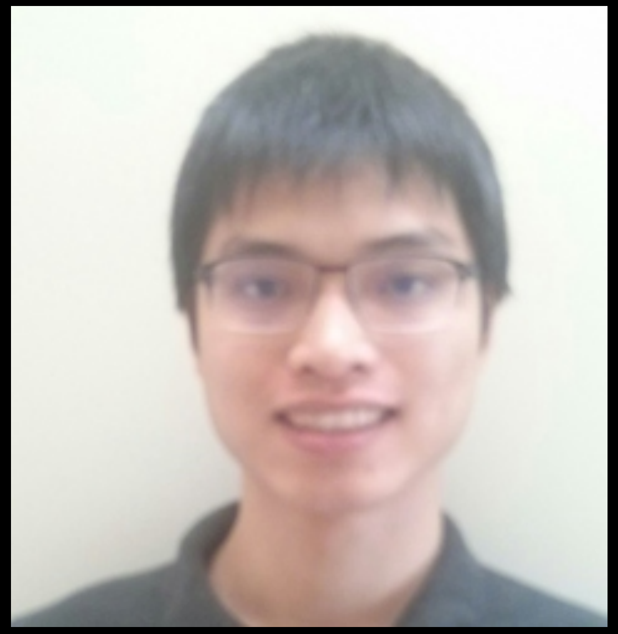
Algorithms and lower bounds for de-Morgan formulas of low-communication leaf gates

Sajin Koroth (Simon Fraser University)

Joint with

Valentine Kabanets
Zhenjian Lu
Dimitrios Myrisiotis
Igor Carboni Oliveira

Outline

- Background
- Circuit model : $Formula[s] \circ \mathscr{G}$
- Prior work
- Results
  - Lower bounds
  - PRG's
  - SAT algorithm's
  - Learning algorithms
- Overview of the lower bound technique

Parallel vs Sequential computation

- Most of linear algebra can be done in parallel
- Gaussian elimination is an outlier
  - Intuitively its an inherently sequential procedure
  - There are theoretical reasons to believe so
  - There is an efficient sequential algorithm

P vs NC^1

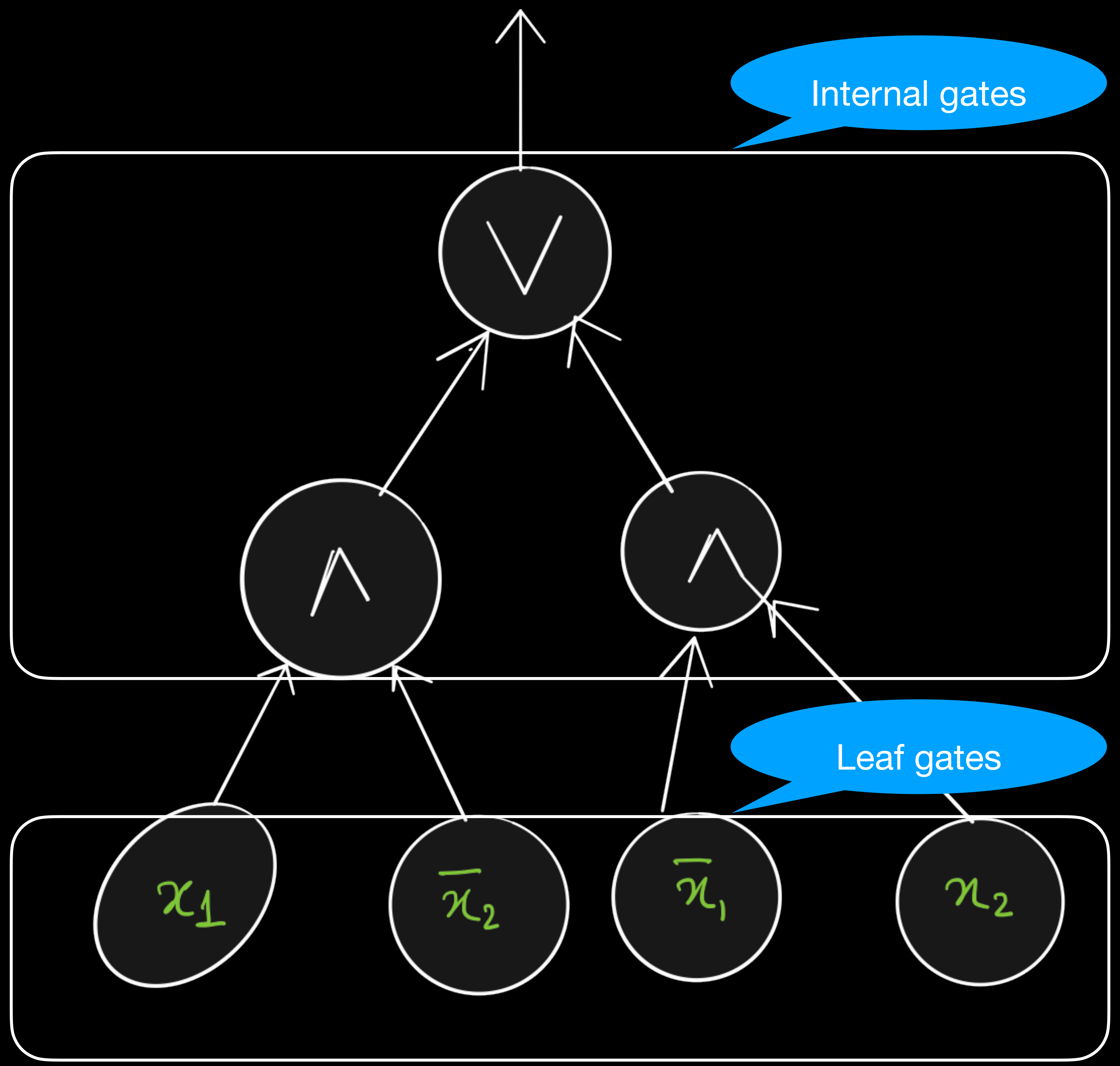
Class P of poly-time solvable problems

Are there problems with efficient sequential algorithms
which do not have efficient parallel algorithms ?

