Gems of TCS

Communication Complexity

Sasha Golovnev

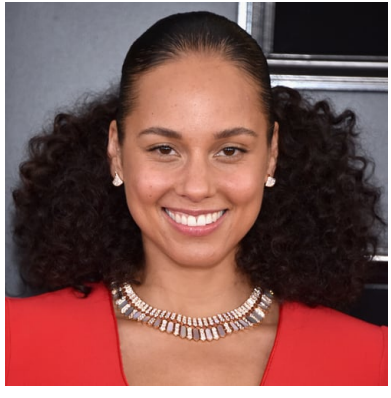
April 8, 2021

Communication Complexity

$F: \{0,1\}^{2n} \rightarrow \{0,1\}$

$F(x_1, x_2, \ldots, x_n, y_1, \ldots, y_n) \quad x_i, y_j \in \{0,1\}$

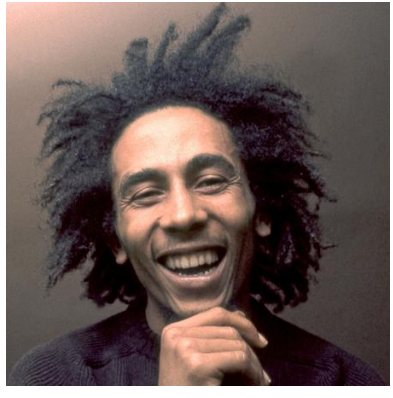
$x_1 \ldots x_n$



$x_1 \quad x_2 \wedge x_3 \quad x_7 \oplus x_9$

$x_1 \wedge y_1 \quad x_7 \quad x_n \oplus x_2$

$y_1 \ldots y_n$



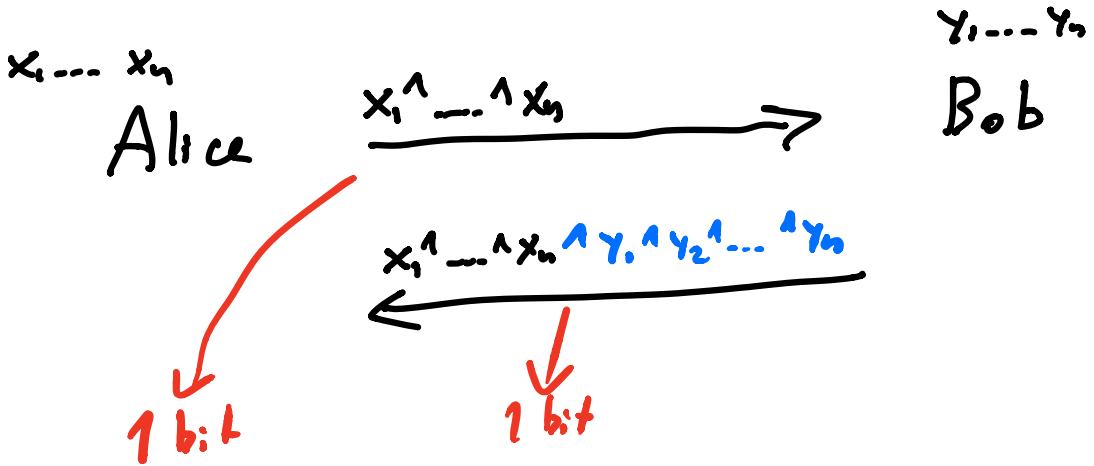
$F(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)$

Minimize total # bits they send to each other

Examples

AND

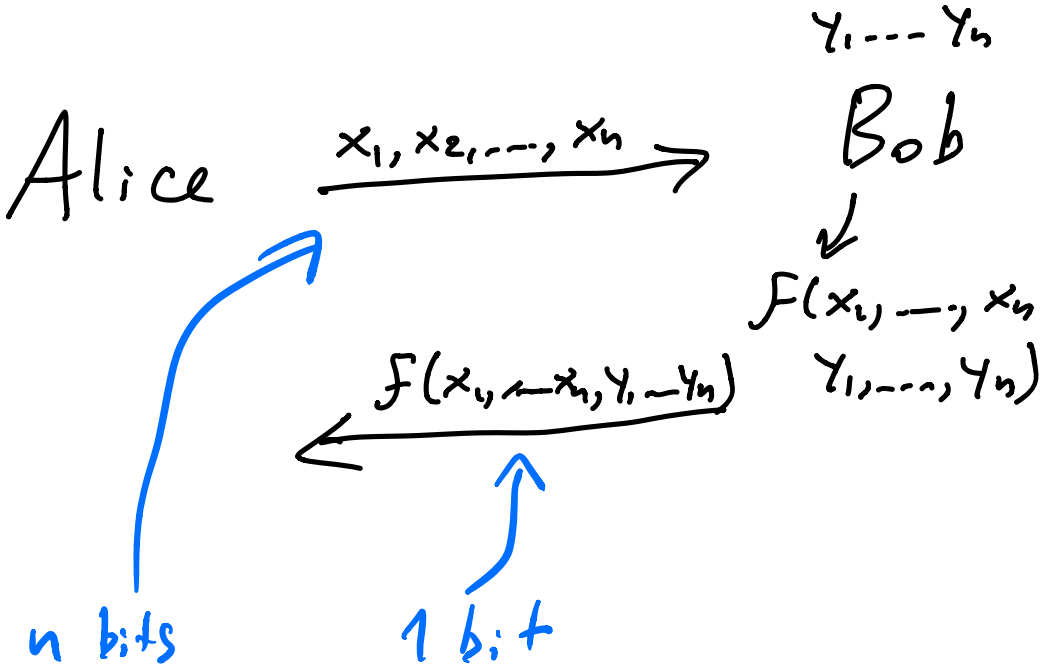
$$f(x_1, \ldots, x_n, y_1, \ldots, y_n) = x_1 \wedge x_2 \wedge \ldots \wedge x_n \wedge y_1 \wedge y_2 \wedge \ldots \wedge y_n$$



