APPENDIX

A MATHEMATICAL PROOFS

A.1 Preliminaries

We derive a lemma to bound the sum of n random variables. This lemma is similar to the Hoeffding bound but cannot be replaced by Hoeffding.

Lemma 1. *Let $X_1, \ldots, X_n$ be n random variables such that*

$$X_i \in \{0, s_i\}, \qquad \Pr(X_i = s_i \mid X_1, \cdots, X_{i-1}) \leqslant p,$$

where $0 \leqslant s_i \leqslant 1$. Let $X = \sum_{i=1}^{n} X_i$, and $\mu = \sum_{i=1}^{n} ps_i = nmp$.

$$Pr\left(X > (1+\Delta)\mu\right) \leqslant e^{-(\Delta-(e-2))nmp}$$

Proof. For any $t > 0$, by using the Markov inequality we have

$$\Pr(X > (1+\Delta)\mu) = \Pr(e^X > e^{(1+\Delta)\mu}) \leqslant \frac{E(e^X)}{e^{(1+\Delta)\mu}}.$$

According to the conditions, we have

$$\begin{aligned} &E(e^X) = E\left(E(e^{\sum_{i=1}^{n} X_i} \mid X_1, \cdots, X_{n-1})\right) \\ =&E\begin{pmatrix} e^{\sum_{i=1}^{n-1} X_i} \cdot \Pr(X_n = 0 \mid X_1, \cdots, X_{n-1}) \\ +e^{s_n+\sum_{i=1}^{n-1} X_i} \cdot \Pr(X_n = s_n \mid X_1, \cdots, X_{n-1}) \end{pmatrix} \\ \leqslant &E\left(e^{\sum_{i=1}^{n-1} X_i}\right) \cdot \left(1 + p\left(e^{s_n} - 1\right)\right) \leqslant \cdots \leqslant \prod_{i=1}^{n} \left(1 + p\left(e^{s_i} - 1\right)\right) \end{aligned}$$

Because of $1 + x < e^x$, we have

$$E(e^X) \leqslant \prod_{i=1}^{n} e^{p(e^{s_i}-1)}.$$

Since for $s_i \leqslant 1$, there is $e^{s_i} - 1 \leqslant (e-1)s_i$, so there is

$$E(e^X) \leqslant e^{\sum_{i=1}^{n} p(e-1)s_i} = e^{(e-1)mnp}.$$

That is

$$Pr(X > (1+\Delta)\mu) \leqslant \frac{e^{(e-1)nmp}}{e^{(1+\Delta)nmp}} = e^{-(\Delta-(e-2))nmp}$$

□

A.2 Definition of Symbols

(1) $\mathcal{S}_i$: $\{e_1, \cdots, e_{N_i}\}$, the set of keys entering the i-th layer, where $N_i = |\mathcal{S}_i|$.

(2) $f_i(e)$: the number of times that key e enters the i-th layer.

(3) $\mathcal{S}_i^0$: $\{e \mid e \in \mathcal{S}_i \wedge \forall i' \leqslant i, f_{i'}(e) \leqslant \frac{\lambda_{i'}}{2}\}$, the set of mice keys.

(4) $\mathcal{S}_i^1$: $\{e \mid e \in \mathcal{S}_i \wedge \exists i' \leqslant i, f_{i'}(e) > \frac{\lambda_{i'}}{2}\}$, the set of elephant keys.

(5) F_i: $\sum_{\{e \in \mathcal{S}_i^0\}} f_i(e)$, the total frequency of mice keys in $\mathcal{S}_i^0$.

(6) C_i: $|\mathcal{S}_i^1|$, the number of elephant keys in $\mathcal{S}_i^1$.

(7) $\mathcal{S}_{i,j}^0$: $\{e \mid e \in \mathcal{S}_i^0 \wedge h(e) = j\}$, the set of mice keys that are mapped to the j-th bucket.

(8) $\mathcal{S}_{i,j}^1$: $\{e \mid e \in \mathcal{S}_i^1 \wedge h(e) = j\}$, the set of elephant keys that are mapped to the j-th bucket.

