Lower Bound Techniques for Multiparty Communication Complexity

Qin Zhang

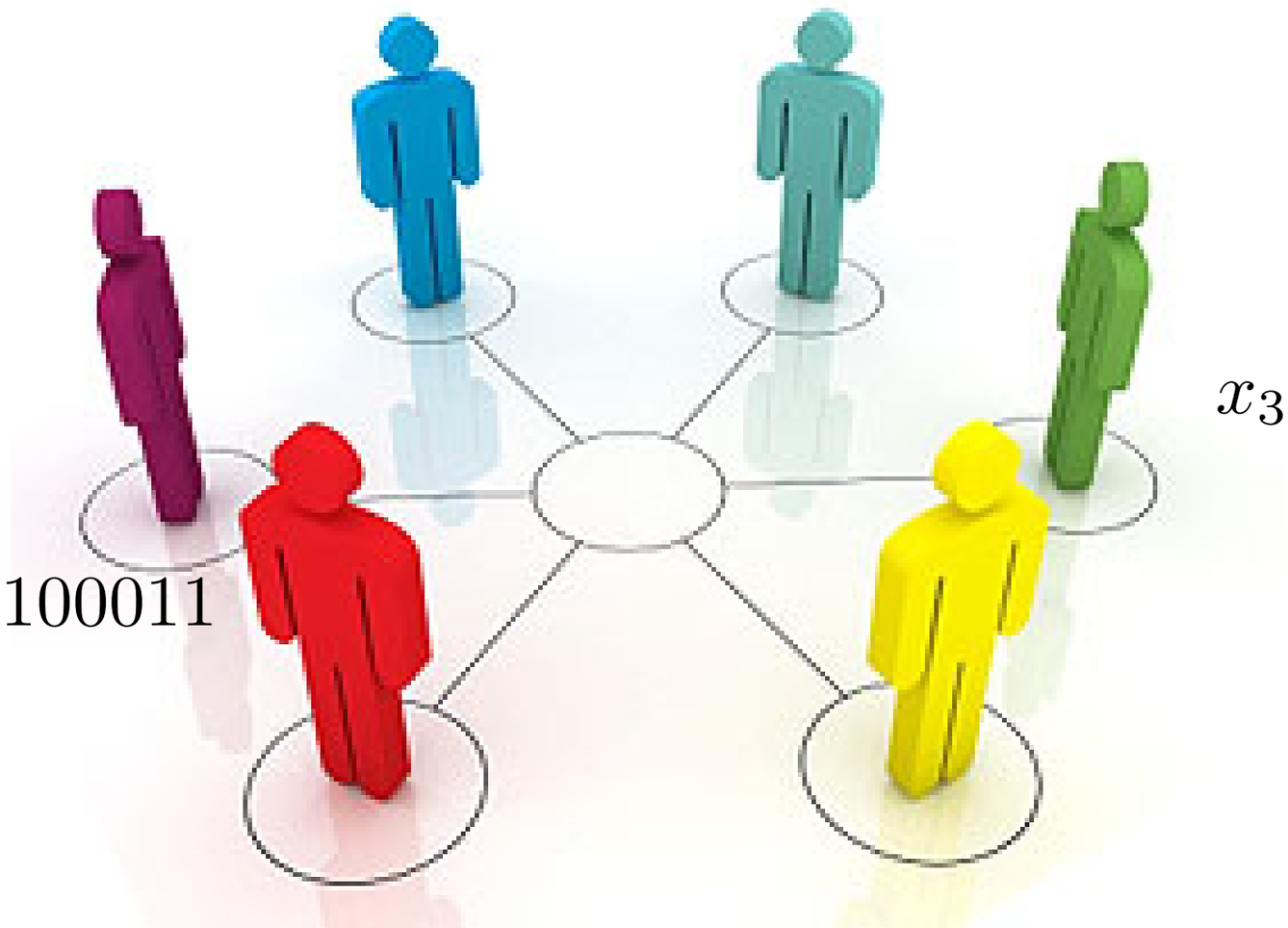
Indiana University Bloomington

Based on works with Jeff Phillips, Elad Verbin and David Woodruff

The multiparty number-in-hand (NIH) comm.

– A model for proving lower bounds

$x_1 = 010011$ $x_2 = 111011$



$x_3 = 111111$

$x_k = 100011$

They want to jointly compute $f(x_1, x_2, \ldots, x_k)$

Goal: minimize total bits of communication

The multiparty number-in-hand (NIH) comm.

– A model for proving lower bounds

$x_1 = 010011$ $x_2 = 111011$

$x_3 = 111111$

$x_k = 100011$

Blackboard (BB): One speaks, everyone hears.

Message-Passing (MP): If x_1 talks to x_2, others cannot hear.

They want to jointly compute $f(x_1, x_2, \ldots, x_k)$

Goal: minimize total bits of communication

The multiparty number-in-hand (NIH) comm.

– A model for proving lower bounds

$x_1 = 010011$ $x_2 = 111011$

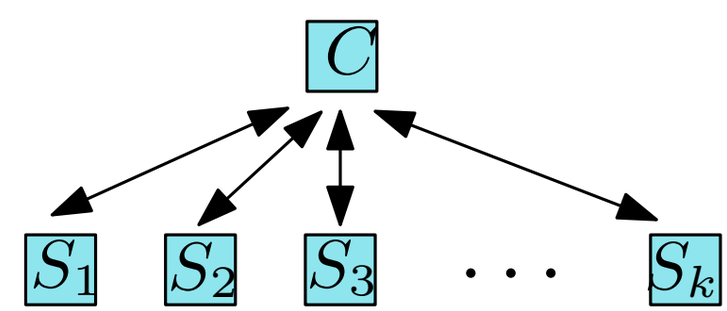
$x_3 = 111111$

$x_k = 100011$

Blackboard (BB): One speaks, everyone hears.

Message-Passing (MP): If x_1 talks to x_2, others cannot hear.

Coordinator



They want to jointly compute $f(x_1, x_2, \ldots, x_k)$

Goal: minimize total bits of communication

Previously, for multiparty NIH comm. model

Well studied? YES and NO.

Previously, for multiparty NIH comm. model

Well studied? YES and NO.

Quite a few papers in BB model:

[Alon et al. '96], [Bar-Yossef et al. '04], ...

Previously, for multiparty NIH comm. model

Well studied? YES and NO.

Quite a few papers in BB model:

[Alon et al. '96], [Bar-Yossef et al. '04], ...

But, Almost nothing in MP model before 2012.

Previously, for multiparty NIH comm. model

Well studied? YES and NO.

Quite a few papers in BB model:

[Alon et al. '96], [Bar-Yossef et al. '04], ...

But, Almost nothing in MP model before 2012.

- **Too easy?**

Previously, for multiparty NIH comm. model

Well studied? YES and NO.

Quite a few papers in BB model:

[Alon et al. '96], [Bar-Yossef et al. '04], ...

But, Almost nothing in MP model before 2012.

- **Too easy?**
- **Lack of motivations?**

Too easy?

- Could be more difficult than BB, since in MP each player observes only a piece of protocol transcript.

Too easy?

- Could be more difficult than BB, since in MP each player observes only a piece of protocol transcript.

- We do not want to analyze a k-dimensional matrix directly.

B \ A	0	1
0	0	0
1	0	1

A communication matrix of the 2-party AND function

Too easy?

- Could be more difficult than BB, since in MP each player observes only a piece of protocol transcript.

- We do not want to analyze a k-dimensional matrix directly.

B \ A	0	1
0	0	0
1	0	1

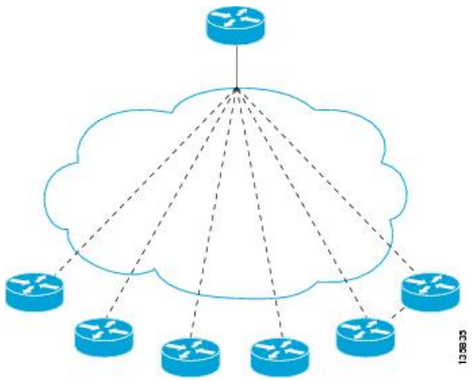
A communication matrix of the 2-party AND function

Even for the 2-party Gap-Hamming problem, where Alice has a n-bit vector and Bob has a n-bit vector, and they want to compute if the hamming distance of these two vectors is bigger than $n/2 + \sqrt{n}$ or less than $n/2 - \sqrt{n}$, it tooks researchers in this area about 10 years and dozens of papers (IW03, Woodruff04, ..., Sherstov12) to fully understand its 2-dim communication matrix!

Lack of motivations?

- **(Continous) Distributed Monitoring**
  communication $\rightarrow$ energy, bandwidth, ...