$CC(f) \leq 2$

Any function $F(x_1 \ldots x_n\, y_1 \ldots y_n)$

$CC(F) \leq n+1$

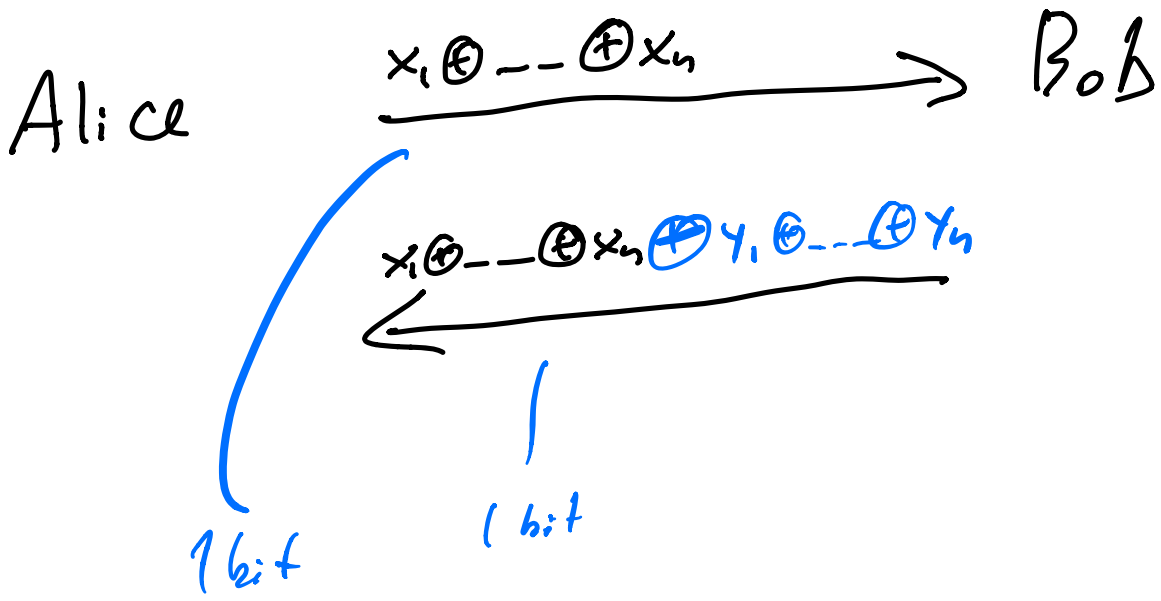


n+1 bits of communication

$CC(F) \leq n+1$

Parity (XOR)

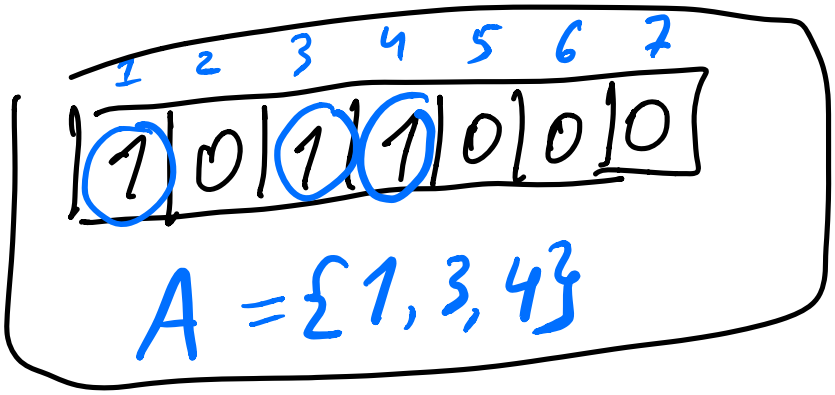
$$f(x_1, \ldots, x_n, y_1, \ldots, y_n) = x_1 \oplus x_2 \oplus \cdots \oplus y_n$$



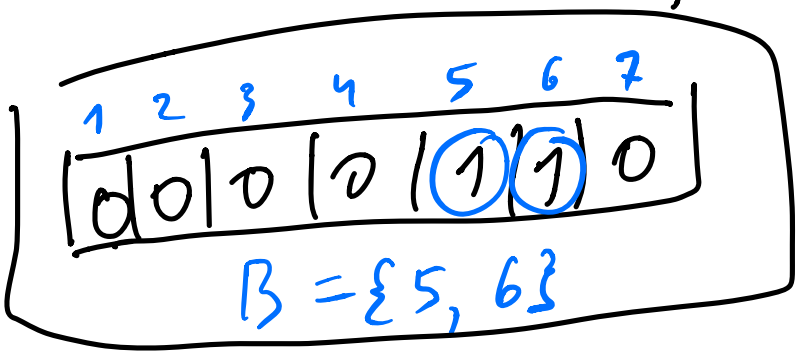
$CC(f) \leq 2$

Median

$x_1 \dots x_n \in \{0,1\}^n$ – indicator vector of subset $A \subseteq \{1,\dots,n\}$



$y_1 \dots y_n \in \{0,1\}^n$ – ind vector of $B \subseteq \{1,\dots,n\}$



$f(x_1,\dots,x_n,y_1,\dots,y_n)$ = Median of $A \cup B$

$A \cup B = \{1,3,4,5,6\}$

$A \cup B = \{1, 3, \boxed{4}, 5, 6\}$

For a set of size S,

Median is the value in this set

greater than $\frac{S-1}{2}$ els of set

smaller than $\frac{S-1}{2}$ els of se-

Comm complexity of Median?