(9) $F_{i,j}$: $\sum_{\{e \in \mathcal{S}_{i,j}^0\}} f_i(e)$, the total frequency of mice keys in $\mathcal{S}_{i,j}^0$.

(10) $C_{i,j}$: $|\mathcal{S}_{i,j}^1|$, the number of elephant keys in $\mathcal{S}_{i,j}^1$.

(11) $\mathcal{P}_{i,k}$: $\{e_1, \cdots, e_k\}$, a subset of $\mathcal{S}_i$ composed of the first k keys.

(12) $f_{i,k}^P$: $\sum_{\{e \in \mathcal{P}_{i,k-1} \cap \mathcal{S}_{i,h(e_k)}^0\}} f_i(e)$, the total frequency of mice keys with a smaller index that conflicts with key e_k.

(13) $c_{i,k}^P$: $\left|\{e \mid e \in \mathcal{P}_{i,k-1} \cap \mathcal{S}_{i,h(e_k)}^1\}\right|$, the number of elephant keys with a smaller index that conflicts with key e_k.

A.3 Properties in One Layer

This section aims to prove that only a small proportion of the keys inserted into the i-th layer will be inserted into the (i+1)-th layer.

Theorem A.1. *(Theorem 1) Let*

$$X_{i,k} = \begin{cases} 0 & C_{i,h(e_k)} = 0 \wedge f_{i,k}^P \leqslant \frac{\lambda_i}{2} \\ f_i(e_k) & C_{i,h(e_k)} = 0 \wedge f_{i,k}^P > \frac{\lambda_i}{2} \\ f_i(e_k) & C_{i,h(e_k)} > 0 \end{cases}, \quad X_i = \sum_{\{e_k \in \mathcal{S}_i^0\}} X_{i,k}.$$

The total frequency of the mice keys in the i-th layer leaving it does not exceed X_i, i.e.,

$$F_{i+1} \leqslant \sum_{\{e \in \mathcal{S}_i^0 \cap \mathcal{S}_{i+1}\}} f_{i+1}(e) \leqslant X_i.$$

Proof. For the mice keys in the j-th bucket of the i-th layer, let the number of times they leave be $F'_{i,j} = \sum_{\{e \in \mathcal{S}_{i,j}^0 \cap \mathcal{S}_{i+1}^0\}} f_{i+1}(e)$. Since a bucket can hold at least λ_i packets of the key, we have:

$$\begin{cases} F'_{i,j} = 0 & C_{i,j} = 0 \wedge F_{i,j} \leqslant \lambda_i \\ F'_{i,j} \leqslant F_{i,j} - \lambda_i & Cs_{i,j} = 0 \wedge F_{i,j} > \lambda_i \\ F'_{i,j} \leqslant F_{i,j} & C_{i,j} > 0 \end{cases}.$$

When $C_{i,j} = 0 \wedge F_{i,j} > \lambda_i$, exists k' satisfies

$$\sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k < k'\}} f_i(e_k) \leqslant \frac{\lambda_i}{2} \leqslant \sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k \leqslant k'\}} f_i(e_k) \leqslant \lambda_i.$$

Then for and only for any $e_k \in \mathcal{S}_{i,j}^0 \wedge k \leqslant k'$, there is $X_{i,k} = 0$, and

$$\begin{aligned} F'_{i,j} &\leqslant \left(\sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k \leqslant k'\}} f_i(e_k) + \sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k > k'\}} f_i(e_k) \right) - \lambda_i \\ &\leqslant 0 + \sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k > k'\}} f_i(e_k) \\ &\leqslant \sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k \leqslant k'\}} X_{i,k} + \sum_{\{e_k \in \mathcal{S}_{i,j}^0 \wedge k > k'\}} X_{i,k}. \end{aligned}$$

Then we have $F'_{i,j} \leqslant \sum_{\{e_k \in \mathcal{S}_{i,j}^0\}} X_{i,k}$, and

$$\sum_{\{e \in \mathcal{S}_i^0 \cap \mathcal{S}_{i+1}\}} f_{i+1}(e) = \sum_{j=1}^{w_i} F'_{i,j} \leqslant \sum_{j=1}^{w_i} \sum_{\{e_k \in \mathcal{S}_{i,j}^0\}} X_{i,k} = X_i.$$

□

Similarly, we have the following lemma.

