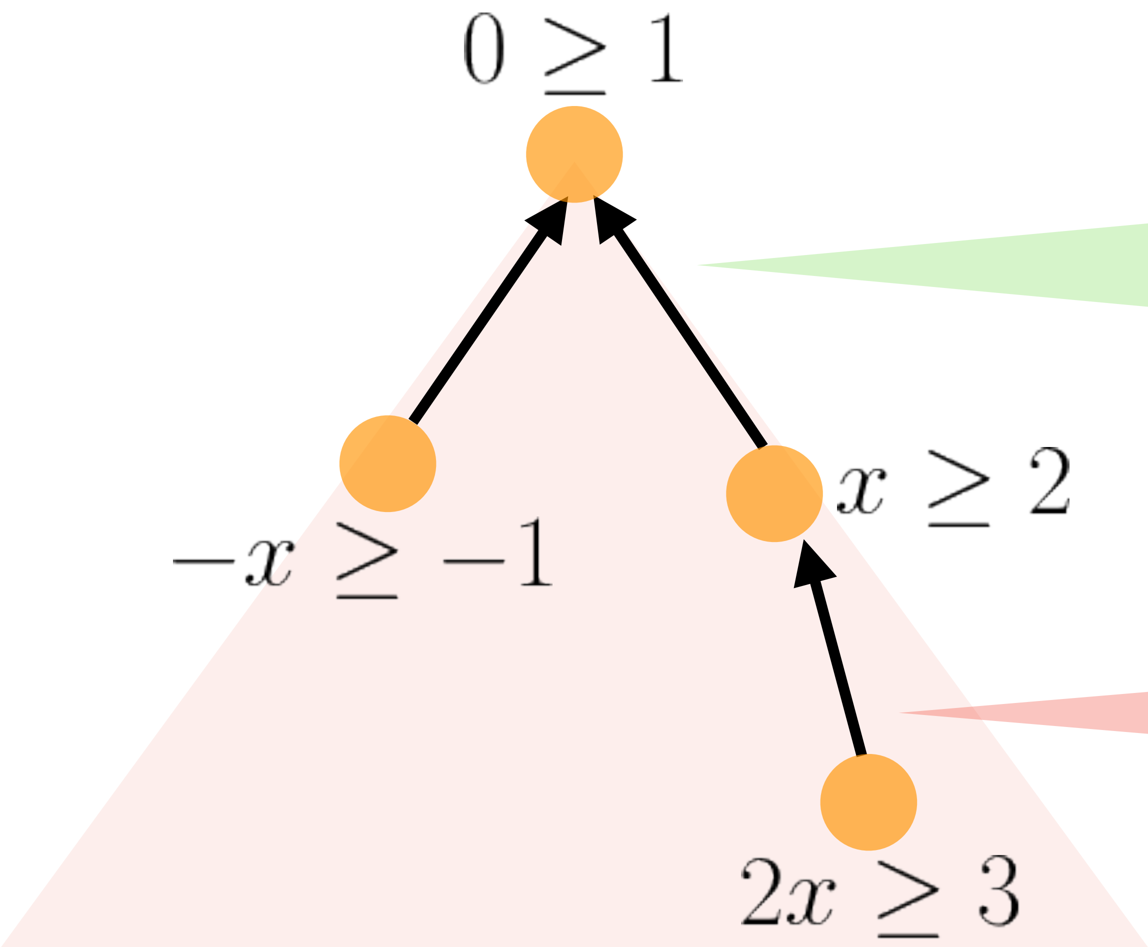
Random θ(log(n))-CNF formulas are Hard for Cutting Planes

Noah Fleming
Department of Computer Science
University of Toronto
Joint work with Denis Pankratov, Toniann Pitassi, and Robert Robere

Cutting Planes



Rules

1) Addition and Multiplication by positive constant

$$\frac{Ax \geq a \qquad Bx \geq b}{c(A+B)x \geq c(a+b)} \text{ for } c \geq 0$$

2) Division with Rounding

$$\frac{dAx \geq a}{Ax \geq \lceil \frac{a}{d} \rceil}$$

Cutting Planes

- Introduced as a method of solving integer linear programming problems [Gomory63, Chvátal73]
- Has short refutations of pigeonhole principle

Feasible interpolation: For any split formula $\mathcal{F}$,

CP-Refutation of $\mathcal{F}$ $\Longrightarrow$ Real Monotone Circuit Computing a related partial function

Split Formula: $A(x, y) \wedge B(y, z)$ on variable sets x, y, z

Example: $\mathrm{Clique}(x, y) \wedge \mathrm{Coloring}(y, z)$

[Pudlak97] Cutting Planes requires an exponential number of lines to refute $\mathrm{Clique}(x, y) \wedge \mathrm{Coloring}(y, z)$.

Random SAT

Random K-CNF:

$\mathcal{F} \sim \mathcal{F}(m, n, k)$: Choose m clauses of width k uniformly at random with replacement from all possible $\binom{n}{k} 2^k$ such clauses

Clause Density $\Delta = m/n$

- Controls whether CNF is satisfiable

Threshold Conjecture: There exists a constant c_k such that for $\mathcal{F} \sim \mathcal{F}(m, n, k)$,

- if $\Delta < c_k$ then $\mathcal{F}$ is satisfiable w.h.p.,
- if $\Delta > c_k$ then $\mathcal{F}$ is unsatisfiable w.h.p.

[Ding,Sly,Sun 15] Resolved for large k

Random SAT

Random K-CNF:

$\mathcal{F} \sim \mathcal{F}(m, n, k)$: Choose m clauses of width k uniformly at random with replacement from all possible $\binom{n}{k} 2^k$ such clauses

- Testbed of hard examples for algorithms in SAT and AI

[Chvátal-Szemerédi]: Random k-CNF formulas $\mathcal{F} \sim \mathcal{F}(m, n, k)$ are w.h.p. hard for Resolution for all $k \geq 3$.

- No efficient Resolution-based algorithms for certifying unsatisfiability of random k-CNF w.h.p.

What about Cutting Planes?

Main Result

$\mathcal{F} \sim \mathcal{F}(m, n, k)$: Choose m clauses of width k uniformly at random with replacement from all possible $\binom{n}{k} 2^k$ such clauses

Theorem: Let $m = n^2 2^k$, $k = \theta(\log n)$ and sample $\mathcal{F} \sim \mathcal{F}(m, n, k)$. With high probability, any Cutting Planes refutation of $\mathcal{F}$ requires $2^{\Omega(n/\log n)}$ lines.

Proved independently by Pavel Pudlák and Pavel Hrubeš

Strategy

Feasible Interpolation: Reduces Cutting Planes refutations of *split formula* to real monotone circuits.

Strategy: Generalize feasible interpolation to work for any unsatisfiable CNF

$\mathcal{F}$: Split formula

CP-Refutation of $\mathcal{F} \Longrightarrow$ Real Monotone Circuit Computing a related partial function

[Pudlak97]

Strategy

Feasible Interpolation: Reduces Cutting Planes refutations of *split formula* to monotone circuits.

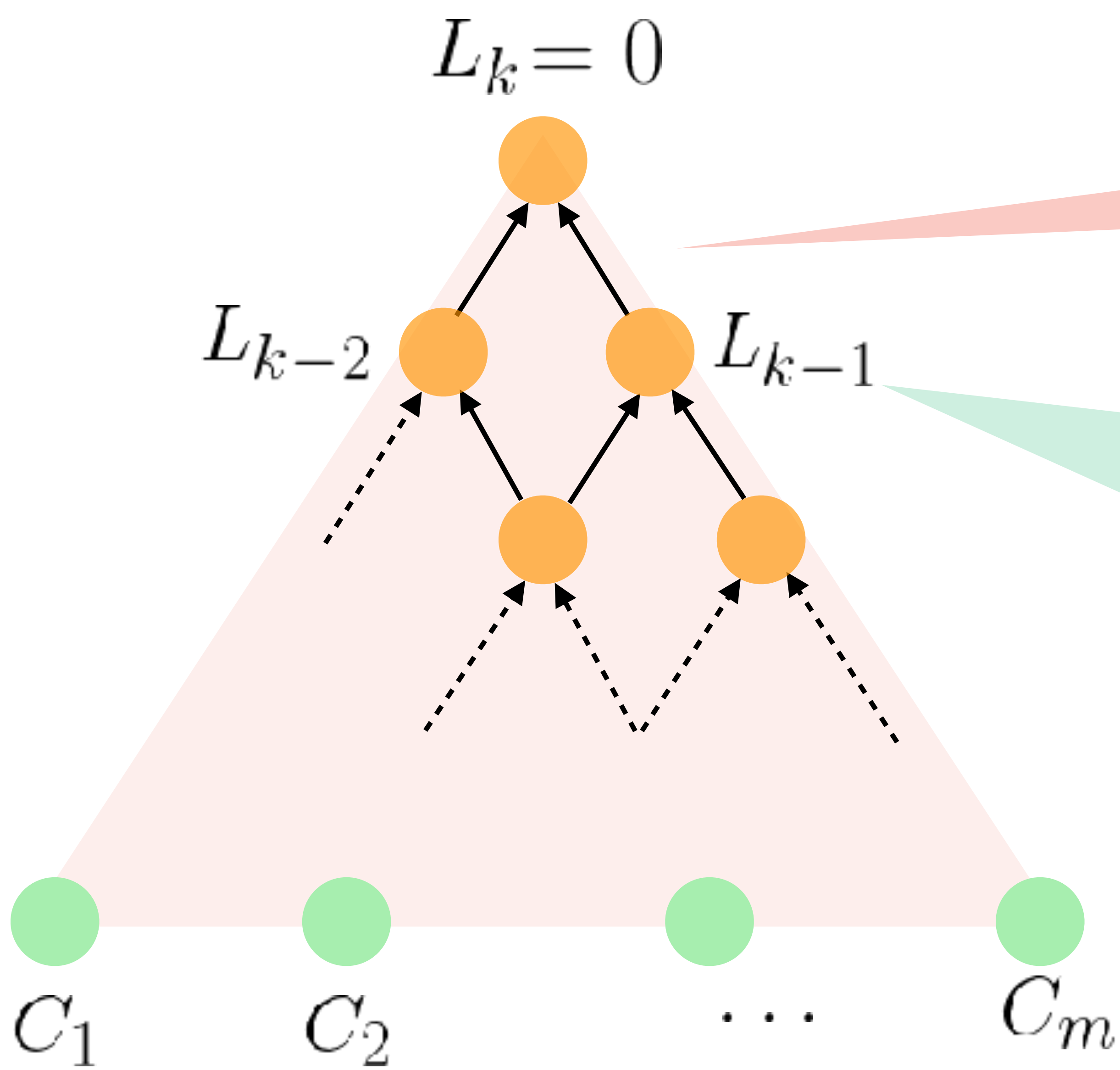
Strategy: Generalize feasible interpolation to work for any unsatisfiable CNF