$CC(P) \leq O(\log^2 n)$

- Alice sends to Bob size of her set $a = |A|$
- Bob sends to Alice $b = |B|$

$S = |A| + |B|$

Alice says I have k els $\leq \frac{n}{2}$

Bob says I have ℓ els $\leq \frac{n}{2}$

$k + \ell < \frac{S-1}{2} \Rightarrow$ look $> \frac{n}{2}$

$k + \ell \geq \frac{S-1}{2} \Rightarrow$ look $\leq \frac{n}{2}$

Comm complexity $\leq$

#rounds $\cdot$ #bits in each round

$\log_2 n$ $\quad$ $2\log_2 n$

$= O(\log^2 n)$

$CC(F) = O(\log n)$

Alice if median of A is $< \frac{n}{2}$
she sends 1 $\leftarrow$

else
she sends 0

Bob if median of B is $< \frac{n}{2}$
he sends 1 $\leftarrow$

else
he sends 0

If both medians are on the same side
$\Rightarrow$ reduce $n \rightarrow n/2$

If medians on different sides
$|A| \rightarrow |A|/2$ $\quad$ $|B| \rightarrow |B|/2$

$$CC(f) \leq \#\text{rounds} \cdot \#\text{bits in round}$$

$$O(\log_2 n) \qquad 2$$

$$\leq O(\log_2 n)$$

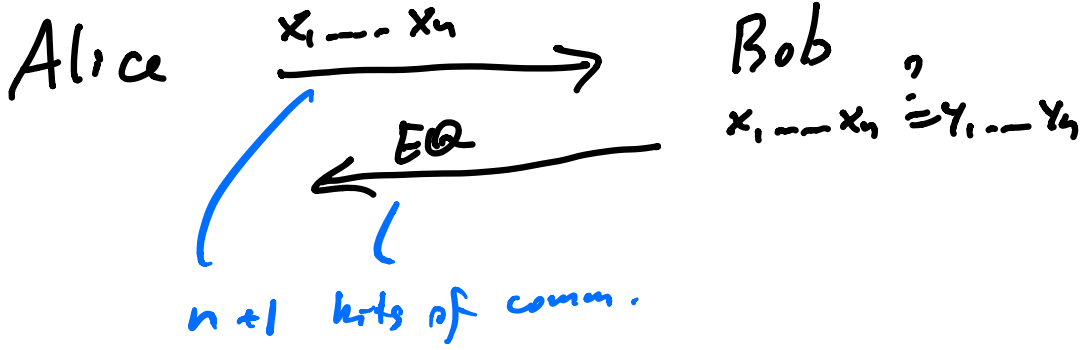
EQUALITY

EQ:

$f(x_1, \ldots, x_n, y_1, \ldots, y_n) = 1$ iff $x_1 \ldots x_n = y_1 \ldots y_n$

$\Leftrightarrow$

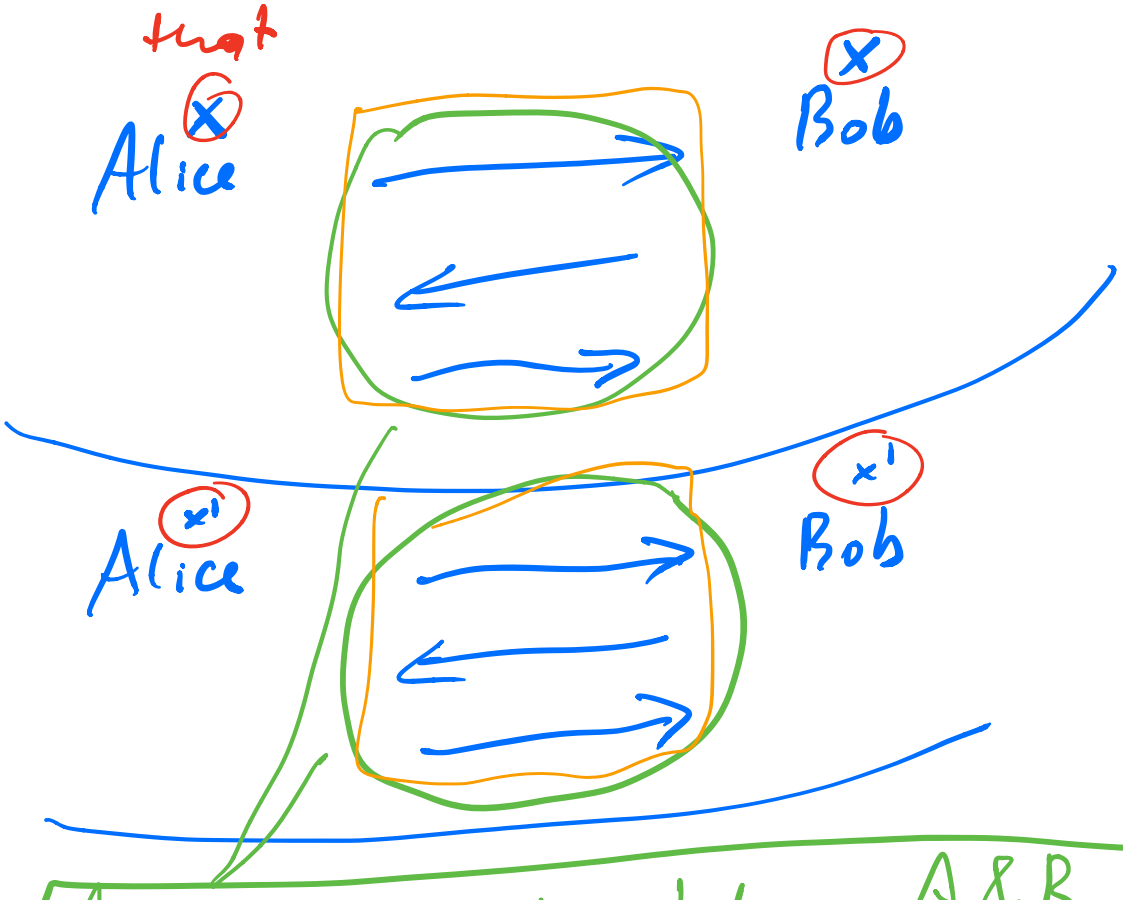
$x_1 = y_1$ AND
$x_2 = y_2$ AND
$\ldots$
$x_n = y_n$