Sensor networks

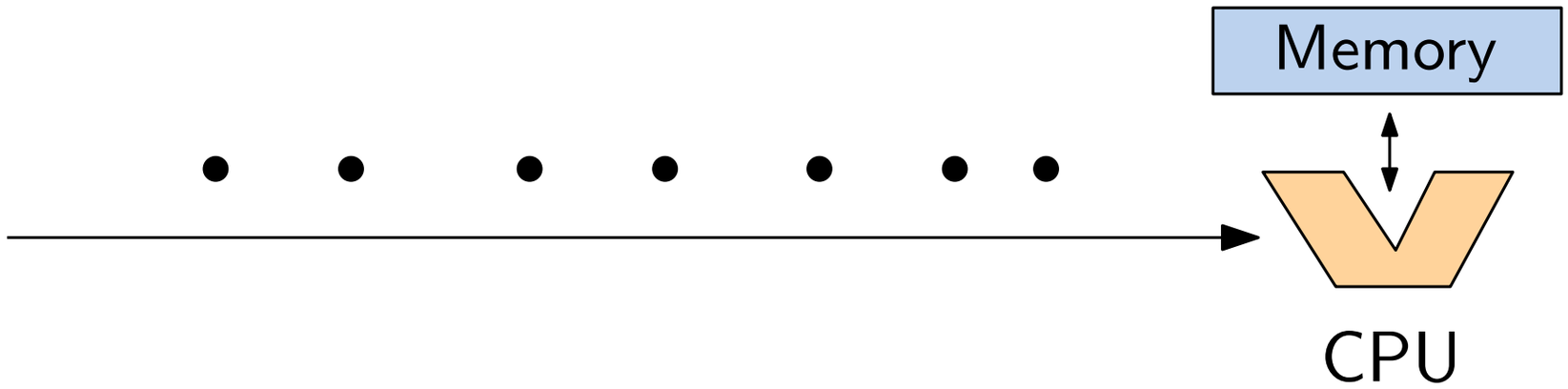


Network routers

Data in the cloud

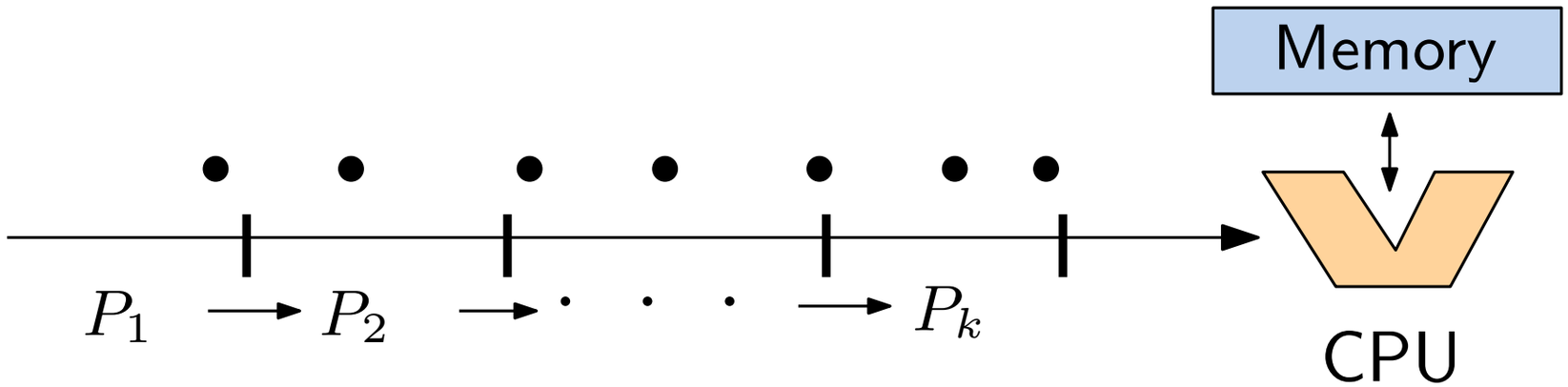
Lack of motivations? (cont.)

- **Data Streams**
  communication $\rightarrow$ space



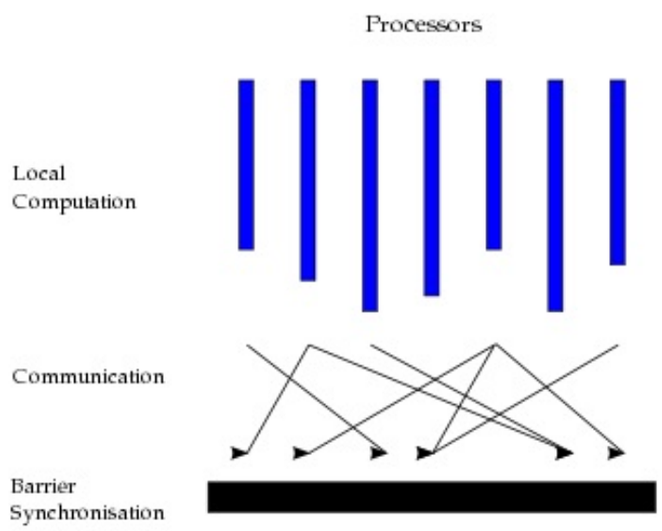
Lack of motivations? (cont.)

- **Data Streams**
  communication $\rightarrow$ space

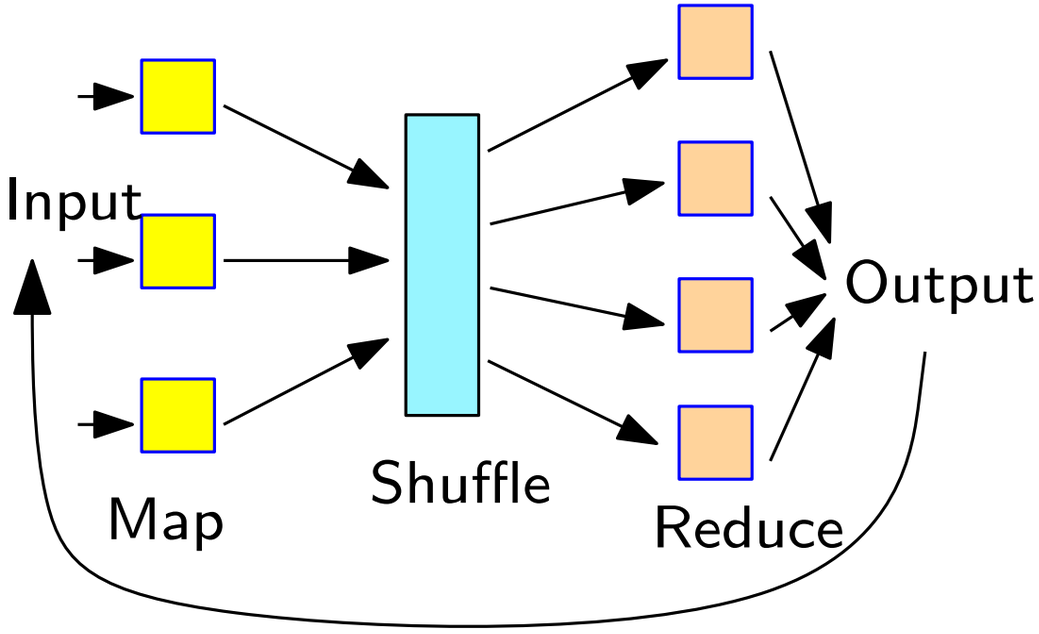


Lack of motivations? (cont.)

- **Distributed Computation for Big Data**
  communication → bandwidth, ...



The BSP model.
E.g., Apache Giraph, Pregel.



The MapReduce model.
E.g., Hadoop, Google MapReduce.

Interesting problems

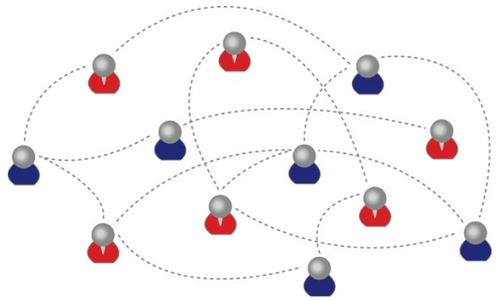
Statistical problems

- Frequency moments F_p
  F_0: #distinct elements
  F_2: size of self-join
- Heavy hitters
- Quantile
- Entropy
- ...

Interesting problems

Statistical problems

- Frequency moments F_p
  F_0: #distinct elements
  F_2: size of self-join
- Heavy hitters
- Quantile
- Entropy
- ...



