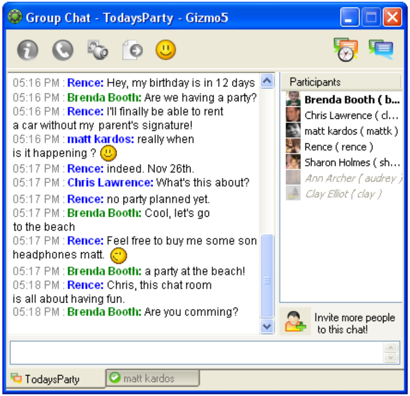
Information Complexity and the Geometry of Communication

T.S. Jayram

IBM Almaden Research Center

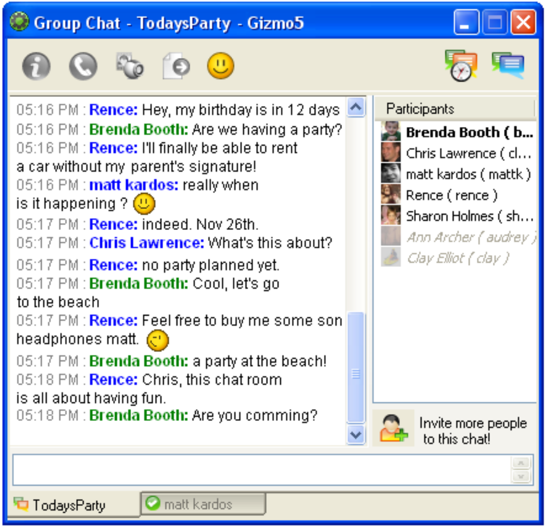
December 19, 2009

Communication



Communication

Communication



"Communication does not signify a problem newly discovered in our times, but a fashion of thinking and a method of analyzing which we apply in the statement of all fundamental problems."

Richard McKeon (1957)

Communication (CS Theory Version)

Multi-party Communication

- There are $t \geq 2$ players $P_1, P_2, \ldots, P_t$
- The input is a tuple $(x_1, x_2, \ldots, x_t)$
- P_i holds x_i (number-in-hand)
- The *protocol* specifies the rules for writing messages (on a blackboard)
- Players can use *private random coins*
- The message transcript depends on both the input and randomness

Communication Complexity

Quantify the *necessary* amount of communication needed to solve a communication problem.

Many ways to measure communication:

1. Length of communication, i.e. #bits in the transcript
2. Rounds of communication
3. Asymmetric communication

How much information about the inputs is contained in Π?

Set Disjointness

- Alice's input is a set $x \subseteq [1..n]$
- Bob's input is a set $y \subseteq [1..n]$
- Communication problem: are the sets are disjoint?

$$\mathrm{Disj}(x, y) = \bigvee_{j=1}^{n} (x_j \wedge y_j)$$

L_∞ estimation

- Inputs are $x, y \in \mathbb{R}^n$
- ℓ_∞ norm $\|a\|_\infty = \max_i |a_i|$
- Threshold t, approximation $\alpha \geq 1$
- Is $\|x - y\|_\infty$ at most t or greater than αt?

GapL_∞ and Dist

For each coordinate j:

1. $|x_j - y_j| \leq t \quad \Longrightarrow \quad \mathrm{Dist}(x_j, y_j) = 0$, or
2. $|x_j - y_j| > \alpha t \quad \Longrightarrow \quad \mathrm{Dist}(x_j, y_j) = 1$

$$\mathrm{GapL}_\infty(x, y) = \bigvee_{j=1}^{n} \mathrm{Dist}(x_j, y_j)$$

Direct Sum

- Given two decision problems f and g:

$$f(x,y) = \bigvee_{j=1}^{n} g(x_j, y_j)$$

- Can we relate the communication complexity of f to that of g?

Direct Sum

Is $CC(f) \geq n \cdot CC(g)$?

- Difficult to prove by directly analyzing the transcript
- Maybe information theory can help ...

Information Theory Notions

Conditional Entropy: $\mathrm{H}(X \mid Y)$ is the *uncertainty* in X *conditioned* on Y

Mutual Information: $\mathrm{I}(X : Y) = \mathrm{H}(X) - \mathrm{H}(X \mid Y)$ is the *reduction* in uncertainty of X when conditioned on Y

Conditional Mutual Information: $\mathrm{I}(X : Y \mid Z) = \mathrm{H}(X \mid Z) - \mathrm{H}(X \mid Y, Z)$

Information Cost/Complexity

Let Π be the transcript of a t-player protocol on input $x_1, x_2, \ldots, x_t$.

