $p'_i = m_i - 1$ and $(W_{m_i}^{*i} \| X_{m_i}^{*i}, \star \| T^{*i}) \in \omega_p \cup \omega_e$ (Type-1 decryption query) or $p'_i < m_i - 1$ but $(W_{p'_i+1}^{*i} \| X_{p'_i+1}^{*i} \notin \omega_p \cup \omega_e$(Type-2 decryption query). Type-1 decryption queries would be straightaway rejected. Type-2 decryption query can be computed based on $\omega_p \cup \omega_e$ until $(W_{p'_i+1}^{*i} \| X_{p'_i+1}^{*i}$ which is fresh. So $f(W_{p'_i+1}^{*i} \| X_{p'_i+1}^{*i})$ is random over a large set. This would ensure with high probability we reject those decryption queries also.

Based on the above observations we perform our analysis of the good transcripts.

Good Transcript Analysis: Now fix a good transcript ω. Let Θ_0 and Θ_1 denote the transcript random variable obtained in the ideal world and real world respectively. As noted before, all the input-output pairs for the underlying permutation are compatible. In the ideal world, all the Y, Z values are sampled uniform at random; the list ω_p is just the partial representation of f; and all the decryption queries are degenerately aborted; whence we get

$$\Pr[\Theta_0 = w] \leq \frac{1}{2^{b\sigma_e}(2^b)_{q_p}}$$

Here σ_e denotes the total number of blocks present in all encryption queries including nonce. In notation $\sigma_e = q_e + \sum_i m_i$.

In the real world, for ω we denote the encryption query, decryption query, and primitive query tuples by ω_e, ω_d and ω_p, respectively. Then, we have

$$\begin{aligned}\Pr[\Theta_1 = \omega] &= \Pr[\Theta_1 = (\omega_e, \omega_p, \omega_d)] \\ &= \Pr[\omega_e, \omega_p] \cdot \Pr[\omega_d \mid \omega_e, \omega_p] \\ &= \Pr[\omega_e, \omega_p] \cdot (1 - \Pr[\neg\omega_d \mid \omega_e, \omega_p]) \\ &\leq \Pr[\omega_e, \omega_p] \cdot \left(1 - \sum_{i \in \mathcal{D}} \Pr[\neg\omega_{d,i} \mid \omega_e, \omega_p]\right) \qquad (1)\end{aligned}$$

Here we have slightly abused the notation to use $\neg\omega_{d,i}$ to denote the event that the i-th decryption query successfully decrypts and and $\neg\omega_d$ is the union $\cup_{i\in\mathcal{D}}\neg\omega_{d,i}$ (i.e. at least one decryption query successfully decrypts). The encryption and primitive queries are mutually permutation compatible, so we have

$$\Pr_{\Theta_1}[\omega_e, \omega_p] = 1/(2^b)_{\sigma_e+q_p} \geq \Pr_{\Theta_0}[\omega_e, \omega_p].$$

Now we show an upper bound $\Pr_{\Theta_1}[\neg\omega_{d,i} \mid \omega_e, \omega_p] \leq \frac{m_i(\sigma_e+q_p)}{2^b-\sigma_e-q_p} + \frac{1}{2^\tau}$ for every type-2 decryption query. Recall that $W^{*i}_{p'_i+1}\|X^{*i}_{p'_i+1}$ is fresh. If $W^{*i}_j\|X^{*i}_j$ is the last input block then $f(W^{*i}_j\|X^{*i}_j) = *\|T^{*i}$ with probability at most $2/2^\tau$ (provided $\sigma_e+q_p \leq 2^{b-1}$ which can be assumed, since otherwise our bound is trivially true). Suppose $W^{*i}_j\|X^{*i}_j$ is not the last block, then the next input block may collide with some encryption or primitive input block with probability at most $\frac{\sigma_e+q_p}{2^b}$. Applying this same argument for all the successive blocks till the last one, we get the probability at most $\frac{m_i(\sigma_e+q_p)}{2^b-\sigma_e-q_p}$, the last block input would be fresh. Hence the probability that the tag matches is at most $2/2^\tau$. Now, by union bound we have

$$\begin{aligned}\Pr[\neg\omega_d \mid \omega_e, \omega_p] &\leq \sum_{i\in\mathcal{D}} \frac{m_i(\sigma_e+q_p)}{2^b-\sigma_e-q_p} + \frac{2}{2^\tau} \\ &\leq \frac{2\sigma_v(\sigma_e+q_p)}{2^b} + \frac{2q_v}{2^\tau} \\ &\leq \frac{4\sigma_v q_p}{2^b} + \frac{2q_v}{2^\tau}.\end{aligned}$$