THEOREM A.2. *Let*

$$Y_{i,k} = \begin{cases} 0 & c^P_{i,k} = 0 \wedge F_{i,h(e_k)} \leqslant \lambda_i, \\ 2 & c^P_{i,k} = 0 \wedge F_{i,h(e_k)} > \lambda_i \\ 2 & c^P_{i,k} > 0. \end{cases}, \qquad Y_i = \sum_{e_k \in \mathcal{S}^1_i} Y_{i,k}.$$

The number of distinct elephant keys in the i-th layer leaving it does not exceed Y_i, i.e.,

$$\left|\mathcal{S}^1_i \cap \mathcal{S}^1_{i+1}\right| \leqslant Y_i.$$

PROOF. For the elephant keys in the j-th bucket of the i-th layer, $\sum_{\{e_k \in \mathcal{S}^1_{i,j}\}} Y_{i,k} < C_{i,j}$ if and only if $C_{i,j} = 1 \wedge F_{i,j} \leqslant \lambda_i$. In this case, the number of collisions in the bucket does not exceed λ_i, and no key enters the $(i+1)$-th layer. Thus we have $|\mathcal{S}^1_{i,j} \cap \mathcal{S}^1_{i+1}| \leqslant \sum_{\{e_k \in \mathcal{S}^1_{i,j}\}} Y_{i,j}$, and

$$\left|\mathcal{S}^1_i \cap \mathcal{S}^1_{i+1}\right| = \sum_{j=1}^{w_i} |\mathcal{S}^1_{i,j} \cap \mathcal{S}^1_{i+1}| \leqslant \sum_{j=1}^{w_i} \sum_{\{e_k \in \mathcal{S}^1_{i,j}\}} Y_{i,j} = Y_i.$$

□

THEOREM A.3. *Let* $W = \frac{4N(R_w R_\lambda)^6}{\Lambda(R_w - 1)(R_\lambda - 1)}$, $\alpha_i = \frac{\|F\|_1}{(R_w R_\lambda)^{i-1}}$, $\beta_i = \frac{\alpha_i}{\frac{\lambda_i}{2}}$, $\gamma_i = (R_w R_\lambda)^{(2^{i-1}-1)}$, *and* $p_i = (R_w R_\lambda)^{-(2^{i-1}+4)}$. *Under the conditions of* $F_i \leqslant \frac{\alpha_i}{\gamma_i}$ *and* $C_i \leqslant \frac{\beta_i}{\gamma_i}$, *we have:*

$$\Pr\left(X_{i,k} > 0 \mid X_{i,1}, \cdots, X_{i,k-1}\right) \leqslant p_i, \qquad \forall e_k \in \mathcal{S}^0_i.$$
$$\Pr\left(Y_{i,k} > 0 \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) \leqslant \frac{3}{4} p_i, \qquad \forall e_k \in \mathcal{S}^1_i.$$

PROOF. By using Markov's inequality, we have

$$\begin{aligned} &\Pr\left(X_{i,k} > 0 \mid X_{i,1}, \cdots, X_{i,k-1}\right) \\ =&\Pr\left(\left(C_{i,h(e_k)} = 0 \wedge f^P_{i,k} > \frac{\lambda_i}{2}\right) \vee C_{i,h(e_k)} > 0 \mid X_{i,1}, \cdots, X_{i,k-1}\right) \\ \leqslant& \Pr\left(C_{i,h(e_k)} > 0 \mid X_{i,1}, \cdots, X_{i,k-1}\right) + \Pr\left(F_{i,h(e_k)} - f_i(e_k) > \frac{\lambda_i}{2} \mid X_{i,1}, \cdots, X_{i,k-1}\right) \\ \leqslant& \frac{E(C_{i,h(e_k)} \mid X_{i,1}, \cdots, X_{i,k-1})}{1} + \frac{E(F_{i,h(e_k)} - f_i(e_k) \mid X_{i,1}, \cdots, X_{i,k-1})}{\frac{\lambda_i}{2}} \\ \leqslant& \frac{C_i}{w_i} + \frac{2F_i}{\lambda_i w_i} \end{aligned}$$