Definition

The information cost of a protocol equals

$$\mathrm{I}(X_1, X_2, \ldots, X_t : \Pi \mid D),$$

where $X_1, X_2, \ldots, X_t$ are *jointly independent conditioned on* D.

Information Complexity $\mathrm{IC}(f)$: the *minimum* information cost of a protocol that *correctly* computes f

[Chakrabarti, Shi, Wirth and Yao], [Bar-Yossef, Jayram, Kumar and Sivakumar]

Direct Sum Theorem

Theorem

Suppose f can be written as an OR of n disjoint instances of g. Then,

$$\mathrm{IC}(f) \geq n \cdot \mathrm{IC}(g).$$

- Since $CC(f) \geq \mathrm{IC}(f)$, we can prove lower bounds on $CC(f)$ by proving lower bounds on $\mathrm{IC}(g)$.
- However, the support of the input distribution of g is on the 0-instances of g!

Leads to interesting connections between information complexity and the geometry of communication.

Part II

Distance Estimation

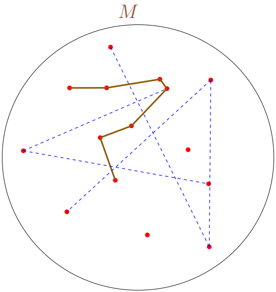
Distance Estimation

A metric space M with distance function $d(\cdot,\cdot)$

- Alice holds $x \in M$; Bob holds $y \in M$
- Threshold t; approximation $\alpha \geq 1$
- Promise problem Dist:

  No instance: $d(x,y) \leq t$ (Close)

  Yes instance: $d(x,y) > \alpha t$ (Far)



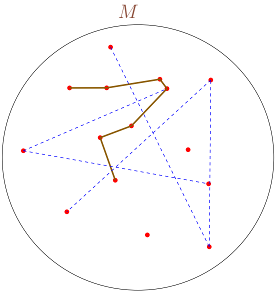
Distance Estimation

A metric space M with distance function $d(\cdot,\cdot)$

- Alice holds $x \in M$; Bob holds $y \in M$
- Threshold t; approximation $\alpha \geq 1$
- Promise problem Dist:

  No instance: $d(x,y) \leq t$ (Close)

  Yes instance: $d(x,y) > \alpha t$ (Far)



Goal

Prove a lower bound on the information complexity of Dist

Metric Embeddings

Notation: Let $\hat{\|}\cdot\hat{\|} = \frac{1}{2}\|\cdot\|^2$.

Let ρ be an embedding of M into Euclidean space ℓ_2

- Suppose $d(x,y) = \hat{\|}\rho(x) - \rho(y)\hat{\|}$, for all $x, y \in M$ (Isometry)
- Then, ρ separates the close and far instances of Dist
- How does one show that such an embedding does not exist?

Poincaré-Type Inequalities

Definition (Poincaré Inequality)

1. distribution η_0 on close instances
2. distribution η_1 on far instances
3. parameter λ

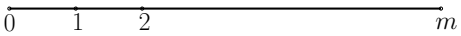
A Poincaré inequality holds for Dist if for all $\rho : M \to \ell_2$:

$$\mathbb{E}_{(x,y)\sim\eta_0} \| \rho(x) - \rho(y) \| \geq \lambda \cdot \mathbb{E}_{(x,y)\sim\eta_1} \| \rho(x) - \rho(y) \|$$

1. $\lambda \geq 1/\alpha \implies$ no isometric embeddings exist
2. with larger λ, can also consider embeddings with distortion
3. with smaller λ, we can still say something interesting!

Examples of Poincaré Inequalities

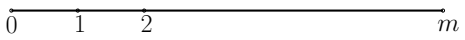
1. Let η_0 = unif. distrib. on $\{0, 1, \ldots, m-1\}$.



$$\mathbb{E}_{x \sim \eta_0} \|\rho(x) - \rho(x+1)\| \geq \frac{1}{m} \cdot \|\rho(0) - \rho(m)\|$$

Examples of Poincaré Inequalities

1. Let η_0 = unif. distrib. on $\{0, 1, \ldots, m-1\}$.

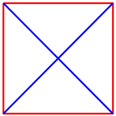


$$\mathbb{E}_{x \sim \eta_0} \|\rho(x) - \rho(x+1)\| \geq \frac{1}{m} \cdot \|\rho(0) - \rho(m)\|$$

