B MICROSCOPESKETCH ERROR ANALYSIS

The symbols frequently used in this section are shown in Table 2.

Table 2: Symbols used in this section.

Notation	Meaning
$\hat{f_i}$	Estimated frequency given by the i^{th} sub-window
f^*	Exact result recorded if there was no rounding
$f(t_1,t_2)$	Frequency during period $(t_1,t_2]$
$f^{(i)}$	True frequency of item e_i in the recent sliding window

B.1 Unilateral Error (Overestimation)

To achieve overestimation only, we round up the pixel counters when zooming out. We apply the overestimation method in Eq. 2 and use the following formula for querying:

$$\hat{f} = S + \sum_{i=n-T}^{n} \hat{f_i} \tag{3}$$

Theorem B.1. *If we use Eq. 3 for querying, $\hat{f} \geqslant f$.*

Proof. For the latest T sub-windows, items falling in this period must be contained in the pixel counters. For the latest $(T+1)^{th}$ sub-window, the frequency must be overestimated because S could be over 0 when the sub-window starts. Also, because we round up the pixel counters when zooming out, the frequency represented by the pixel counters could be overestimated. Therefore, the reported frequency must be no less than the true frequency. □

B.2 Unilateral Error (Underestimation)

To achieve underestimation only, we round down the pixel counters when zooming out. Let i_0 be the minimum $i \geqslant 1$ such that $\hat{f}_{n-T+i} \neq 0$. We apply the underestimation method from Eq. 2 and use the following formula for querying.

$$\hat{f} = S + \sum_{i=n-T+1}^{n} \hat{f_i} - \hat{f_{i_0}} \tag{4}$$

Theorem B.2. *If we use Eq. 4 for querying, $\hat{f} \leqslant f$.*

Proof. During insertion, instead of directly increasing pixel counters, we increase the standard counter S first. As a result, when starting a new sub-window, S leads to overestimation because it may not be 0. Let i_0 be the minimum $i \geqslant 1$ such that $\hat{f}_{n-T+i} \neq 0$. If the overestimation caused by S occurs, it must be recorded in $P[(n-T+i_0) \bmod (T+2)]$. Also, because we round down the pixel counters when zooming out, the frequency recorded in the pixel counters can only be lower than the true frequency during the corresponding period.

We'll get underestimation if we eliminate the overestimated part, by removing $\hat{f_{i_0}}$ from the estimation $S + \sum_{i=n-T+1}^{n} \hat{f_i}$. Therefore, $\hat{f} \leqslant f$, *i.e.*, our MicroSketch achieves underestimation by Eq. 4. □

B.3 Unbiased Rounding Error

Here unbiased rounding error implies that rounding does not produce bias, *i.e.*, the frequency is given by MicroSketch is an unbiased estimate of the result if we do not do the rounding when zooming out. To achieve unbiased rounding error, we apply the optimization of unbiased rounding from Section 3.3. When zooming out, we round it up with a probability of $P \bmod c$; otherwise we round up it.

Theorem B.3. *Given an item e, let $\hat{f}$ be the estimated frequency given by our MicroSketch and f^* the exact result that should be recorded if there was no rounding. $E(\hat{f}) = f^*$, i.e., our MicroSketch has a property of unbiased rounding error.*

Proof. In a sub-window, the estimated frequency will change when we zoom out and do rounding. Let P be the number recorded in the pixel counter and Z the zooming counter before halving. We do rounding only when P is odd. The pixel counter after halving P' has a probability of $\frac{P \bmod c}{c}$ to be $\frac{P-P \bmod c}{c}+1$ and a probability of $\frac{1-P \bmod c}{c}$ to be $\frac{P-P \bmod c}{c}$. At the same time, Z is increased by 1. The expectation is $E(P' \cdot c^{Z+1}) = \frac{P \bmod c}{c} \cdot (\frac{P-P \bmod c}{c}+1) \cdot c^{Z+1}$ $+ \frac{1-P \bmod c}{c} \cdot (\frac{P-P \bmod c}{c}) \cdot c^{Z+1} = P \cdot c^Z$, which indicates that the expectation will not change after we halve the pixel counter.

