Communication complexity of sparse set disjointness and exists-equal

Mert Sağlam (U. Washington)
and
Gábor Tardos (Rényi Institute, Budapest)

Simons Institute workshop, April 21, 2015, Berkeley, CA
Model considered

- 2-party
  Alice and Bob computing $f(x, y)$

- randomized
  with joint random source

\[ \forall x, y: \Pr[\text{output} = f(x, y)] > \frac{3}{4} \]
Model considered

• 2-party
  Alice and Bob computing $f(x,y)$

• randomized
  with joint random source

∀ $x,y$: $\Pr[\text{output} = f(x,y)] > 3/4$

Goals

1. Minimize total communication: $|m_1| + |m_2| + \ldots + |m_k|$  
2. Minimize # of rounds: $k$
main examples

- Set disjointness $D_n : x, y \subseteq H; \quad x \cap y = \emptyset \quad |H| = n$

Can we meet next week?
main examples

• Set disjointness $D_n$: $x, y \subseteq H; \quad x \cap y = \emptyset \quad |H| = n$
• Equality $E_n$: $x, y \in \{0,1\}^n; \quad x = y ?$

Are our copies of Harry Potter and the Sourcerer’s Stone identical?
main examples

• Set disjointness $D_n : x,y \subseteq H; \quad x \cap y = \emptyset$ ?
• Equality $E_n : x,y \in \{0,1\}^n; \quad x = y$ ?

Well understood:
• $D_n$ requires $\Omega(n)$ bits of communication (Kalyanasundaram-Schnitger)
• 2 bits / 1 round enough for $E_n$
main examples

• Set disjointness $D_n: x, y \subseteq H; \ x \cap y = \emptyset$ ?
• Equality $E_n: x, y \in \{0,1\}^n; \ x = y$ ?

Well understood:
• $D_n$ requires $\Omega(n)$ bits of communication (Kalyanasundaram-Schnitger)
• 2 bits / 1 round enough for $E_n$

We consider variants of these problems
set disjointness

\[ D_n : x, y \subseteq H; \quad x \cap y = \emptyset \quad |H| = n \]
sparse set disjointness

$SD_{k,n}: x, y \subseteq H; \quad x \cap y = \emptyset \quad |H| = n; \quad |x|, |y| \leq k; \quad k \ll n$
sparse set disjointness

\( SD_{k,n} \): \( x, y \subseteq H; \quad x \cap y = \emptyset \quad |H| = n; \quad |x|, |y| \leq k; \quad k \ll n \)

containments:

\[
D_k \leq SD_{k,n} \leq D_n
\]

Complexity of \( SD_{k,n} \) is \( \Omega(k), \quad O(n), O(k \log n), O(k \log k) \).
sparse set disjointness

$SD_{k,n}: x, y \subseteq H; \quad x \cap y = \emptyset \ ? \quad |H| = n; \quad |x|, |y| \leq k; \quad k << n$

containments:

$$D_k \leq SD_{k,n} \leq D_n$$

Complexity of $SD_{k,n}$ is $\Omega(k)$, $O(n)$, $O(k \log n)$, $O(k \log k)$.

Thm [Håstad-Wigderson (2007)]: It is $\Theta(k)$. 
Håstad-Wigderson protocol for $SD_{k,n}$

- $O(k)$ bits (optimal)
- $O(\log k)$ rounds
- constant error proved
Håstad-Wigderson protocol for $SD_{k,n}$

- $O(k)$ bits (optimal)
- $O(\log k)$ rounds
- Constant error proved

Sağlam - T improvement

- $O(k)$ bits
- $\log^* k$ rounds
- Exponentially small error
Håstad-Wigderson protocol for $SD_{k,n}$

- $O(k)$ bits (optimal)
- $O(\log k)$ rounds
- Constant error proved
- $0/1$ output

Sağlam - T improvement

- $O(k)$ bits
- $\log^* k$ rounds
- Exponentially small error
- Outputs the actual intersection of the input sets
Håstad-Wigderson protocol for $SD_{k,n}$

• $O(k)$ bits (optimal)
• $O(\log k)$ rounds
• constant error proved
• 0/1 output

Sağlam - T improvement

• $O(k)$ bits
• $\log^* k$ rounds
• exponentially small error
• outputs the actual intersection of the input sets
• $r$-round $O(k \log^{(r)} k)$ bit protocol for $r < \log^* k$
• optimality proof for all $r$
Håstad-Wigderson protocol

$x \subseteq H$  \hspace{1cm} A. picks random $S_0$ with $x \subseteq S_0 \subseteq H$