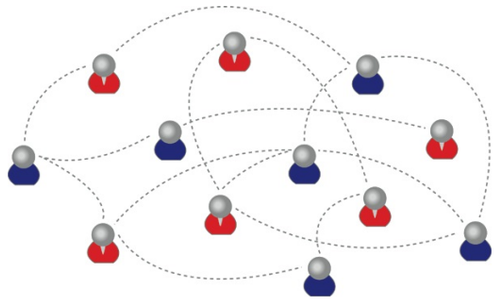
Graph problems

- Connectivity
- Bipartiteness
- Counting triangles
- Matching
- Minimum spanning tree
- ...

Interesting problems

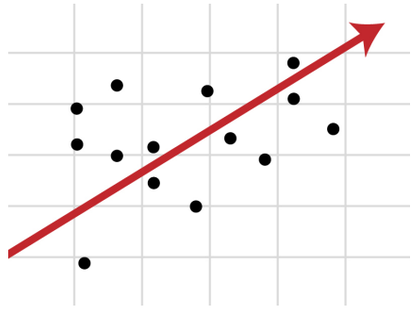
Statistical problems

- Frequency moments F_p
  F_0: #distinct elements
  F_2: size of self-join
- Heavy hitters
- Quantile
- Entropy
- ...



Graph problems

- Connectivity
- Bipartiteness
- Counting triangles
- Matching
- Minimum spanning tree
- ...



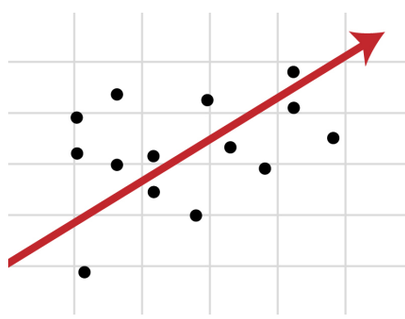
Numerical linear algebra

- L_p regression
- Low-rank approximation
- ...

Interesting problems

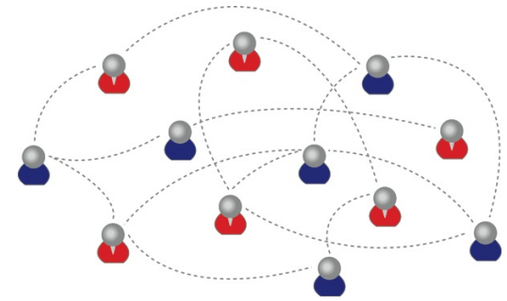
Statistical problems

- Frequency moments F_p
  F_0: #distinct elements
  F_2: size of self-join
- Heavy hitters
- Quantile
- Entropy
- ...



Numerical linear algebra

- L_p regression
- Low-rank approximation
- ...



Graph problems

- Connectivity
- Bipartiteness
- Counting triangles
- Matching
- Minimum spanning tree
- ...

DB queries

- Conjuntive queries

Distributed learning

- Classifiers

...

Geometry problems

This talk

- We will talk lower bounds.

  Most lower bounds match the upper bounds.

- Introduce two general techniques that work in both BB model and MP model for proving lower bounds, with examples.

- Conclude with several future directions.

Our new ideas and techniques

- Ideas
  1. Reduce a k-player problem to a 2-player problem.
  2. Decompose a complicated problem to several simpler problems

Our new ideas and techniques

- Ideas
  1. Reduce a k-player problem to a 2-player problem.
  2. Decompose a complicated problem to several simpler problems
- In fact, we are

  exploring the underlying patterns of communication matrices

Our new ideas and techniques

- Ideas
  1. Reduce a k-player problem to a 2-player problem.
  2. Decompose a complicated problem to several simpler problems
- In fact, we are

  exploring the underlying patterns of communication matrices
- Concrete techniques
  1. **Symmetrization**
     With Phillips and Verbin, 2012
  2. **Composition**
     With Woodruff, 2012

Our new ideas and techniques

- Ideas
  1. Reduce a k-player problem to a 2-player problem.
  2. Decompose a complicated problem to several simpler problems
- In fact, we are

  exploring the underlying patterns of communication matrices
- Concrete techniques
  1. **Symmetrization** Example: Graph Connecitivy
     With Phillips and Verbin, 2012
  2. **Composition** Example: Distinct Elements
     With Woodruff, 2012

Communication complexity

I choose not to introduce formal definitions/concepts of communication complexity. Just want to mention Yao's lemma:

We will prove lower bounds for randomized protocols. To do that we will usually pick some input distribution, and then prove lower bounds for any deterministic protocol that succeeds w.pr 0.99 under that input distribution.

1st Technique: Symmetrization

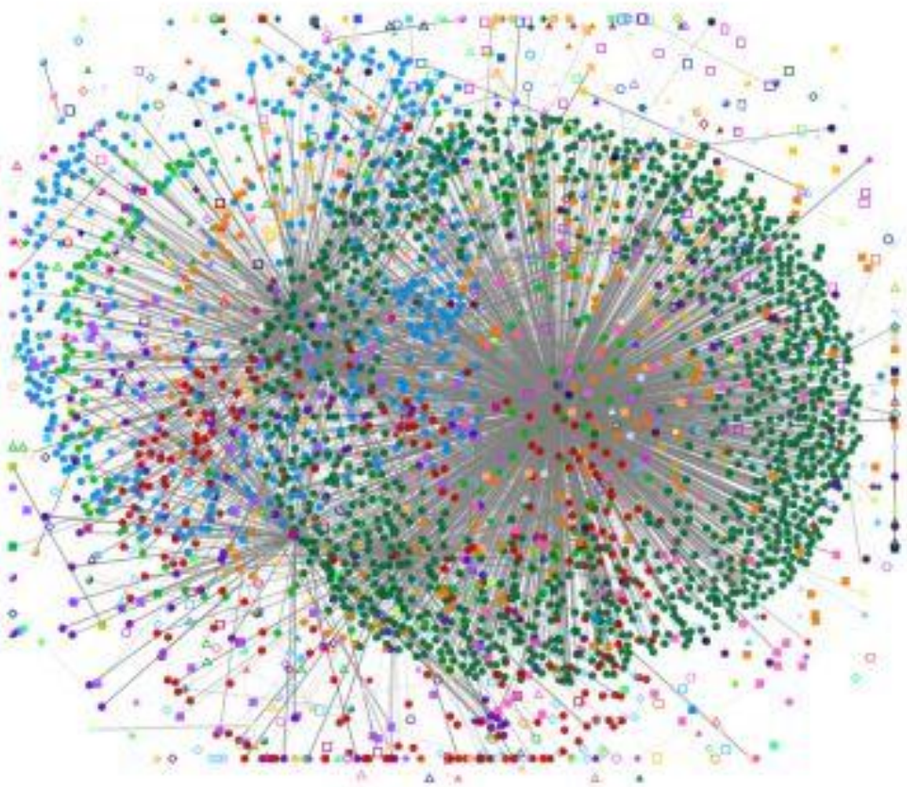
Example: graph connectivity

k sites each holds a set of edges of a graph

Goal: compute whether the graph is connected.

A trivial UB: $O(nk \log n)$ bits.

We prove that this is tight up to a $\log$ factor.



Set disjointness (2-DISJ)

$X \subseteq \{1, \ldots, n\}$

$$X \cap Y = \emptyset?$$

$Y \subseteq \{1, \ldots, n\}$

Set disjointness (2-DISJ)