2. Hamming cube $\{0,1\}^d$:

- η_0 / η_1 = unif. distrib. on the edges/diagonals

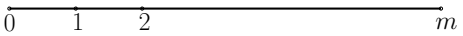
$$\mathbb{E}_{(x,y) \sim \eta_0} \|\rho(x) - \rho(y)\| \geq \frac{1}{d} \cdot \mathbb{E}_{(x,y) \sim \eta_1} \|\rho(x) - \rho(y)\|$$



- Implies that the Hamming cube does not embed into ℓ_1

Examples of Poincaré Inequalities

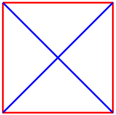
1. Let $\eta_0 =$ unif. distrib. on $\{0, 1, \ldots, m-1\}$.



$$\mathbb{E}_{x\sim\eta_0}\|\rho(x)-\rho(x+1)\| \geq \frac{1}{m}\cdot\|\rho(0)-\rho(m)\|$$

2. Hamming cube $\{0,1\}^d$:

- $\eta_0/\eta_1 =$ unif. distrib. on the edges/diagonals



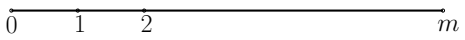
$$\mathbb{E}_{(x,y)\sim\eta_0}\|\rho(x)-\rho(y)\| \geq \frac{1}{d}\cdot\mathbb{E}_{(x,y)\sim\eta_1}\|\rho(x)-\rho(y)\|$$

- Implies that the Hamming cube does not embed into ℓ_1

3. [Andoni,Krauthgamer] proved a Poincaré inequality for edit distance and used it to prove communication vs. approximation tradeoff.

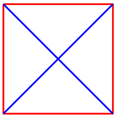
Examples of Poincaré Inequalities

1. Let $\eta_0 =$ unif. distrib. on $\{0, 1, \ldots, m-1\}$.



$$\mathbb{E}_{x \sim \eta_0} \|\rho(x) - \rho(x+1)\| \geq \frac{1}{m} \cdot \|\rho(0) - \rho(m)\|$$

2. Hamming cube $\{0,1\}^d$:



- $\eta_0/\eta_1 =$ unif. distrib. on the edges/diagonals

$$\mathbb{E}_{(x,y) \sim \eta_0} \|\rho(x) - \rho(y)\| \geq \frac{1}{d} \cdot \mathbb{E}_{(x,y) \sim \eta_1} \|\rho(x) - \rho(y)\|$$

- Implies that the Hamming cube does not embed into ℓ_1

3. [Andoni,Krauthgamer] proved a Poincaré inequality for edit distance and used it to prove communication vs. approximation tradeoff.

4. Expanders: Poincaré inequality with $\lambda =$ normalized spectral gap.

Information Complexity of Dist

Theorem ([Andoni, Jayram and Patrascu])

Suppose Dist *satisfies a Poincaré inequality w.r.t. distributions* η_0, η_1, *and parameter* λ. *Then,*

$$\mathrm{IC}(\mathrm{Dist}) \geq \frac{\lambda}{8}$$

History

- [Bar-Yossef, Jayram, Kumar and Sivakumar] proved this result for $M = \mathbb{R}$.
  - Yields space lower bounds for estimating ℓ_∞ in a data stream

History

- [Bar-Yossef, Jayram, Kumar and Sivakumar] proved this result for $M = \mathbb{R}$.
  - Yields space lower bounds for estimating ℓ_∞ in a data stream
- [Jayram and Woodruff] proved this result for the Hamming cube
  - Yields space lower bounds estimating cascaded norms $L_k \circ L_0$

History

- [Bar-Yossef, Jayram, Kumar and Sivakumar] proved this result for $M = \mathbb{R}$.
  - Yields space lower bounds for estimating ℓ_∞ in a data stream
- [Jayram and Woodruff] proved this result for the Hamming cube
  - Yields space lower bounds estimating cascaded norms $L_k \circ L_0$
- Our result implies improved lower bounds for sketching edit distance: $\Omega(\log\log d) \Longrightarrow \Omega(\log d)$.