$CC(f) \leq n+1$

Thm $CC(EQ) \geq n$

Proof: Assume $x \neq x' \in \{0,1\}^n$ such that



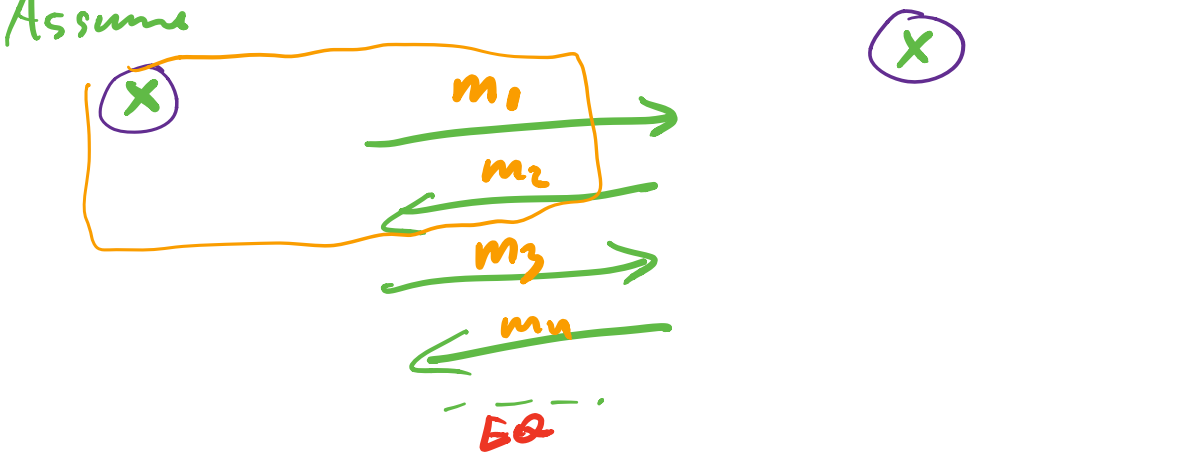
Assume communication between A & B same for these inputs
We'll arrive at contradiction ←

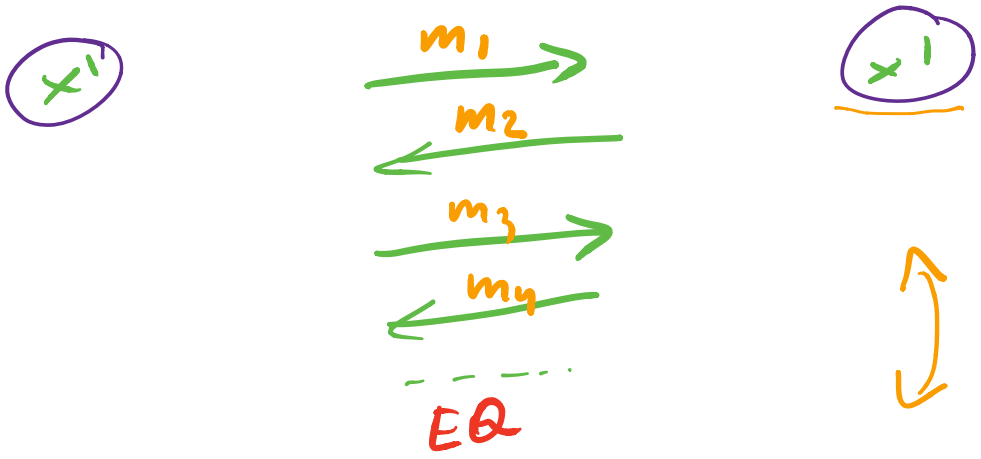
There are 2^n different inputs
Alice × Bob

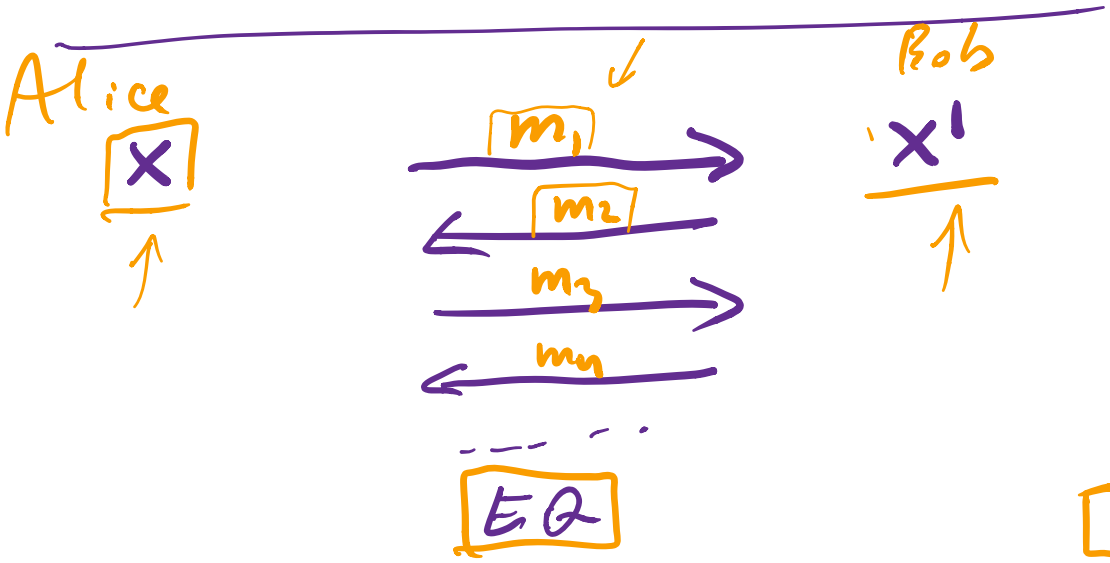
There must be 2^n different "communication transcripts" between $\Rightarrow |protocol| \geq n$

Proving queen statement.