We have Theorem 3 follows from Equation 1, Lemma 1 and Theorem 2.

6 Bounding Multichain

Theorem 4. *We have,*

$$\mu_{k,t} \leq \mathsf{mcoll}(t, 2^\tau) + \mathsf{mcoll}(t, 2^r) + k \cdot \mathsf{mcoll}'(t^2, 2^b)$$

Observation We have if $v_i \xrightarrow{x} v_j$ and $v_i \xrightarrow{x} v_k$ then $v_j = v_k$. Similarly if $v_i \underset{x}{\rightarrow} v_j$ and $v_i \underset{x}{\rightarrow} v_k$ then $v_j = v_k$. and hence if $v_i \xrightarrow[y]{x} v_j$ and $v_i \xrightarrow[y]{x} v_k$ then

$v_j = v_k$.

More Notations: Let $W^{fwd,a}$ denote the size of the set $\{i : dir_i = +, \lfloor v_i \rfloor_\tau = a\}$ and $W^{fwd} = \max_a W^{fwd,a}$. This denotes the maximum number of multicollision in the τ- least significant bits of forward query responses.

Similarly define $W^{bck,a} = |\{i : dir_i = -, \lfloor u_i \rfloor_r = a\}|$ and $W^{bck} = \max_a W^{bck,a}$. This denotes the maximum number of multicollisions in the r- least significant bits of backward query responses.

Now Let $W^{mitm,a} = |\{(i,j) : v_i \xrightarrow{a} v_j \text{ or } v_i \underset{a}{\rightharpoonup} v_j\}|$ and $W^{mitm} = \max_a W^{mitm,a}$.

Lemma 2.

$$W_k \leq W^{fwd} + w^{bck} + k.W^{mitm}$$

Proof. Let $p = W_k$ and $\{\mathcal{W}_1, \ldots, \mathcal{W}_p\}$ be k-chains such that:

$$\forall 1 \leq i \leq p \mathcal{W}_i : v_0^i \xrightarrow[x_k]{(x_1,\ldots,x_{k-1})} v_k^i \text{ and}$$

$$\forall 1 \leq i \leq p; \lfloor v_0^i \rfloor_r = u; \lfloor v_k^i \rfloor_\tau = v.$$

Define

$$\omega_p^0 = |\{\mathcal{W}_i \in \{\mathcal{W}_1, \ldots, \mathcal{W}_p\} \mid (u_0^i, v_0^i, -) \in \theta\}|$$

$$\omega_p^{k+1} = |\{\mathcal{W}_i \in \{\mathcal{W}_1, \ldots, \mathcal{W}_p\} \mid (u_k^i, v_k^i, +) \in \theta\}|$$

$$\omega_p^j = |\{\mathcal{W}_i \in \{\mathcal{W}_1, \ldots, \mathcal{W}_p\} \mid (u_{j-1}^i, v_{j-1}^i, +) \in \theta \text{ and } (u_j^i, v_j^i, -) \in \theta\}| \, \forall 1 \leq j <= k$$

Then clearly By union bound ;

$$W_k \leq \omega_p^0 + \omega_p^{k+1} + \sum_{j=1}^{k} \omega_p^j$$

Now by defininition of $W^{fwd}, W^{bck}, W^{mitm}$ we have $\omega_p^0 \leq W^{fwd}; \omega_p^{k+1} \leq W^{bck}, \omega_p^j \leq W^{mitm} \forall 1 \leq j \leq k$. □