Alice sends $S_0$

“$x \cap y \subseteq S_0$”
Håstad-Wigderson protocol

$x \subseteq H$

A. picks random $S_0$ with $x \subseteq S_0 \subseteq H$

**Alice sends $S_0$**

“$x \cap y \subseteq S_0$”

$y \subseteq H$

$x \cap y \subseteq y \cap S_0$
Håstad-Wigderson protocol

\[ x \subseteq H \]

A. picks random \( S_0 \) with \( x \subseteq S_0 \subseteq H \)

Alice sends \( S_0 \)

"\( x \cap y \subseteq S_0 \)"

\[ y \subseteq H \]

\[ x \cap y \subseteq y \cap S_0 \]

B. picks random \( S_1 \) with \( y \cap S_0 \subseteq S_1 \subseteq H \)

Bob sends \( S_1 \)

"\( x \cap y \subseteq S_1 \)"
Håstad-Wigderson protocol

\[ x \subseteq H \]

A. picks random \( S_0 \) with \( x \subseteq S_0 \subseteq H \)

\[ \text{Alice sends } S_0 \]

\[ "x \cap y \subseteq S_0" \]

\[ y \subseteq H \]

\[ x \cap y \subseteq y \cap S_0 \]

B. picks random \( S_1 \) with \( y \cap S_0 \subseteq S_1 \subseteq H \)

\[ \text{Bob sends } S_1 \]

\[ "x \cap y \subseteq S_1" \]

\[ x \cap y \subseteq x \cap S_1 \]
Håstad-Wigderson protocol

A. picks random $S_0$ with $x \subseteq S_0 \subseteq H$

Alice sends $S_0$

“$x \cap y \subseteq S_0$”

B. picks random $S_1$ with $y \cap S_0 \subseteq S_1 \subseteq H$

Bob sends $S_1$

“$x \cap y \subseteq S_1$”

A. picks random $S_2$ with $x \cap S_1 \subseteq S_2 \subseteq H$

Alice sends $S_2$
Håstad-Wigderson protocol

1. Alice sends $S_0$
   
   $$x \cap y \subseteq S_0$$

2. Bob sends $S_1$
   
   $$x \cap y \subseteq S_1$$

3. Alice sends $S_2$
   
   $$x \cap y \subseteq S_2$$

   etc.
Håstad-Wigderson protocol

Alice sends $S_0$

Bob sends $S_1$

Alice sends $S_2$

etc.

Stop and output "disjoint" if current set is empty, otherwise output "intersect" when $O(\log k)$ rounds or $O(k)$ bits are used up.
Håstad-Wigderson protocol

Alice sends $S_0$
Bob sends $S_1$
Alice sends $S_2$

\(\text{etc.}\)

Stop and output "disjoint" if current set is empty,
otherwise output "intersect" when the $O(\log k)$ rounds or $O(k)$ bits are used up.
Analysis of the Håstad-Wigderson protocol

• How to send a random set containing $x$?
  $w_1, w_2, \ldots$ random sets from joint random source
  Send **index** $\min\{j \mid w_j \supseteq x\}$.
  $E[\# \text{ of bits sent}] \approx |x|$
Analysis of the Håstad-Wigderson protocol

- How to send a random set containing $x$?
  - $w_1, w_2, …$ random sets from joint random source
  - Send index $\min\{j | w_j \supseteq x\}$.
  - $E[\text{# of bits sent}] \approx |x|$

- 1-sided error (no error for intersecting sets).
Analysis of the Håstad-Wigderson protocol

• How to send a random set containing $x$?
  
  $w_1, w_2, \ldots$ random sets from joint random source
  
  Send index $\min\{j | w_j \supseteq x\}$.
  
  $E[\text{# of bits sent}] \approx |x|$

• 1-sided error (no error for intersecting sets).

• If $x$ and $y$ are disjoint, then size of current set halves in expectation in every round.
Analysis of the Håstad-Wigderson protocol

• How to send a random set containing $x$?
  $w_1, w_2, \ldots$ random sets from joint random source
  Send index $\min\{j | w_j \supseteq x\}$.
  $E[\# \text{ of bits sent}] \approx |x|$

• 1-sided error (no error for intersecting sets).