$$\begin{aligned} &\Pr\left(Y_{i,k} > 0 \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) \\ =&\Pr\left(\left(c^P_{i,k} = 0 \wedge F_{i,h(e_k)} > \lambda_i\right) \vee c^P_{i,k} > 0 \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) \\ \leqslant& \Pr\left(C_{i,h(e_k)} - 1 > 0 \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) + \Pr\left(F_{i,h(e_k)} > \lambda_i \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) \\ \leqslant& \frac{E(C_{i,h(e_k)} - 1 \mid Y_{i,1}, \cdots, Y_{i,k-1})}{1} + \frac{E(F_{i,h(e_k)} \mid Y_{i,1}, \cdots, Y_{i,k-1})}{\lambda_i} \\ \leqslant& \frac{C_i}{w_i} + \frac{F_i}{\lambda_i w_i}. \end{aligned}$$

Recall that $w_i = \lceil \frac{W(R_w - 1)}{R^i_w} \rceil$ and $\lambda_i = \frac{\Lambda(R_\lambda - 1)}{R^i_\lambda}$, under the conditions of $F_i \leqslant \frac{\alpha_i}{\gamma_i}$ and $C_i \leqslant \frac{\beta_i}{\gamma_i}$, we have

$$\begin{aligned} &\Pr\left(X_{i,k} > 0 \mid X_{i,1}, \cdots, X_{i,k-1}\right) \\ \leqslant& \frac{\beta_i}{\gamma_i w_i} + \frac{2\alpha_i}{\gamma_i \lambda_i w_i} = \frac{4\alpha_i}{\gamma_i \lambda_i w_i} \leqslant \frac{1}{(R_w R_\lambda)^{2^{i-1}+4}} = p_i. \\ &\Pr\left(Y_{i,k} > 0 \mid Y_{i,1}, \cdots, Y_{i,k-1}\right) \\ \leqslant& \frac{\beta_i}{\gamma_i w_i} + \frac{\alpha_i}{\gamma_i \lambda_i w_i} = \frac{3\alpha_i}{\gamma_i \lambda_i w_i} \leqslant \frac{3}{4(R_w R_\lambda)^{2^{i-1}+4}} \leqslant \frac{3}{4} p_i. \end{aligned}$$

□

THEOREM A.4. *(Theorem 2) Let* $W = \frac{4N(R_w R_\lambda)^6}{\Lambda(R_w - 1)(R_\lambda - 1)}$, $\alpha_i = \frac{\|F\|_1}{(R_w R_\lambda)^{i-1}}$, $\beta_i = \frac{\alpha_i}{\frac{\lambda_i}{2}}$, $\gamma_i = (R_w R_\lambda)^{(2^{i-1}-1)}$, *and* $p_i = (R_w R_\lambda)^{-(2^{i-1}+4)}$. *Under the conditions of* $F_i \leqslant \frac{\alpha_i}{\gamma_i}$ *and* $C_i \leqslant \frac{\beta_i}{\gamma_i}$, *we have*

$$\Pr\left(X_i > (1+\Delta)\frac{p_i \alpha_i}{\gamma_i}\right) \leqslant \exp\left(-(\Delta - (e-2))\frac{2p_i\alpha_i}{\lambda_i \gamma_i}\right).$$

and

$$\Pr\left(Y_i > (1+\Delta)\frac{3}{2}\frac{p_i \beta_i}{\gamma_i}\right) \leqslant \exp\left(-(\Delta - (e-2))\frac{3p_i\beta_i}{4\gamma_i}\right).$$