Modeled as circuits

Circuit complexity

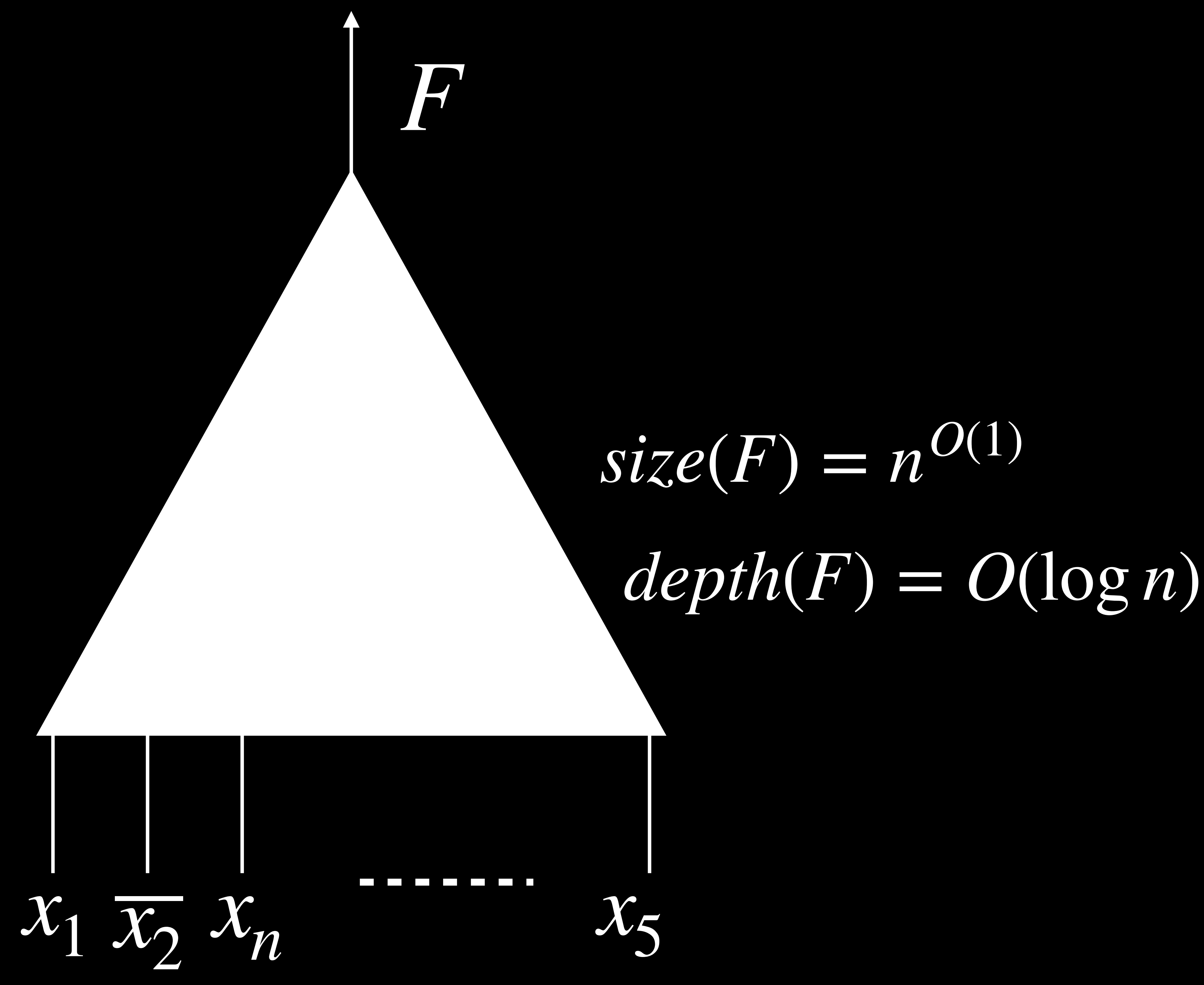
- Complexity parameters :
  - Size : # of gates
  - Depth : length of the longest path from root to leaf
  - Fan in : 2, Fan out
- Formulas :
  - Underlying DAG is a tree
  - No reuse of computation
  - Depth = log (Size)



Circuit complexity

Class NC^1 = Poly-Size Formulas

- Efficient parallel computation (formally CREW PRAM):
  - Polynomially many processors
  - Logarithmic computation time



