$F[i] \gtreqless 0$:

# ALLOWING NEGATIVE COUNTERS

IDEA $F[i] = 5 - 3 - 2 + 7 - 1 + 7 - 1 - 1$

$= \underbrace{(5+7+7)}_{F^+[i]} - \underbrace{(3+2+1+1+1)}_{F^-[i]}$

$$F^+[i] \quad - \quad F^-[i]$$

$$\tilde{F}[i] = \tilde{F}^+[i] - \tilde{F}^-[i]$$

Keep 2 count-min sketches: one for $F^+$ and one for $F^-$

The analysis for $F^+$ or $F^-$ is the same as before when approximated with $\tilde{F}^+$ and $\tilde{F}^-$.

$i_1, i_2, \ldots, i_s, \ldots$

$\langle i_1, z_1 \rangle, \langle i_2, x_2 \rangle, \ldots, \langle i_s, z_s \rangle.$

$$\textcircled{\triangle} \quad \|F\| = \sum_{i=1}^{n} |F[i]| = \|F^+\| + \|F^-\|$$

$$\|F^+\| \to X^+_{ji}, \quad \|F^-\| \to X^-_{ji}$$

observations

$$\textcircled{\times} \quad |X_{ji}| = |\underbrace{X^+_{ji}}_{\geq 0}| + |\underbrace{X^-_{ji}}_{\geq 0}| = X^+_{ji} + X^-_{ji}$$

$$\Pr\left[|X_{ji}| > \varepsilon \|F\|\right] \underset{\text{Markov's inequality}}{<} \frac{E\left[|X_{ji}|\right]}{\varepsilon \|F\|} \overset{\otimes}{=} \frac{\overbrace{E[X^+_{ji}]}^{\frac{\varepsilon}{e}\|F^+\|} + \overbrace{E[X^-_{ji}]}^{\frac{\varepsilon}{e}\|F^-\|}}{\varepsilon \|F\|}$$

$$\overset{\triangle}{=} \frac{\frac{\varepsilon}{e}\cancel{\|F\|}}{\cancel{\varepsilon}\cancel{\|F\|}} = \frac{1}{e}$$

we can bound the error in this way
the rest of the analysis is unchanged.

# Interval queries

$\langle i_1, z_1 \rangle \ \langle i_2, z_1 \rangle \cdots \cdots \boxed{\langle i_s, z_s \rangle} \cdots \cdots$

$F[1] \ F[2] \cdots F[i_s] \cdots F[n]$

$\Downarrow$

$+z_s$

Simple query: $F[i]$

$x = y = i$

Interval query: $F[x] + F[x+1] + \cdots + F[y] \triangleq F_{xy} = \sum_{i=x}^{y} F[i]$