PROOF. According to Theorem A.3,

$$\Pr\left(\frac{X_{i,k}}{\frac{\lambda_i}{2}} = \frac{f_i(e_k)}{\frac{\lambda_i}{2}} \mid \frac{X_{i,1}}{\frac{\lambda_i}{2}}, \cdots, \frac{X_{i,k-1}}{\frac{\lambda_i}{2}}\right) \leqslant p_i.$$
$$\Pr\left(\frac{Y_{i,k}}{2} = 1 \mid \frac{Y_{i,1}}{2}, \cdots, \frac{Y_{i,k-1}}{2}\right) \leqslant \frac{3}{4} p_i.$$

According to Lemma 1,

$$\Pr\left(X_i > (1+\Delta)\frac{p_i\alpha_i}{\gamma_i}\right) \leqslant \Pr\left(X_i > (1+\Delta)p_iF_i \mid F_i = \frac{\alpha_i}{\gamma_i}\right)$$
$$= \Pr\left(\sum_{\{e_k\in\mathcal{S}_i^0\}}\frac{X_{i,k}}{\frac{\lambda_i}{2}} > (1+\Delta)p_i\sum_{\{e_k\in\mathcal{S}_i^0\}}\frac{f_i(e_k)}{\frac{\lambda_i}{2}} \mid F_i = \frac{\alpha_i}{\gamma_i}\right)$$
$$\leqslant \exp\left(-(\Delta-(e-2))\frac{\alpha_i}{\gamma_i\frac{\lambda_i}{2}}p_i\right) = \exp\left(-(\Delta-(e-2))\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right).$$
$$\Pr\left(Y_i > (1+\Delta)\frac{3}{2}\frac{p_i\beta_i}{\gamma_i}\right) \leqslant \Pr\left(Y_i > (1+\Delta)\frac{3}{2}p_iC_i \mid C_i = \frac{\beta_i}{\gamma_i}\right)$$
$$= \Pr\left(\sum_{\{e_k\in\mathcal{S}_i^1\}}\frac{Y_{i,k}}{2} > (1+\Delta)\frac{3}{4}p_i\sum_{\{e_k\in\mathcal{S}_i^1\}}\frac{2}{2} \mid C_i = \frac{\beta_i}{\gamma_i}\right)$$
$$\leqslant \exp\left(-(\Delta-(e-2))\frac{\beta_i}{\gamma_i}\frac{3}{4}p_i\right) = \exp\left(-(\Delta-(e-2))\frac{3p_i\beta_i}{4\gamma_i}\right).$$

□

Theorem A.5. *(Theorem 3) Let $R_wR_\lambda \geqslant 2$, $W = \frac{4N(R_wR_\lambda)^6}{\Lambda(R_w-1)(R_\lambda-1)}$, $\alpha_i = \frac{\|F\|_1}{(R_wR_\lambda)^{i-1}}$, $\beta_i = \frac{\alpha_i}{\frac{\lambda_i}{2}}$, $\gamma_i = (R_wR_\lambda)^{(2^{i-1}-1)}$, and $p_i = (R_wR_\lambda)^{-(2^{i-1}+4)}$. We have*

$$\Pr\left(F_{i+1} > \frac{\alpha_{i+1}}{\gamma_{i+1}} \mid F_i \leqslant \frac{\alpha_i}{\gamma_i} \wedge C_i \leqslant \frac{\beta_i}{\gamma_i}\right)$$
$$\leqslant \exp\left(-(9-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right).$$
$$\Pr\left(C_{i+1} > \frac{\beta_{i+1}}{\gamma_{i+1}} \mid F_i \leqslant \frac{\alpha_i}{\gamma_i} \wedge C_i \leqslant \frac{\beta_i}{\gamma_i}\right)$$
$$\leqslant \exp\left(-(5-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right) + \exp\left(-(\frac{11}{3}-e)\frac{3p_i\beta_i}{4\gamma_i}\right).$$

Proof. According to settings, we have

$$p_i\frac{\alpha_i}{\gamma_i} = \frac{\|F\|_1}{(R_wR_\lambda)^{(2^i+i+2)}} \leqslant \frac{1}{8}\frac{\alpha_{i+1}}{\gamma_{i+1}}$$
$$p_i\frac{\beta_i}{\gamma_i} = p_i\frac{\alpha_i}{\gamma_i\frac{\lambda_i}{2}} \leqslant \frac{1}{8}\frac{\alpha_{i+1}}{\gamma_{i+1}\frac{\lambda_{i+1}}{2}} = \frac{1}{8}\frac{\beta_{i+1}}{\gamma_{i+1}}.$$