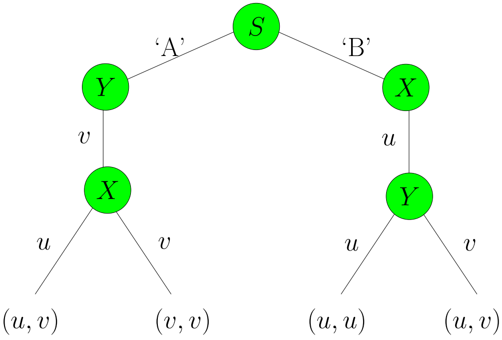
Input Distribution

1. Define (X, Y, D) such that X and Y are independent conditioned on D
2. $\mathrm{Dist}(X, Y) = 0$

Input Distribution

1. Define (X, Y, D) such that X and Y are independent conditioned on D
2. $\mathrm{Dist}(X, Y) = 0$

$$D = (S, T) \qquad T = (u, v) \sim \eta_0$$



Transcript Wave Function

Let $\pi(x, y) =$ prob. distrib. over transcripts on a fixed input (x, y)

- $\pi(x, y)_\tau =$ prob. that the transcript equals τ
- $\pi(x, y)$ belongs to the unit simplex

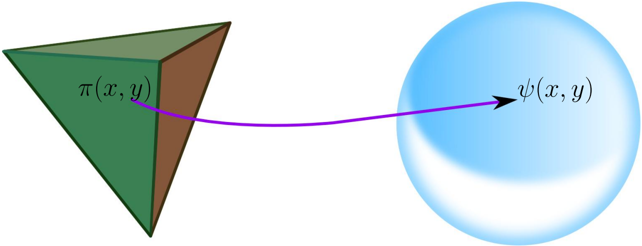
Transcript Wave Function

Let $\pi(x,y)$ = prob. distrib. over transcripts on a fixed input (x,y)

- $\pi(x,y)_\tau$ = prob. that the transcript equals τ
- $\pi(x,y)$ belongs to the unit simplex

Definition (Transcript wave function $\psi(x,y)$)

- $\psi(x,y)_\tau = \sqrt{\pi(x,y)_\tau}\ \forall\tau$.
- $\psi(x,y)$ belongs to the *positive orthant* of the *unit sphere*



Hellinger Distance

Definition (Hellinger Distance)

The square of the Hellinger distance between two transcript wave functions $\psi(x, y)$ and $\psi(x', y')$ is defined as $\hat{\|}\psi(x, y) - \psi(x', y')\hat{\|}$.

Hellinger Distance

Definition (Hellinger Distance)

The square of the Hellinger distance between two transcript wave functions $\psi(x,y)$ and $\psi(x',y')$ is defined as $\|\psi(x,y)-\psi(x',y')\|$.

Proposition (Information Cost to Hellinger distance)

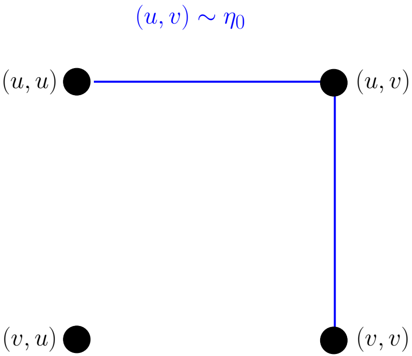
$$\mathrm{I}(X,Y:\Pi \mid D) \geq \frac{1}{2}\cdot \mathbb{E}_{(u,v)\sim\eta_0} \|\psi(u,u)-\psi(u,v)\| + \|\psi(u,v)-\psi(v,v)\|$$

Proof

$\mathrm{I}(X, Y : \Pi \mid S, T)$

Proof

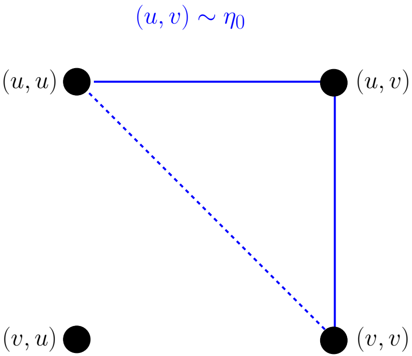
$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \frac{1}{2} \cdot \mathbb{E}_{\eta_0} \hat{\|} \psi(u,u) - \psi(u,v) \hat{\|} \\
&\qquad\qquad + \hat{\|} \psi(u,v) - \psi(v,v) \hat{\|}
\end{aligned}
$$



(Information Cost to Hellinger distance)

Proof

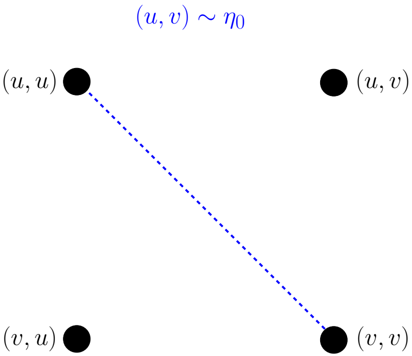
$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \frac{1}{2} \cdot \mathbb{E}_{\eta_0} \|\psi(u,u) - \psi(u,v)\|^2 \\
&\qquad\qquad + \|\psi(u,v) - \psi(v,v)\|^2 \\
&\quad \geq \frac{1}{4} \cdot \mathbb{E}_{\eta_0} \|\psi(u,u) - \psi(v,v)\|^2
\end{aligned}
$$