• If $x$ and $y$ are disjoint, then size of current set halves in expectation in every round.
  Current set will be empty in $O(\log k)$ expected rounds.
  Expected total $\#$ of bits sent: $O(k)$. 

protocol with fewer rounds

- **biased** random sets: contains each element independently with probability $p << 1/2$
- small $p$: **quicker** decrease in set size: $k \rightarrow pk$
  - longer message for a random set containing $x$: $|x|\log(1/p)$
protocol with fewer rounds

- **biased** random sets: contains each element independently with probability $p \ll 1/2$
- small $p$: **quicker** decrease in set size: $k \rightarrow pk$
  longer message for a random set containing $x$: $|x|\log(1/p)$
- optimal tradeoff:
  $$1/p_{i+1} = 2^{1/p_i}$$
Protocol for finding the intersection

Problem: “current set” $x_i$ contains the true intersection $x \cap y$ — remains large throughout protocol
Sending random set containing $x_i$ is infeasible
Protocol for finding the intersection

Problem: “current set” $x_i$ contains the true intersection $x \cap y$ — remains large throughout protocol
Sending random set containing $x_i$ is infeasible

Good: $|x_i - (x \cap y)|$ is expected to decrease just as fast,
$x_i - x_{i+2} = x_i - S_{i+1}$ probably very small
Protocol for finding the intersection

**Problem:** “current set” $x_i$ contains the true intersection $x \cap y$ — remains large throughout protocol
Sending random set containing $x_i$ is infeasible

**Good:** $|x_i - (x \cap y)|$ is expected to decrease just as fast,
$x_i - x_{i+2} = x_i - S_{i+1}$ probably very small

**Solution:** no need for Bob to give full message $S_{i+1}$,
enough to send a few bits — Alice chooses $S_{i+1}$ to minimize $|x_i - S_{i+1}|$
Equality problem

Is \( x = y \)?

\( O(1) \) bits enough in single round
(with joint random source)
Exists-equal problem

Is $x_1 = y_1$ or $x_2 = y_2$ or ...

$\exists x_k = y_k$?
Exists-equal problem

\[
\begin{aligned}
x_1, x_2, \ldots, x_k, & \quad & y_1, y_2, \ldots, y_k \\
\text{Is} & \quad x_1 = y_1 \text{ or} \\
& \quad x_2 = y_2 \text{ or} \\
& \quad \ldots \\
& \quad x_k = y_k ?
\end{aligned}
\]

Special case of Sparse Set Disjointness:

\[
\{(1,x_1),(2,x_2),\ldots,(k,x_k)\} \cap \{(1,y_1),(2,y_2),\ldots,(k,y_k)\} \neq \emptyset ?
\]
Exists-equal problem

\[ x_1, \, x_2, \, \ldots, \, x_k \quad \text{and} \quad y_1, \, y_2, \, \ldots, \, y_k \]

Is \[ x_1 = y_1 \] or \[ x_2 = y_2 \] or \[ \ldots \]
\[ x_k = y_k \]?

**Special case of Sparse Set Disjointness:**

\[ \{(1,x_1),(2,x_2),\ldots,(k,x_k)\} \cap \{(1,y_1),(2,y_2),\ldots,(k,y_k)\} \neq \emptyset \]

- \( O(k) \) bits in \( \log^*k \) rounds suffice.
- Or \( O(k \log^r k) \) bits in \( r \) rounds.

This is optimal for any \( r \).
Single equality

10 bits in single round
solves with <0.1% error

OR of $k$ equalities

solve with 45% error
in single round:

need $\Omega(k \log k)$ bits
Single equality

10 bits in single round solves with <0.1% error

OR of $k$ equalities

solve with 45% error in single round:

need $\Omega(k \log k)$ bits

best is to solve each equality separately with $O(1/k)$ error
Lower bound proofs

• for Exists-equal (sparse set disjointness follows)
Lower bound proofs

• for Exists-equal (sparse set disjointness follows)
• elementary argument for single round
• round-elimination argument in general

Need: special form of isoperimetric inequality on the Hamming cube $[t]^n$

What is the most "compact" set $H \subseteq [t]^n$ with $|H|$ fixed?

Obtained through "shrinking"
Lower bound proofs

• for Exists-equal (sparse set disjointness follows)
• elementary argument for single round
• round-elimination argument in general

Need: special form of isoperimetric inequality on the Hamming cube $[t]^k$

What is the most “compact” set $H \subseteq [t]^k$ with $|H|$ fixed?
Lower bound proofs

• for Exists-equal (sparse set disjointness follows)
• elementary argument for single round
• round-elimination argument in general

Need: special form of isoperimetric inequality on the Hamming cube \([t]^k\)

What is the most “compact” set \(H \subseteq [t]^k\) with \(|H|\) fixed?

compact \(\approx \) small \(E_x[\log |B_x \cap H|] \)
Lower bound proofs

- for Exists-equal (sparse set disjointness follows)
- elementary argument for single round
- round-elimination argument in general

**Need:** special form of *isoperimetric inequality* on the Hamming cube \([t]^k\)

What is the most “compact” set \(H \subseteq [t]^k\) with \(|H|\) fixed?

**compact** \(\approx\) small \(E_x[\log|B_x \cap H|]\)

Obtained through “shifting”
Thank You