Recall that $C_{i+1} = |\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^0| + |\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^1|$, and

$$|\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^0| \leqslant \frac{\sum_{\{e\in\mathcal{S}_i^0\cap\mathcal{S}_{i+1}\}} f_i(e)}{\frac{\lambda_{i+1}}{2}} \leqslant \frac{X_i}{\frac{\lambda_{i+1}}{2}}$$

Let $\Gamma_i = \left(F_i \leqslant \frac{\alpha_i}{\gamma_i} \wedge C_i \leqslant \frac{\beta_i}{\gamma_i}\right)$, according to Theorem A.1 and Theorem 2,

$$\Pr(F_{i+1} > \frac{\alpha_{i+1}}{\gamma_{i+1}} \mid \Gamma_i) \leqslant \Pr\left(X_i > 8p_i\frac{\alpha_i}{\gamma_i} \mid \Gamma_i\right)$$
$$\leqslant \exp\left(-(9-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right).$$

According to Theorem A.2 and Theorem 2,

$$\Pr(C_{i+1} > \frac{\beta_{i+1}}{\gamma_{i+1}} \mid \Gamma_i)$$
$$= \Pr(|\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^0| + |\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^1| > \frac{\beta_{i+1}}{\gamma_{i+1}} \mid \Gamma_i)$$
$$\leqslant \Pr(|\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^0| > \frac{\beta_{i+1}}{2\gamma_{i+1}} \vee |\mathcal{S}_{i+1}^1 \cap \mathcal{S}_i^1| > \frac{\beta_{i+1}}{2\gamma_{i+1}} \mid \Gamma_i)$$
$$\leqslant \Pr(\frac{X_i}{\frac{\lambda_{i+1}}{2}} > \frac{\beta_{i+1}}{2\gamma_{i+1}} \mid \Gamma_i) + \Pr(Y_i > \frac{\beta_{i+1}}{2\gamma_{i+1}} \mid \Gamma_i)$$
$$\leqslant \Pr(X_i > 4p_i\frac{\alpha_i}{\gamma_i} \mid \Gamma_i) + \Pr(Y_i > 4p_i\frac{\beta_i}{\gamma_i} \mid \Gamma_i)$$
$$\leqslant \exp\left(-(5-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right) + \exp\left(-(\frac{11}{3}-e)\frac{3p_i\beta_i}{4\gamma_i}\right).$$

□

A.4 Space and Time Complexity

Theorem A.6. *(Theorem 4) Let $R_wR_\lambda \geqslant 2$, $W = \frac{4N(R_wR_\lambda)^6}{\Lambda(R_w-1)(R_\lambda-1)}$, $\alpha_i = \frac{\|F\|_1}{(R_wR_\lambda)^{i-1}}$, $\beta_i = \frac{\alpha_i}{\frac{\lambda_i}{2}}$, $\gamma_i = (R_wR_\lambda)^{(2^{i-1}-1)}$, and $p_i = (R_wR_\lambda)^{-(2^{i-1}+4)}$. For given Λ and $\Delta < \frac{1}{4}$, let d be the root of the following equation*

$$\frac{R_\lambda^d}{(R_wR_\lambda)^{(2^d+d)}} = \Delta_1\frac{\Lambda}{N}\ln(\frac{1}{\Delta}).$$

And use an SpaceSaving of size $\Delta_2\ln(\frac{1}{\Delta})$ (as the $(d+1)$-layer), then

$$\Pr\left(\forall \text{ item } e, \left|\hat{f}(e) - f(e)\right| \leqslant \Lambda\right) \geqslant 1-\Delta,$$

where

$$\Delta_1 = 2R_w^2R_\lambda^2(R_\lambda-1), \qquad \Delta_2 = 3\left(\frac{R_wR_\lambda^2}{R_\lambda-1}\right)\Delta_1 = 6R_w^3R_\lambda^4.$$