Therefore, $E(\hat{f_i}) = f_i$, where f_i is the frequency without rounding during the i^{th} sub-window. Since $\hat{f}$ is a linear combination of the estimated frequency of pixel counters, we have $E(\hat{f}) = f^*$, *i.e.*, MicroSketch has a property of unbiased rounding error. □

In subsection B.4 we will further show that MicroSketch is unbiased under specific assumption. For those structures where items are inserted into fixed counters, if we use MicroSketches to replace each counter, the bias will not increase further since MicroSketch is unbiased. Particularly, if the estimate by the original structure is unbiased, then the adapted structure also achieves unbiased results.

Below we show the variance of MicroSketch when we apply linear approximation from Eq. 2 and use unbiased rounding.

Theorem B.4. *Let Z be the number recorded in the zooming counter when querying and R_i be $(f_i \bmod c^Z)$. The variance of $\hat{f}$ satisfies*

$$Var(\hat{f}) = \sum_{i=n-T+1}^{n} R_i \cdot (c^Z - R_i) + p^2 \cdot R_{n-T} \cdot (c^Z - R_{n-T}) \tag{5}$$

Proof. Since randomness comes from rounding, f_i and Z are not random. Thus, R_i is also a fixed value. Let P_0 be $\frac{f_i - R_i}{c^Z}$. The number recorded in the pixel counter $P[i \bmod (T+2)]$ is either P_0 or P_0+1. Hence, $\hat{f_i} = P[i \bmod (T+2)] \times c^Z = f_i - R_i$ or $f_i - R_i + c^Z$. Since $E(\hat{f_i}) = f_i$, $\hat{f_i}$ is $(f_i - R_i + c^Z)$ with a probability $\frac{R_i}{c^Z}$, and $(f_i - R_i)$ with a probability $1 - \frac{R_i}{c^Z}$. The variance of $\hat{f_i}$ satisfies

$$\begin{aligned} Var(\hat{f_i}) &= E(\hat{f_i} - E(\hat{f_i}))^2 = \frac{R_i}{c^Z} \cdot (c^Z - R_i)^2 + (1 - \frac{R_i}{c^Z}) \cdot R_i^2 \\ &= R_i \cdot (c^Z - R_i) \end{aligned} \tag{6}$$

The rounding in all pixel counters is independent of each other. Therefore, we have $Var(\hat{f}) = Var(S + \sum_{i=n-T+1}^{n} \hat{f_i} + p \cdot \hat{f}_{n-T}) = \sum_{i=n-T+1}^{n} R_i \cdot (c^Z - R_i) + p^2 \cdot R_{n-T} \cdot (c^Z - R_{n-T})$. □

B.4 Error Bound

Below, we introduce the error caused by the sliding window and do further analysis. For the current time t, we assume that the increment of the frequency is stable at the edge of the sliding window, *i.e.*, the frequency is proportional to the duration. Let v be the frequency in normalized time at the edge of the sliding window.

First, we prove that $\hat{f}$ is unbiased, *i.e.*, $E(\hat{f})=f$. From previous analysis, $E(\hat{f}-f^*)=0$. Let $f(t_1,t_2)$ be the frequency during period $(t_1,t_2]$. f is the frequency in a window. We have $f=f(t-W,t)$ $=f(t-Tw,t-t \bmod w-(T-1)w)+f(t-t \bmod w-(T-1)w,t)$. Since f^* is the exact result that should be recorded if there was no rounding, we can get that $f^*=f(t-t \bmod w-(T-1)w,t)+$ $p\cdot f(t-t \bmod w-Tw,t-t \bmod w-(T-1)w)$. Thus, we have $E(f^*-f)=p\cdot f(t-t \bmod w-Tw,t-t \bmod w-(T-1)w)-$ $f(t-Tw,t-t \bmod w-(T-1)w)=p\cdot wv-(w-t \bmod w)v=0$.

Therefore, under our assumption we have $E(\hat{f})-f=E(\hat{f}-f^*)-E(f^*-f)=0$, *i.e.*, the estimated frequency by MicroSketch is unbiased. Similarly to the previous proof, we have the variance $Var(\hat{f})=\sum_{i=n-T+1}^{n}R_i\cdot(c^Z-R_i)+p^2\cdot R_{n-T}\cdot(c^Z-R_{n-T})$. According to Chebyshev inequality, we get the error bound of MicroSketch: for $\epsilon\geqslant 0$, we have

$$\begin{aligned}&Pr\left[\left|\hat{f}-f\right|\geqslant\epsilon\right]\leqslant\frac{1}{\epsilon^2}Var(\hat{f})\\&=\frac{1}{\epsilon^2}\left(\sum_{i=n-T+1}^{n}R_i\cdot(c^Z-R_i)+p^2\cdot R_{n-T}\cdot(c^Z-R_{n-T})\right)\end{aligned}\tag{7}$$

