Problem: →

at time t (i, S_i)

$\Rightarrow C_i \rightarrow C_i + S_i$ at time t.

→ We are only allowed $<< O(N)$ memory sketch S data structure.

$C = [\ \ |\ \ |\ \ |\ \ |\ \cdots\ |\ \ ]$

1 2 3 4 N

Goal: →

→ What is the count of comp. i. C_i ?

Update(S, i, S_i)

float queryCounts(S, i)

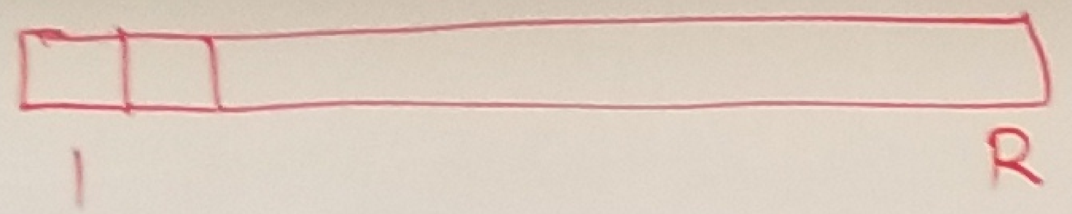
for $t = 1 \rightarrow T$

update(S, i, S_i)

Count Sketch

float queryCount(S,i)

Count - Sketch

S → 

$O(R) << O(N)$

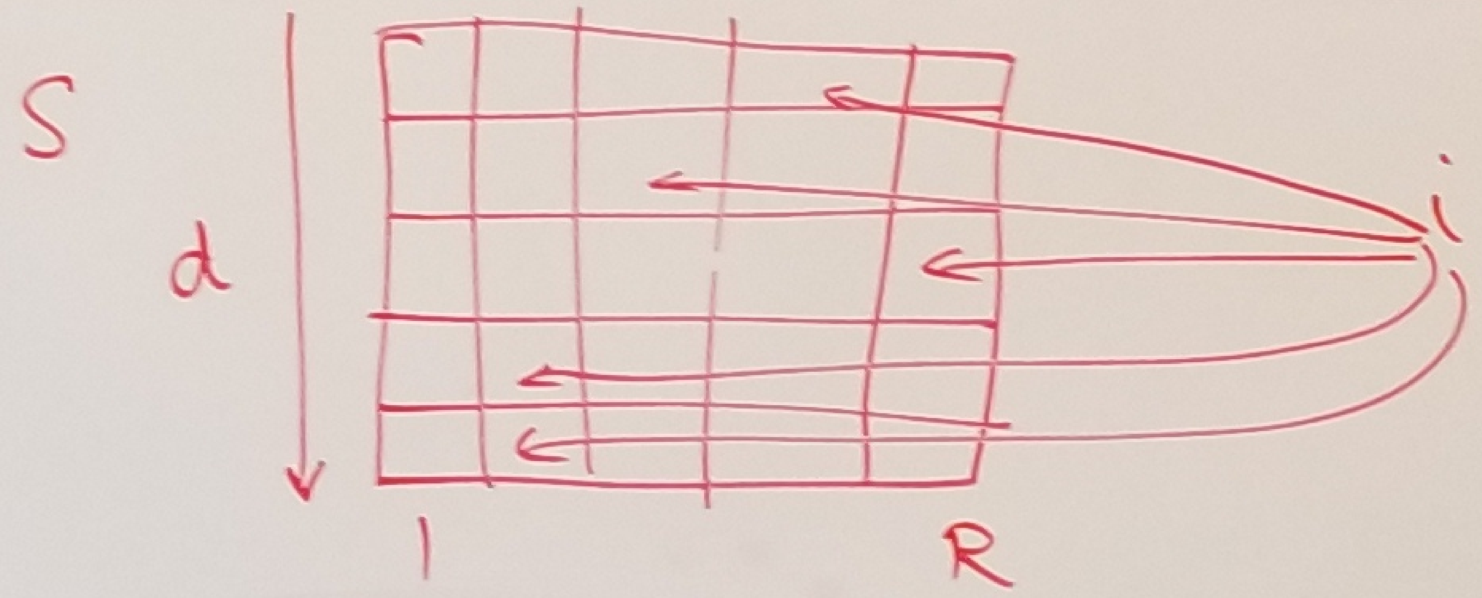
Update(S, i, δ_i)

$S[h(i)] \mathrel{+}= \delta_i$

$$E[S(h(i))] = E\left[c_i + \sum_{j \neq i} c_j \mathbb{1}_{\{h(j) = h(i)\}}\right]$$