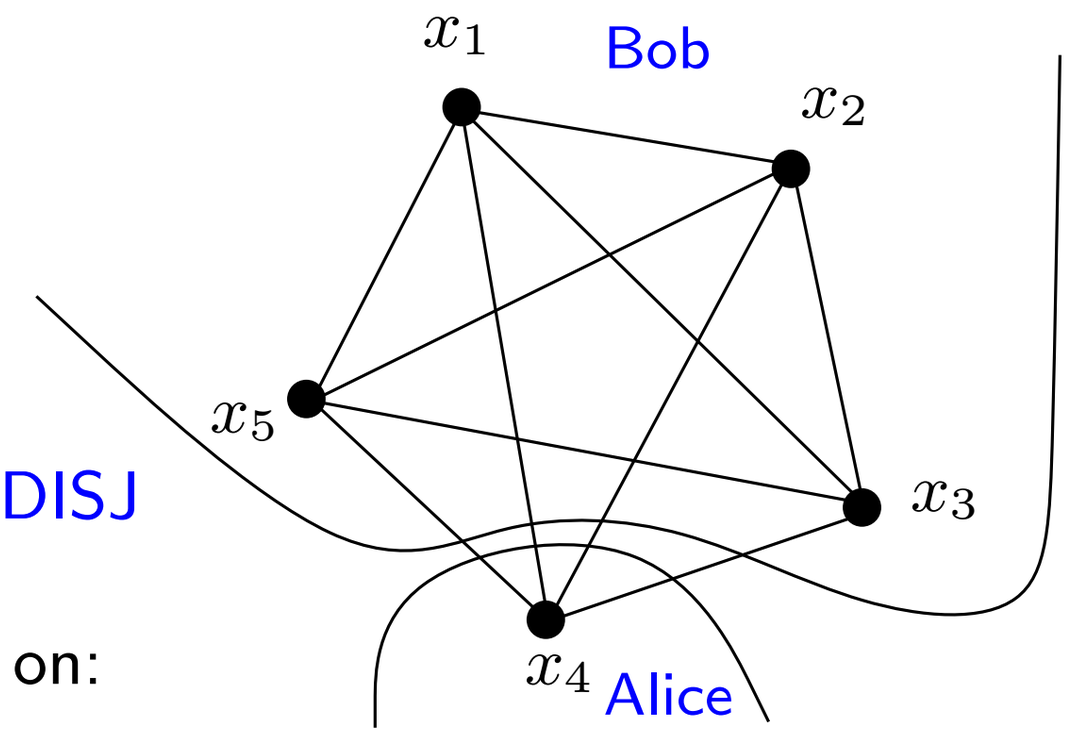
$X \cap Y = \emptyset?$

$X \subseteq \{1, \ldots, n\}$ $\qquad$ $Y \subseteq \{1, \ldots, n\}$

Exists a hard distribution μ_β, under which
$|X \cap Y| = 1$ (YES instance) w.p. β and
$|X \cap Y| = 0$ (NO instance) w.p. $1 - \beta$.

Lemma: (Generalization of [Razborov '90, BJKS '04])
$\mathbf{E}[\mathrm{CC}_{\mu_\beta}^{\beta/100}(\text{2-DISJ})] = \Omega(n)$

2-DISJ $\Rightarrow$ k-CONN

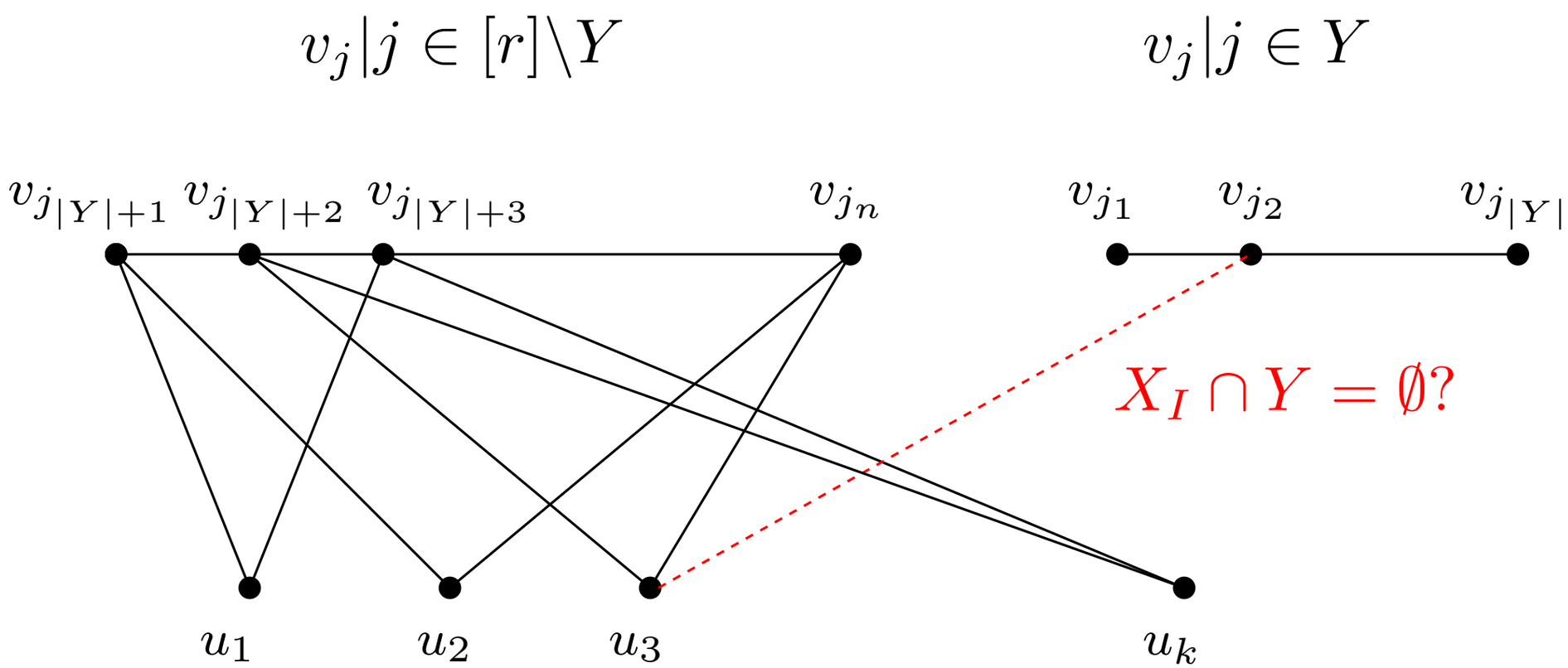


Alice and Bob want to solve the 2-DISJ $(X, Y) \sim \mu_\beta$. (set $\beta = 1/k^2$)
$\Rightarrow$ running a protocol for k-CONN on:

Alice plays a random site S_I with input $X_I = X$.
Bob plays other $k-1$ sites, and constructs inputs for them using Y: it choose $X_i \sim \mu_\beta | Y$ $(i \neq I)$.
Note: $X_1, \ldots, X_k$ are i.i.d. conditioned on Y.

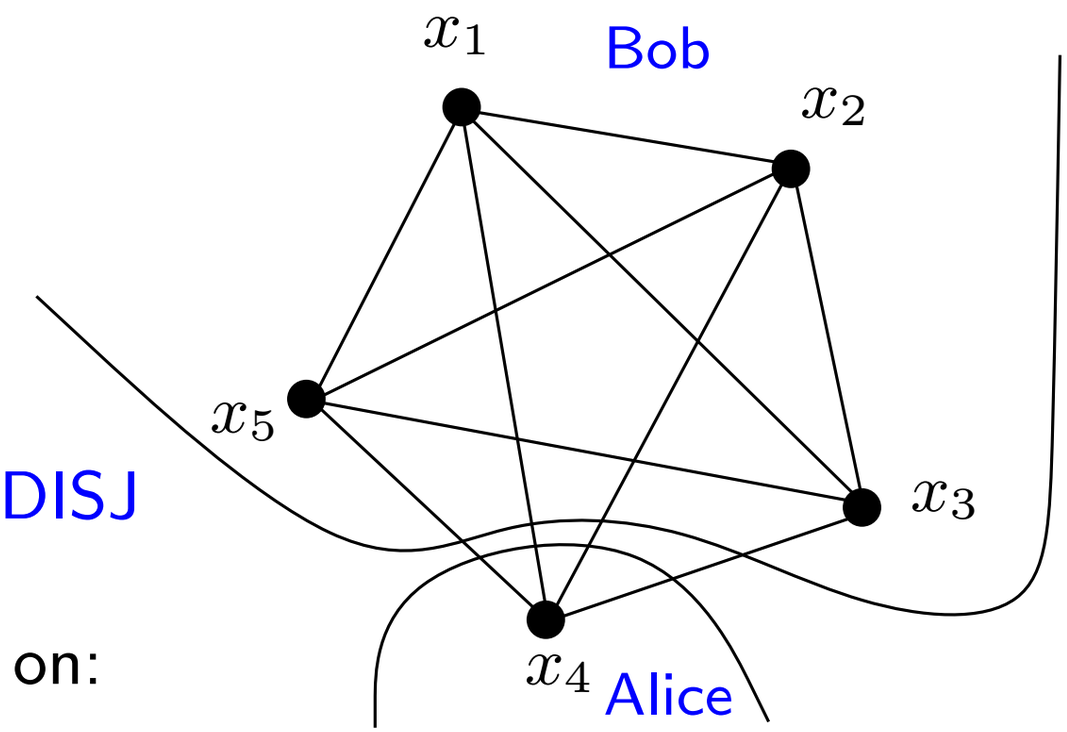
Vertices $u_1, \ldots, u_k$, $v_1, \ldots, v_n$. An edge (u_i, v_j) exists iff $X_{i,j} = 1$.
Let ν be the distribution of the graph.

2-DISJ $\Rightarrow$ k-CONN (cont.)