(Cauchy-Schwartz + Triangle Inequality)

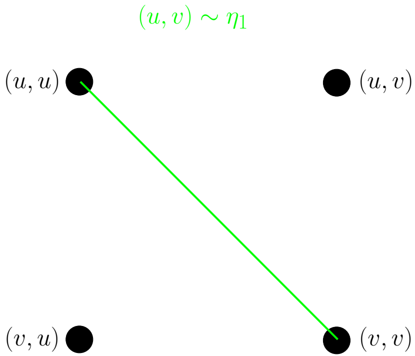
Proof

$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \tfrac{1}{2} \cdot \mathbb{E}_{\eta_0} \hat{\|}\psi(u,u) - \psi(u,v)\hat{\|} \\
&\qquad\qquad + \hat{\|}\psi(u,v) - \psi(v,v)\hat{\|} \\
&\quad \geq \tfrac{1}{4} \cdot \mathbb{E}_{\eta_0} \hat{\|}\psi(u,u) - \psi(v,v)\hat{\|}
\end{aligned}
$$



Proof

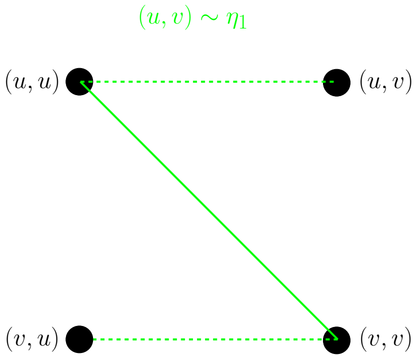
$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \frac{1}{2} \cdot \mathbb{E}_{\eta_0} \left\| \psi(u,u) - \psi(u,v) \right\| \\
&\qquad\qquad + \left\| \psi(u,v) - \psi(v,v) \right\| \\
&\quad \geq \frac{1}{4} \cdot \mathbb{E}_{\eta_0} \left\| \psi(u,u) - \psi(v,v) \right\| \\
&\quad \geq \frac{\lambda}{4} \cdot \mathbb{E}_{\eta_1} \left\| \psi(u,u) - \psi(v,v) \right\|
\end{aligned}
$$



(Poincaré inequality: $\rho(u) = \psi(u,u)$)

Proof

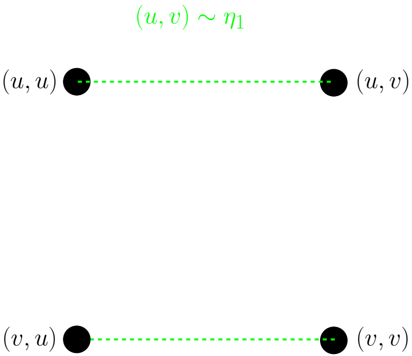
$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \frac{1}{2} \cdot \mathbb{E}_{\eta_0} \|\psi(u,u) - \psi(u,v)\| \\
&\qquad\qquad + \|\psi(u,v) - \psi(v,v)\| \\
&\quad \geq \frac{1}{4} \cdot \mathbb{E}_{\eta_0} \|\psi(u,u) - \psi(v,v)\| \\
&\quad \geq \frac{\lambda}{4} \cdot \mathbb{E}_{\eta_1} \|\psi(u,u) - \psi(v,v)\| \\
&\quad \geq \frac{\lambda}{8} \cdot \mathbb{E}_{\eta_1} (\|\psi(u,u) - \psi(u,v)\| \\
&\qquad\qquad + \|\psi(v,u) - \psi(v,v)\|)
\end{aligned}
$$



(Pythogorean Property of Communication Protocols [BJKS])