Proof. (Theorem 4)

$$\mathsf{Ex}\left[W^{bck}\right] = \mathsf{Ex}\left[\mathsf{mc}_{t,2^r}\right] \leq \mathsf{mcoll}(t, 2^r) \leq \frac{rt}{2^r}$$

$$\mathsf{Ex}\left[W^{fwd}\right] = \mathsf{Ex}\left[\mathsf{mc}_{t,2^\tau}\right] \leq \mathsf{mcoll}(t, 2^\tau) \leq \frac{\tau t}{2^\tau}$$

$$\mathsf{Ex}\left[W^{mitm}\right] = \mathsf{Ex}\left[\mathsf{mc}_{t^2,2^b}\right] \leq \mathsf{mcoll}'(t^2, 2^b) \leq \frac{bt^2}{2^b}.$$

□

References

1. Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. Sponge functions. In *ECRYPT hash workshop*, volume 2007(9). Citeseer, 2007.
2. Avik Chakraborti, Nilanjan Datta, Mridul Nandi, and Kan Yasuda. Beetle family of lightweight and secure authenticated encryption ciphers. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pages 218–241, 2018.
3. Bishwajit Chakraborty and Mridul Nandi. Orange. NIST LwC Round 1 Candidate, 2019. `https://csrc.nist.gov/CSRC/media/Projects/Lightweight-Cryptography/documents/round-1/spec-doc/orange-spec.pdf`.
4. Jian Guo, Thomas Peyrin, and Axel Poschmann. The photon family of lightweight hash functions. In *Annual Cryptology Conference*, pages 222–239. Springer, 2011.
5. Jacques Patarin. The "coefficients h" technique. In *International Workshop on Selected Areas in Cryptography*, pages 328–345. Springer, 2008.

A Multicollision results

A.1 Expected multicollition in an uniform sample

Let $\mathsf{X}_1, \ldots, \mathsf{X}_q \leftarrow_{\$} \mathcal{D}$ where $|\mathcal{D}| = N$. For notational simplicity, we write $\log_2 N$ as n. We denote the maximum multicollision random variable for the sample as $\mathsf{mc}_{q,N}$. More precisely, $\mathsf{mc}_{q,N} = \max_a |\{i : \mathsf{X}_i = a\}|$. For any integer $\rho \geq 2$,

$$\begin{aligned}
\Pr[\mathsf{mc}_{q,N} \geq \rho] &\leq \sum_{a \in \mathcal{D}} \Pr[|\{i : \mathsf{X}_i = a\}| \geq \rho] \\
&\leq N \cdot \frac{\binom{q}{\rho}}{N^\rho} \\
&\leq N \cdot \frac{q^\rho}{N^\rho \rho!} \\
&\leq N \cdot \left(\frac{qe}{\rho N}\right)^\rho
\end{aligned}$$

We justify the inequalities in the following way: The first inequality is due to the union bound. If there are at least ρ indices for which X_i takes value a, we can choose the first ρ indices in $\binom{q}{\rho}$ ways. This justifies the second inequality. The last inequality follows from the simple observation that $e^\rho \geq \rho^\rho/\rho!$. Thus, we have

$$\Pr[\mathsf{mc}_{q,N} \geq \rho] \leq N \cdot \left(\frac{qe}{\rho N}\right)^\rho. \tag{2}$$

For any positive integer valued random variable Y bounded by q:

$$\begin{aligned}
\mathsf{Ex}[\mathsf{Y}] &\leq \sum_{i=0}^{q} Pr[Y \geq i] \\
&\leq (\rho - 1) + \sum_{i=\rho}^{q} Pr[Y \geq i] \\
&\leq (\rho - 1) + \rho \sum_{j=1}^{\lceil \frac{q}{\rho} \rceil} Pr[Y \geq j \cdot \rho] \\
&\leq (\rho - 1) + \rho \sum_{j=1}^{\lceil \frac{q}{\rho} \rceil} N \cdot \left(\frac{qe}{j \cdot \rho N}\right)^{j \cdot \rho} \text{ substituting Eq (2)} \\
&\leq (\rho - 1) + \rho N \sum_{j=1}^{\lceil \frac{q}{\rho} \rceil} \left(\frac{qe}{\rho N}\right)^{j \cdot \rho}
\end{aligned}$$

