Disjointness is hard in the multi-party number-on-the-forehead model

Troy Lee
Rutgers University

Adi Shraibman
Weizmann Institute of Science

A brief history of disjointness

- Set intersection: Alice holds $x \in \{0,1\}^n$, Bob $y \in \{0,1\}^n$. Do they share a common element?

A brief history of disjointness

- Set intersection: Alice holds $x \in \{0,1\}^n$, Bob $y \in \{0,1\}^n$. Do they share a common element?
- Deterministic communication complexity n bits

A brief history of disjointness

- Set intersection: Alice holds $x \in \{0,1\}^n$, Bob $y \in \{0,1\}^n$. Do they share a common element?

- Deterministic communication complexity n bits

- Nondeterministic complexity is $O(\log n)$.

A brief history of disjointness

- Set intersection: Alice holds $x \in \{0,1\}^n$, Bob $y \in \{0,1\}^n$. Do they share a common element?

- Deterministic communication complexity n bits

- Nondeterministic complexity is $O(\log n)$.

- Randomized complexity $\Theta(n)$ [KS87, Raz92]

A brief history of disjointness

- Set intersection: Alice holds $x \in \{0,1\}^n$, Bob $y \in \{0,1\}^n$. Do they share a common element?
- Deterministic communication complexity n bits
- Nondeterministic complexity is $O(\log n)$.
- Randomized complexity $\Theta(n)$ [KS87, Raz92]
- Quantum complexity $\Theta(\sqrt{n})$ [lower Raz03, upper AA03]

Number-on-the-forehead model

- k-players, input $x_1, \ldots, x_k$. Player i knows everything but x_i.
- Large overlap in information makes showing lower bounds difficult.
- Lower bounds have application to powerful models like circuit complexity and complexity of proof systems.
- Best lower bounds are of the form $n/2^k$. Bound of $n/2^{2k}$ for generalized inner product function $\oplus(x_1 \wedge \ldots \wedge x_k)$ [BNS89].

Disjointness in the number-on-the-forehead model

- Best lower bound $\Omega(\frac{\log n}{k-1})$, and best upper bound $O(kn/2^k)$ [lower Tes02, BPSW06, upper Gro94].

Disjointness in the number-on-the-forehead model

- Best lower bound $\Omega(\frac{\log n}{k-1})$, and best upper bound $O(kn/2^k)$ [lower Tes02, BPSW06, upper Gro94].

- Kushilevitz and Nisan: "The only technique from two-party complexity that generalizes to multiparty complexity is the discrepancy method."

Disjointness in the number-on-the-forehead model

- Best lower bound $\Omega(\frac{\log n}{k-1})$, and best upper bound $O(kn/2^k)$ [lower Tes02, BPSW06, upper Gro94].

- Kushilevitz and Nisan: "The only technique from two-party complexity that generalizes to multiparty complexity is the discrepancy method." For disjointness, discrepancy can only show bounds of $O(\log n)$.

Disjointness in the number-on-the-forehead model

- Best lower bound $\Omega(\frac{\log n}{k-1})$, and best upper bound $O(kn/2^k)$ [lower Tes02, BPSW06, upper Gro94].

- Kushilevitz and Nisan: "The only technique from two-party complexity that generalizes to multiparty complexity is the discrepancy method." For disjointness, discrepancy can only show bounds of $O(\log n)$.

- Researchers have studied restricted models—bound of $n^{1/3}$ for three players where first player speaks and dies [BPSW06]. Bound of $n^{1/k}/k^k$ in one-way model [VW07].

Our results

- We show disjointness requires randomized communication

$$\Omega\left(\frac{n^{1/(k+1)}}{2^{2^k}}\right)$$

in the general k-party number-on-the-forehead model.

Our results

- We show disjointness requires randomized communication

$$\Omega\left(\frac{n^{1/(k+1)}}{2^{2^k}}\right)$$

  in the general k-party number-on-the-forehead model.

- Chattopadhyay and Ada independently obtained similar results

Our results

- We show disjointness requires randomized communication

$$\Omega\left(\frac{n^{1/(k+1)}}{2^{2^k}}\right)$$

  in the general k-party number-on-the-forehead model.

- Chattopadhyay and Ada independently obtained similar results

- Separates multiparty communication complexity versions of NP and BPP for up to $k = \log\log n - O(\log\log\log n)$ many players.

Application to proof systems

- As linear and semidefinite programming are some of the most sophisticated algorithms we have developed, natural to see how they fare on NP-complete problems.

- One way to formalize this is through proof complexity: for example cutting planes, Lovász-Schrijver proof systems.

Application to proof systems

- As linear and semidefinite programming are some of the most sophisticated algorithms we have developed, natural to see how they fare on NP-complete problems.

- One way to formalize this is through proof complexity: for example cutting planes, Lovász-Schrijver proof systems.

- Beame, Pitassi, and Segerlind show that lower bounds on NOF disjointness imply lower bounds for a very general class of proof systems, including the above [BPS06].