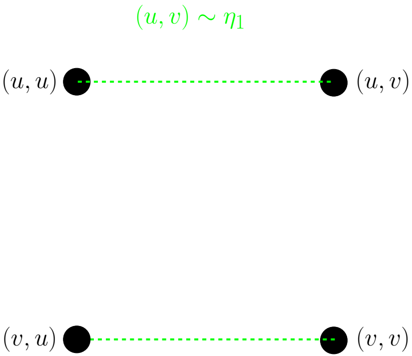
Proof

$$
\begin{aligned}
&\mathrm{I}(X, Y : \Pi \mid S, T) \\
&\quad \geq \frac{1}{2} \cdot \mathbb{E}_{\eta_0} \hat{\|} \psi(u,u) - \psi(u,v) \hat{\|} \\
&\qquad\qquad + \hat{\|} \psi(u,v) - \psi(v,v) \hat{\|} \\
&\quad \geq \frac{1}{4} \cdot \mathbb{E}_{\eta_0} \hat{\|} \psi(u,u) - \psi(v,v) \hat{\|} \\
&\quad \geq \frac{\lambda}{4} \cdot \mathbb{E}_{\eta_1} \hat{\|} \psi(u,u) - \psi(v,v) \hat{\|} \\
&\quad \geq \frac{\lambda}{8} \cdot \mathbb{E}_{\eta_1} (\hat{\|} \psi(u,u) - \psi(u,v) \hat{\|} \\
&\qquad\qquad + \hat{\|} \psi(v,u) - \psi(v,v) \hat{\|})
\end{aligned}
$$



Proof

$$
\begin{aligned}
\mathrm{I}(X, Y : \Pi \mid S, T) &\\
\geq \tfrac{1}{2} \cdot \mathbb{E}_{\eta_0} &\|\psi(u,u) - \psi(u,v)\| \\
&\quad + \|\psi(u,v) - \psi(v,v)\| \\
\geq \tfrac{1}{4} \cdot \mathbb{E}_{\eta_0} &\|\psi(u,u) - \psi(v,v)\| \\
\geq \tfrac{\lambda}{4} \cdot \mathbb{E}_{\eta_1} &\|\psi(u,u) - \psi(v,v)\| \\
\geq \tfrac{\lambda}{8} \cdot \mathbb{E}_{\eta_1} &(\|\psi(u,u) - \psi(u,v)\| \\
&\quad + \|\psi(v,u) - \psi(v,v)\|) \\
\geq \tfrac{\lambda}{8} &
\end{aligned}
$$



(Soundness)

Part III

The And problem

The t-party And problem

- Player P_i holds a *single* bit x_i, $i = 1..t$
- Promise:

  Yes instance: *every* player's bit is set to 1

  No instance: *at most one* player's bit is set to 1

History

- Introduced in [Bar-Yossef, Jayram, Kumar and Sivakumar]
  - For proving multi-party set disjointness [Alon, Matias, and Szegedy] lower bounds via a direct-sum theorem
  - Yields space lower bounds for frequency moments in data streams

History

- Introduced in [Bar-Yossef, Jayram, Kumar and Sivakumar]
  - For proving multi-party set disjointness [Alon, Matias, and Szegedy] lower bounds via a direct-sum theorem
  - Yields space lower bounds for frequency moments in data streams
- [BJKS] proved a lower bound of $\Omega(1/t^2)$
  - Improved to $\Omega(1/t^{1+\varepsilon})$, for any $\varepsilon > 0$, for one round communication

History

- Introduced in [Bar-Yossef, Jayram, Kumar and Sivakumar]
  - For proving multi-party set disjointness [Alon, Matias, and Szegedy] lower bounds via a direct-sum theorem
  - Yields space lower bounds for frequency moments in data streams
- [BJKS] proved a lower bound of $\Omega(1/t^2)$
  - Improved to $\Omega(1/t^{1+\varepsilon})$, for any $\varepsilon > 0$, for one round communication
- [Chakrabarti, Khot, and Sun] gave a $\Omega(1/(t \log t))$ bound for general communication and a $\Omega(1/t)$ bound for 1-round communication

Result

Theorem ([Gronemeier])

The information complexity of And *for general communication is* $\Omega(1/t)$.

Input Distribution

Conditioning: Choose index D uniformly from $[1..t]$.

Input:
1. Choose X_D uniformly from $\{0, 1\}$
2. Set $X_j = 0$ for all $j \neq D$

Input Distribution

Conditioning: Choose index D uniformly from $[1..t]$.