Below we consider the error bound of MicroSketch-CM with d MicroSketches arrays and the size of each array is m. Let $e_1,e_2,\cdots,e_M$ be the items inserted into MicroSketch-CM and their true frequency from time $(t-W)$ to t is $f^{(1)},f^{(2)},\cdots,f^{(M)}$, and let $\boldsymbol{f}=(f^{(1)},f^{(2)},\cdots,f^{(M)})$. For an item e_i, let $\hat{f}^{(i)}$ be the estimated frequency by our MicroSketch-CM and $\hat{f}_{CM}^{(i)}$ be the frequency in the recent sliding window,*i.e.*, from time $(t-W)$ to t, by a corresponding CM sketch .

For the error bound of the estimated frequency of e_i, we can consider it as two parts:

$$\begin{aligned}Pr\left[\left|\hat{f}^{(i)}-f^{(i)}\right|\geqslant\epsilon\|\boldsymbol{f}\|_1\right]&\leqslant Pr\left[\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]\\&+Pr\left[\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\leqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1,\left|\hat{f}^{(i)}-\hat{f}_{CM}^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]\end{aligned}\tag{8}$$

From the theorem about the error bound of the CM sketch[22], when $m\geqslant\frac{2e}{\epsilon}$ we have $Pr\left[\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]<e^{-d}$.

The estimation is given by the MicroSketch with the minimal value among the d mapped MicroSketches. In that specific MicroSketch, from Eq. 7 we can get that $Pr\left[\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]$ $\leqslant\frac{4}{\epsilon^2\|\boldsymbol{f}\|_1^2}\left(\sum_{i=n-T+1}^{n}R_i\cdot(c^Z-R_i)+p^2\cdot R_{n-T}\cdot(c^Z-R_{n-T})\right)$ $\leqslant\frac{4}{\epsilon^2\|\boldsymbol{f}\|_1^2}\left(T\cdot\left(\frac{c^Z}{2}\right)^2+p^2\cdot\left(\frac{c^Z}{2}\right)^2\right)\leqslant\frac{1}{\epsilon^2\|\boldsymbol{f}\|_1^2}(T+1)\cdot c^{2Z}$.

When $\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\leqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1$, we have $c^Z\cdot c^{l-1}\leqslant c^Z\cdot\max P\leqslant\hat{f}_{CM}^{(i)}\leqslant f^{(i)}+\frac{\epsilon}{2}\|\boldsymbol{f}\|_1$, here P is the number recorded in the pixel counters and l is the number of bits used for each pixel counter. Thus, we can get that $Z\leqslant\left\lfloor\ln\left(f^{(i)}+\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right)/\ln c\right\rfloor-l+1$.

Let $Z_0=\left\lfloor\ln\left(f^{(i)}+\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right)/\ln c\right\rfloor-l+1$. Then we have

$$\begin{aligned}&Pr\left[\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\leqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1,\left|\hat{f}^{(i)}-\hat{f}_{CM}^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]\\\leqslant&Pr\left[\left|\hat{f}^{(i)}-\hat{f}_{CM}^{(i)}\right|\geqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\middle|\left|\hat{f}_{CM}^{(i)}-f^{(i)}\right|\leqslant\frac{\epsilon}{2}\|\boldsymbol{f}\|_1\right]\\\leqslant&\sup_{Z\leqslant Z_0}\frac{1}{\epsilon^2\|\boldsymbol{f}\|_1^2}(T+1)\cdot c^{2Z}\leqslant\frac{1}{\epsilon^2\|\boldsymbol{f}\|_1^2}(T+1)\cdot c^{2Z_0}\end{aligned}\tag{9}$$

Therefore, for any $\epsilon\geqslant 0$, when $m\geqslant\frac{2e}{\epsilon}$ we can get the error bound of our MicroSketch-CM

$$Pr\left[\left|\hat{f}^{(i)}-f^{(i)}\right|\geqslant\epsilon\|\boldsymbol{f}\|_1\right]\leqslant e^{-d}+\frac{1}{\epsilon^2\|\boldsymbol{f}\|_1^2}(T+1)\cdot c^{2Z_0}\tag{10}$$