In formula, $depth(F) = O(\log size(F))$

Circuit complexity

P vs NC^1 rephrased

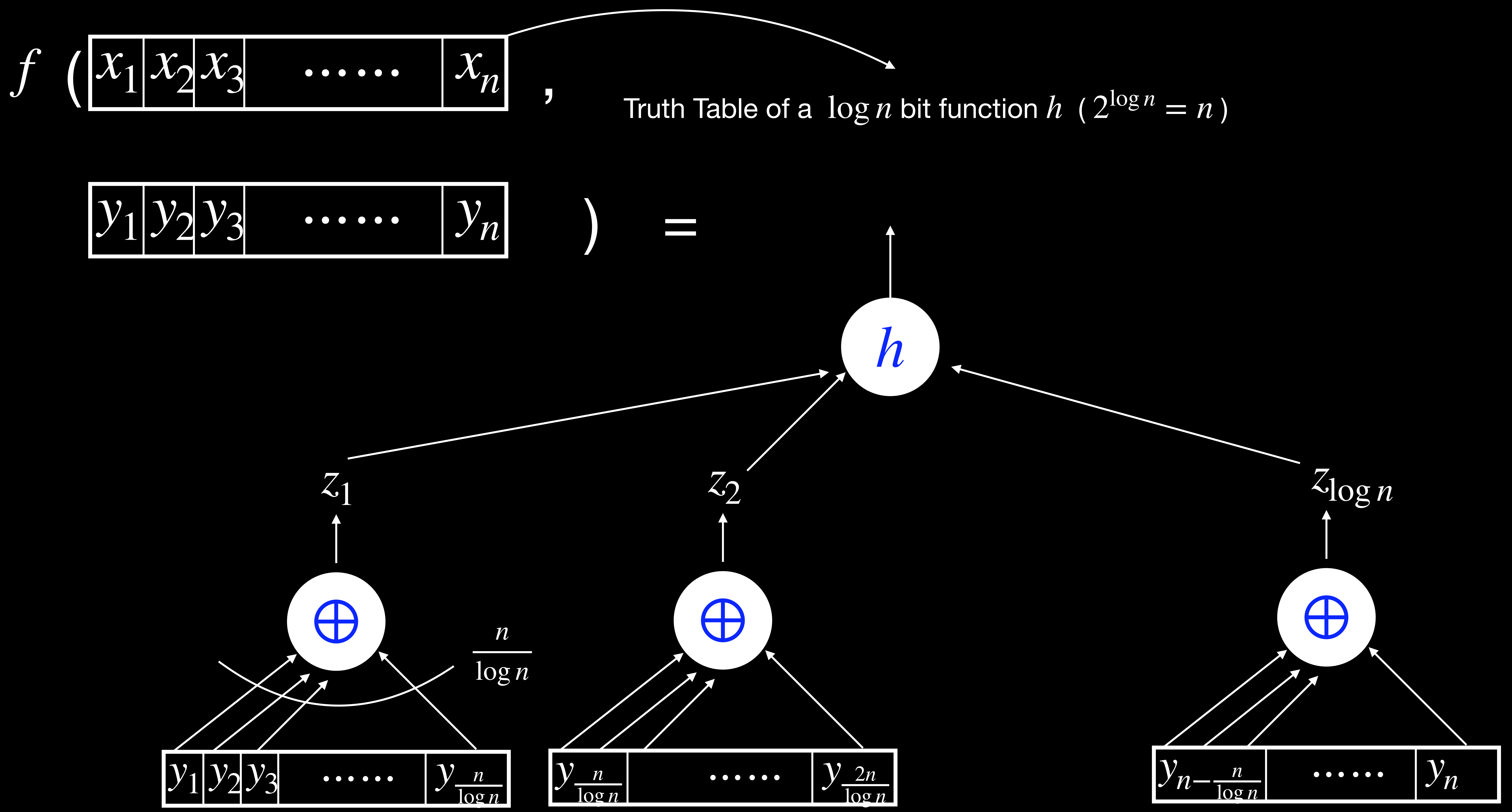
- A Boolean function f (candidates: Perfect matching, Gaussian elimination etc)
  - That can be computed in poly-time ($f \in P$)
  - Any de-Morgan formula computing it has super-poly size ($f \notin NC^1$)

P vs NC^1
State of the art

- Andreev'87 : $\Omega(n^{2.5-o(1)})$ for a function in P called the Andreev function
- Also, Andreev'87 : $\Omega(n^{1+\Gamma-o(1)})$, where Γ is the shrinkage exponent
- Paterson and Zwick'93 : $\Gamma \geq 1.63$
- Hastad'98 (breakthrough) : $\Gamma \geq 2 - o(1)$
- Tal'14 : $\Gamma = 2$
- Best l.b. for Andreev's function (Tal'14) : $\Omega\left(\frac{n^3}{\log^2 n \log\log n}\right)$
- Best l.b. for a function in P (Tal'16) : $\Omega\left(\frac{n^3}{\log n (\log\log n)^2}\right)$