Input:
1. Choose X_D uniformly from $\{0,1\}$
2. Set $X_j = 0$ for all $j \neq D$

1. $X_1, X_2, \ldots, X_t$ are jointly independent conditioned on D
2. $X_1 \wedge X_2 \wedge \cdots \wedge X_t = 0$

Proof

Information Cost to Hellinger distance:

$$\mathrm{I}(X_1, X_2, \ldots, X_t : \Pi \mid D) \geq \frac{1}{t} \sum_{i=1}^{t} \left\| \psi(\emptyset) - \psi(\{i\}) \right\|$$

Note: bit-vectors of length t $\equiv$ subsets of $[1..t]$

Lemma

Let $t = 2^k$.

$$\sum_{i=1}^{t} \left\| \psi(\emptyset) - \psi(\{i\}) \right\| \geq \left\| \psi(\emptyset) - \psi([1..t]) \right\| \cdot \prod_{\ell=1}^{k} \left(1 - \frac{1}{2^\ell}\right),$$

Proof

Information Cost to Hellinger distance:

$$\mathrm{I}(X_1, X_2, \ldots, X_t : \Pi \mid D) \geq \frac{1}{t} \sum_{i=1}^{t} \hat{\|}\psi(\emptyset) - \psi(\{i\})\hat{\|}$$

Note: bit-vectors of length $t \equiv$ subsets of $[1..t]$

Lemma

Let $t = 2^k$.

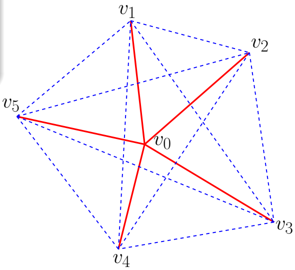
$$\sum_{i=1}^{t} \hat{\|}\psi(\emptyset) - \psi(\{i\})\hat{\|} \geq \hat{\|}\psi(\emptyset) - \psi([1..t])\hat{\|} \cdot \prod_{\ell=1}^{k} \left(1 - \frac{1}{2^\ell}\right),$$

1. $\hat{\|}\psi(\emptyset) - \psi([1..t])\hat{\|} = \Omega(1)$ (soundness)
2. $\prod_{\ell=1}^{k}\left(1 - \frac{1}{2^\ell}\right) = 0.288788\ldots$ as $k \to \infty$ (digital search tree constant)

Star Property—a Negative-Type Inequality

Proposition

$$\sum_{i=1}^{s} \| v_0 - v_i \| \geq \frac{1}{s} \sum_{1 \leq i < j \leq s} \| v_i - v_j \|$$



Star Property—a Negative-Type Inequality

Proposition

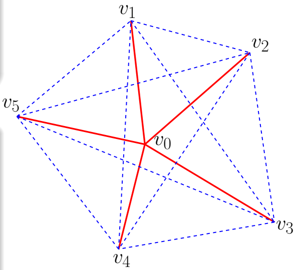
$$\sum_{i=1}^{s} \|v_0 - v_i\| \geq \frac{1}{s} \sum_{1 \leq i < j \leq s} \|v_i - v_j\|$$

Proof.

Rewrite inequality as

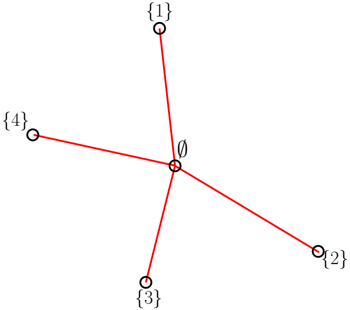
$$\Big\| \sum_{0 \leq i \leq s} b_i v_i \Big\| \geq 0,$$

where $b_0 = s$ and $b_1 = b_2 = \cdots = b_s = -1$. $\square$



Proof of Technical Lemma

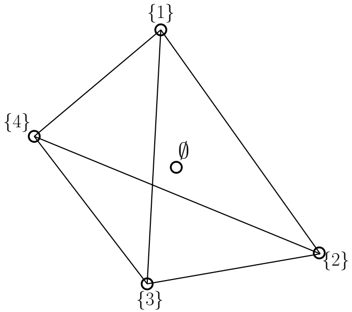
$$\sum_{i=1}^{2^k} \|\psi(\emptyset) - \psi(\{i\})\|$$



Proof of Technical Lemma

$$\sum_{i=1}^{2^k} \|\psi(\emptyset) - \psi(\{i\})\|$$

$$\geq \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \|\psi(\{i\}) - \psi(\{j\})\|$$

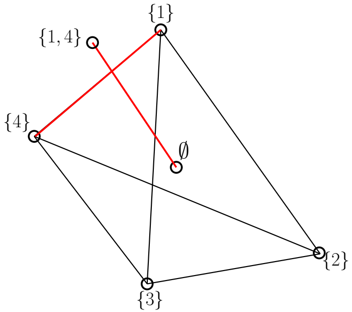


(Star property)