Proof. Recall that $\Gamma_i = \left(F_i \leqslant \frac{\alpha_i}{\gamma_i} \wedge C_i \leqslant \frac{\beta_i}{\gamma_i}\right)$, When all conditions Γ_i (including Γ_{d+1}) are true, we have

$$C_{d+1} \leqslant \frac{\beta_{d+1}}{\gamma_{d+1}} = \frac{2NR_\lambda^{d+1}}{(R_wR_\lambda)^{(2^d+d-1)}(R_\lambda-1)\Lambda} = \left(\frac{2R_wR_\lambda^2}{R_\lambda-1}\right)\Delta_1\ln(\frac{1}{\Delta}).$$
$$F_{d+1} \leqslant \frac{\alpha_{d+1}}{\gamma_{d+1}} = \frac{\lambda_{d+1}}{2}\frac{\beta_{d+1}}{\gamma_{d+1}} = \lambda_{d+1}\left(\frac{R_wR_\lambda^2}{R_\lambda-1}\right)\Delta_1\ln(\frac{1}{\Delta}).$$

Since we use an SpaceSaving of size $\Delta_2\ln(\frac{1}{\Delta}) > C_{d+1}$, it can record all elephant keys without error, and the estimation error for mice keys does not exceed

$$\frac{F_{d+1}}{\Delta_2\ln(\frac{1}{\Delta}) - C_{d+1}} \leqslant \frac{\lambda_{d+1}\left(\frac{R_wR_\lambda^2}{R_\lambda-1}\right)\Delta_1\ln(\frac{1}{\Delta})}{\Delta_2\ln(\frac{1}{\Delta}) - \left(\frac{2R_wR_\lambda^2}{R_\lambda-1}\right)\Delta_1\ln(\frac{1}{\Delta})} = \lambda_{d+1}$$

Therefore, for any item e,

$$\left|\hat{f}(e) - f(e)\right| = \sum_{i=1}^{d}\lambda_i \leqslant \sum_{i=1}^{\infty}\frac{\Lambda(R_\lambda-1)}{R_\lambda^i} = \Lambda$$

Next, we deduce the probability that at least one condition Γ_i is false. Note that

$$\left.\begin{array}{l}(\frac{11}{3}-e)\frac{3p_i\beta_i}{4\gamma_i}\\ (9-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\\ (5-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\end{array}\right\}\geqslant\frac{p_i\alpha_i}{\lambda_i\gamma_i}.$$

Then According to Theorem A.5, we have

$$\Pr\left(\neg\left(\bigwedge_{i=1}^{d}\Gamma_{i+1}\right)\right)=\Pr\left(\bigvee_{i=1}^{d}\neg\Gamma_{i+1}\right)=\Pr\left(\bigvee_{i=1}^{d}\left(\bigwedge_{j=1}^{i}\Gamma_j\wedge\neg\Gamma_{i+1}\right)\right)$$

$$\leqslant\sum_{i=1}^{d}\Pr\left(\Gamma_i\wedge\neg\Gamma_{i+1}\right)\leqslant\sum_{i=1}^{d}\Pr\left(\neg\Gamma_{i+1}\mid\Gamma_i\right)$$

$$\leqslant\sum_{i=1}^{d}\left(\begin{array}{r}\exp\left(-(\frac{11}{3}-e)\frac{3p_i\beta_i}{4\gamma_i}\right)\\ +\exp\left(-(9-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right)+\exp\left(-(5-e)\frac{2p_i\alpha_i}{\lambda_i\gamma_i}\right)\end{array}\right)$$

$$\leqslant\sum_{i=1}^{d}3\exp\left(-\frac{p_i\alpha_i}{\lambda_i\gamma_i}\right).$$