$\mathcal{F}$: Any unsatisfiable CNF

CC-Refutation of $\mathcal{F} \Longleftrightarrow$ Monotone Circuit Computing a related partial function

CC Refutations

Unsatisfiable $\mathcal{F}(X,Y) = C_1(x,y) \wedge \ldots \wedge C_m(x,y)$ over partition $X \cup Y$



Inference Rules:

Any Sound Inference

Lines:

$$L_i : \{0,1\}^n \to \{0,1\}$$

such that L_i has a small communication protocol over partition $X \cup Y$

Size: Number of lines

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

↓

Communication protocol
for the search problem

↓

Communication protocol
for the Karchmer-Wigderson game

↓

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

Communication protocol
for the search problem

Communication protocol
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

Communication protocol
for the search problem

Communication protocol
for the Karchmer-Wigderson game

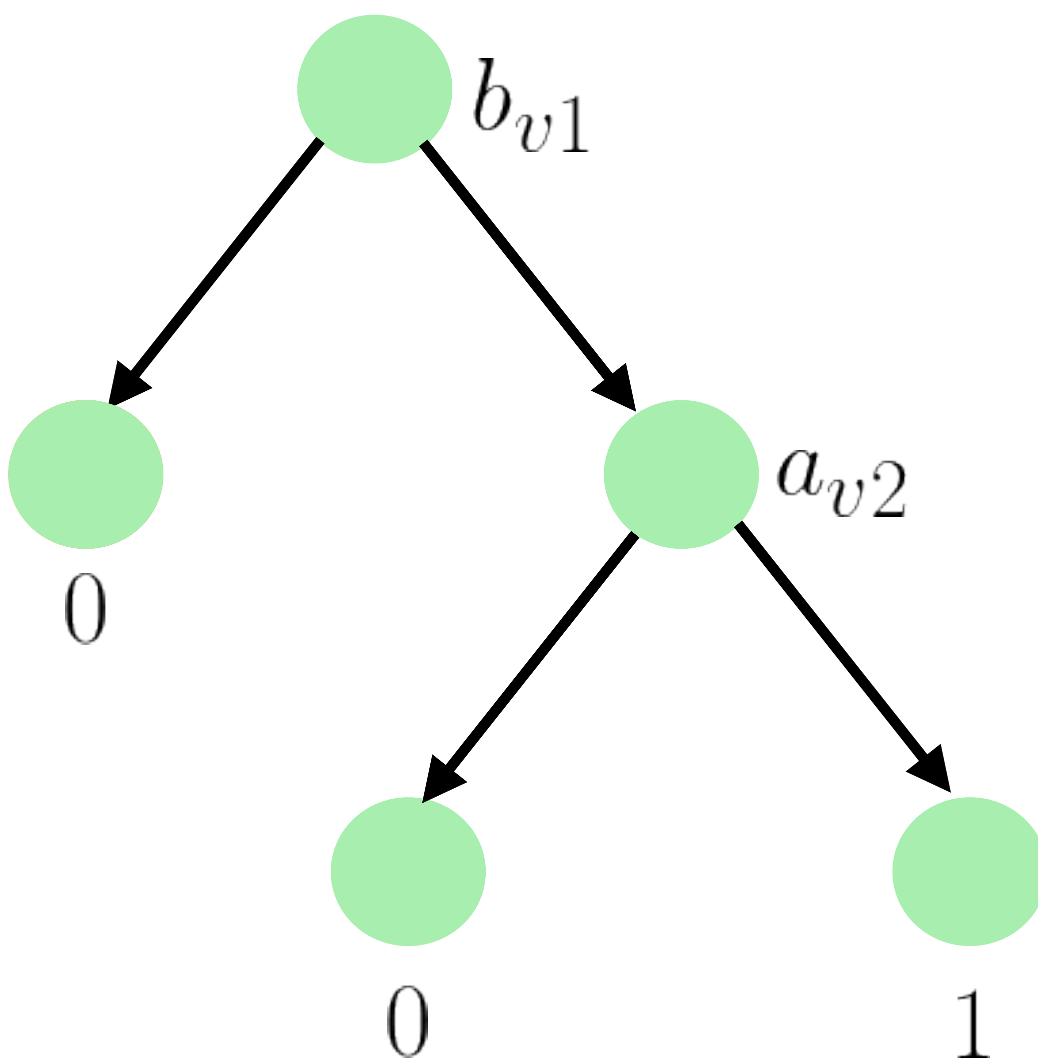
Monotone circuit
separating minterms from maxterms

Intuition for DAG-like Protocol

Deterministic CC Protocol
computing $f(x, y)$

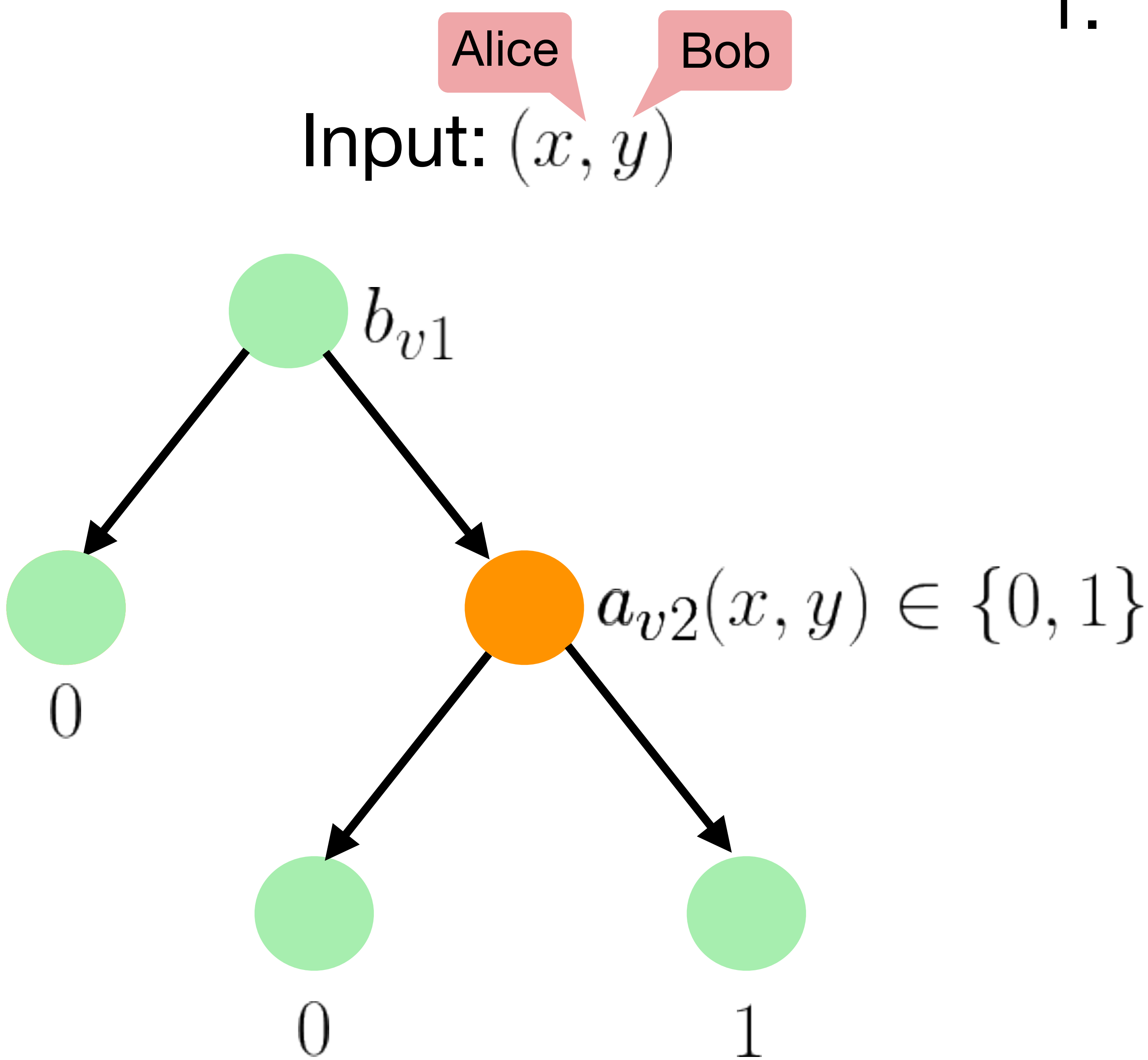
Alice Bob

Input: (x, y)