Assume



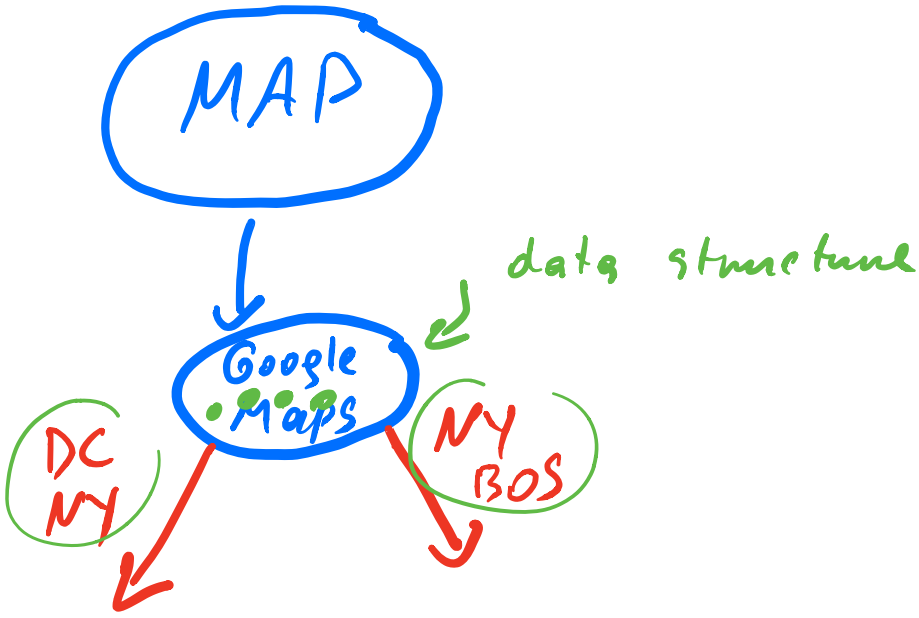




□

Comm complexity

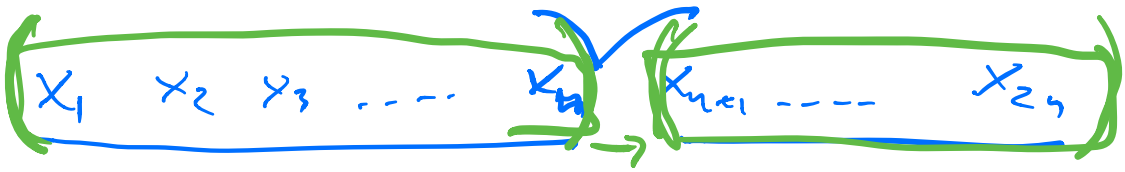
Data Structure



≡ minimum amount of data MAP needs to share with APP to find shortest path

≡ comm MAP ⟷ AP

Big Data Alg / Streaming algs



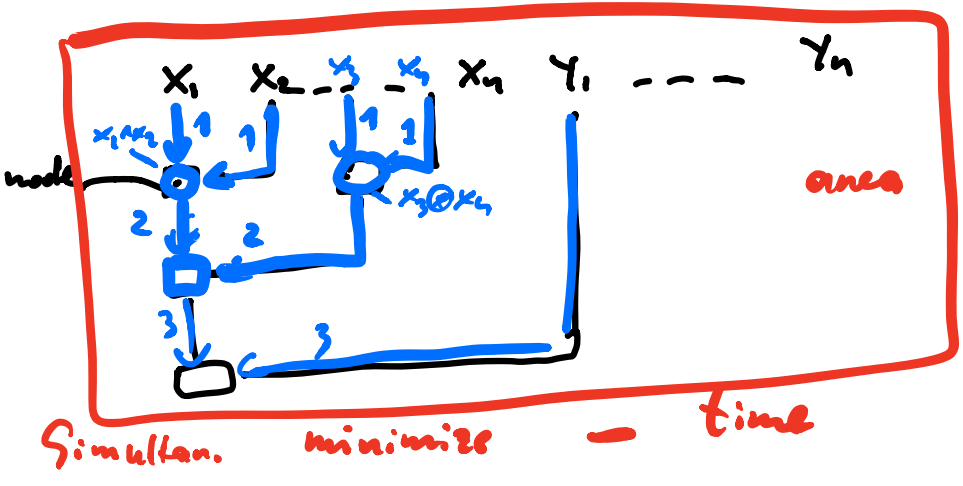
≡ comm complexity

VLSI Circuits

Very Large Scale Integration

x_1 x_2 ---- x_n y_1 ---- y_n

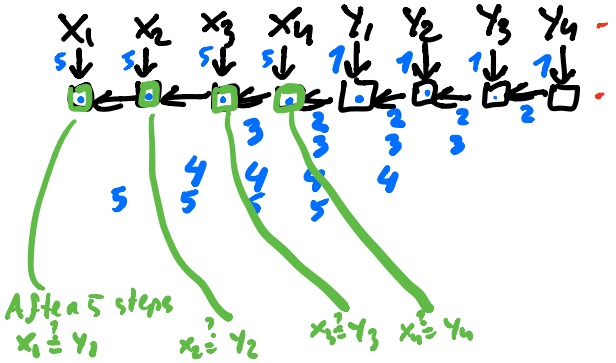
EQ - each time IF/LOOP



Simultan. minimize — time

— area

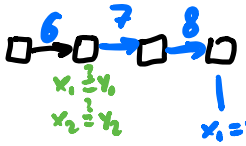
Small chip for EQ



↕ – min wire length

AREA $\approx 2n$

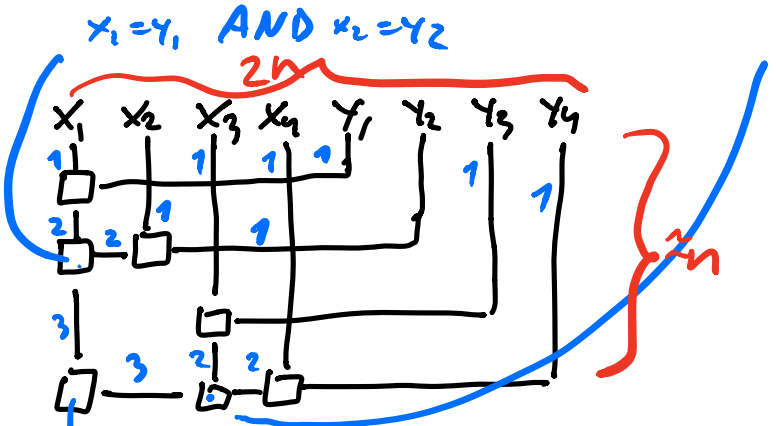
TIME $\approx 2n$



$x_1 = y_1$ AND $x_2 = y_2$ AND $x_3 = y_3$ AND $x_4 = y_4$

$EQ(x_1 x_2 x_3 x_4, y_1 y_2 y_3 y_4)$

Fast chip for EQ



$x_3=y_3$ AND $x_4=y_4$

2nd step: $x_5=y_5$ AND $x_6=y_6$

$\vdots$

$x_{n-1}=y_{n-1}$ AND $x_n=y_n$

$(x_1=y_1)$ AND $(x_2=y_2)$ AND $(x_3=y_3)$ AND $(x_4=y_4)$