Now if $\left(\frac{qe}{\rho N}\right) < 1$ then we have

$$\sum_{j=1}^{\lceil \frac{q}{\rho} \rceil} \left(\frac{qe}{\rho N}\right)^{j \cdot \rho} \leq \sum_{j=1}^{\infty} \left(\frac{qe}{\rho N}\right)^{j \cdot \rho} \leq \frac{\left(\frac{qe}{\rho N}\right)^{\rho}}{1 - \left(\frac{qe}{\rho N}\right)^{\rho}}$$

Hence if $\left(\frac{qe}{\rho N}\right) < 1$

$$\mathsf{Ex}[\mathsf{Y}] \leq (\rho - 1) + \rho N \cdot \frac{\left(\frac{qe}{\rho N}\right)^{\rho}}{1 - \left(\frac{qe}{\rho N}\right)^{\rho}} \tag{3}$$

Using Eq. (2), and Eq. (3) we can prove the following results for the expected value of maximum multicollision. We write $\mathsf{mcoll}(q, N)$ to denote $\mathsf{Ex}[\mathsf{mc}_{q,N}]$.

Proposition 1. $\mathsf{mcoll}(q, N) < \begin{cases} \frac{4n}{\log n} & \text{if } q = N, n \geq 16 \\ 4n & \text{if } q = nN, n \geq 4 \\ 4n\lceil \frac{q}{nN} \rceil & \text{if } q \geq nN, n \geq 4 \\ 4 \log q & \text{if } q < N, n \geq 16 \end{cases}$

Proof. First let $q = N$. Substituting q in Eq. 3 we have

$$\mathsf{Ex}[Y] \leq (\rho - 1) + \rho N \cdot \frac{\left(\frac{e}{\rho}\right)^{\rho}}{1 - \left(\frac{e}{\rho}\right)^{\rho}}$$

Now Let $\rho = \frac{4n}{logn}, n \geq 16$ Then $\frac{e}{\rho} < \frac{1}{2}$ and Hence $1 - \left(\frac{e}{\rho}\right)^{\rho} > \frac{e}{\rho}$. Hence

$$\mathsf{Ex}[Y] < (\rho - 1) + \rho N \cdot \left(\frac{e}{\rho}\right)^{\rho - 1}$$

Now by substituting the value of ρ in $\rho N \cdot \left(\frac{e}{\rho}\right)^{\rho-1}$ and by taking logarithm it can be easily shown that $\rho N \cdot \left(\frac{e}{\rho}\right)^{\rho-1} \leq 1$ and hence $\mathsf{Ex}[Y] < \frac{4n}{logn}, n \geq 16$.

Let $q = nN, \rho = 4n$. Substituting q, ρ in Eq. 3 we have

$$\mathsf{Ex}[Y] \leq (4n-1) + 4nN \cdot \frac{\left(\frac{e}{4}\right)^{4n}}{1-\left(\frac{e}{4}\right)^{4n}}$$

Now let $n \geq 4$ then we have $4n \leq N$ and hence

$$4nN \cdot \frac{\left(\frac{e}{4}\right)^{4n}}{1-\left(\frac{e}{4}\right)^{4n}} \leq N^2 \cdot \frac{\left(\frac{e}{4}\right)^{4n}}{1-\left(\frac{e}{4}\right)^{4n}}$$

Notice that for $n \geq 2$ we have

$$\frac{\left(\frac{e}{4}\right)^{4n}}{1-\left(\frac{e}{4}\right)^{4n}} < \left(\frac{e}{4}\right)^{\frac{18}{5}n} = \left[\left(\frac{e}{4}\right)^{\frac{18}{5}}\right]^n \leq \left(\frac{1}{4}\right)^n = \frac{1}{N^2}$$

The first ineqality follows from the facts that for $n \geq 2$

$1-\left(\frac{e}{4}\right)^{4n} \geq 1-\left(\frac{e}{4}\right)^{8} > 9/10$ and $\left(\frac{e}{4}\right)^{\frac{2}{5}n} \leq \left(\frac{e}{4}\right)^{\frac{4}{5}} < \frac{3}{4} \implies 1-\left(\frac{e}{4}\right)^{4n} > \left(\frac{e}{4}\right)^{\frac{2}{5}n}$

Hence $\mathsf{Ex}[Y] < 4n, n \geq 4$.