Intuition for DAG-like Protocol

Deterministic CC Protocol computing $f(x, y)$

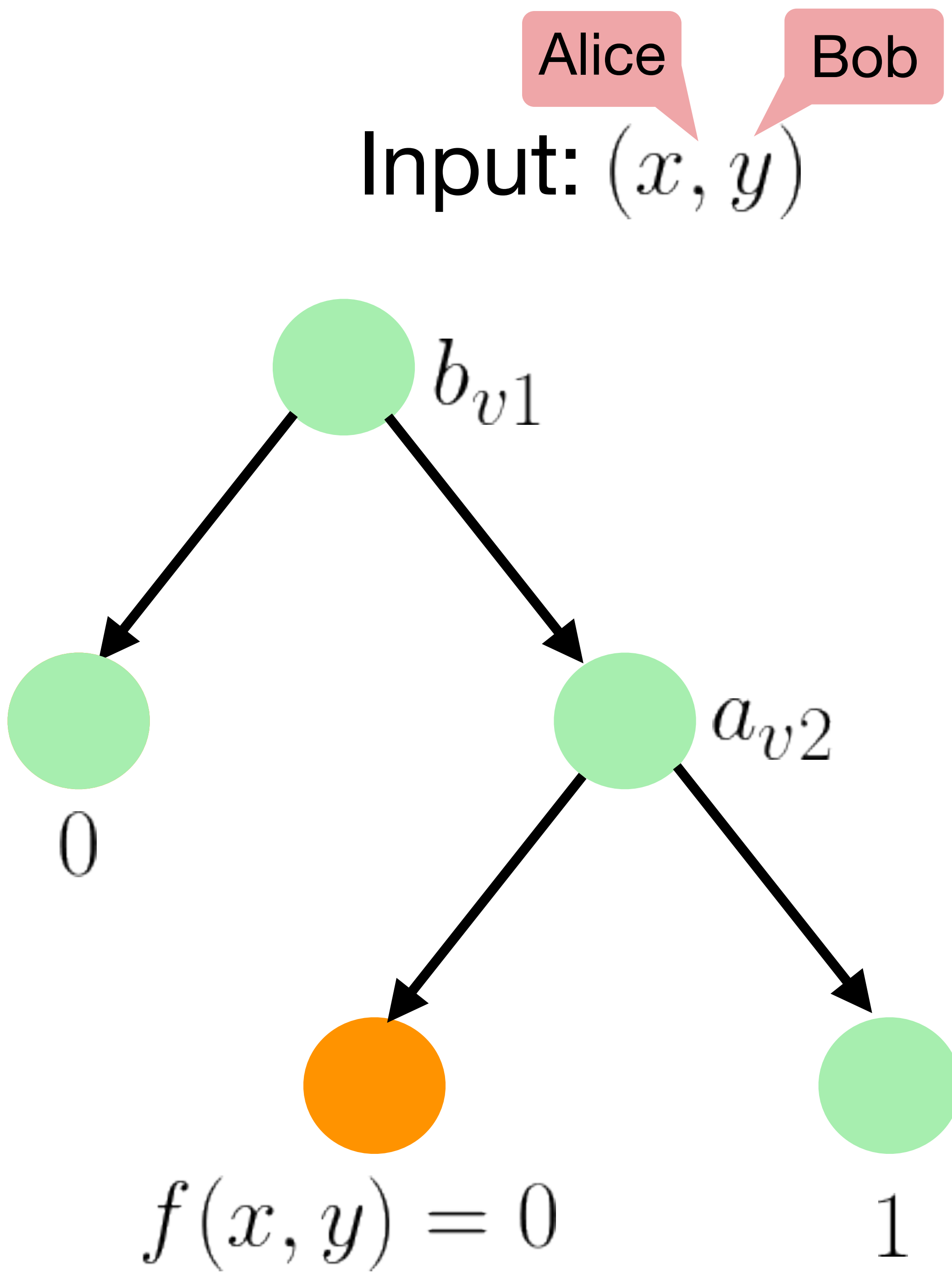


Properties

1. Non-leaf: exactly one child consistent with (x, y), players can efficiently determine which.

Intuition for DAG-like Protocol

Deterministic CC Protocol computing $f(x, y)$

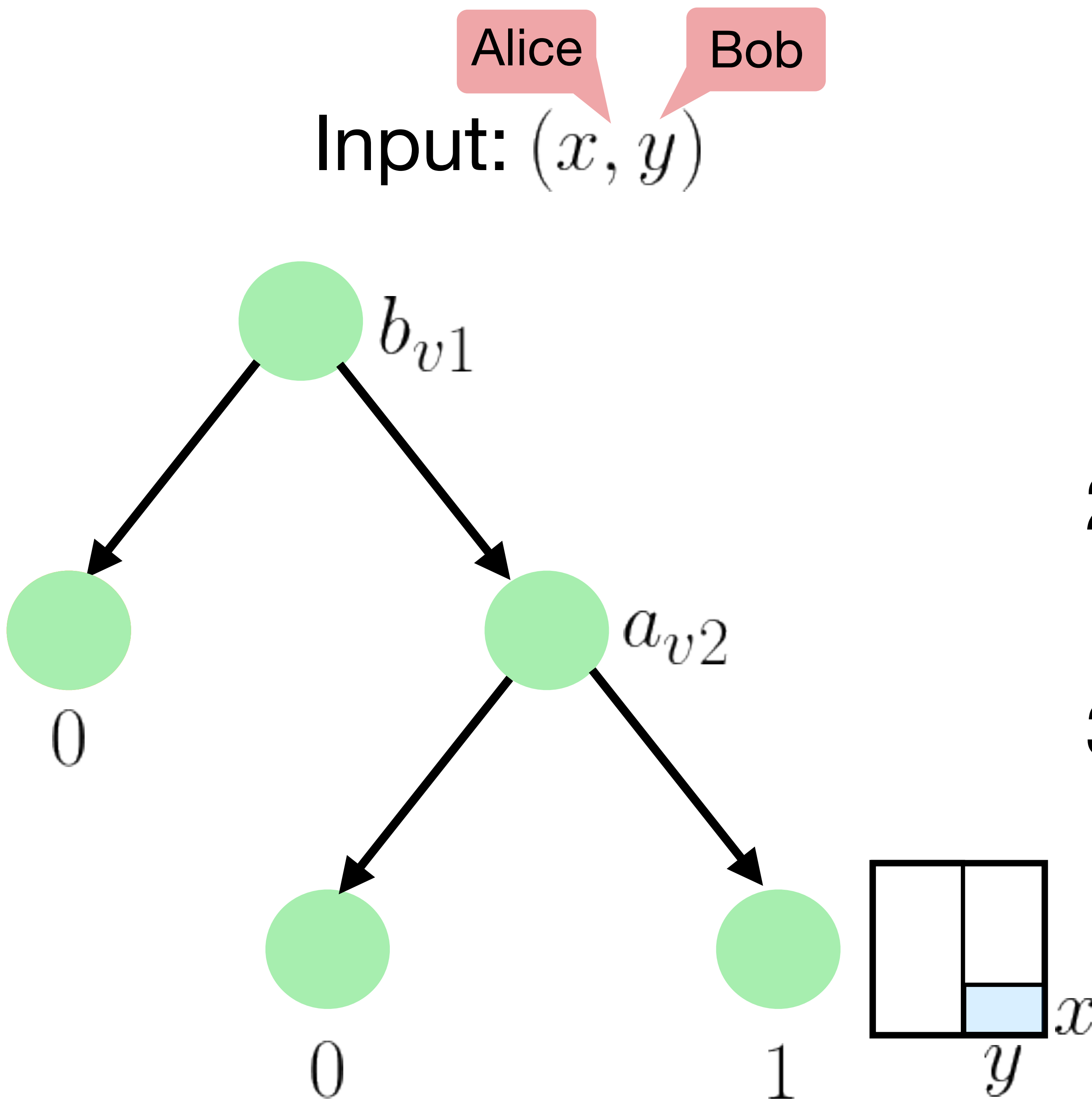


Properties

1. Non-leaf: exactly one child consistent with (x, y), players can efficiently determine which
2. Leaf: labelled with α s.t. $f(x, y) = \alpha$

Intuition for DAG-like Protocol

Deterministic CC Protocol computing $f(x, y)$



Properties

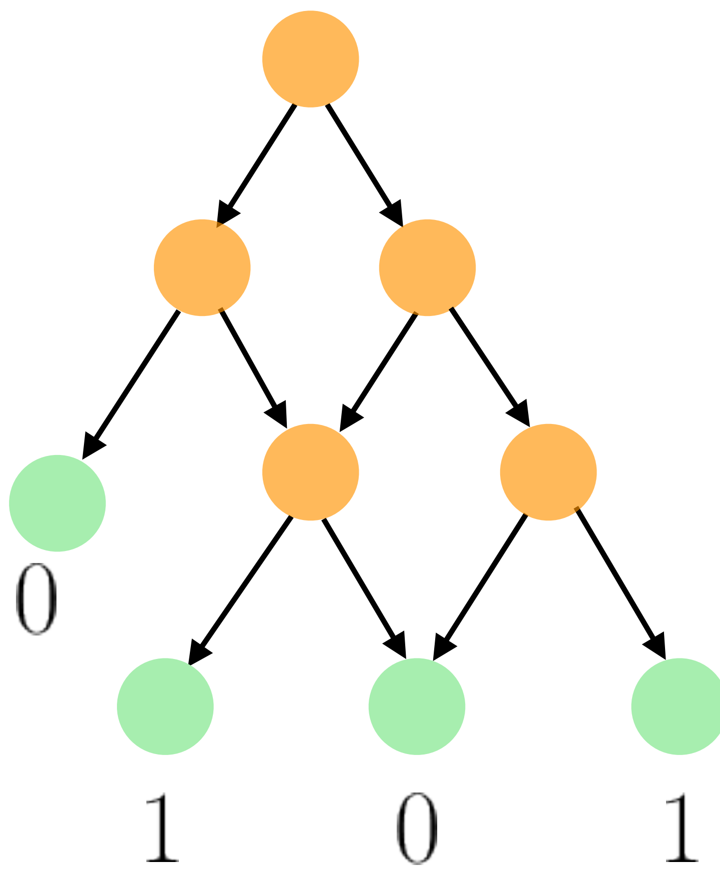
1. Non-leaf: exactly one child consistent with (x, y), players can efficiently determine which
2. Leaf: labelled with α s.t. $f(x, y) = \alpha$
3. For every node, players can efficiently check if they can reach this node on input (x, y)

