$\hat{f}_i$ that $Var(\hat{f}_i) = E_{s(e_j)\in\{1,-1\}}\left(\left(\sum_{e_j\in S_e, j\neq i} f_j \cdot s(e_j)\right)\cdot s(e_i)\right)^2 = E_{s(e_j)\in\{1,-1\}}\left(\sum_{e_j\in S_e, j\neq i} f_j \cdot s(e_j)\right)^2$

From the analysis in 5.1, we find that $s(e_i)$ and whether e_i is error-free is independent. Thus, the cross terms have the same chance to be 1 and -1, so the expectation of their sum is 0. Therefore, we have $Var(\hat{f}_i) = E_{s(e_j)\in\{1,-1\}}\left(\sum_{e_j\in S_e, j\neq i}(f_j)^2\right) \leqslant \sum_{e_j\neq e_i}(f_j)^2$ □

According to the variance, we can derive an error bound for $\|f\|_2$.

Theorem D.2. *Let $l = \frac{e}{\epsilon^2}$, then $P\left(\left|\hat{f}_i - f_i\right| \geqslant \epsilon \|f\|_2\right) \leqslant \frac{1}{e}$*

Proof. Based on Chebyshev's theorem, we can get that $P\left(\left|\hat{f}_i - f_i\right| \geqslant \sqrt{e\sum_{e_j\neq e_i}(f_j)^2}\right) \leqslant \frac{Var(\hat{f}_i)}{\left(\sqrt{e\sum_{e_j\neq e_i}(f_j)^2}\right)^2} \leqslant \frac{1}{e}$.

For items in $B[h(e_i)]$, we can have an estimation that $\sum_{e_j}(f_j)^2 = \frac{1}{l}(\|f\|_2)^2$. Therefore, we can get

$P\left(\left|\hat{f}_i - f_i\right| \geqslant \epsilon\|f\|_2\right) \leqslant P\left(\left|\hat{f}_i - f_i\right| \geqslant \epsilon\sqrt{l\cdot\sum_{h(e_j)=h(e_i)} f_j^2}\right) \leqslant P\left(\left|\hat{f}_i - f_i\right| \geqslant \sqrt{e\sum_{e_j\neq e_i} f_j^2}\right) \leqslant \frac{1}{e}$ □

We can find that this bound is relatively loose because it also takes effect on the items in the Waving Counter. However, for items in the Heavy Part, $\sqrt{l\cdot\sum_{e_j\neq e_i}(\Delta_i f_j)^2}$ is often much smaller than $\|f\|_2$.

We can also derive an error bound of $\|f\|_1$.

Theorem D.3. *Let $l = \frac{e}{\epsilon}$, we have*

$$P\left(\left|\hat{f}_i - f_i\right| \geqslant \epsilon\|f\|_1\right) \leqslant \frac{1}{e}$$

Proof. For our WavingSketch, we have

$$\mathbb{E}\left[\left|\hat{f}_i - f_i\right|\right] = \mathbb{E}\left[\left|\sum_{e_j\neq e_i} f_j\cdot s(e_j)\right|\right] \leqslant \mathbb{E}\left[\left|\sum_{e_j\neq e_i} f_j\right|\right] \leqslant \frac{\epsilon\|f\|_1}{e}$$

By the Markov inequality,

$$P\left(\left|\hat{f}_i - f_i\right| \geqslant \epsilon\|f\|_1\right) \leqslant P\left(\left|\hat{f}_i - f_i\right| \geqslant e\mathbb{E}\left[\left|\hat{f}_i - f_i\right|\right]\right) \leqslant \frac{1}{e}$$

□

D.2 Parameter Analysis

We analyze the influence of parameters in WavingSketch. We use $c = dl$ to denote the number of cells in WavingSketch. Then we show that for fixed c, how d influences the performance of our WavingSketch.

Theorem D.4. *Let e_i be the i_{th} most frequent item in the data stream. The probability that its frequency f_i is among top-d largest frequencies in bucket $B[h(e_i)]$ is at least $1 - \frac{d^d}{d!}\cdot\left(\frac{i-1}{c}\right)^d$*

Proof. Let P_i be the probability that $B[h(e_i)]$ contains at least d items whose frequency is higher than e_i. When $i \leqslant d$, $P_i = 0$. So we only need to discuss the case that $i > d$. When $i > d$, we have $P_i \leqslant \binom{i-1}{d}\cdot\left(\frac{1}{l}\right)^d \leqslant \frac{d^d}{d!}\cdot\left(\frac{i-1}{c}\right)^d$. Therefore, the probability that f_i is among top-d largest frequencies in bucket $B[h(e_i)]$ is at least $1 - \frac{d^d}{d!}\cdot\left(\frac{i-1}{c}\right)^d$. □

We can find that, when i decreases, P_i decreases sharply, which indicates that the probability that e_i is top-d items in $B[h(e_i)]$ becomes much higher. According to Stirling's approximation,

$$1 - \frac{d^d}{d!}\cdot\left(\frac{i-1}{c}\right)^d \approx 1 - \frac{1}{\sqrt{2\pi d}}\cdot\left(\frac{e(i-1)}{c}\right)^d \tag{3}$$

We can also find that, when $i < \frac{c}{e} + 1$, the probability that e_i is top-d items in $B[h(e_i)]$ increases with d increasing.