Tree-like semantically entailed proof systems

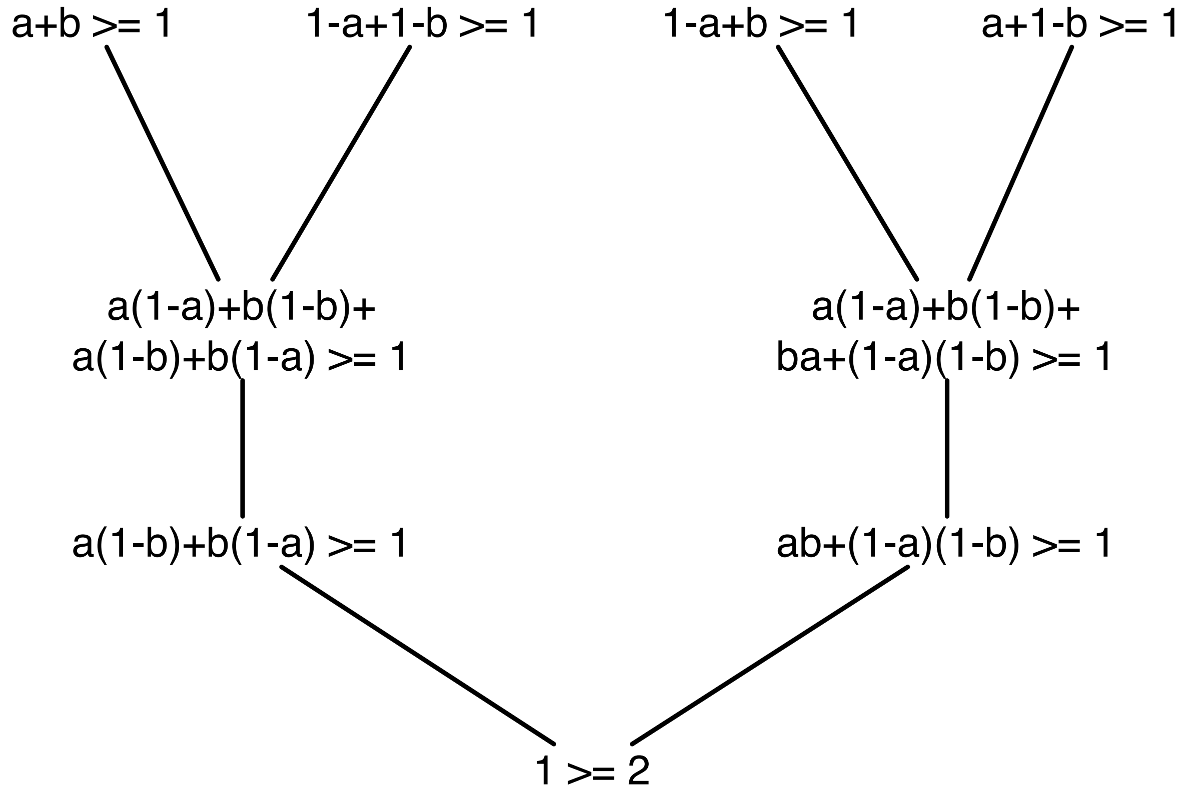
- Say trying to show a CNF formula ϕ is not satisfiable

- Refutation is a binary tree with nodes labeled by degree d polynomial inequalities and derives $0 \geq 1$.

- Axioms are clauses of ϕ, represented as inequalities.

- Derivation rule is Boolean soundness: if every 0/1 assignment which satisfies f and g also satisfies h, then one may conclude h from f, g.

Example: $(a \vee b) \wedge (\neg a \vee \neg b) \wedge (\neg a \vee b) \wedge (a \vee \neg b)$



Application to proof systems

- Via [BPS06] and our results on disjointness, we obtain subexponential lower bounds on the size of tree-like degree d semantically entailed proofs needed to refute certain CNFs for any constant d.

- Examples: cutting planes ($d = 1$), Lovász-Schrijver systems ($d = 2$).

- Exponential bounds known for (general) cutting planes [Pud97] and tree-like Lovász-Schrijver systems [KI06], but rely heavily on specific properties of these proof systems. Even for $d = 2$ no nontrivial bounds were known on semantically entailed proof systems.

Discrepancy method: two-party

- Recall for two players, letting $A[x, y] = (-1)^{f(x,y)}$.

$$\mathrm{disc}_P(A) = \max_{\substack{x \in \{0,1\}^{|X|} \\ y \in \{0,1\}^{|Y|}}} |x^T (A \circ P) y|$$

Discrepancy method: two-party

- Recall for two players, letting $A[x,y] = (-1)^{f(x,y)}$.

$$\begin{aligned}\mathrm{disc}_P(A) &= \max_{\substack{x \in \{0,1\}^{|X|} \\ y \in \{0,1\}^{|Y|}}} |x^T (A \circ P) y| \\ &= \max_C |\langle A \circ P, C \rangle|\end{aligned}$$

where C is a combinatorial rectangle.

Cylinder intersections

- Analog of combinatorial rectangle in multiparty case is a cylinder intersection

- Action of player i does not depend on x_i. Described by a function $\phi^i(x_1, \ldots, x_k)$ invariant under setting of x_i.