CC-Games (PLS games [Razborov95])

CC-Game Computing $f(x, y)$

Alice Bob

Input: (x, y)



Satisfying:

1. Non-leaf: exactly one child consistent with (x, y), players can efficiently determine which
2. Leaf: labelled with α s.t. $f(x, y) = \alpha$
3. For every node, players can efficiently check if they can reach this node on input (x, y)

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

Communication protocol
for the search problem

Communication protocol
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

CC-game
for the search problem

CC-game
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

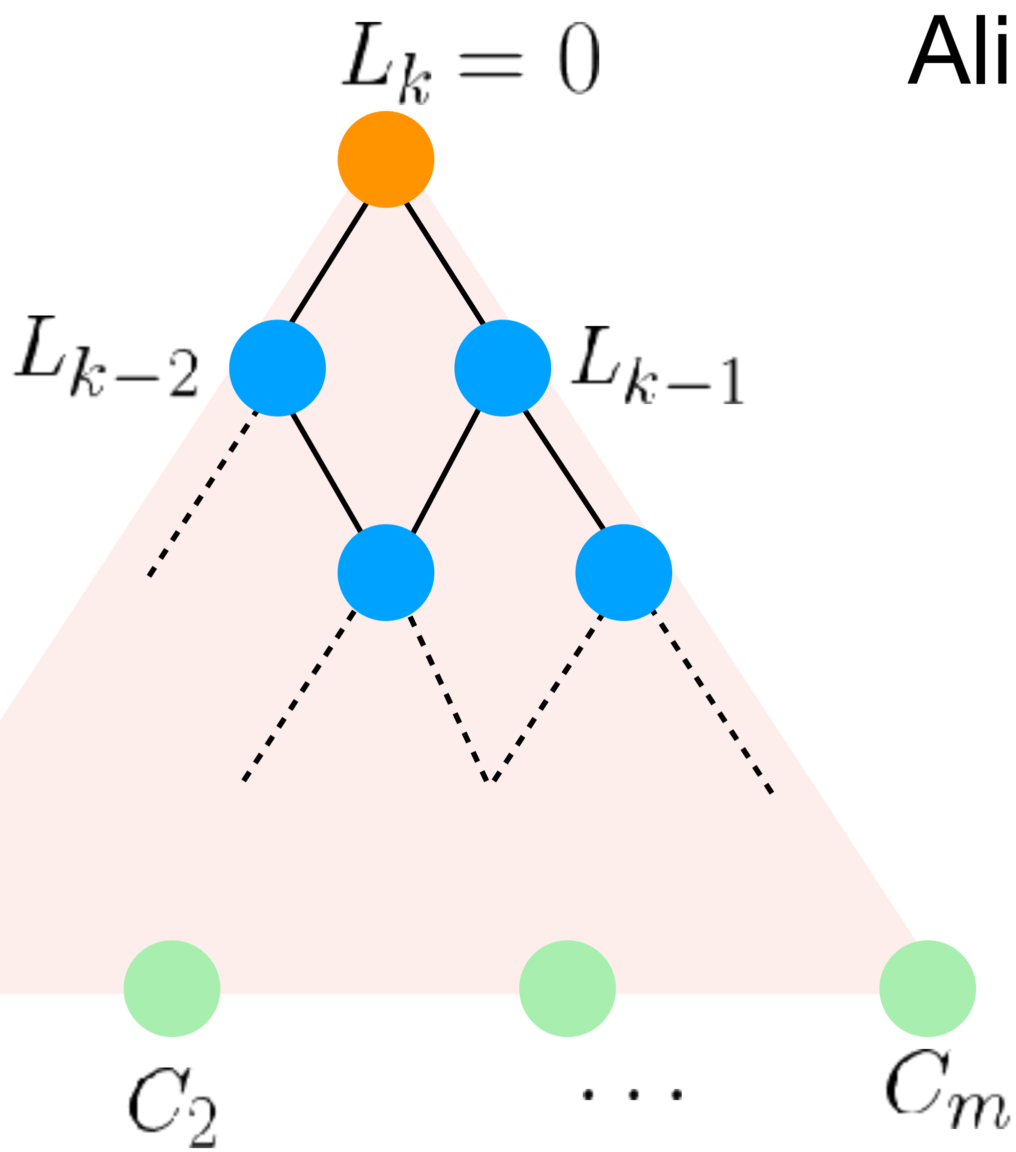
CC-Refutation $\mathcal{F}(X,Y)$ ⇨ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

Unsatisfiable $\mathcal{F}(X,Y) = C_1(x,y) \wedge \ldots \wedge C_m(x,y)$ over partition $X \cup Y$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x,y), output $i \in [m]$ such that $C_i(x,y) = 0$

CC-refutation of $\mathcal{F}$:

Alice: x Bob: y

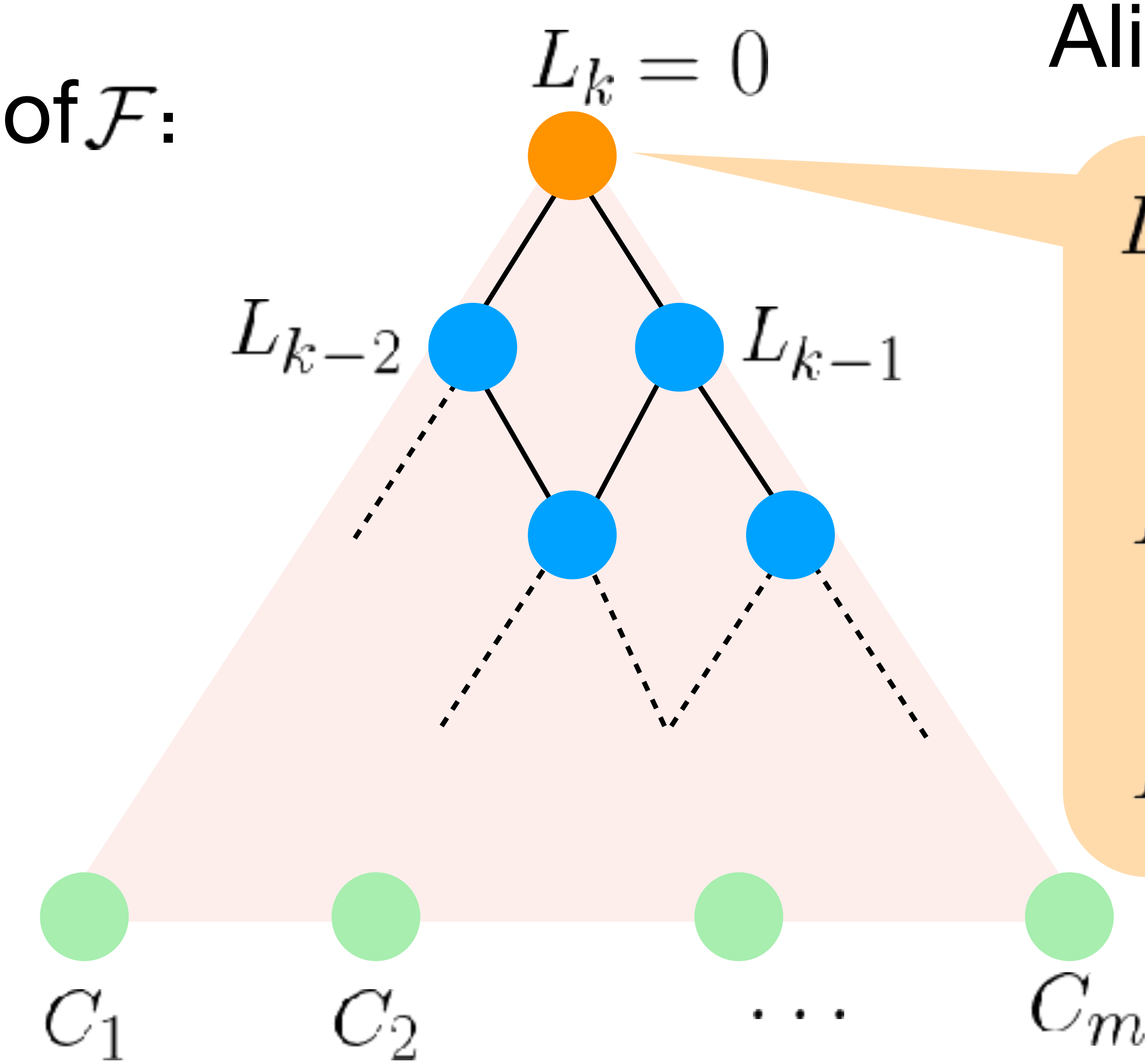


CC-Refutation $\mathcal{F}(X,Y) \Rightarrow$ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x,y), output $i \in [m]$ such that $C_i(x,y) = 0$

Alice: x Bob: y

CC-refutation of $\mathcal{F}$:



$L_{k-1}, L_{k-2} \models L_k$

Therefore

$L_{k-1}(x,y) = 0$

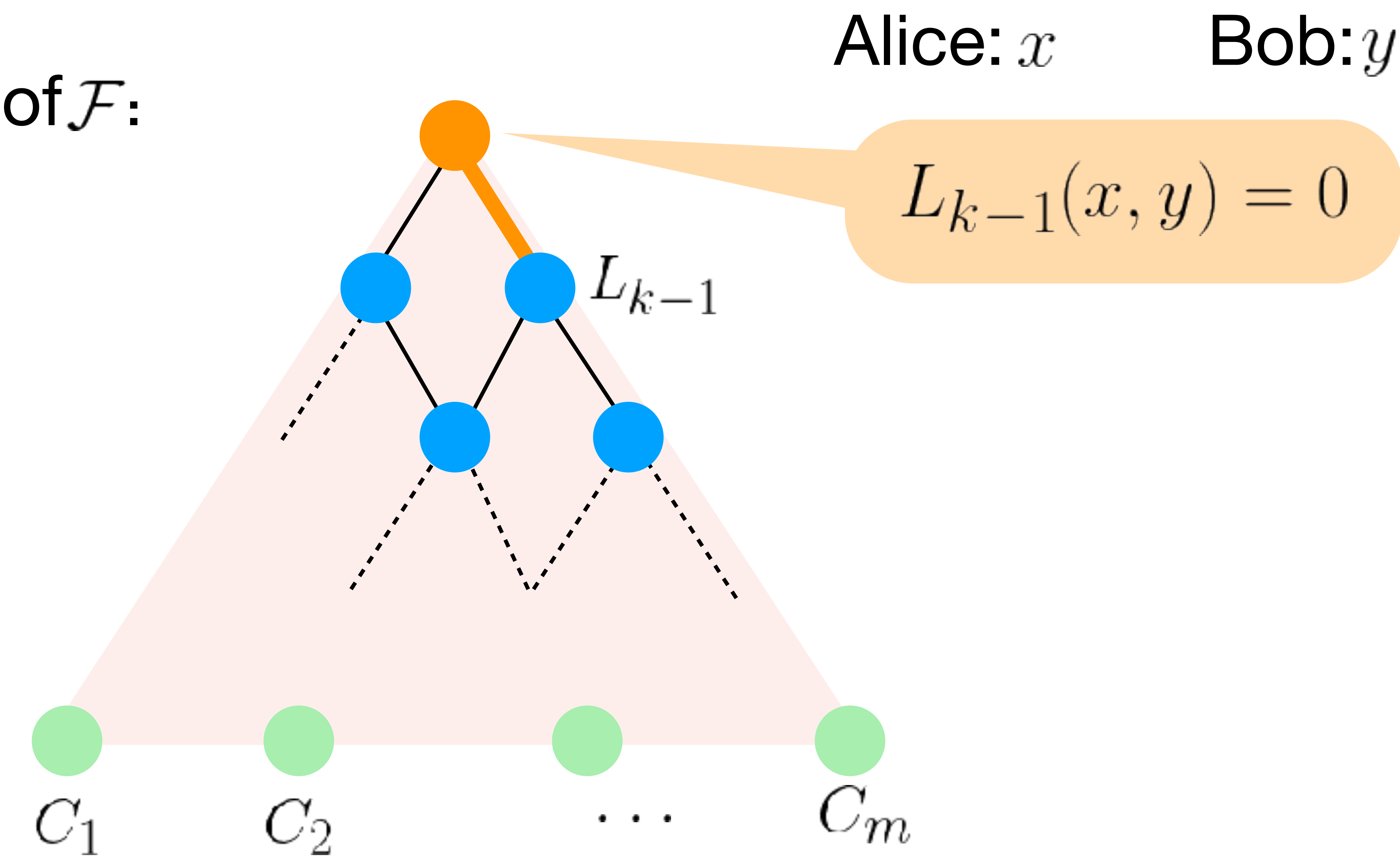
OR

$L_{k-2}(x,y) = 0$

CC-Refutation $\mathcal{F}(X,Y) \Rightarrow$ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x,y), output $i \in [m]$ such that $C_i(x,y)=0$

CC-refutation of $\mathcal{F}$:

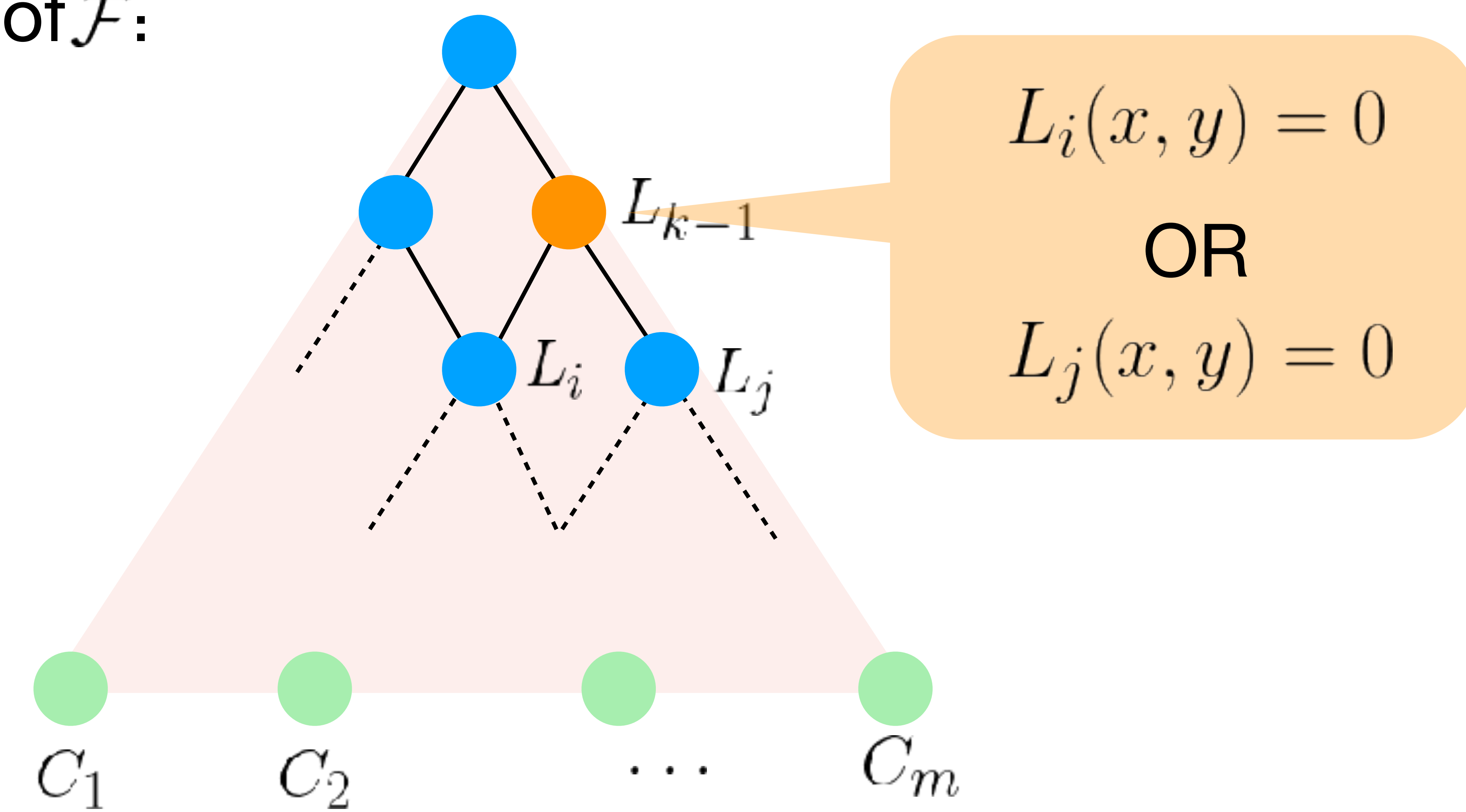


CC-Refutation $\mathcal{F}(X, Y)$ ⇨ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x, y), output $i \in [m]$ such that $C_i(x, y) = 0$

CC-refutation of $\mathcal{F}$:

Alice: x Bob: y

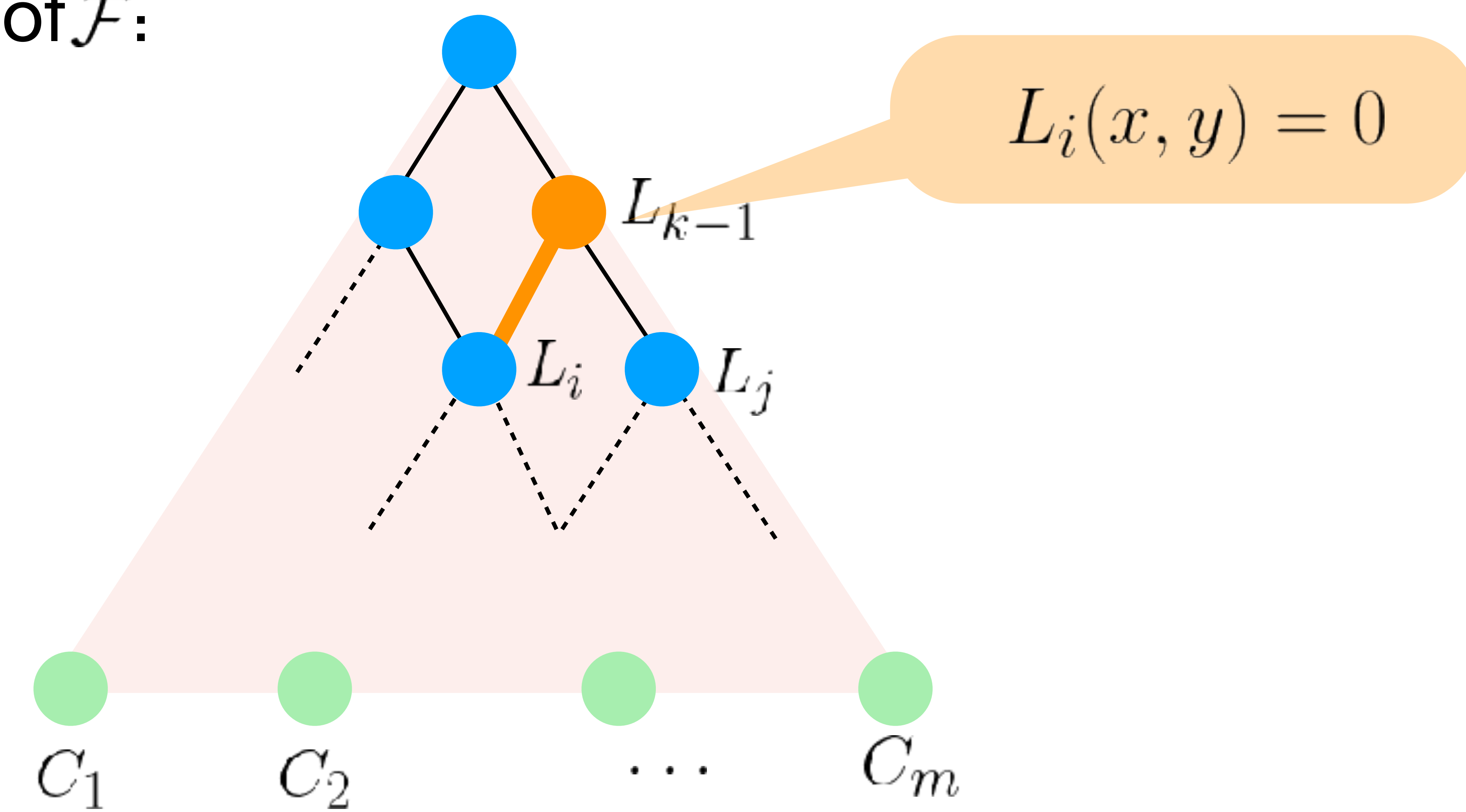


CC-Refutation $\mathcal{F}(X, Y) \Rightarrow$ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x, y), output $i \in [m]$ such that $C_i(x, y) = 0$

Alice: x Bob: y

CC-refutation of $\mathcal{F}$:

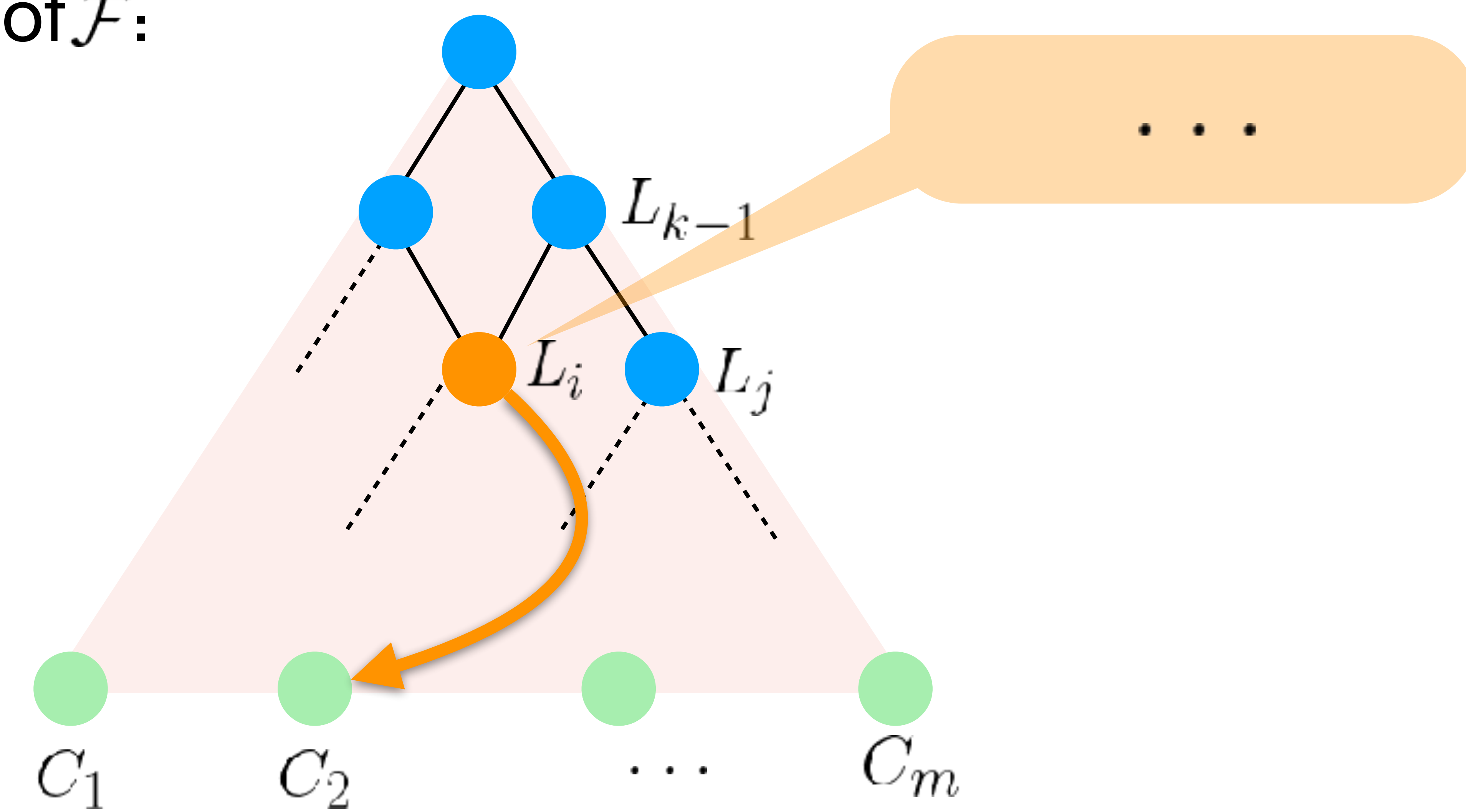


CC-Refutation $\mathcal{F}(X,Y) \Rightarrow$ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x,y), output $i \in [m]$ such that $C_i(x,y) = 0$

Alice: x Bob: y

CC-refutation of $\mathcal{F}$:

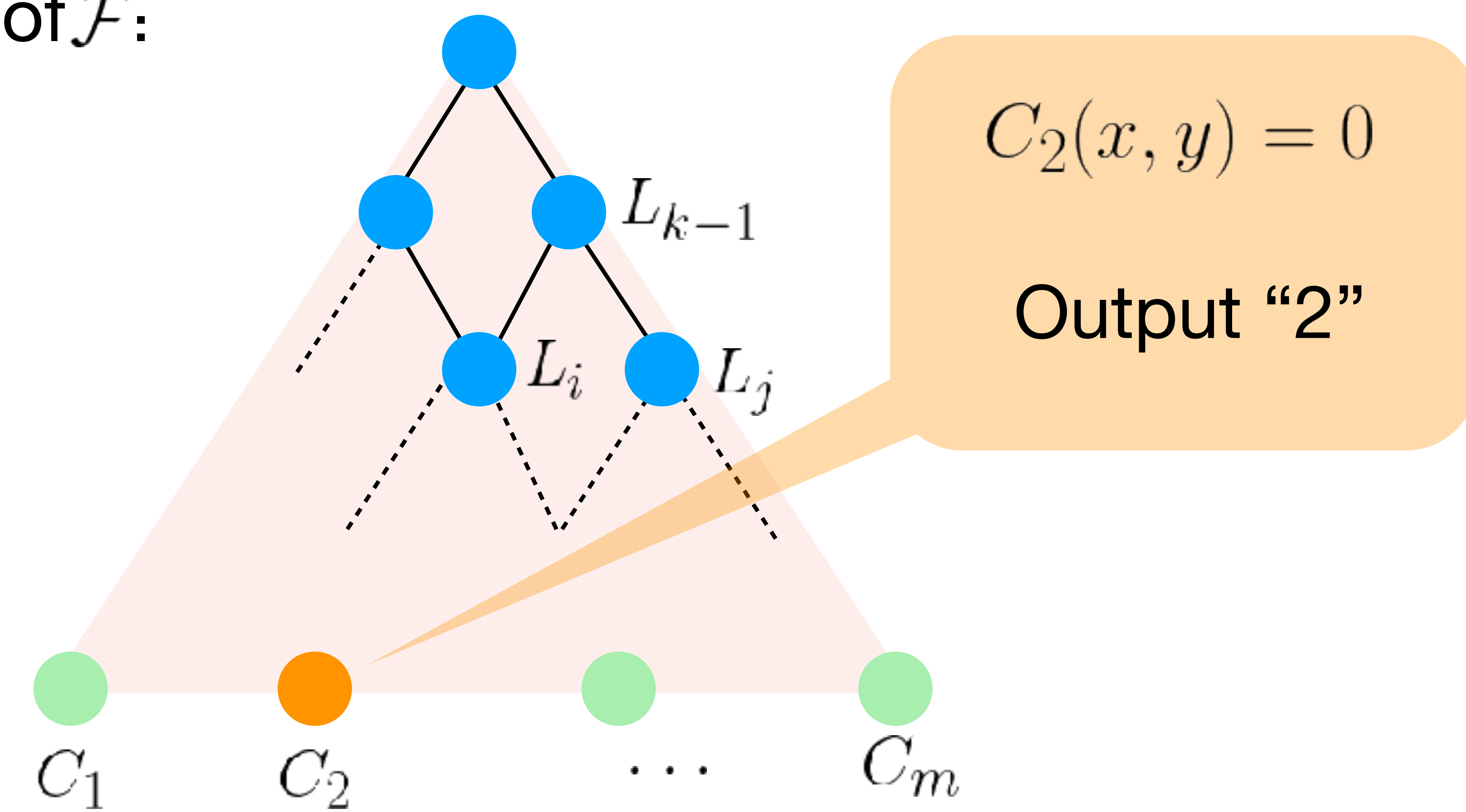


CC-Refutation $\mathcal{F}(X, Y)$ ⇨ Protocol $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\mathrm{Search}_{X,Y}(\mathcal{F})$: Given truth assignment (x, y), output $i \in [m]$ such that $C_i(x, y) = 0$