Proof of Technical Lemma

$$\sum_{i=1}^{2^k} \|\psi(\emptyset) - \psi(\{i\})\|$$

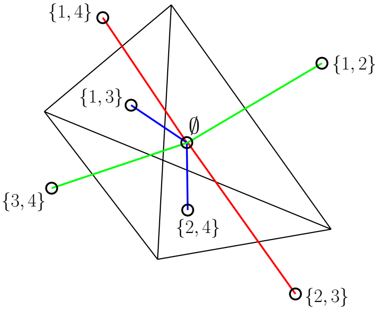
$$\geq \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \|\psi(\{i\}) - \psi(\{j\})\|$$



(Cut-and-paste [BJKS]: $\|\psi(A) - \psi(B)\| = \|\psi(A \cap B) - \psi(A \cup B)\|$)

Proof of Technical Lemma

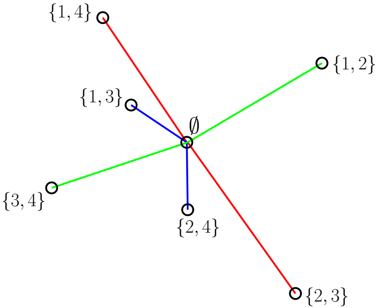
$$
\begin{aligned}
&\sum_{i=1}^{2^k} \|\psi(\emptyset) - \psi(\{i\})\| \\
&\geq \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \|\psi(\{i\}) - \psi(\{j\})\| \\
&= \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \|\psi(\emptyset) - \psi(\{i, j\}\|
\end{aligned}
$$



(Cut-and-paste [BJKS]: $\|\psi(A) - \psi(B)\| = \|\psi(A \cap B) - \psi(A \cup B)\|$)

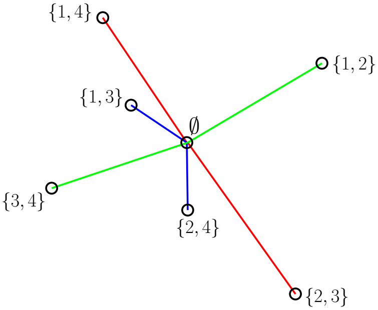
Proof of Technical Lemma

$$\sum_{i=1}^{2^k} \hat{\|}\psi(\emptyset) - \psi(\{i\})\hat{\|}$$
$$\geq \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \hat{\|}\psi(\{i\}) - \psi(\{j\})\hat{\|}$$
$$= \frac{1}{2^k} \sum_{1 \leq i < j \leq 2^k} \hat{\|}\psi(\emptyset) - \psi(\{i, j\})\hat{\|}$$



Proof of Technical Lemma

$$
\begin{aligned}
&\sum_{i=1}^{2^k} \hat{\|}\psi(\emptyset)-\psi(\{i\})\hat{\|} \\
&\geq \frac{1}{2^k} \sum_{1\leq i<j\leq 2^k} \hat{\|}\psi(\{i\})-\psi(\{j\})\hat{\|} \\
&= \frac{1}{2^k} \sum_{1\leq i<j\leq 2^k} \hat{\|}\psi(\emptyset)-\psi(\{i,j\})\hat{\|} \\
&\quad\vdots \\
&\geq \hat{\|}\psi(\emptyset)-\psi([1..t])\hat{\|}\cdot \prod_{\ell=1}^{k}\left(1-\frac{1}{2^\ell}\right)
\end{aligned}
$$



(Induction)

Part IV

Future Directions

Conclusion

- The techniques exploit the geometry of Hellinger distance and communication protocols
- Applications:
  - Lower bounds for data streams [Jayram and Woodruff]
  - Sketching edit distance [Andoni, Jayram and Patrascu]
  - Communication complexity of functions in AC^0 (aka And-Or trees) [Jayram,Kopparty and Raghavendra]
- What about number-on-forehead?

Conclusion

- The techniques exploit the geometry of Hellinger distance and communication protocols
- Applications:
  - Lower bounds for data streams [Jayram and Woodruff]
  - Sketching edit distance [Andoni, Jayram and Patrascu]
  - Communication complexity of functions in AC^0 (aka And-Or trees) [Jayram,Kopparty and Raghavendra]
- What about number-on-forehead?
  - Consider multi-party set-disjointness
  - Best bounds of the form $\Omega(n^{1/k}/f(k))$ [Lee and Shraibman; Chattopadhyay and Ada; Beame and Huynh-Ngoc]
  - Can one hope to prove a $\Omega(n/f(k))$ bound?

A Possible Approach

- Define information cost as $\sum_i \mathrm{I}(X_i : \Pi \mid X_{-i}, R_{-i}, D)$
- Direct sum works with this definition!
- Big unknown: the information complexity of And?

Questions/Comments?