(u_i, v_j) exists if and only if $X_{i,j} = 1$

2-DISJ $\Rightarrow$ k-CONN



Alice and Bob want to solve the 2-DISJ $(X, Y) \sim \mu_\beta$. (set $\beta = 1/k^2$)
$\Rightarrow$ running a protocol for k-CONN on:

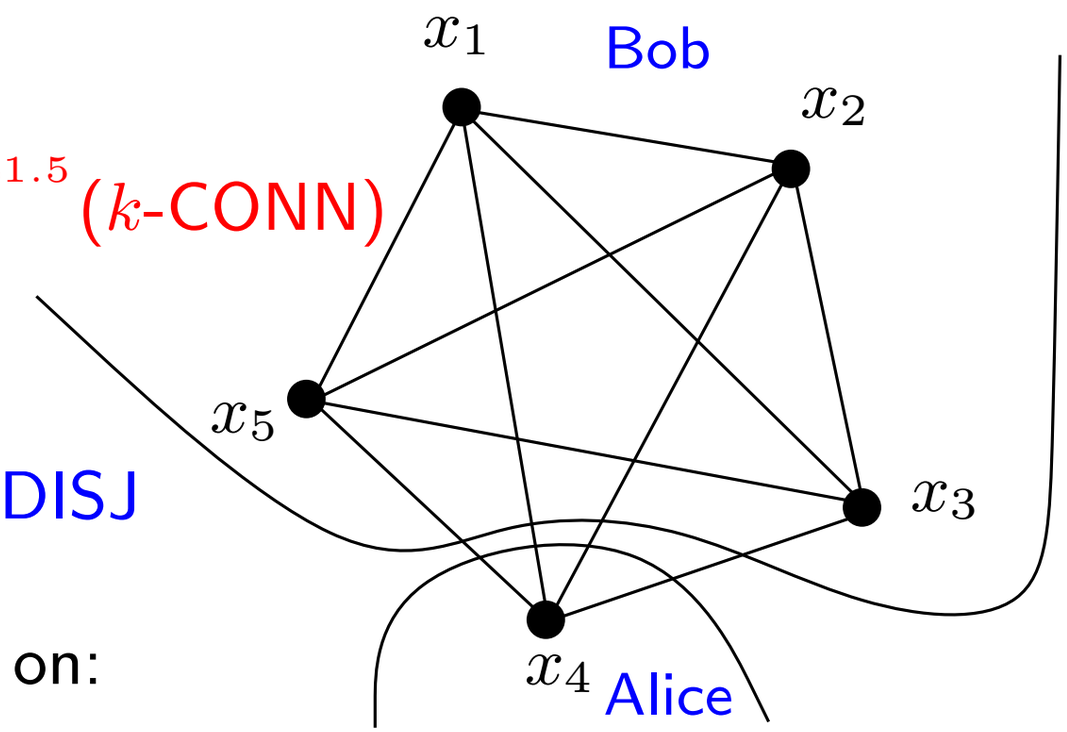
Alice plays a random site S_I with input $X_I = X$.
Bob plays other $k-1$ sites, and constructs inputs for them using Y: it choose $X_i \sim \mu_\beta | Y$ $(i \neq I)$.
Note: $X_1, \ldots, X_k$ are i.i.d. conditioned on Y.

Vertices $u_1, \ldots, u_k$, $v_1, \ldots, v_n$. An edge (u_i, v_j) exists iff $X_{i,j} = 1$.
Let ν be the distribution of the graph.

Correctness: Solving k-CONN w.pr. $1 - \beta^{1.5}$ under ν also solves 2-DISJ w.pr. $1 - \beta/100$ under μ_β.

2-DISJ $\Rightarrow$ k-CONN

$$\mathbf{E}[\mathsf{CC}^{\beta/100}_{\mu_\beta}(\text{2-DISJ})] = O\left(\tfrac{1}{k}\right) \cdot \mathsf{CC}^{\beta^{1.5}}_{\nu}(k\text{-CONN})$$



Alice and Bob want to solve the 2-DISJ $(X, Y) \sim \mu_\beta$. (set $\beta = 1/k^2$)
$\Rightarrow$ running a protocol for k-CONN on:

Alice plays a random site S_I with input $X_I = X$.
Bob plays other $k-1$ sites, and constructs inputs for them using Y: it choose $X_i \sim \mu_\beta | Y$ $(i \neq I)$.
Note: $X_1, \ldots, X_k$ are i.i.d. conditioned on Y.

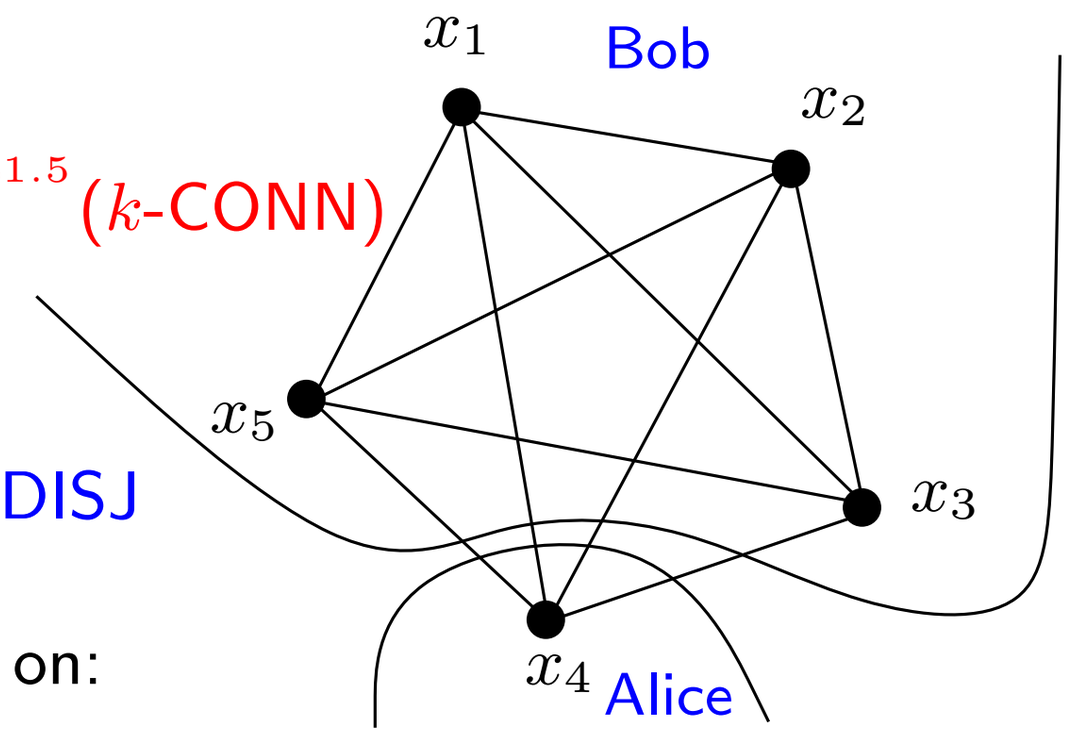
Vertices $u_1, \ldots, u_k$, $v_1, \ldots, v_n$. An edge (u_i, v_j) exists iff $X_{i,j} = 1$.
Let ν be the distribution of the graph.

Correctness: Solving k-CONN w.pr. $1 - \beta^{1.5}$ under ν also solves 2-DISJ w.pr. $1 - \beta/100$ under μ_β.

2-DISJ $\Rightarrow$ k-CONN

$$\mathbf{E}[\mathsf{CC}^{\beta/100}_{\mu_\beta}(\text{2-DISJ})] = O\left(\tfrac{1}{k}\right) \cdot \mathsf{CC}^{\beta^{1.5}}_{\nu}(k\text{-CONN})$$

$\Omega(n)$



Alice and Bob want to solve the 2-DISJ $(X, Y) \sim \mu_\beta$. (set $\beta = 1/k^2$)
$\Rightarrow$ running a protocol for k-CONN on:

Alice plays a random site S_I with input $X_I = X$.
Bob plays other $k-1$ sites, and constructs inputs for them using Y: it choose $X_i \sim \mu_\beta | Y$ $(i \neq I)$.
Note: $X_1, \ldots, X_k$ are i.i.d. conditioned on Y.