Query(S, i)

return $S[h(i)]$



Update(S, i, δ_i)

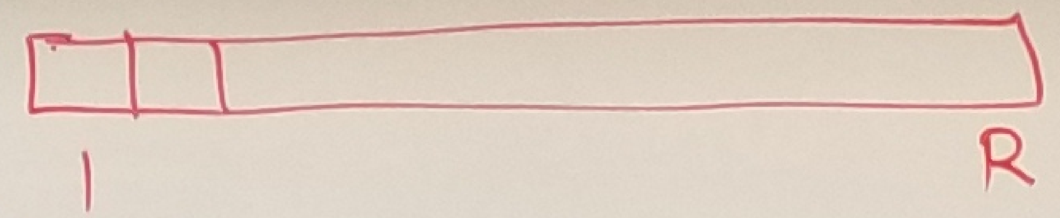
for $j = 1$ to d

$S[j][h_j(i)] \mathrel{+}= \delta_i$

Query(S, i)

return $\min_j S[j][h_j(i)]$

Count " Sketch.

$S \rightarrow$ 

$O(R) << O(N)$

$h : [1 - N] \longrightarrow [1 - R]$

$g : [1 - N] \rightarrow \{+1, -1\}$

$E(g(i)) = 0 \; ; \; E\left[1_{\{h(i)=h(j)\}_{i \neq j}} \right] = 1/R$

$g(i) \times S(h(i))$

Update (S, i, δ_i)

$S[h(i)] \mathrel{+}= \delta_i \times g(i)$

$$= \left[\sum_j 1_{\{h(i)=h(j)\}} \times c_j \times g(j) \right] \times g(i)$$

Query (S, i)

return $S[h(i)] \times g(i)$

$$= c_i \times g(i)^2 + \sum_{j \neq i} 1_{\{h(j)=h(i)\}} \times c_j \times g(j) \times g(i)$$

$E(g(i) \times S(h(i))]$

$$= c_i \, E[g(i)^2] + \sum_{j \neq i} c_j \times E(1_{\{\}}) \times \cancel{E(g(j))} \times \cancel{E(g(i))}$$

$$= c_i$$

$$g(i) \times S(h(i)) = \hat{c}_i$$

$$\boxed{E(\hat{c}_i) = c_i}$$

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$$

$$(a+b+c)^2 = (a+b+c)(a+b+c)$$

$$\mathrm{Var}(\hat{c}_i) = E\left[\left(g(i) \times S[h(i)] - c_i\right)^2\right] =$$

$$= E\left[\left(c_i \times g(i)^2 + \sum_{j \neq i} 1_j \times c_j \times g(j) \times g(i) - c_i\right)^2\right] = E\left[\left(\sum_{j \neq i} 1_j \times c_j \times g(j) \times g(i)\right)^2\right]$$

$$= E\left[\sum_{j \neq i} 1_j^2 \times c_j^2 \times g(j)^2 \times g(i)^2 + \sum_{\substack{j \neq i \\ k \neq i}} 1_j 1_k \times c_j \times c_k \times g(j) \times g(k) \times g(i) \times g(i)\right]$$

$$= E\left[\sum_{j \neq i} 1_j \times c_j^2 + \cancel{\sum_{\substack{j \neq i \\ k \neq i}} 1_j 1_k \times c_j c_k \times g(j) \times g(k)}\right] = \sum_{j \neq i} c_j^2 \times E(1_j) = \frac{1}{R} \sum_{j \neq i} c_j^2$$

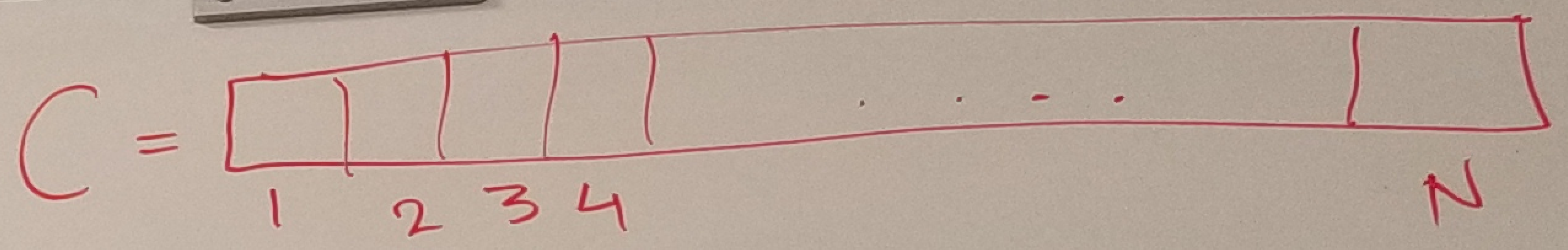
$$\mathrm{Var}(\hat{c}_i) \le \frac{1}{R}\sum c_j^2 = \frac{\|c\|_2^2}{R}$$