- Cylinder intersection $C = \phi^1(x_1, \ldots, x_k) \cdots \phi^k(x_1, \ldots, x_k)$ where each ϕ^i is a $0/1$ valued function which does not depend on x_i.

- A successful c-bit NOF protocol decomposes communication tensor into 2^c many monochromatic cylinder intersections.

Discrepancy method: multi-party

- In the multiparty case, $A[x_1, \ldots, x_k] = (-1)^{f(x_1,\ldots,x_k)}$ becomes communication tensor

$$\mathrm{disc}_P(A) = \max_{\substack{C \\ \text{cylinder intersection}}} |\langle A \circ P, C \rangle|$$

- Function is hard if discrepancy is small: $R_{1/3}(A) = \Omega(1/\mathrm{disc}(A))$ where $\mathrm{disc}(A) = \min_P \ \mathrm{disc}_P(A)$.

Rewriting discrepancy

$$\frac{1}{\mathrm{disc}(A)} = \max_{\substack{P \\ \ell_1(P)=1, P \geq 0}} \frac{|\langle A, A \circ P \rangle|}{\mathrm{disc}_P(A)}$$

Rewriting discrepancy

$$\frac{1}{\mathrm{disc}(A)} = \max_{\substack{P \\ \ell_1(P)=1, P \geq 0}} \frac{|\langle A, A \circ P \rangle|}{\mathrm{disc}_P(A)}$$

$$= \max_{P : P \geq 0} \frac{|\langle A, A \circ P \rangle|}{\mathrm{disc}_P(A)}$$

Rewriting discrepancy

$$
\begin{aligned}
\frac{1}{\operatorname{disc}(A)} &= \max_{\substack{P \\ \ell_1(P)=1, P\geq 0}} \frac{|\langle A, A \circ P\rangle|}{\operatorname{disc}_P(A)} \\
&= \max_{P: P\geq 0} \frac{|\langle A, A \circ P\rangle|}{\operatorname{disc}_P(A)} \\
&= \max_{Q: A\circ Q\geq 0} \frac{|\langle A, Q\rangle|}{\mu^*(Q)}
\end{aligned}
$$

where we define $\operatorname{disc}_P(A) = \mu^*(A \circ P)$.

Norm based approach

- Dropping restriction on sign of Q arrive exactly at definition of dual norm:
$$\mu(A) = \max_Q \frac{|\langle A, Q \rangle|}{\mu^*(Q)}$$

Norm based approach

- Dropping restriction on sign of Q arrive exactly at definition of dual norm:
$$\mu(A) = \max_Q \frac{|\langle A, Q \rangle|}{\mu^*(Q)}$$

- This remains a lower bound on communication complexity—If A correlates with Q and Q is hard, then A must be hard as well.

Norm based approach

- Dropping restriction on sign of Q arrive exactly at definition of dual norm:
$$\mu(A) = \max_Q \frac{|\langle A, Q\rangle|}{\mu^*(Q)}$$

- This remains a lower bound on communication complexity—If A correlates with Q and Q is hard, then A must be hard as well.

- For two parties, μ norm is equal to γ_2 norm, up to constant factors.

- Difficult part of showing lower bounds is how to choose Q.

Pattern matrix method

- Pattern matrix method of [She07, She08], and generalization to multiparty case by [Cha07], reduces high dimensional task of choosing Q to a one-dimensional task.

- Focus on a structured subtensor A of disjointness $\mathrm{OR}(x_1 \wedge \ldots \wedge x_k)$.

- Choose Q to be similarly structured subtensor of $q(x_1 \wedge \ldots \wedge x_k)$. This structure gives $\langle A, Q \rangle \sim \langle \mathrm{OR}, q \rangle$.

- Degree/Discrepancy theorem: if q has pure high degree, $\mu^*(Q)$ (or discrepancy) will be small.

Pattern matrix method

- Pattern matrix method of [She07, She08], and generalization to multiparty case by [Cha07], reduces high dimensional task of choosing Q to a one-dimensional task.

- Focus on a structured subtensor A of disjointness $\mathrm{OR}(x_1 \wedge \ldots \wedge x_k)$.

- Choose Q to be similarly structured subtensor of $q(x_1 \wedge \ldots \wedge x_k)$. This structure gives $\langle A, Q \rangle \sim \langle \mathrm{OR}, q \rangle$.

- Degree/Discrepancy theorem: if q has pure high degree, $\mu^*(Q)$ (or discrepancy) will be small. Use the original (and still only) technique of [BNS92] to upper bound multiparty discrepancy.

Conclusion

- Beame and Huynh-Ngoc recently show a bound of $n^{\Omega(1/k)}/2^{O(k)}$ on complexity of an AC^0 function. By reduction they get non-trivial bounds on disjointness for up to $(\log n)^{1/3}$ players.

- They use a stronger property of the function, going beyond just its approximate degree.

- Follow-up work [LSS08] extends γ_2 to the multiparty case to lower bound multiparty quantum communication. We show that k-party μ and γ_2 are related up to multiplicative factor 2^k and can thus transfer bounds shown here and by discrepancy method to the quantum case.