Cubic formula lower bounds

Andreev's function



Cubic formula lower bounds

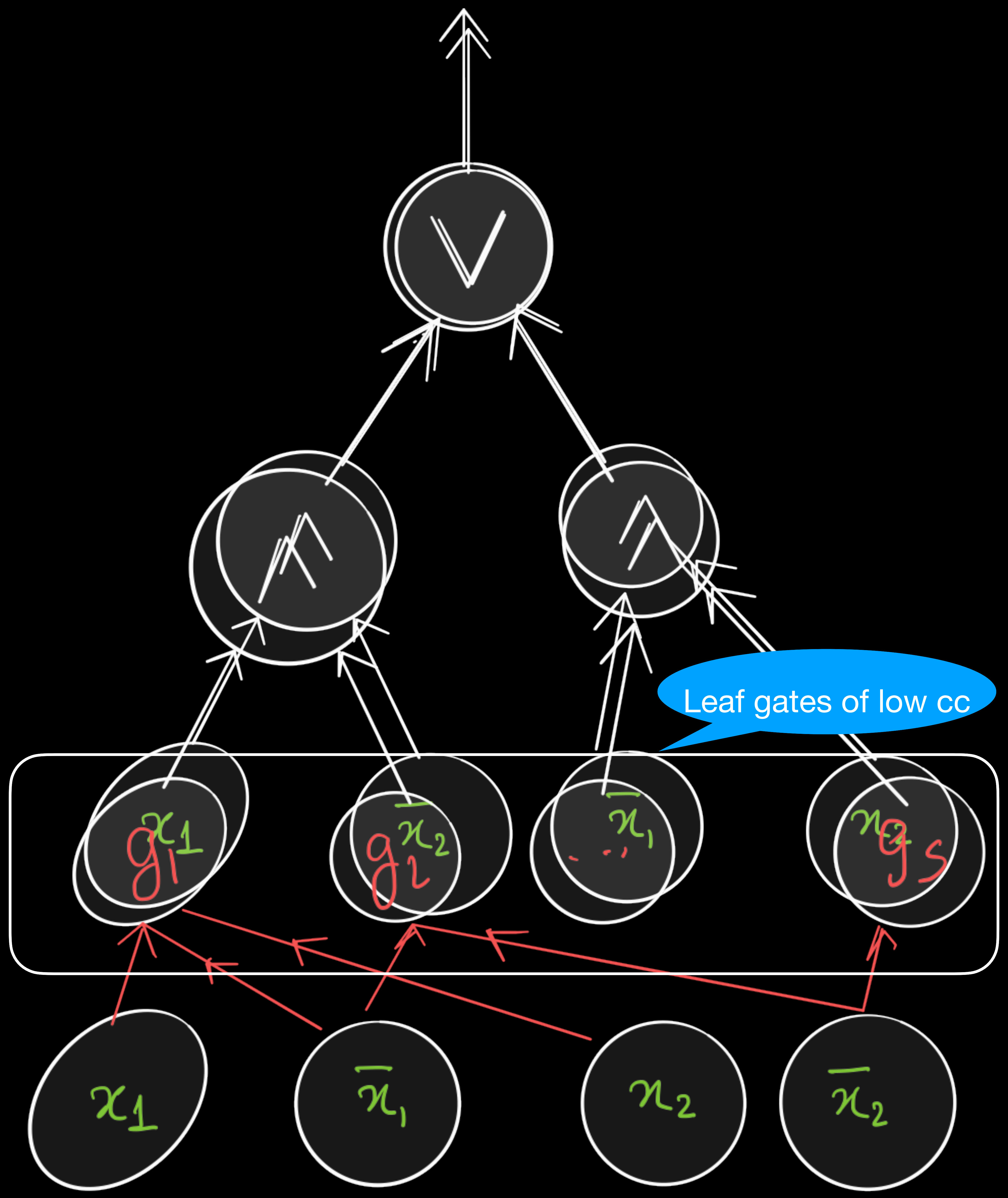
Hastad's result

- (Tal'14) : $\Omega\left(\frac{n^3}{\log^2 n \log\log n}\right)$
- Doesn't work if there are parity gates at bottom

Our Model

Augmenting de-Morgan formulas

- de-Morgan formulas : leaf gates, input literals
- Our model : leaf gates, low communication functions



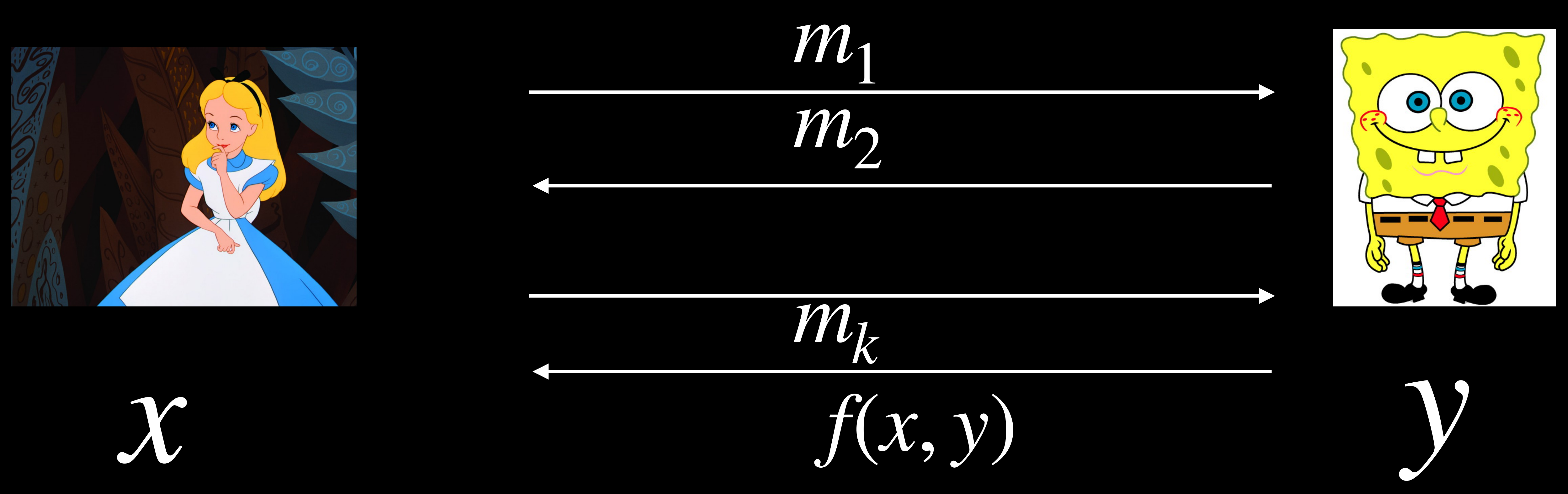
Our model

Reformulation

- $Formula[s] \circ \mathcal{G}$
  - Size s de-Morgan formula
  - $\mathcal{G}$: A family of Boolean functions
  - Leaf gates are functions $g \in \mathcal{G}$
- Our model :
  - $\mathcal{G}$ - low communication complexity Boolean functions
  - $s = \tilde{O}(n^2)$