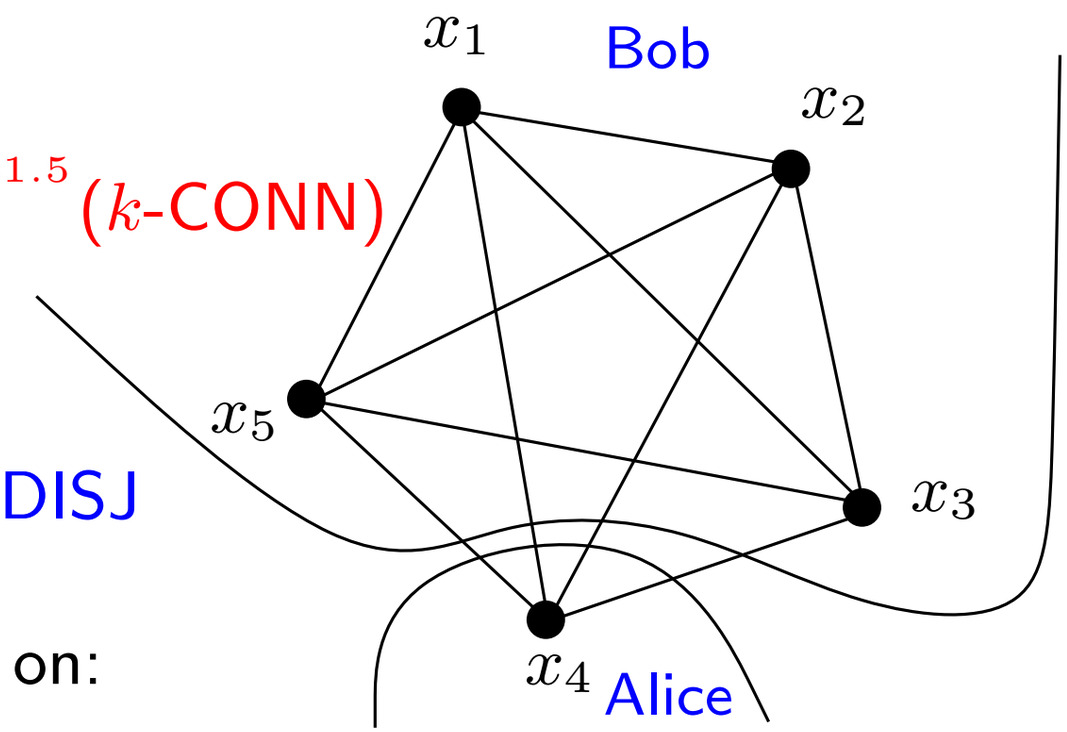
Vertices $u_1, \ldots, u_k$, $v_1, \ldots, v_n$. An edge (u_i, v_j) exists iff $X_{i,j} = 1$.
Let ν be the distribution of the graph.

Correctness: Solving k-CONN w.pr. $1 - \beta^{1.5}$ under ν also solves 2-DISJ w.pr. $1 - \beta/100$ under μ_β.

2-DISJ $\Rightarrow$ k-CONN

$$\mathbf{E}[\mathsf{CC}^{\beta/100}_{\mu_\beta}(\text{2-DISJ})] = O\left(\tfrac{1}{k}\right) \cdot \mathsf{CC}^{\beta^{1.5}}_{\nu}(k\text{-CONN})$$

$\Omega(n)$ $\qquad$ $\Omega(nk)$



Alice and Bob want to solve the 2-DISJ $(X, Y) \sim \mu_\beta$. (set $\beta = 1/k^2$)
$\Rightarrow$ running a protocol for k-CONN on:

Alice plays a random site S_I with input $X_I = X$.
Bob plays other $k-1$ sites, and constructs inputs for them using Y: it choose $X_i \sim \mu_\beta | Y$ $(i \neq I)$.
Note: $X_1, \ldots, X_k$ are i.i.d. conditioned on Y.

Vertices $u_1, \ldots, u_k$, $v_1, \ldots, v_n$. An edge (u_i, v_j) exists iff $X_{i,j} = 1$.
Let ν be the distribution of the graph.

Correctness: Solving k-CONN w.pr. $1 - \beta^{1.5}$ under ν also solves 2-DISJ w.pr. $1 - \beta/100$ under μ_β.

A useful message

We can show trivial UBs are tight for many statistical and graph problems in the MP model (e.g., #distinct elements, bipartiteness, cycle/triangle-freeness, ...), when exact answers are wanted. (Woodruff, Z, 2013a)

Approximation is often necessary if we want comm. efficient protocols.

A useful message

We can show trivial UBs are tight for many statistical and graph problems in the MP model (e.g., #distinct elements, bipartiteness, cycle/triangle-freeness, ...), when exact answers are wanted. (Woodruff, Z, 2013a)

Approximation is often necessary if we want comm. efficient protocols.

Example:

Computing the diameter of a graph exactly needs $\Omega(kn^2)$ bits (when $m = n^2$) in the MP model.

However, there is a protocol that computes the diameter up to an additive error of 2 using $\tilde{O}(kn^{3/2})$ bits.

Other problems using symmetrization

- k-bitwise-MAJ
- k-bitwise-OR/AND/XOR

  With applications in

  - Distinct elements reporting
  - ε-kernels (geometry)
  - Heavy-hitters
- Testing bipartiteness,

  Testing cycle

  Testing triangle-freeness,

  ...

Other problems using symmetrization

- k-bitwise-MAJ
- k-bitwise-OR/AND/XOR

  With applications in

  - Distinct elements reporting
  - ε-kernels (geometry)
  - Heavy-hitters
- Testing bipartiteness,

  Testing cycle

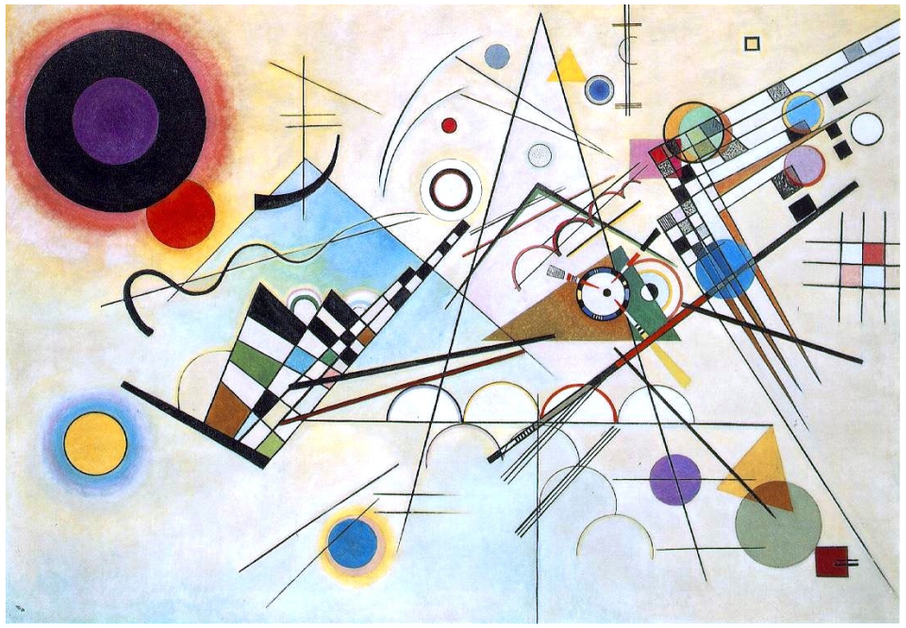
  Testing triangle-freeness,

  ...

Key:

Find the right 2-player problem for reduction

2nd Technique: Composition



Example – the F_0 (distinct elements) problem

k sites each holds a set X_i ($i \in \{1, 2, \ldots, k\}$).

Goal: compute #distinct elements$(\cup_{i=1}^{k} X_i)$ up to a $(1+\varepsilon)$-approx.

Example – the F_0 (distinct elements) problem

k sites each holds a set X_i ($i \in \{1, 2, \ldots, k\}$).

Goal: compute #distinct elements$(\cup_{i=1}^{k} X_i)$ up to a $(1+\varepsilon)$-approx.

Current best UB: $\tilde{O}(k/\varepsilon^2)$
[Cormode, Muth, Yi. 2008]

Holds in MP model

Example – the F_0 (distinct elements) problem

k sites each holds a set X_i ($i \in \{1, 2, \ldots, k\}$).

Goal: compute #distinct elements($\cup_{i=1}^{k} X_i$) up to a $(1+\varepsilon)$-approx.

Current best UB: $\tilde{O}(k/\varepsilon^2)$
[Cormode, Muth, Yi. 2008]

Holds in MP model

Previous LB: $\Omega(k)$

And $\Omega(1/\varepsilon^2)$
Direct reduction from Gap-Hamming

Example – the F_0 (distinct elements) problem

k sites each holds a set X_i ($i \in \{1, 2, \ldots, k\}$).

Goal: compute #distinct elements($\cup_{i=1}^{k} X_i$) up to a $(1+\varepsilon)$-approx.

Current best UB: $\tilde{O}(k/\varepsilon^2)$
[Cormode, Muth, Yi. 2008]
Holds in MP model

Previous LB: $\Omega(k)$

And $\Omega(1/\varepsilon^2)$
Direct reduction from Gap-Hamming

New LB: $\Omega(k/\varepsilon^2)$. [Woodruff, Z. 2012, 2013b]. Holds in MP model. Better UB in the BB model

The proof framework

Step 1: Find two modular problems k-GAP-MAJ and 2-DISJ of simpler structures s.t.