$$\Pr(|X - E(x)| > a) < \frac{\mathrm{Var}(X)}{a^2}$$

$$\sigma = \sqrt{\mathrm{Var}}$$

$$\Rightarrow \Pr(|X - E(x)| > k\sigma) < \frac{1}{k^2}$$

$$\Pr\left(\text{Error} > \sqrt{\frac{3}{R}}\,\|c\|_2\right) \le \frac{1}{3}$$



Problem: →

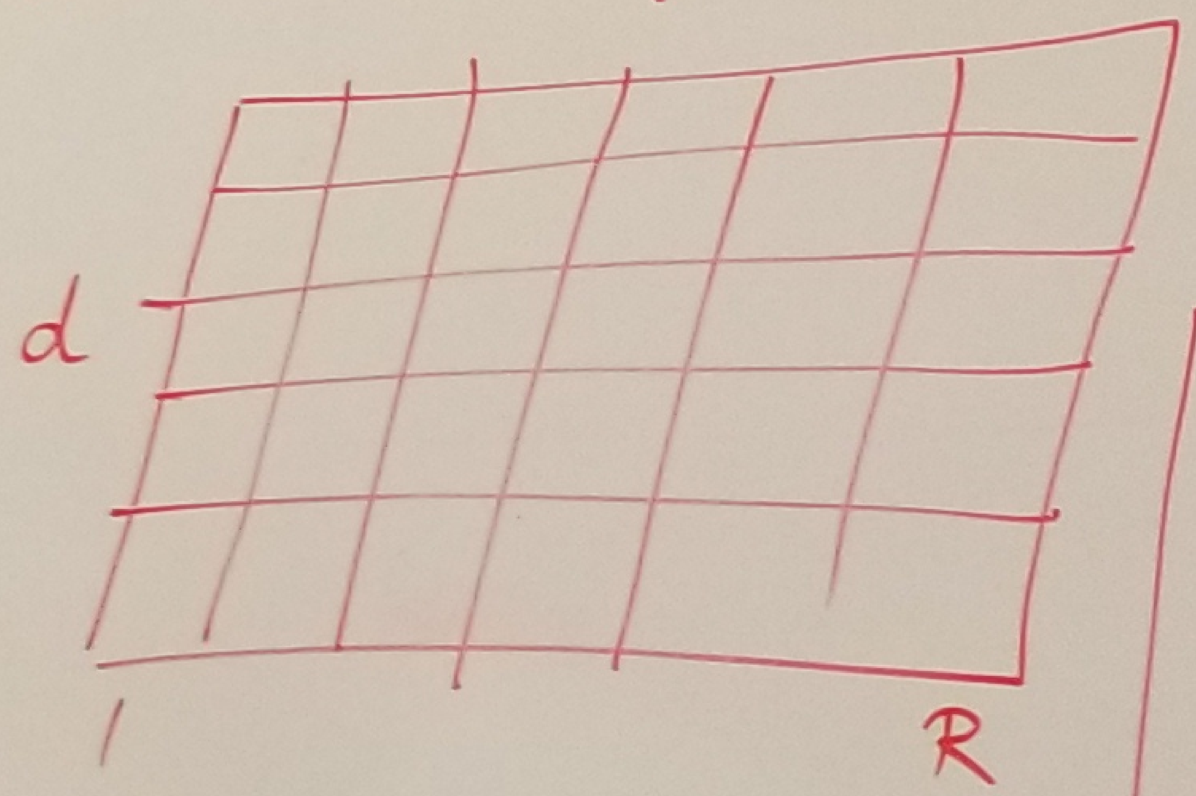
at time t c_i,

⇒ $c_i \to c_i + \delta_i$ at time t.

→ We are only allowed << O(N) memory sketch S data structure.

$h_1, h_2 \dots h_d : [1-N] \to [1-R]$

$g_1, g_2 \dots g_d : [1-N] \to \{+1, -1\}$



Update (S, i, δ_i)

for $j = 1 \to d$

$S[j][h_j(i)] \mathrel{+}= \delta_i \times g_j(i)$

Query (S, i)

return $\underset{j}{\text{median}} \left[S[j][h_j(i)] \times g(i) \right]$

$c_i - \frac{3}{\sqrt{R}}\|c\|_2 \le \hat{c}_i \le c_i + \frac{3}{\sqrt{R}}\|c\|_2$

with prob $1-\delta$

$d = \log(1/\delta)$