# LOWER BOUND

Consider the Count-Min Sketch problem $\Rightarrow$

$\Omega(1/\varepsilon)$ words required

CONCEPTUAL STEPS

① INDEX problem in communication complexity

② Reduction from INDEX to Count-Min Sketches

## ① Preliminaries

Entropy $H(p) = -p\log_2 p - (1-p)\log_2(1-p)$ # bits of an event occurring with probability $p$

INDEX problem

Alice has a sequence $x \in \{0,1\}^n$

Bob has an index $i \in [n]$ (i.e. a sequence of $\log_2 n$ bits)

one-round protocol (i.e. only one message can be sent)

Bob wants to compute $x[i]$ with probability $p > \frac{1}{2}$

[$p = \frac{1}{2}$ is easy; why?]

[two rounds is easy: why?]

Q: how many bits Alice should send in the message?
(recall: only one msg can be exchanged)

A: # needed bits is $\geq n(1 - H(p))$

## ②

Given an instance $(x, i)$ of INDEX, choose $\varepsilon$ s.t. $n = \frac{1}{\varepsilon}$ where $|X| = n$

Define $F[1..\frac{1}{\varepsilon}]$: $F[i] = \begin{cases} 2 & \text{if } x[i] = 1 \\ 0 & \text{if } x[i] = 0 \end{cases}$

NOTE: choose $x$ s.t. #1s = #0s in $X$: there are $\binom{n}{n/2}$ such strings

Hence, $||F|| = 2 \times \frac{n}{2} = \frac{1}{\varepsilon}$

Count-Min Sketch gives $\tilde{F}[i]$ s.t. $F[i] \le \tilde{F}[i] \le F[i] + \underbrace{\varepsilon ||F||}_{=1}$

Thus, the approximate value satisfy $F[i] \le \tilde{F}[i] \le F[i] + 1$

FACT: since $F[i]$ is either 0 or 2, this implies that we can determine with certainty $F[i]$ from $\tilde{F}[i]$! Thus, $X[i]$...

INDEX $\Rightarrow$ the space required to store $\tilde{F}[1..\frac{1}{\varepsilon}]$ is $\Omega(n) = \Omega(1/\varepsilon)$ words

Q.E.D.

## PROBLEM

Count-Min Sketch assumes $\varepsilon$ to be a constant... instead the above lower bound uses $\varepsilon = \frac{1}{n}$ non-const.
Find a way to use $\varepsilon$ = constant!