F_0 can be thought as a composition of them.

Step 2: Analyze the complexities of k-GAP-MAJ and 2-DISJ.

Step 3: Compose two modular problems so that:

$$\text{Complexity}(F_0) = \text{Complexity}(k\text{-GAP-MAJ}) \times \text{Complexity}(2\text{-DISJ}).$$

The proof framework

Step 1: Find two modular problems k-GAP-MAJ and 2-DISJ of simpler structures s.t.

F_0 can be thought as a composition of them.

Step 2: Analyze the complexities of k-GAP-MAJ and 2-DISJ.

Step 3: Compose two modular problems so that:

$$\text{Complexity}(F_0) = \text{Complexity}(k\text{-GAP-MAJ}) \times \text{Complexity}(2\text{-DISJ}).$$

The proof presented here is for **a special case when** $k = 2/\varepsilon^2$.

k-GAP-MAJ

k sites each holds a bit Z_i chosen uniform at random from $\{0,1\}$

Goal: compute

$$k\text{-GAP-MAJ}(Z_1, Z_2, \ldots, Z_k) = \begin{cases} 0, & \text{if } \sum_{i\in[k]} Z_i \le k/2 - \sqrt{k}, \\ 1, & \text{if } \sum_{i\in[k]} Z_i \ge k/2 + \sqrt{k}, \\ \text{don't care}, & \text{otherwise,} \end{cases}$$

k-GAP-MAJ

k sites each holds a bit Z_i chosen uniform at random from $\{0,1\}$

Goal: compute

$$k\text{-GAP-MAJ}(Z_1, Z_2, \ldots, Z_k) = \begin{cases} 0, & \text{if } \sum_{i\in[k]} Z_i \le k/2 - \sqrt{k}, \\ 1, & \text{if } \sum_{i\in[k]} Z_i \ge k/2 + \sqrt{k}, \\ \text{don't care}, & \text{otherwise,} \end{cases}$$

Lemma: Any protocol Π that computes k-GAP-MAJ correctly w.p. 0.9999 has to learn $\Omega(k)$ Z_i's well, that is,

$$H(Z_i \mid \Pi) \le H_b(1/100) \text{ for } \Omega(k) \text{ of } i.$$

In other words: $I(\Pi; Z_1, \ldots, Z_k) = \Omega(k)$.

Set disjointness (2-DISJ)

$X \subseteq \{1, \ldots, n\}$

$$X \cap Y = \emptyset?$$

$Y \subseteq \{1, \ldots, n\}$

Set disjointness (2-DISJ)

$X \subseteq \{1, \ldots, n\}$

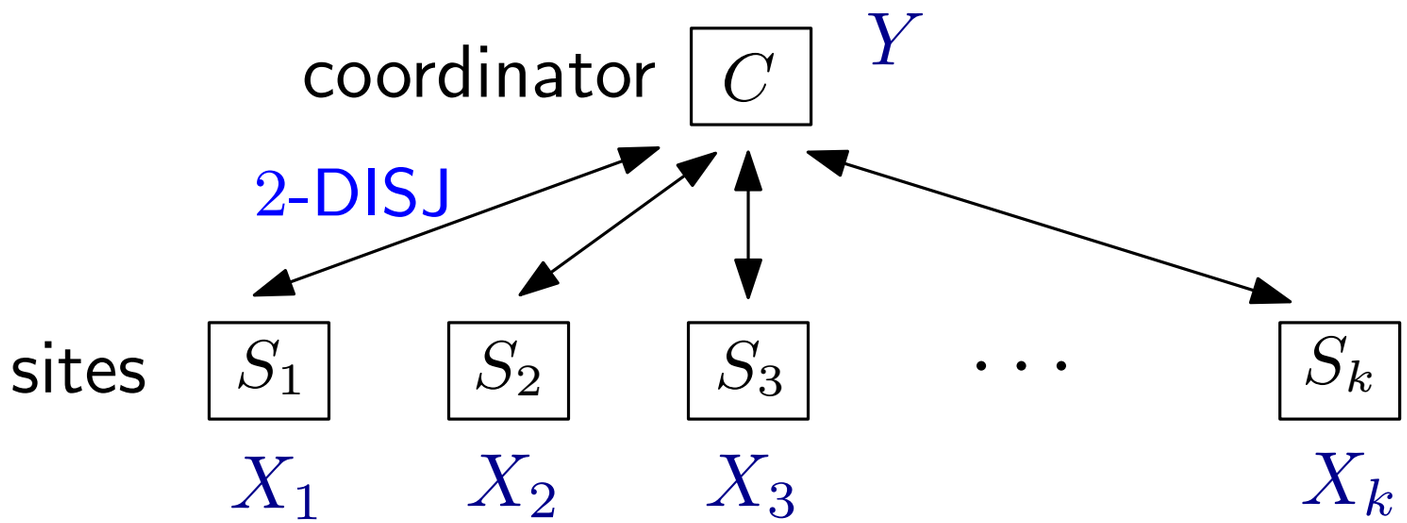
$X \cap Y = \emptyset?$

$Y \subseteq \{1, \ldots, n\}$

Exists a hard distribution μ_β, under which
$|X \cap Y| = 1$ (YES instance) w.p. β and
$|X \cap Y| = 0$ (NO instance) w.p. $1 - \beta$.

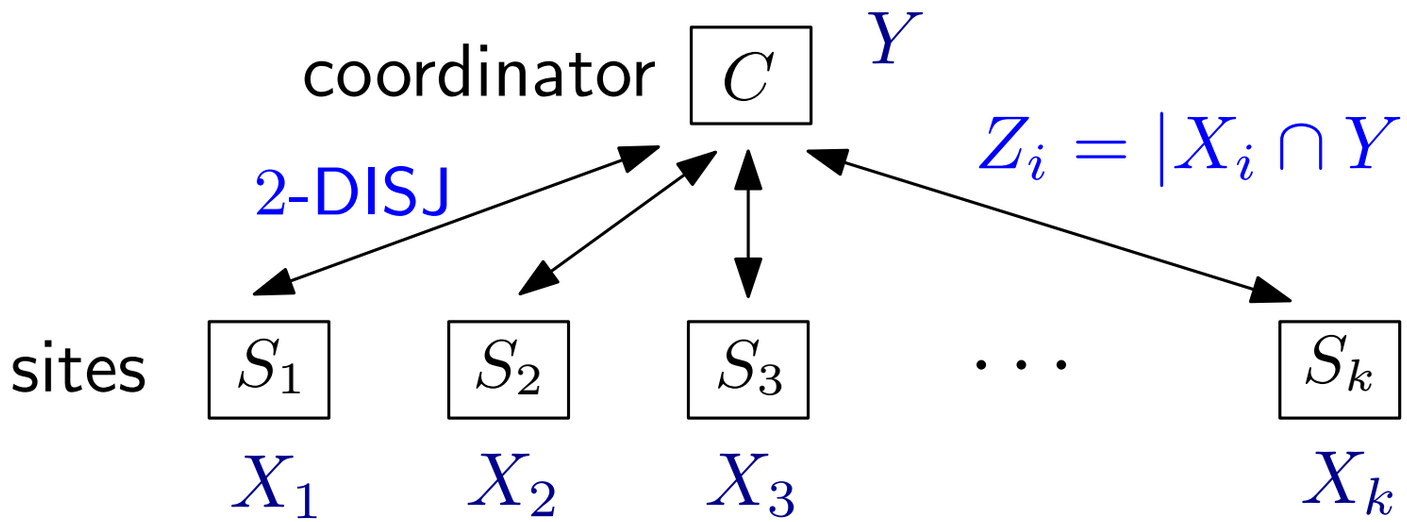
Lemma: (Generalization of [Razborov '90, BJKS '04])
$\mathbf{E}[\mathrm{CC}_{\mu_\beta}^{\beta/100}(\text{2-DISJ})] = \Omega(n)$