Alice: x Bob: y

CC-refutation of $\mathcal{F}$:



Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

CC-game
for the search problem

CC-game
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

CC-game
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

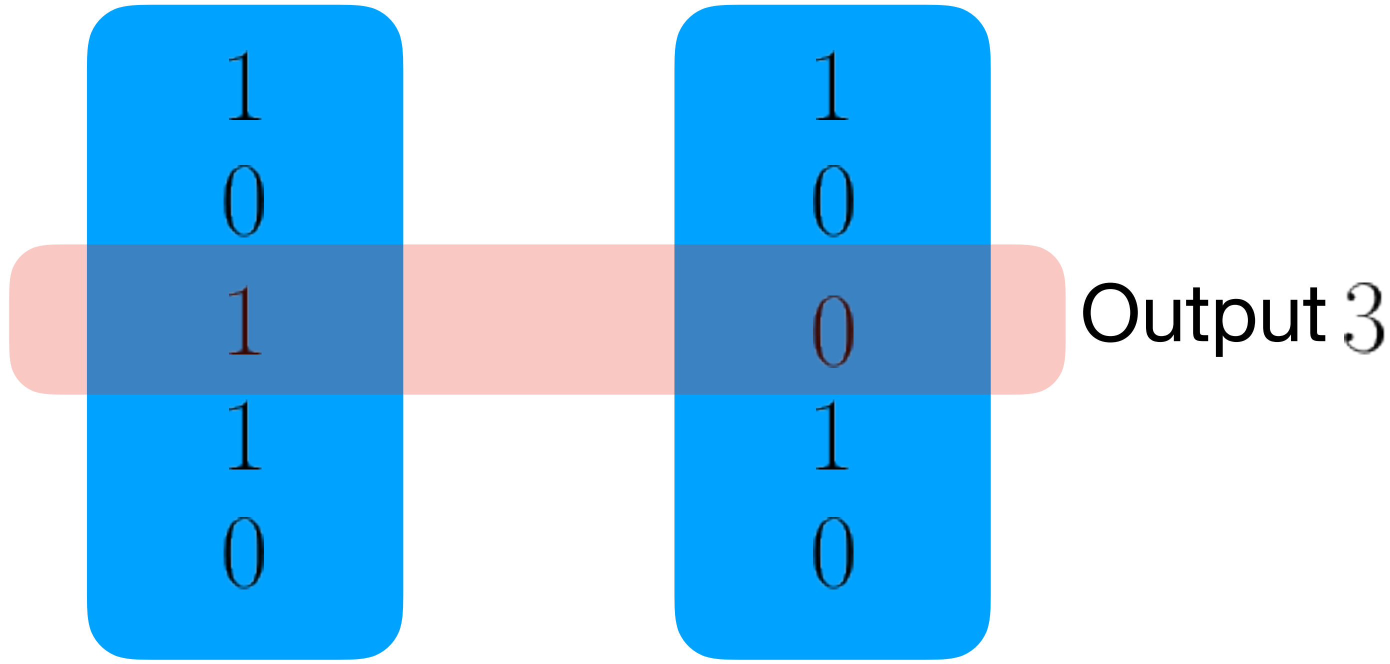
CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

CC-game
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Monotone Karchmer-Wigderson Game

Alice: $x \in \{0,1\}^n$ s.t. $f(x) = 1$ Bob: $y \in \{0,1\}^n$ s.t. $f(y) = 0$



Output: $i \in [n]$ such that $x_i = 1, y_i = 0$

[Razborov 95]: For any partial function monotone function $f : \{0,1\}^n \to \{0,1,*\}$,

monotone CKT size(f) = CC-Game size($KW^+(f)$)

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

CC-game
for the Karchmer-Wigderson game

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X,Y)$

↓

CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

CC-game for KW^+(minterms, maxterms)

↓

Monotone circuit
separating minterms from maxterms

Strategy

CC-Refutation of $\mathcal{F}(X, Y)$

CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

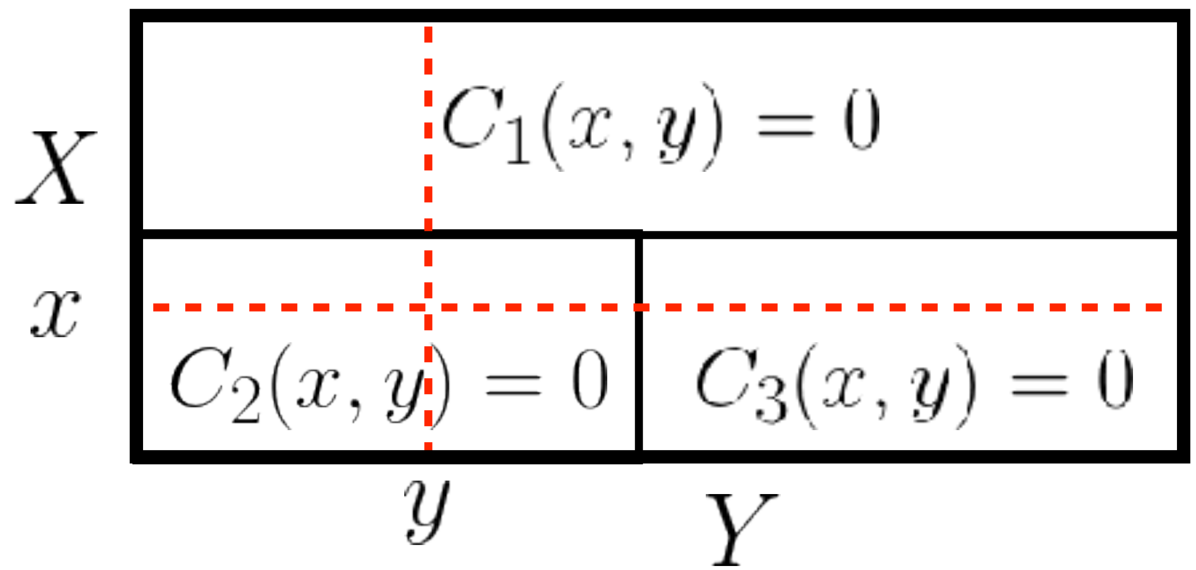
CC-game for KW^+(minterms, maxterms)

Monotone circuit
separating minterms from maxterms

Unsatisfiability Certificate

Unsatisfiable $\mathcal{F}(X,Y) = C_1(x,y) \wedge C_2(x,y) \wedge C_3(x,y)$ over partition $X \cup Y$

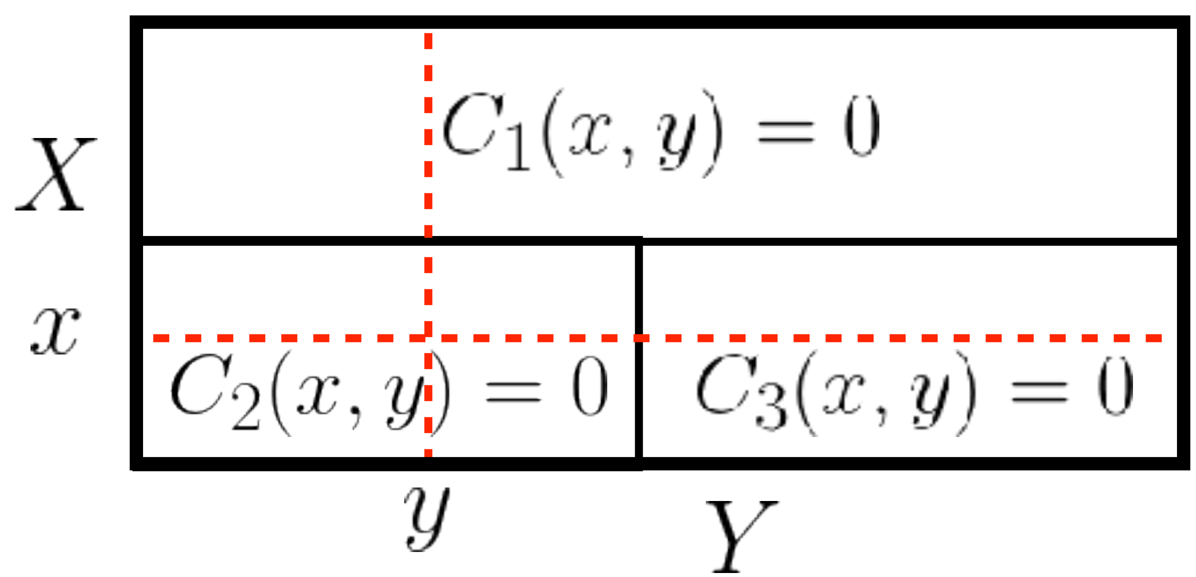
Inputs to $\mathrm{Search}_{X,Y}(\mathcal{F})$ Alice: x Bob: y



Unsatisfiability Certificate

Unsatisfiable $\mathcal{F}(X,Y) = C_1(x,y) \wedge C_2(x,y) \wedge C_3(x,y)$ over partition $X \cup Y$

Inputs to $\mathrm{Search}_{X,Y}(\mathcal{F})$ Alice: x Bob: y



Corresponding inputs to monotone KW game:

		Example	General
Alice:	$x \to \mathcal{U}(x)$	$\mathcal{U}(x) = [0, 1, 1]$	$\mathcal{U}(x)_i = 1 \iff C_i \restriction_X (x) = 0$
Bob:	$y \to \mathcal{V}(y)$	$\mathcal{V}(y) = [0, 0, 1]$	$\mathcal{V}(y)_i = 0 \iff C_i \restriction_Y (y) = 0$

Unsatisfiability Certificate

Alice: $x \to \mathcal{U}(x)$ | $\mathcal{U}(x)_i = 1 \iff C_i \restriction_X (x) = 0$

Bob: $y \to \mathcal{V}(y)$ | $\mathcal{V}(y)_i = 0 \iff C_i \restriction_Y (y) = 0$

Resulting partial function (unsatisfiability certificate):

Minterms: (abuse of notation)

$\mathcal{U}$: The set of outputs of the map $\mathcal{U}$ over all x

Maxterms:

$\mathcal{V}$: The set of outputs of the map $\mathcal{V}$ over all y

Equivalently [HP17],

Unsatisfiability Certificate: $z \in \{0,1\}^m$

$$\mathrm{Certificate}_{\mathcal{F}}(z) = \begin{cases} 1 & \text{if } \{C_i \restriction_X : i \in [m] \setminus z\} \text{ is satisfiable,} \\ 0 & \text{if } \{C_i \restriction_Y : i \in [m]\} \text{ is satisfiable,} \\ * & \text{otherwise} \end{cases}$$

Strategy

CC-Refutation of $\mathcal{F}(X,Y)$

$\Downarrow$

CC-game for $\mathrm{Search}_{X,Y}(\mathcal{F})$

$\Downarrow$

CC-game for KW^+(minterms, maxterms)

$\Downarrow$

Monotone circuit separating $\mathcal{U}$ from $\mathcal{V}$

Theorem: Let $\mathcal{F}$ be an unsatisfiable CNF and $X \cup Y$ be any partition of the variables of $\mathcal{F}$.