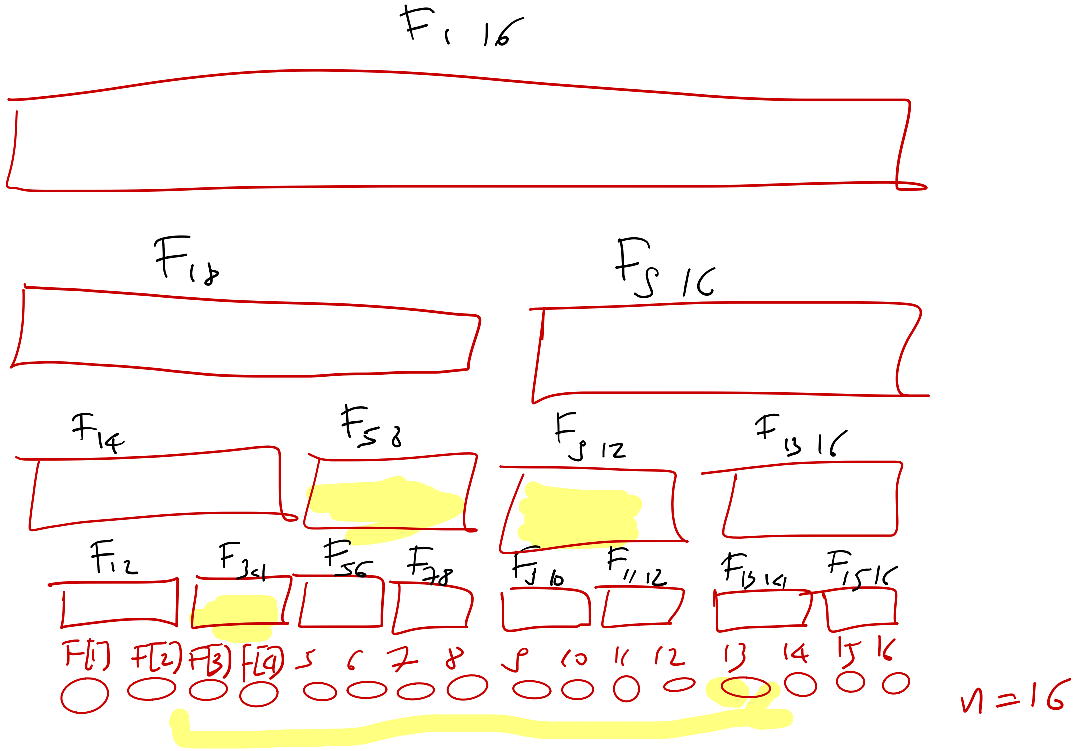
Wrong way: $\tilde{F}_{xy} = \sum_{i=x}^{y} \tilde{F}[i]$

① expensive, cost $O(y-x+1)$ time

② error is additive?

dyadic intervals

$$F_{3\,13} = \underbrace{F_{3\,4} + F_{5\,8} + F_{9\,12} + F_{13}}_{\le 2 \lg n \text{ addends}}$$



**obs** There are $n + \frac{n}{2} + \frac{n}{4} + \ldots < 2n$ dyadic intervals and thus many counters $F_{ab}$

**obs** When $F[i_s] \mathrel{+}= z_s$ is executed, we also perform $F_{ab} \mathrel{+}= z_s$ for each interval such that $a \le i_s \le b$ (we change $\log n$ counters at most)

**obs** Given a query $F_{xy}$, the latter can be expressed as the sum of $\le 2 \log n$ dyadic intervals (indeed, if there were 3 consecutive intervals on the same level, then we could replace two of them by an interval of double size on the upper level)

$$F_{xy} = \sum_{a,b} F_{ab}$$

We use count-min sketches on the counters $F_{ab} \rightsquigarrow \tilde{F}_{ab}$, thus

$$\boxed{\tilde{F}_{xy} = \sum_{a,b} \tilde{F}_{ab}}$$

① cost is $O(\lg n)$ time

② error is bounded

$$\tilde{F}_{xy} \leq F_{xy} + 2\varepsilon \lg n \, \|F\| \quad \text{with prob.} > 1-\delta$$

Let's see ②

Define $X_{ji}^{ab}$ as the garbage for $F_{ab}$, that is

$$\tilde{F}_{ab} = F_{ab} + X_{ji}^{ab}$$

Let $X_{xy} = \sum_{ab} X_{ji}^{ab}$

Observe that $\tilde{F}_{xy} = F_{xy} + X_{xy}$ and that $E[X_{xy}] = \frac{2\varepsilon \lg N}{e} \|F\|$

Then $\Pr[X_{xy} > 2\varepsilon \lg N \|F\|] < \frac{E[X_{xy}]}{2\varepsilon \lg N \|F\|} = \frac{1}{e}$

The rest is identical! $\square$

Note Using the individual errors on each $X_{ji}^{ab}$ does not work!
That is, $\Pr[X_{ji}^{ab} > \varepsilon \|F\|]$ does not imply the above result.