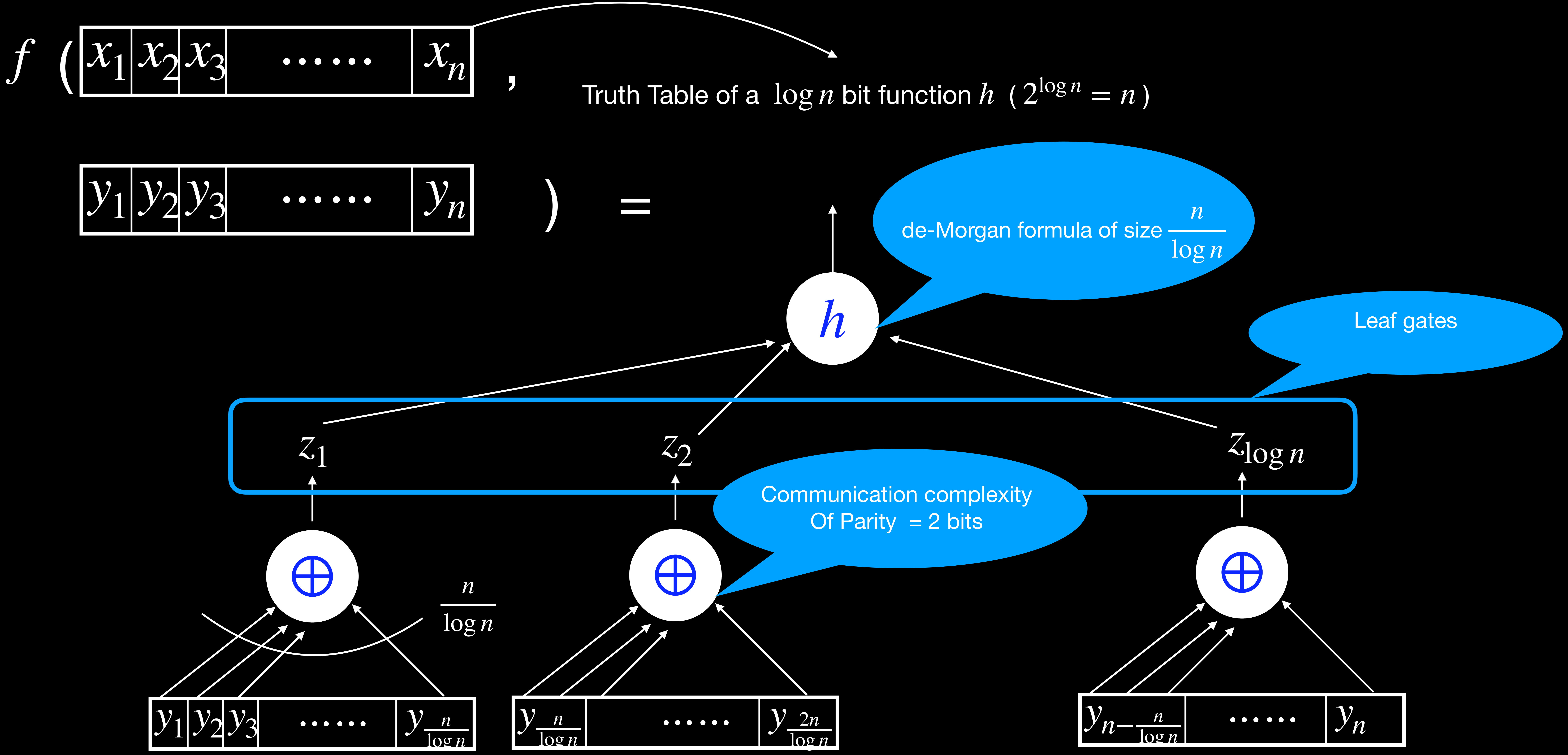
Communication complexity

- Yao's 2-party model
  - Input divided into 2 parts x, y
  - Goal : compute $f(x, y)$ with minimal communication