CC-Refutation size $\mathcal{F}(X,Y) \approx$ Monotone CKT size ($\mathcal{U}$, $\mathcal{V}$)

Monotone Circuit Lower Bound

$\mathcal{F} \sim \mathcal{F}(m, n, k)$: Choose m clauses of width k uniformly at random with replacement from all possible $\binom{n}{k}2^k$ such clauses

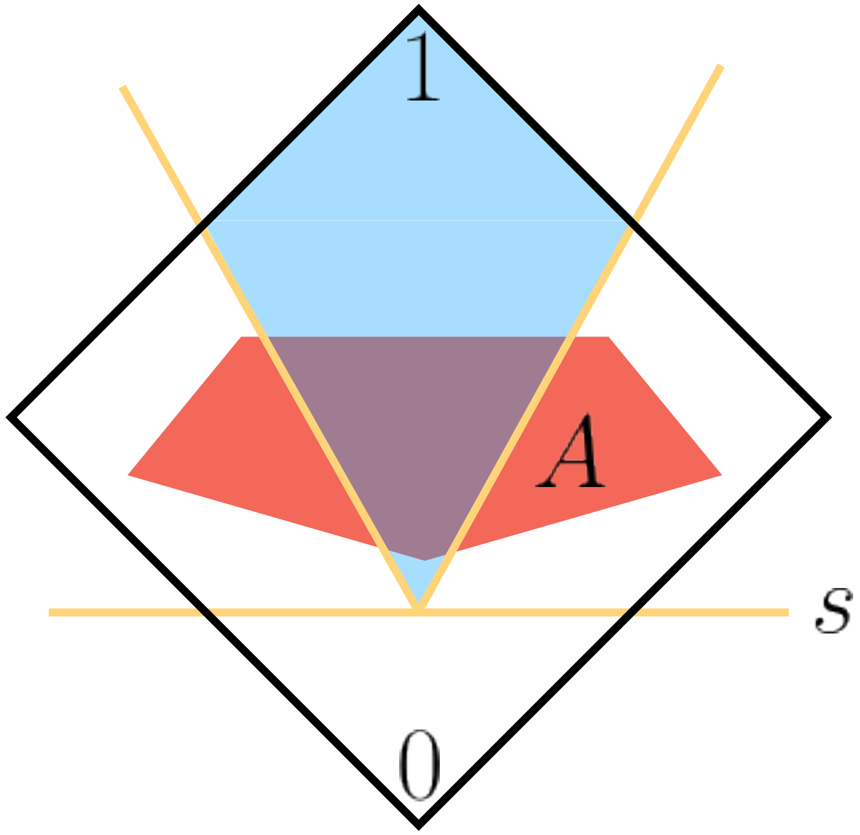
Theorem: Let $m = n^2 2^k$, $k = \theta(\log n)$, and sample $\mathcal{F} \sim \mathcal{F}(m, n, k)$. W.h.p, any monotone circuit separating $\mathcal{U}$ from $\mathcal{V}$ requires $2^{\Omega(n/\log n)}$ gates.

Symmetric Method

Spread out Measure: $A_b(s, A) = \max_{I \subseteq [n]: |I| = s} |\{x \in A : \forall i \in I, x_i = b\}|$

$A_b(s, A)$ small if no set of s variables that, set to b, agrees with a lot of strings in A

$A_1(s, A)$:



Symmetric Method

Spread out Measure: $A_b(s, A) = \max_{I \subseteq [n] : |I| = s} |\{x \in A : \forall i \in I, x_i = b\}|$

Let $f : \{0,1\}^n \to \{0,1\}$ be a partial monotone function

[Jukna99] Any monotone circuit computing f requires at least

$$\min\left\{\frac{|U| - (s-1)A_1(1, U)}{(s-1)^s A_1(s, U)}, \frac{|V|}{(s-1)^s A_0(s, V)}\right\}$$

gates, for any s. Where $U = f^{-1}(1)$, $V = f^{-1}(0)$.

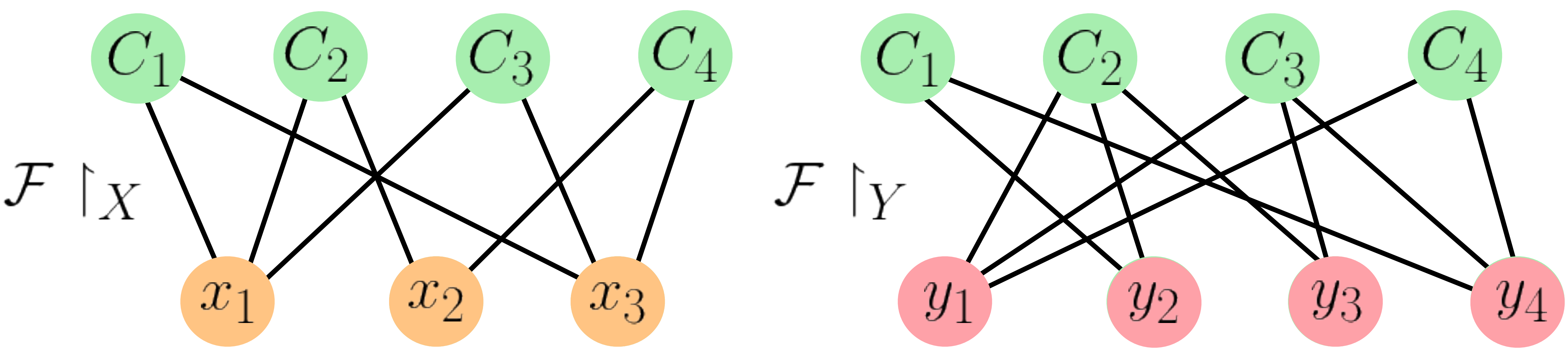
Monotone Circuit Lower Bound

Good expansion properties $\Longrightarrow A_b(s, A)$ is small,

$\mathcal{F} \sim \mathcal{F}(m, n, k)$ is expanding w.h.p.!

Problem

Need $\mathcal{U}$ and $\mathcal{V}$ to be expanding.



That is, we need $\mathcal{F} \restriction_X$ and $\mathcal{F} \restriction_Y$ to be expanding with respect to the fixed variable partition $X \cup Y$.

Balanced Random CNF

Temporary Solution: Change Distribution

$\text{Balanced} - \mathcal{F}(m, n, 2k)$:

1. Sample $\mathcal{F}_X \sim \mathcal{F}(m, n, k)$ on X-variables, $\mathcal{F}_X = C_1(x) \wedge \ldots \wedge C_m(x)$
2. Sample $\mathcal{F}_Y \sim \mathcal{F}(m, n, k)$ on Y-variables, $\mathcal{F}_Y = C_1(y) \wedge \ldots \wedge C_m(y)$

Output: $\mathcal{F}(X, Y) = (C_1(x) \vee C_1(y)) \wedge \ldots \wedge (C_m(x) \vee C_m(y))$

$\mathcal{F} \restriction_X$ and $\mathcal{F} \restriction_Y$ both expanding!

Theorem: Let $m = n^2 2^k$, $k = \theta(\log n)$, and sample $\text{Balanced} - \mathcal{F}(m, n, 2k)$. W.h.p, any monotone circuit separating $\mathcal{U}$ from $\mathcal{V}$ requires $2^{\Omega(n/\log n)}$ gates.

Random CNF

Strategy: Reduce to balanced case!

Sample $\mathcal{F} \sim \mathcal{F}(m, n, k)$, show existence of partition $X \cup Y$, such that

1. Most of the clauses of $\mathcal{F}$ are balanced w.r.t. X and Y,
2. There exists a large set of assignments $\mathcal{A}$ to the X-variables and $\mathcal{B}$ to the Y-variables which satisfy all of the unbalanced clauses.

Apply Symmetric Method of Approximations to $(\mathcal{U}(\mathcal{A}), \mathcal{V}(\mathcal{B}))$

Theorem: Let $m = n^2 2^k$, $k = \theta(\log n)$ and sample $\mathcal{F} \sim \mathcal{F}(m, n, k)$. With high probability, any Cutting Planes refutation of $\mathcal{F}$ requires $2^{\Omega(n/\log n)}$ lines.

Conclusion

- First exponential lower bound on the size of Cutting Planes refutations of random $\theta(\log n)$-CNFs

- Lower bound for random k-CNF for k = constant?
  -Improve symmetric method of approximations, $(s-1)^s$ term in the denominator kills us!

- Cutting Planes lower bound for Tseitin formulas?
  -**Technique incapable of handling Tseitin formulas!**
  - $O(n)$ upper bound on Tseitin in CC.

Thanks!