4th step: $x_1 \dots x_8$ $y_1 \dots y_8$

5 x_1 x_{16} $y_1 \dots y_{16}$

$\log_2 n$ steps $x_1 \dots x_n \stackrel{?}{=} y_1 \dots y_n$

Time $\approx \log_2 n$

AREA $\approx 2n^2$

For EQ

Time $= 2n$

Area $= 2n$

Time $= \log_2 n$

Area $= 2n^2$

Can we have Time $\approx \log n$ and Area $\approx n$?

NO! Prove using comm complexity

Thm: If $\exists$ chip time T area A

$\Rightarrow CC(f) \leq (\sqrt{A}+1) \cdot T$

Cor $CC(EQ) \geq n$

$n \leq CC(EQ) \leq (\sqrt{A}+1) \cdot T$

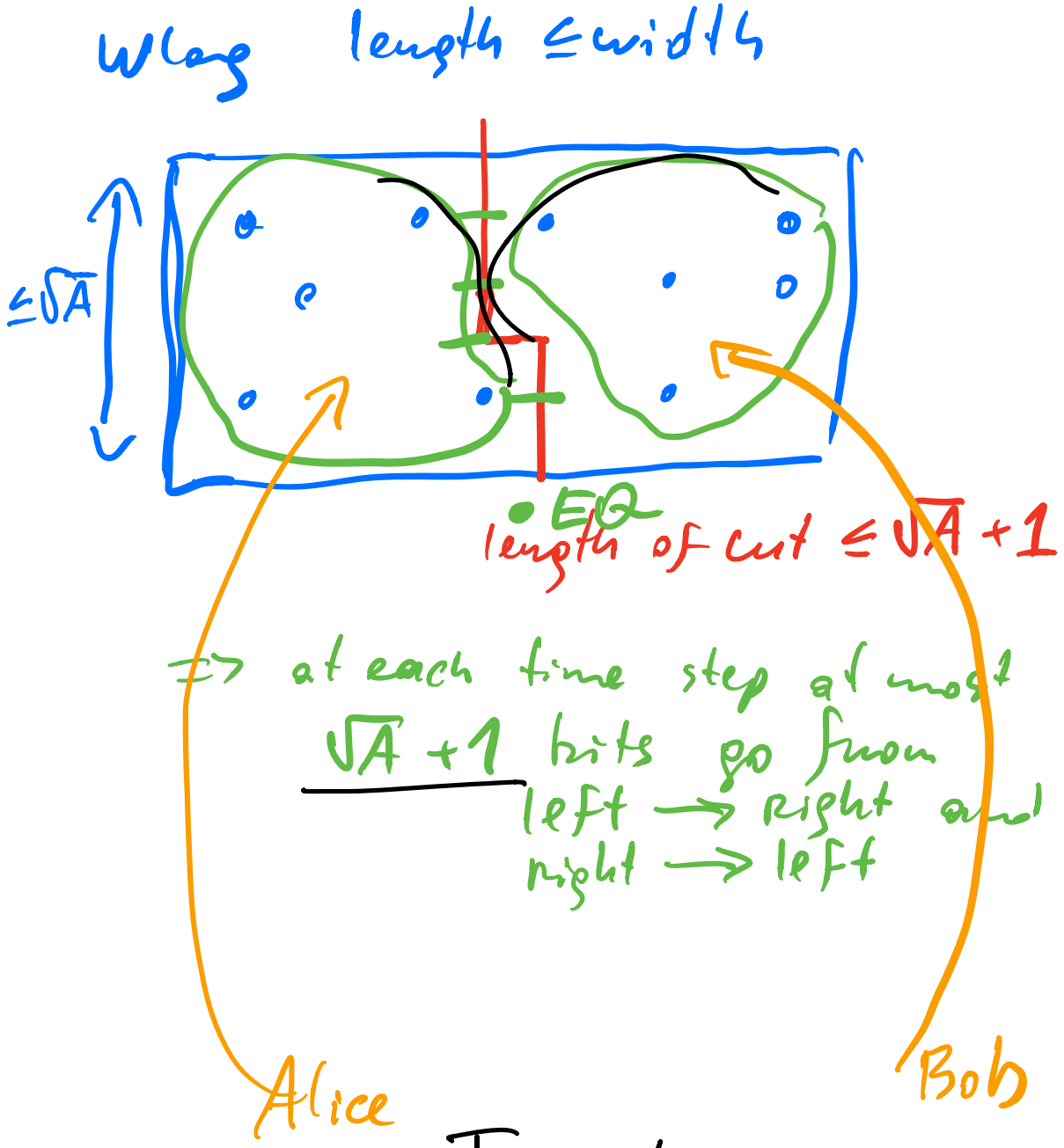
For any chip computing EQ:

$(\sqrt{A}+1) \cdot T \geq n$

In particular, fast chip with $T \approx \log n \Rightarrow$

$\sqrt{A} \geq \Omega\left(\frac{n}{\log n}\right)$

$\Rightarrow A \geq \Omega\left(\frac{n^2}{\log n}\right)$



T rounds
at each round send $\sqrt{A}+1$

In the end we compute EQ

$$CC(EQ) \leq (\sqrt{A}+1) \cdot T \qquad \square$$

Randomized Communication

What if Alice and Bob can use randomness?