When $q \geq nN$, we can group them into $\lceil q/nN \rceil$ samples each of size exactly nN (we can add more samples if required). This would prove the result when $q \geq nN$.

Finally, when $q < N$, we can simply bound $\mathsf{Ex}[\mathsf{mc}_{q,N}] < 4 \log q$.

When $n \geq 16$, for all q, we can write the bounds into one single form:

$$\mathsf{mcoll}(q, N) < nq/N \tag{4}$$

A.2 Expected Maximum Multicollision in a Non-uniform Random Sample

Now we bound expectation of maximum multicollision in a sample $\mathsf{X}_1, \ldots, \mathsf{X}_q$ (can be arbitrarily dependent) which is not completely uniform random. However, it satisfies the following property for all distinct $i_1, \ldots, i_\rho$ for any integer $\rho \geq 2$:

$$\Pr(\mathsf{X}_{i_1} = a, \cdots \mathsf{X}_{i_\rho} = a) \leq \frac{1}{N'^r} \tag{5}$$

Then, we can actually perform the same analysis as before. For any integer $\rho \geq 2$, it can be shown that

$$\Pr[\mathsf{mc}_{q,N} \geq \rho] \leq N \cdot \left(\frac{qe}{\rho N'}\right)^\rho \tag{6}$$

Using it, we can prove the following results for expected value of maximum multicollision.

Proposition 2. $\mathsf{Ex}\left[\mathsf{mc}_{q,N}\right] < \begin{cases} 4\log q & \text{if } q < N' \\ \frac{4n}{\log n} & \text{if } N' \le q < N'n \\ \frac{4q}{N'} & \text{if } q \ge N'n \end{cases}$

In the non-random case, we denote $\mathsf{Ex}\left[\mathsf{mc}_{q,N}\right]$ by $\mathsf{mcoll}'(q,N)$ As before, when $n \ge 16$, we have

$$\mathsf{mcoll}'(q,N) \le nq/N' \tag{7}$$

B Bounding bad events(Proof of Lemma 1)

bounding $\Pr\left[\mathtt{B1}\right]$: Fix $i \in (q_p]$.Since K is randomly chosen, probability of $(U^i, V^i) \in \omega_p$ s.t. $\lfloor U^i \rfloor_\kappa = K$ is bounded by $\frac{1}{2^\kappa}$. Hence bounding over all i, we have

$$\Pr\left[\mathtt{B1}\right]] \le \frac{q_p}{2^\kappa}$$

.

bounding $\Pr\left[\mathtt{B2}\right]$: This event can be analysed by dividing into the following cases

Case 1. $\exists i,j,a; Z_j^i \| Y_j^i = V_a$. Encryption after primitive query : This case can be bounded by probability at most $\frac{1}{2^b}$. Running over q_p many primitive queries and σ_e many blocks we have

$$\Pr\left[Case1\right] \le \frac{q_p \cdot \sigma_e}{2^b}$$

Case 2. $\exists i,j,a; Z_j^i \| Y_j^i = V_a, dir_a = +$. Encryption before primitive query This can be bounded by probability at most $\frac{1}{2^b - a + 1}$ Running over σ_e many encryption blocks and q_f many a indices we have

$$\Pr\left[Case2\right] \le \frac{q_f \cdot \sigma_e}{2^b - a + 1}$$

Case 3. $\exists i,j,a; Z_j^i \| Y_j^i = V_a, dir_a = -$. Encryption before primitive query Here the adversary has access to Y_j^i as it has already been released. Let Φ_{out} denote the number of multi-collisions in Y_j^i.