Our model

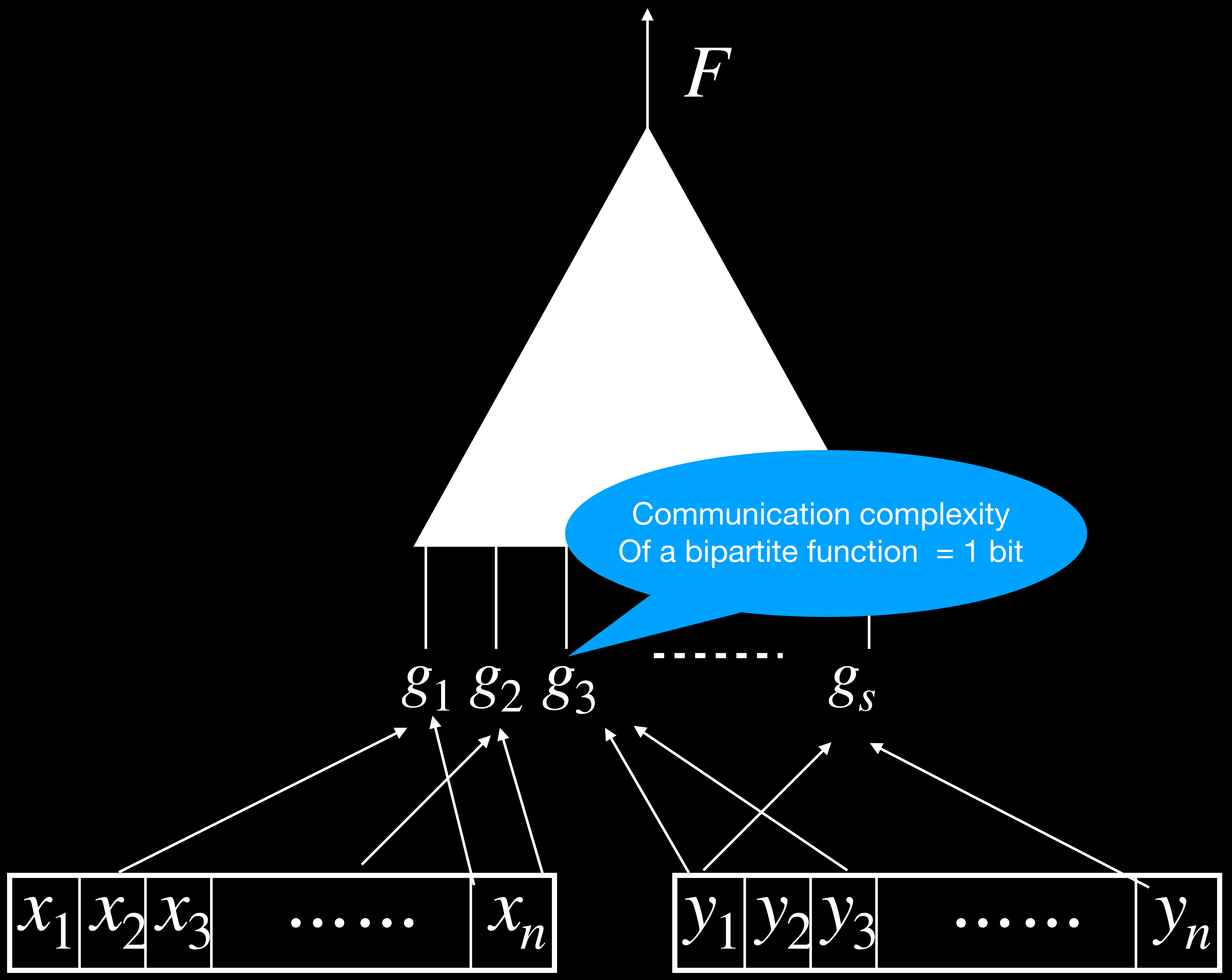
Complexity of Andreev's function



Our model

Prior work - Bipartite Formulas

- Input is divided into two parts, x, y
- Every leaf can gate can access any Boolean function of either x or y but not both
- Models a well known measure - graph complexity
- Tal'16: Bipartite formula complexity of IP_n is $\tilde{\Omega}(n^2)$
- Earlier methods could not do super linear



Our model

Connection to Hardness Magnification

- $MCSP_N[k]$: Given the truth table of a function f on n bits ($N = 2^n$)
  - Yes : if f has a circuit of size at most k
  - No : otherwise
  - Meta computational problem with connections to Crypto, learning theory, circuit complexity etc
- OPS'19:
  - If there exists an ϵ such that $MCSP_N[2^{o(n)}]$ is not in $Formula[N^{1+\epsilon}] \circ XOR$
  - then, $NP \notin NC^1$

Our model

Connection to PRG for polytopes

- Polytope : AND of LTF's
- LTF : $sign(w_1x_1 + \ldots + w_nx_n - \theta)$
  - $w_1, \ldots, w_n, \theta \in \mathbb{R}$
  - Ex : $3x_1 + 4x_2 + 5x_7 \geq 12$
  - Nisan'94 : Randomized communication complexity $O(\log n)$
- PRG's for polytopes : Approximate volume computation

Our model

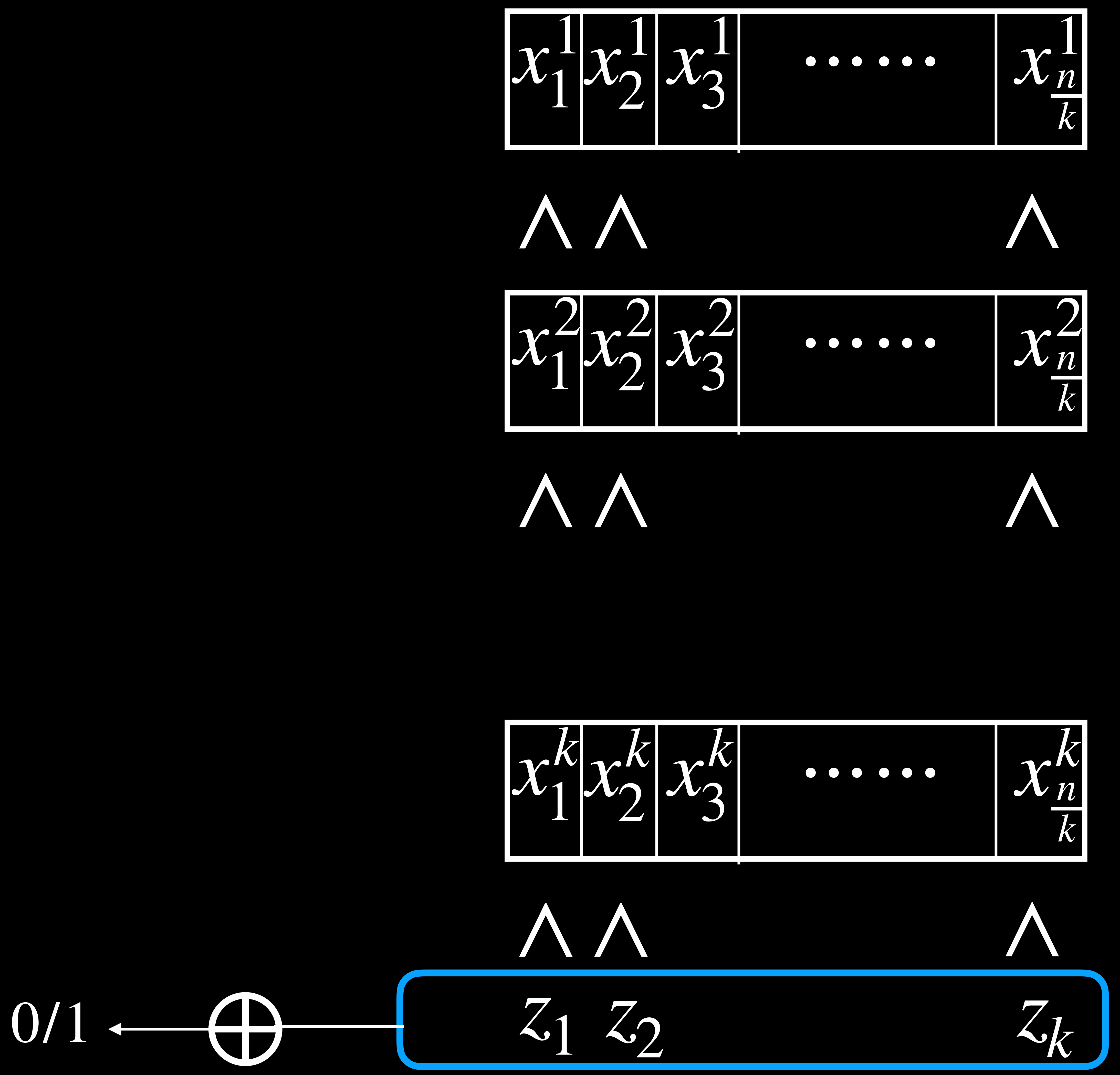
Interesting low communication bottom gates