Can we solve EQ much more efficiently?

YES! Lecture 3 Cloud Sync

Cloud Sync

- Synchronize local files to the cloud

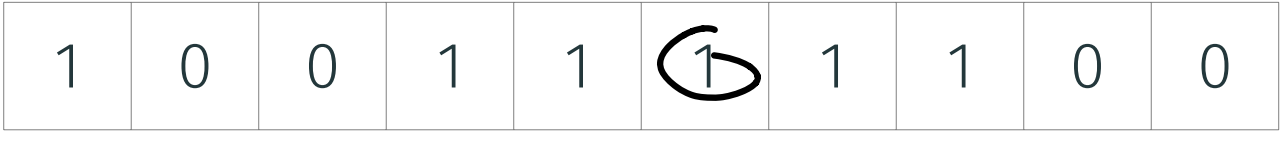
Cloud Sync

- Synchronize local files to the cloud

- Has file been changed? File length: n bits

Randomized Protocol

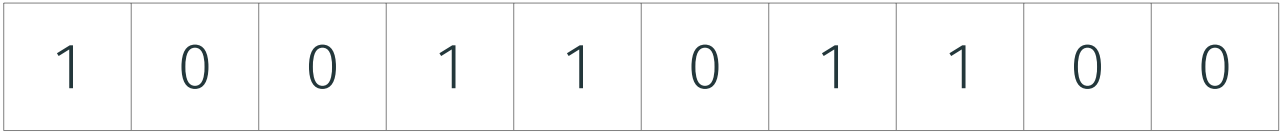
local file



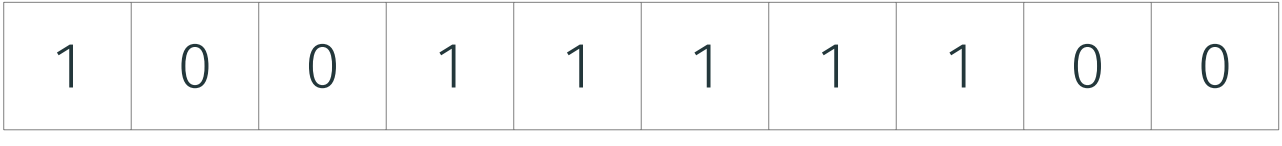
cloud file

Randomized Protocol

local file



$a \in \{0, \ldots, 2^n - 1\}$



cloud file

Randomized Protocol

local file Alice

1	0	0	1	1	0	1	1	0	0

$a \in \{0, \ldots, 2^n - 1\}$

$b \in \{0, \ldots, 2^n - 1\}$

1	0	0	1	1	1	1	1	0	0

cloud file Bob

Randomized Protocol

local file

1	0	0	1	1	0	1	1	0	0

$a \in \{0, \ldots, 2^n - 1\}$

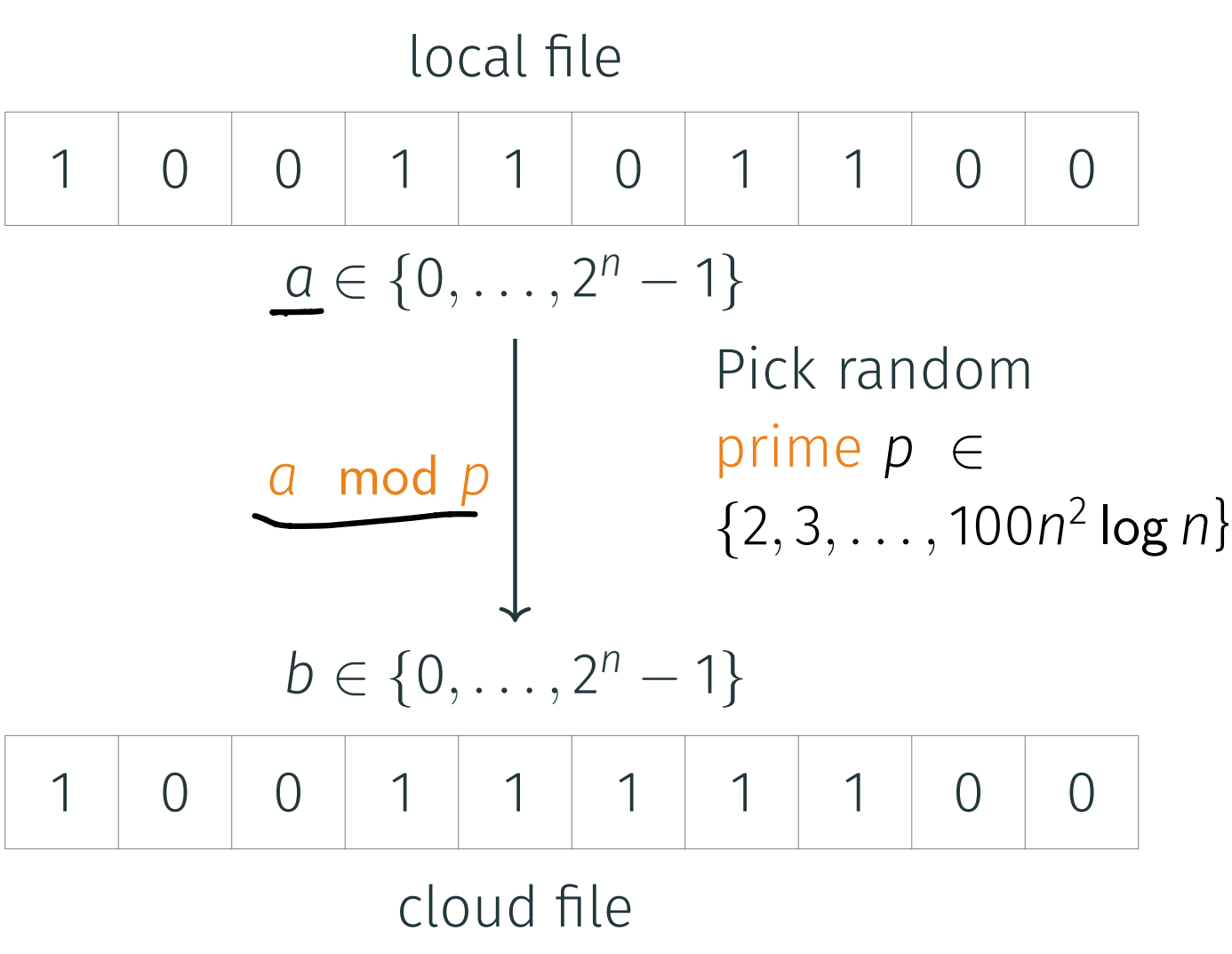
Pick random prime $p \in \{2, 3, \ldots, 100n^2 \log n\}$