$$\begin{aligned}
\Pr\left[Case3\right] &= \sum_{\Phi_{out}} \Pr\left[Case3 \wedge \Phi_{out}\right] \\
&= \sum_{\Phi_{out}} \Pr\left[Case3 | \Phi_{out}\right] \cdot \Pr\left[\Phi_{out}\right] \\
&\le \sum_{\Phi_{out}} \frac{\Phi_{out} . q_b}{2^c} \Pr\left[\Phi_{out}\right] \\
&\le \frac{q_p}{2^c} \sum_{\Phi_{out}} \Phi_{out} \Pr\left[\Phi_{out}\right] \\
&\le Ex[\Phi_{out}] \cdot \frac{q_p}{2^c} = \frac{q_p \cdot \mathsf{mcoll}(\sigma_e, 2^r)}{2^c}
\end{aligned}$$

Since the three Cases are mutually exclusive, we have,

$$\Pr\left[\mathtt{B2}\right] \leq \frac{2 \cdot q_p \cdot \sigma_e}{2^b} + \frac{q_p \cdot \mathsf{mcoll}(\sigma_e, 2^r)}{2^c}$$

bounding $\Pr\left[\mathtt{B3} \neg \mathtt{B1}\right]$: Case 1: $\exists i, j, a$, $W_j^i \| X_j^i = U_a$, encryption after primitive: This case can be bounded by probability at most $1/2^b$, as Y_{j-1}^i and Z_{j-1}^i are chosen uniformly at random and hence X_j^i and W_j^i are determined randomly . We have at most σ_e many (i, j) pairs and q_p many a indices. Thus this can be bounded by at most $\sigma_e q_p / 2^b$.
Case 2: $\exists i, j, a$, $W_j^i \| X_j^i = U_a$, $dir_a = -$, encryption before primitive: This case can be bounded by probability at most $1/(2^b - a + 1)$. We have at most σ_e many (i, j) pairs and q_b many a indices. Thus this can be bounded by at most $\sigma_e q_b / (2^b - a + 1)$.
Case 3: $\exists i, j, a$, $W_j^i \| X_j^i = U_a$, $dir_a = +$, encryption before primitive: Let Φ_{in} denote the number of multicollisions on X_j^i.

With a similar analysis on the multicollision of output values, we have $\Pr\left[\text{Case 3}\right] \leq \mathsf{Ex}\left[\Phi\right]_{in} \frac{q_b}{2^c}$. Since the three cases are mutually exclusive, we have

$$\Pr[\mathtt{B3} \neg \mathtt{B1}] \leq \frac{2\sigma_e q_p}{2^b} + \frac{q_p \mathsf{mcoll}(\sigma_e, 2^r)}{2^c}.$$

Bounding Pr[B4]: The probability of this event can be bounded in a straightforward manner by at most $\sigma_e(\sigma_e - 1)/2^{b+1}$.
Bounding Pr[B5]: This event is similar to B4, and the probability is bounded by at most $\sigma_e(\sigma_e - 1)/2^{b+1}$.
Bounding Pr[B6]: Note that after the i-th online query the adversary knows the following values;
$Y_{j-1}^i, X_j^i, Z_{j-1}^i \oplus \alpha Z_{j-2}^i = Z_{j-1}^i \oplus \alpha^j Z_{-1}^i \ \forall 1 \leq j \leq m_i - 1; Y_{m_i-1}^i, W_{m_i-1}^i,$
$\alpha^{\delta_{m_i}^i} Z_{m_i-1} \oplus \alpha Z_{m_i-2} = \alpha^{\delta_{m_i}^i} Z_{m_i-1} \oplus \alpha^{m_i} Z_{-1}^i, T.$
Case 1. $p_i = m_i - 1, j' = m_{i'}; W_{m_i}^i \| X_{m_i}^i = W_{m_{i'}}^{i'} \| X_{m_{i'}}^{i'}$: The values of $W_{m_i}^i \| X_{m_i}^i$ and $W_{m_{i'}}^{i'} \| X_{m_{i'}}^{i'}$ upto r-most significant bits can be matched by adjusting $\lfloor C_{p_i+1}^{*i} \rfloor_r = \lfloor C_{m_{i'}}^{i'} \rfloor_r \oplus Y_{m_i-1}^{*i} \oplus Y_{m_{i'}-1}^{i'}$ Now We have $\lfloor W_{m_i}^i \| X_{m_i}^i \rfloor_c = \alpha^{\delta_{m_i}^i} Z_{m_i-1}^i \oplus \lceil C_{m_i}^{*i} \rceil_c$ and $\lfloor W_{m_{i'}}^{i'} \| X_{m_{i'}}^{i'} \rfloor_c = \alpha^{\delta_{m_{i'}}^{i'}} Z_{m_{i'}-1}^{i'} \oplus \lceil C_{m_{i'}}^{i'} \rceil_c$. Hence Case 1 happens if and only if