Note that

$$\begin{aligned}\exp\left(-\frac{p_d\alpha_d}{\lambda_d\gamma_d}\right)&=\exp\left(-\frac{NR_\lambda^d}{(R_wR_\lambda)^{(2^d+d+2)}\Lambda(R_\lambda-1)}\right)\\&=\exp\left(-\frac{1}{R_w^2R_\lambda^2(R_\lambda-1)}\Delta_1\ln(\frac{1}{\Delta})\right)\\&=\Delta^{\left(\frac{1}{R_w^2R_\lambda^2(R_\lambda-1)}\Delta_1\right)}=\Delta^2.\end{aligned}$$

Since $\Delta\leqslant 1$, and the monotonicty of $\exp\left(-\frac{p_d\alpha_d}{\lambda_d\gamma_d}\right)$, we have

$$\begin{aligned}\exp\left(-\frac{p_i\alpha_i}{\lambda_i\gamma_i}\right)&=\exp\left(-\frac{p_{i+1}\alpha_{i+1}}{\lambda_{i+1}\gamma_{i+1}}\cdot R_w^{(2^i+1)}R_\lambda^{(2^i)}\right)\\&\leqslant\exp\left(-\frac{p_{i+1}\alpha_{i+1}}{\lambda_{i+1}\gamma_{i+1}}\right)^{R_wR_\lambda}\leqslant\exp\left(-\frac{p_{i+1}\alpha_{i+1}}{\lambda_{i+1}\gamma_{i+1}}\right)^2\\&\leqslant\Delta^2\exp\left(-\frac{p_{i+1}\alpha_{i+1}}{\lambda_{i+1}\gamma_{i+1}}\right)\end{aligned}$$

Therefore, we have

$$\sum_{i=1}^{d}3\exp\left(-\frac{p_i\alpha_i}{\lambda_i\gamma_i}\right)\leqslant 3\sum_{i=1}^{d}\Delta^{2i}\leqslant\left(\frac{3\Delta}{1-\Delta^2}\right)\Delta\leqslant\Delta.$$

In other words,

$$\Pr\left(\forall\ \mathit{item}\ e,\left|\hat{f}(e)-f(e)\right|\leqslant\Lambda\right)\geqslant 1-\Delta,$$

which leads to a weaker conclusion,

$$\forall\ \mathit{item}\ e,\Pr\left(\left|\hat{f}(e)-f(e)\right|\leqslant\Lambda\right)\geqslant 1-\Delta.$$

□

Theorem A.7. *Using the same settings as Theorem A.6, the space complexity of the algorithm is $O(\frac{N}{\Lambda}+\ln(\frac{1}{\Delta}))$, and the time complexity of the algorithm is amortized $O(1+\Delta\ln\ln(\frac{N}{\Lambda}))$.*

Proof. Recall that d is the root of the equation

$$\frac{R_\lambda^d}{(R_wR_\lambda)^{(2^d+d)}}=\Delta_1\frac{\Lambda}{N}\ln(\frac{1}{\Delta}),$$

which means $d=O\left(\ln\ln(\frac{N}{\Lambda})\right)$. Therefore, total space used by the data structure is

$$\begin{aligned}\sum_{i=1}^{d}w_i+\Delta_1\ln(\frac{1}{\Delta})=&\sum_{i=1}^{d}\lceil\frac{W(R_w-1)}{R_w^i}\rceil+O(\ln(\frac{1}{\Delta}))\\ \leqslant&\frac{4N(R_wR_\lambda)^6}{\Lambda(R_w-1)(R_\lambda-1)}+d+O(\ln(\frac{1}{\Delta}))\\ =&O(\frac{N}{\Lambda}+\ln(\frac{1}{\Delta}))\end{aligned}$$

Next, we analyze the time complexity. When all condition Γ_i are true, for a new item $e\notin\mathcal{S}_1$, the probability that it enters the $(i+1)$-th layer from the i-th layer is $\frac{\frac{F_i}{\frac{\lambda_2}{2}}+C_i}{w_i}\leqslant p_i$. Thus the time complexity of insert item e does noes exceed

$$(1-\Delta)\cdot(1+\sum_{i=1}^{d}p_i)+\Delta\cdot d=O(1+\Delta\ln\ln(\frac{N}{\Lambda})).$$

□