- Bipartite functions
- Parities
- LTF’s (Linear threshold functions)
- PTF’s (Polynomial threshold functions)

Our results

Target function - Generalized inner product

- Generalization of binary inner product
- $IP_n(x, y) = \sum_{i \in [n]} x_i y_i$
- $GIP_n^k(x^1, x^2, \ldots, x^k) = \sum_{i \in [n/k]} \prod_{j \in [k]} x_i^j$



Our results

Lower bound

- Let GIP_n^k be computed on average by $F \in Formula[s] \circ \mathcal{G}$,
  - That is, $\Pr_x[F(x) = GIP_n^k(x)] \geq 1/2 + \epsilon$
  - Then, $s = \Omega\left(\dfrac{n^2}{k^2 \cdot 16^k \cdot R^k_{\epsilon/2n^2}(\mathcal{G}) \cdot \log^2(1/\epsilon)}\right)$
  - $R^k_{\epsilon/2n^2}(\mathcal{G})$: Randomized communication of $\mathcal{G}$ with error $\epsilon/2n^2$ in the number on forehead communication complexity model

Our results

MCSP lower bounds

- If $MCSP_N[2^{cn}]$ is computed $Formula[s] \circ XOR$, then $s = \tilde{O}(n^2)$
- Contrast : OPS'19:
  - If there exists an ϵ such that $MCSP_N[2^{o(n)}]$ is not in $Formula[N^{1+\epsilon}] \circ XOR$
  - then, $NP \notin NC^1$
- Our techniques cannot handle $MCSP_N[2^{o(n)}]$

Our results

PRG

- A pseudo random generator G is said to ϵ fool a function class $\mathcal{F}$ if
  - $\left| \Pr_{z \in \{0,1\}^{l(n)}} \left[f(G(z)) = 1 \right] - \Pr_{x \in \{0,1\}^n} \left[f(x) = 1 \right] \right| \le \epsilon$
  - f is any function from $\mathcal{F}$
  - $G : \{0,1\}^{l(n)} \to \{0,1\}^n$
  - z is the seed, $l(n) \lll n$
  - Smaller the seed length compared to n the better

Our results

PRG

- Parities at the bottom can make things harder.
    - AC^0 best known PRG seed length $poly(\log n)$
    - $AC^0 \circ XOR$ best known only $(1 - o(1))n$

Our results
PRG

- There is a PRG that ϵ-fools $Formula[s] \circ XOR$
  - Seed length : $O(\sqrt{s} \cdot \log s \cdot \log(1/\epsilon) + \log n)$
  - Seed length is optimal, unless lower bound can be improved

Our results
PRG

- Natural generalization to $Formula[s] \circ \mathcal{G}$
- There is a PRG that ϵ-fools $Formula[s] \circ \mathcal{G}$
  - Seed length : $n/k + O(\sqrt{s} \cdot (R^{k-NIH}_{\epsilon/6s}(\mathcal{G}) + \log s) \cdot \log(1/\epsilon) + \log k) \cdot \log k$
  - Number in hand

Our results

PRG - Corollaries

- (Ours + Vio15) : There is a PRG
  - Seed length : $O(n^{1/2} \cdot m^{1/4} \cdot \log n \cdot \log(n/\epsilon))$
  - ϵ-fools intersection of m halfspaces over $\{0,1\}^n$
  - Our results beats earlier results when $m = O(n)$ and $\epsilon \leq 1/n$

Our results

PRG - Corollaries

- There is a PRG
  - Seed length : $O(n^{1/2} \cdot s^{1/4} \cdot \log n \cdot \log(n/\epsilon))$
  - ϵ-fools $Formula[s] \circ SYM$
  - First of its kind
  - Blackbox counting algorithm (Whitebox due to CW19)

Our results

SAT Algorithm

- Given circuit class $\mathscr{C}$
  - Circuit SAT : Given $C \in \mathscr{C}$, is there an x, $C(x) = 1$
  - #Circuit SAT : Given $C \in \mathscr{C}$, how many x, $C(x) = 1$

Our results

SAT Algorithm

- Randomized #SAT algorithm for $Formula[s] \circ \mathcal{G}$
  - Running time 2^{n-t}
  - $$t = \Omega\left(\left(\frac{n}{\sqrt{s}\cdot \log^2 s \cdot R^2_{1/3}(\mathcal{G})}\right)^{1/2}\right)$$

    $\log n$ for LTFs
  - First of its kind #SAT for unbounded depth Boolean circuits with PTF's at the bottom

Our results
Learning algorithm

- There is PAC-learning algorithm
  - Learns $Formula[n^{2-\gamma}] \circ XOR$
  - Accuracy : ϵ, Confidence : δ
  - Time complexity : $poly(2^{n/\log n}, 1/\epsilon, \log(1/\delta))$
- $Formula[n^{2-\gamma}]$ can be learned in $2^{o(n)}$ [Rei11]
- Crypto connection:
  - $MOD_3 \circ XOR$ is assumed to compute PRFs (BIP+18)
  - If true, $Formula[n^{2.8}] \circ XOR$ can't be learned in $2^{o(n)}$ time