$$\lceil C_{m_{i'}}^{i'} \rceil_c = \alpha^{\delta_{m_{i'}}^{i'}} Z_{m_{i'}-1}^{i'} \oplus \alpha^{\delta_{m_i}^i} Z_{m_i-1}^i \oplus \lceil C_{m_i}^{*i} \rceil_c == \alpha^{m_i} Z_{-1}^i \oplus \alpha^{m_{i'}} Z_{-1}^{i'} \oplus \lceil C_{m_i}^{*i} \rceil_c \oplus A$$

Where A is some known value. Now if $N^{*i} \neq N^{i'}$ we have $Z_{-1}^{i'}, Z_{-1}^i$ are chosen independently at uniformly random, hence, we have probability that the above holds is at most $\frac{1}{2^c}$. If $N^{*i} = N^{i'}$ then we must have $m_i \neq m_{i'}$ and hence since Z_{-1}^i is chosen at uniformly random, we have $\alpha^{m_i} Z_{-1}^i \oplus \alpha^{m_{i'}} Z_{-1}^i$ is uniformly random. Hence the probability is again at most $\frac{1}{2^c}$. Varying over all $i \in \mathcal{D}$ and $i' \in \mathcal{E}$ we have

$$\Pr\left[\text{Case 1}\right] \leq \frac{q_v q_e}{2^c}$$

Case 2. $p_i = m_i - 1, j' < m_{i'}; W^i_{m_i} \| X^i_{m_i} = W^{i'}_{j'} \| X^{i'}_{j'}$:

We have $W^i_{m_i} \| X^i_{m_i} = (\lfloor C^{*i}_{p_i+1} \rfloor_r \oplus Y^{*i}_{m_i-1}) \| (\alpha^{m_i} Z^i_{-1} \oplus \lceil C^{*i}_{m_i} \rceil_c \oplus A)$

$W^{i'}_{j'} \| X^{i'}_{j'} = (\alpha^{j'} Z^{i'}_{-1} \oplus \lceil C^{i'}_{j'} \rceil_c \oplus B) \| (\lfloor C^{i'}_{j'} \rfloor_r \oplus Y^{i'}_{j'-1})$. Where A and B are known values.

If $r = c = \frac{b}{2}$ it can be seen that Case 2 holds iff $(\alpha^{m_i} Z^i_{-1} \oplus \lceil C^{*i}_{m_i} \rceil_c) = (\lfloor C^{i'}_{j'} \rfloor_r \oplus Y^{i'}_{j'-1})$ and $(\lfloor C^{*i}_{p_i+1} \rfloor_r \oplus Y^{*i}_{m_i-1}) = (\alpha^{j'} Z^{i'}_{-1} \oplus \lceil C^{i'}_{j'} \rceil_c)$ both holds. If $N^{*i} \neq N^{i'}$ we have $Z^{i'}_{j'-1}, Z^i_{m_i-1}$ are chosen independently uniformly at random, we have for fix i, i', j' the probablity is bounded by $\frac{1}{2^{2c}}$.

If $N^{*i} = N^{i'}$, We have since Z^i_{-1} is chosen uniformly at random and since both the equations need to hold independently we have again the probability is bounded by $\frac{1}{2^{2c}}$.