The proof sketch



$(X_i, Y) \sim \mu_{1/2}$ (hard input dist. for DISJ) for each $i \in [k]$

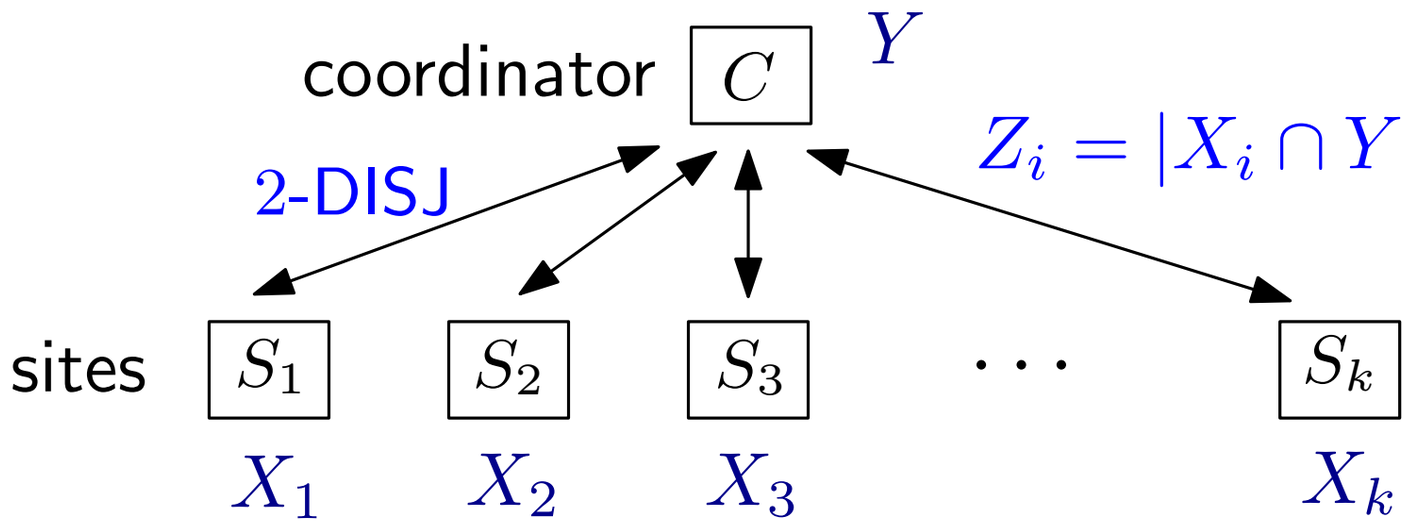
The proof sketch



$$Z_i = |X_i \cap Y| \begin{cases} 1 & \text{w.p. } 1/2 \\ 0 & \text{w.p. } 1/2 \end{cases}$$

$(X_i, Y) \sim \mu_{1/2}$ (hard input dist. for DISJ) for each $i \in [k]$

The proof sketch



$$Z_i = |X_i \cap Y| \begin{cases} 1 & \text{w.p. } 1/2 \\ 0 & \text{w.p. } 1/2 \end{cases}$$

$(X_i, Y) \sim \mu_{1/2}$ (hard input dist. for DISJ) for each $i \in [k]$

$$F_0(X_1, X_2, \ldots, X_k) \iff k\text{-GAP-MAJ}(Z_1, Z_2, \ldots, Z_k)$$

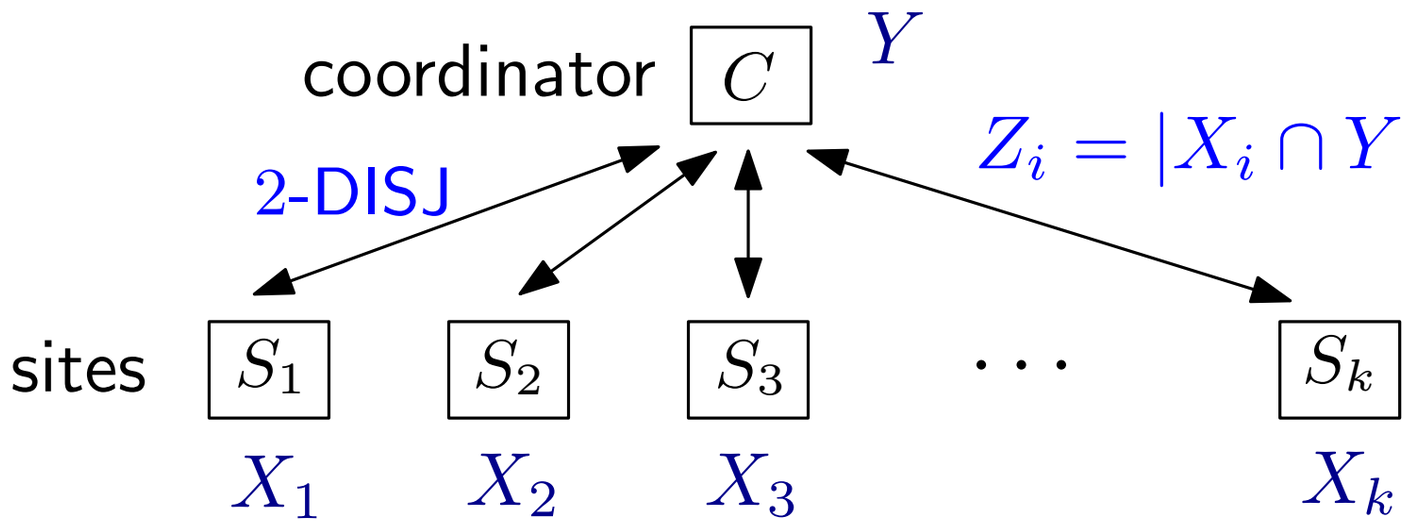
$(Z_i = |X_i \cap Y|)$

$\implies$ learn $\Omega(k)$ Z_i's well

$\implies$ need $\Omega(nk)$ bits

(using an embedding argument)

The proof sketch



$$Z_i = |X_i \cap Y| \begin{cases} 1 & \text{w.p. } 1/2 \\ 0 & \text{w.p. } 1/2 \end{cases}$$

$(X_i, Y) \sim \mu_{1/2}$ (hard input dist. for DISJ) for each $i \in [k]$

$F_0(X_1, X_2, \ldots, X_k) \iff k$-GAP-MAJ$(Z_1, Z_2, \ldots, Z_k)$

$(Z_i = |X_i \cap Y|)$

$\implies$ learn $\Omega(k)$ Z_i's well

$\implies$ need $\Omega(nk)$ bits

(using an embedding argument)

For the reduction set universe $n = \Theta(1/\varepsilon^2)$ we get $\Omega(k/\varepsilon^2)$ LB.

Q.E.D.

Other problems using composition

- Frequency moments F_p $(p > 1)$
- Entropy
- ℓ_p
- Heavy-hitters
- Quantile

Other problems using composition

- Frequency moments F_p $(p > 1)$
- Entropy
- ℓ_p
- Heavy-hitters
- Quantile

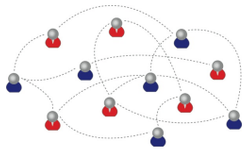
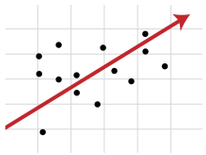
Key:

Find the right modular problems

Furture directions

Future directions – Problem space

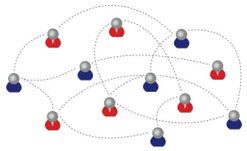
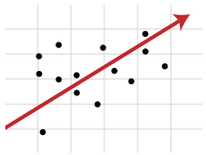
- The current understandings


	statistics	graph	linear algebra	DB queries
Exact computation	well	medium		limited
Approximation	well	limited	limited	limited

Future directions – Problem space

- The current understandings


	statistics	graph	linear algebra	DB queries
Exact computation	well	medium		limited
Approximation	well	limited	limited	limited

Concretely, e.g.,

What is the CC of computing a const approx of the *size* of the matching?

Future directions – Problem space (cont.)

- Communication complexity of distributed property testing.

  Communication complexity of distributed property testing can also be used for proving LBs for traditional property testing, by extending a framework by Blais et al. 2011.

Future directions – Problem space (cont.)

- Communication complexity of distributed property testing.

  Communication complexity of distributed property testing can also be used for proving LBs for traditional property testing, by extending a framework by Blais et al. 2011.

- Communication complexity of relations
  (instead of functions, useful for database queries)

Future directions – Problem space (cont.)

- Communication complexity of distributed property testing.

  Communication complexity of distributed property testing can also be used for proving LBs for traditional property testing, by extending a framework by Blais et al. 2011.

- Communication complexity of relations
  (instead of functions, useful for database queries)

- Need to understand more primitive problems (building blocks), e.g., breadth-first search (BFS), pointer-chasing, etc.

  So far building blocks are quite limited: mainly 2-DISJ, 2-GAP-HAMMING, k-GAP-MAJ, k-OR and their variants.

Future directions – Model/Technique space

- Can we relax or generalize the symmetrization technique?

  E.g., do not require "totally symmetric distributions"

Future directions – Model/Technique space

- Can we relax or generalize the symmetrization technique?

  E.g., do not require "totally symmetric distributions"

- Consider round complexity (practical needs).

- Consider players with limited space (practical needs).

The end

THANK YOU

Questions?