Lower bound technique

Outline

- GIP_n^k cannot even be weakly approximated by low communication complexity functions
- Weakness of $Formula[s] \circ \mathscr{G}$: Size s formula can be "approximated" by degree $\sqrt{s}$ polynomial
- GIP_n^k is weakly approximated by a collection of leaf gates

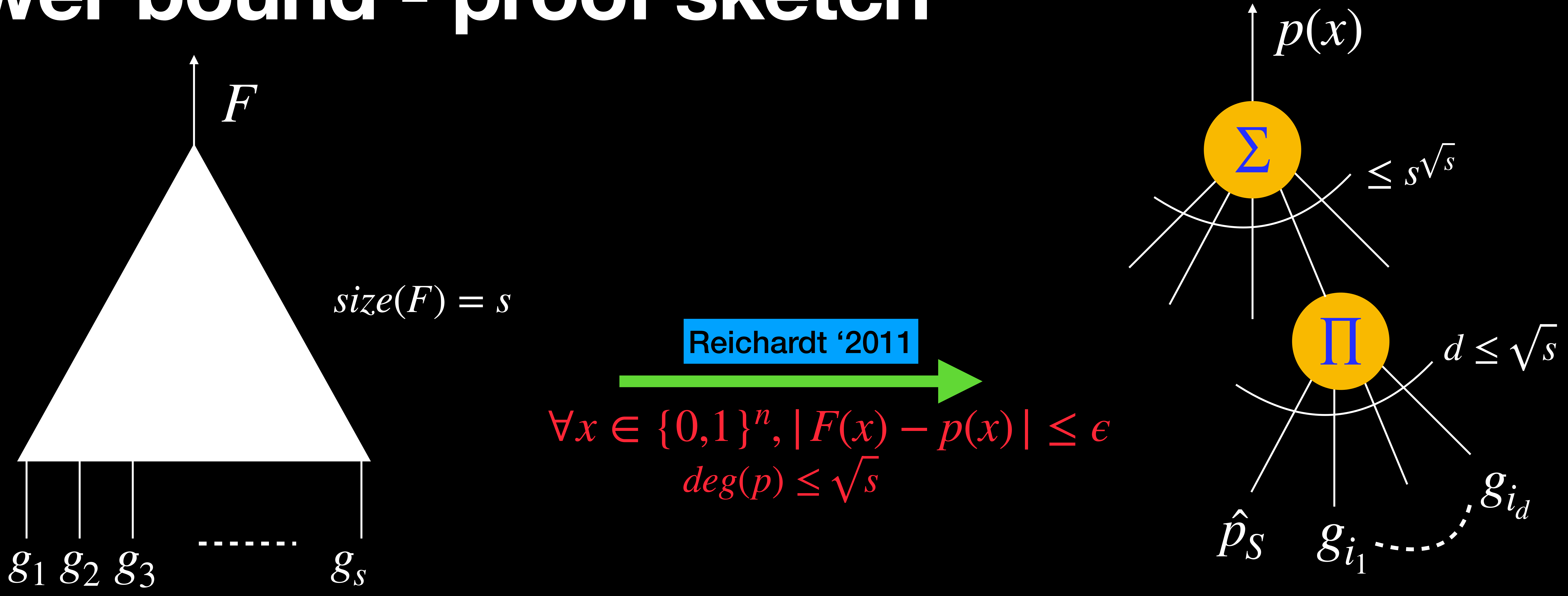
Lower bound technique
Part I

- GIP_n^k cannot even be weakly approximated by low communication complexity functions
- In the number on forehead model
    - Protocol computes GIP_n^k with error ϵ (uniform distribution)
    - Then commn.comp $> n/4^k - \log(1/(1-2\epsilon))$

Lower bound technique
Part II

- Weakness of $Formula[s] \circ \mathcal{G}$: Size s formula can be "approximated" by degree $\sqrt{s}$ polynomial
- Reichardt'11 : Approximation of Boolean formulas by Polynomials
    - $F(y_1, \ldots, y_m)$ be a formula of size s
    - There is a real polynomial $p(y_1, \ldots, y_m)$ of degree $O(\sqrt{s})$
    - For every $y \in \{0,1\}^m, |F(a) - p(a)| \leq 1/10$
- Fact : For any $0 < \epsilon < 1, \widetilde{deg}_\epsilon(f) \leq \widetilde{deg}(f) \cdot \log(1/\epsilon)$
- Corollary : For any formula F of size $s, \widetilde{deg}_\epsilon(F) \leq \sqrt{s} \cdot \log(1/\epsilon)$

Lower bound - proof sketch



- F correlates well (ϵ) with p
- F correlates well $(\frac{1}{s\sqrt{s}})$ with a monomial $(\hat{p}_S \prod_{j \in [S], |S| \leq \sqrt{s}} g_{i_j})$
- Since each g_i has low communication complexity, so does $\prod_{j \in [S], |S| \leq \sqrt{s}} g_{i_j}$
- F correlated well with the target function f, thus it correlates well with the monomial (a low communication function) !!!!!!!

Limitations of our approach

- To get better lower bounds, find a smaller degree approximating polynomial
- Approximate degree bound of Reichardt ($\sqrt{s}$) cannot be improved
  - AND_n function can be computed by a size n de-Morgan formula
  - Approximate degree of AND_n is $\theta(\sqrt{n})$

Future directions

- Extend lower bounds to $Formula[s] \circ \mathcal{G}$ when $s = \omega(n^2)$
- Design a PRG of seed length $n^{o(1)}$ and error $\epsilon \leq 1/n$ for intersection of n half spaces
- Learn $Formula[s] \circ XOR$ in time $2^{\tilde{O}(\sqrt{s})}$

Thank you

Questions?