Now varying over all $i \in \mathcal{D}, i' \in \mathcal{E}, j' \in (m'_i]$ we have

$$\Pr[\text{case 2}] \leq \frac{q_v \sigma_e}{2^{2c}}$$

Case 3. $p_i < m_i - 1, j' = m_{i'}; W^i_{m_i} \| X^i_{m_i} = W^{i'}_{j'} \| X^{i'}_{j'}$: This can be bounded in the same way as in Case 2. by

$$\Pr[\text{case 3}] \leq \frac{q_v q_e}{2^{2c}}$$

Case 4. $p_i < m_i - 1, j' < m_{i'}; W^i_{p_i+1} \| X^i_{p_i+1} = W^{i'}_{j'} \| X^{i'}_{j'}$:

$X^i_{m_i}$ and $X^{i'}_{j'}$ can be matched by adjusting $\lfloor C^{*i}_{p_i+1} \rfloor_r = \lfloor C^{i'}_{j'} \rfloor_r \oplus Y^{*i}_{p_i} \oplus Y^{i'}_{j'-1}$

Now $W^i_{m_i}$ and $W^{i'}_{j'}$ matches iff

$$\lceil C^{i'}_{j'} \rceil_c = Z^i_{p_i} \oplus Z^{i'}_{j'-1} \oplus \lceil C^{*i}_{p_i+1} \rceil_c = \alpha^{p_i+1} Z^i_{-1} \oplus \alpha^{j'} Z^{i'}_{-1} \oplus \lceil C^{*i}_{p_i+1} \rceil_c \oplus A$$

Where A is some known value.

Now We have if $N^{*i} \neq N^{i'}$ then Z^i_{-1} and $Z^{i'}_{-1}$ are independent and chosen uniformly at random. If $N^{*i} = N^{i'}$ then we must have $p_i + 1 \neq j'$ and hence $\alpha^{p_i+1} Z^i_{-1} \oplus \alpha^{j'} Z^i_{-1}$ is uniformly random. Hence, the probability that the above happens in the i-th query can be bounded by $\frac{\sigma_e}{2^c}$ and hence,

$$\Pr[\text{ Case 4}] \leq \frac{q_v \sigma_e}{2^c}$$

Since all the above cases are mutually exclusive we have

$$\Pr[\mathsf{B6}] \leq \frac{4\sigma_e \sigma_v}{2^c}$$

Bounding Pr[B7]: Let $W_k(\omega_{p'})$ denote the k-length multi-chain induced by ω_p. Suppose the event holds for the i-th decryption query and $N^{*i} = N^{i'}$. So $Z^{i'}_{p_i} \| Y^{i'}_{p_i}$ must be the starting node of the multi-chain. Since $Z^{i'}_{p_i}$ can be chosen randomly and independent of ω_p we have the probability to hold B7 in the i-th decryption query is at most $\frac{W_{m_i}}{2^c}$. So by union bound $\Pr[\mathsf{B7} | \omega_p] \leq \sum_{i \in \mathcal{D}} \frac{W_{m_i}}{2^c}$. Hence

$$\Pr[\mathsf{B7}] \leq \sum_{i \in \mathcal{D}} \frac{\mu_{m_i,q_p}}{2^c}$$

Bounding $\Pr[\mathsf{B8}]$: This event corresponds to the case when the first non-trivial decryption query block matches a primitive query and after following some partial chain matches an encryption query block. The probability of this event happening in the i-th decryption query is at most $\frac{q_p}{2^c} \times \frac{m_i^* \Phi_{in}}{2^c}$. Taking expectation we obtain

$$\Pr[\mathsf{B8}] \leq \frac{q_p \sigma_v \mathsf{mcoll}(\sigma_e, 2^r)}{2^{2c}}$$

Lemma 1 can be proved by adding all the probabilities and bounding $\mathsf{mcoll}(\sigma_e, 2^r)$ by $\frac{r\sigma_e}{2^r}, \forall r \geq 16$.
Theorem 1 can be proved by using the bound on μ_{σ_v,q_p} from Theorem 4 (See Appendix) and plugging it in Theorem 3.