$b \in \{0, \ldots, 2^n - 1\}$

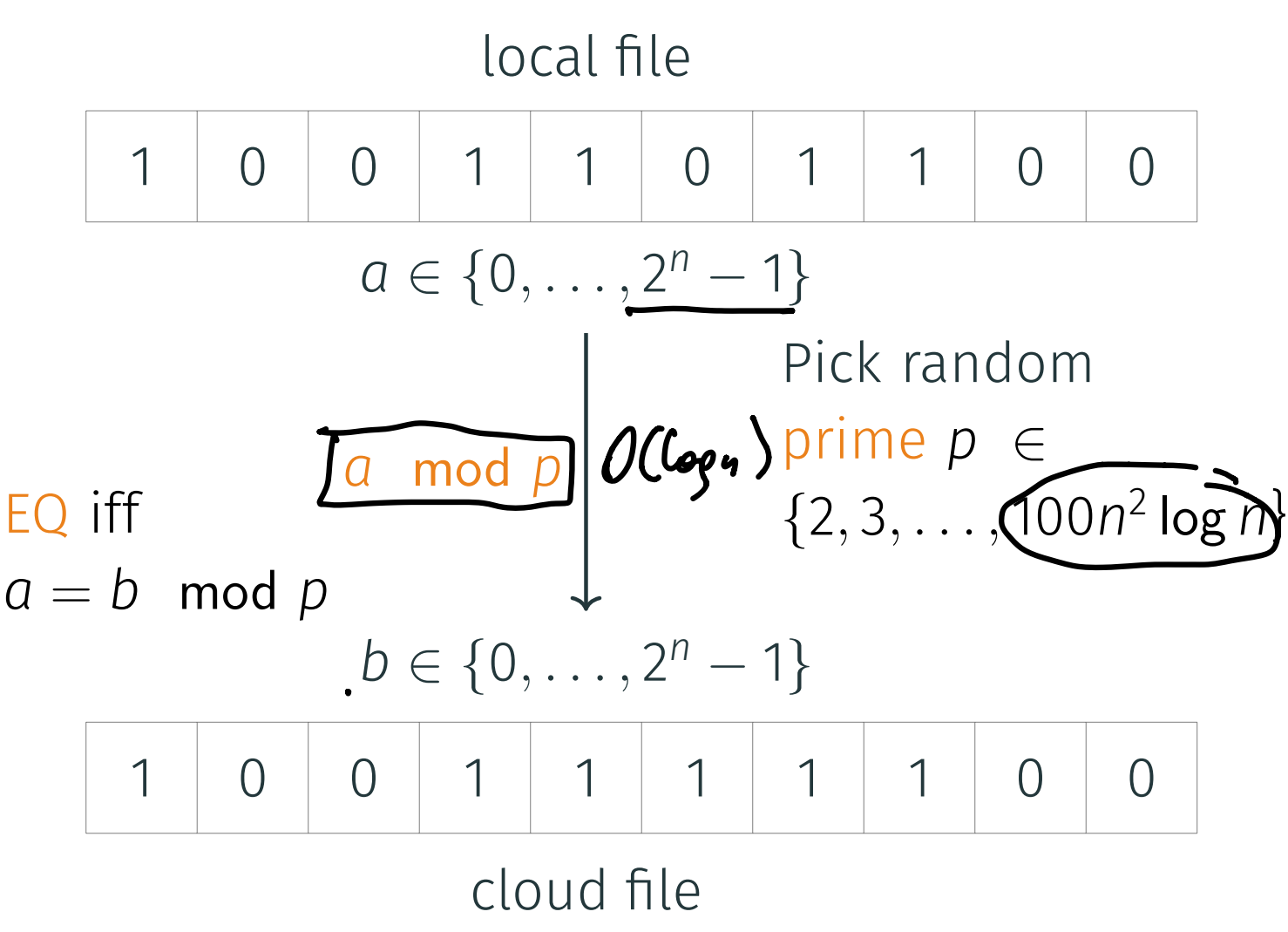
1	0	0	1	1	1	1	1	0	0

cloud file

Randomized Protocol



Randomized Protocol



$CC(\widetilde{EQ}) \geq n$

Randomized $CC(EQ) \leq O(\log n)$

Disjointness