Streaming Lower Bounds for Approximating MAX-CUT

Michael Kapralov[1]

[1]EPFL

(Based on joint works with Sanjeev Khanna, Madhu Sudan and Ameya Velingker)

Graphs a common abstraction for representing real world data:

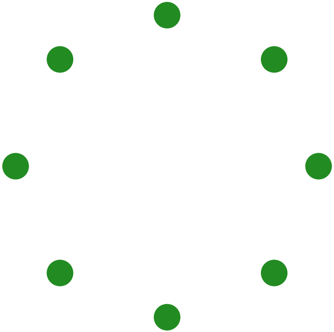
- social networks (Facebook, Twitter)
- web topologies
- interaction graphs
- …

Modern graphs are often too large to fit into memory of a compute node

Need graph analysis primitives that use very little space

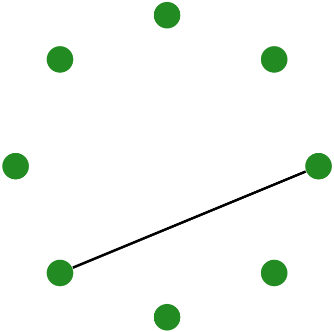
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



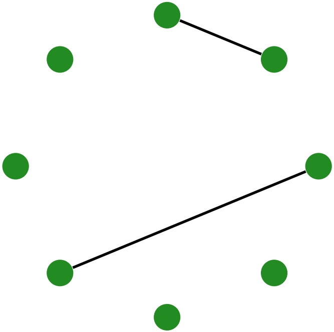
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



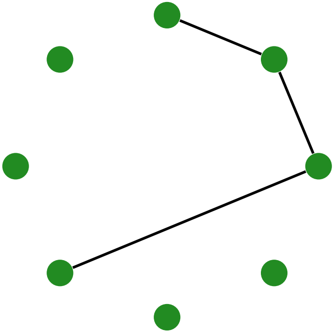
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



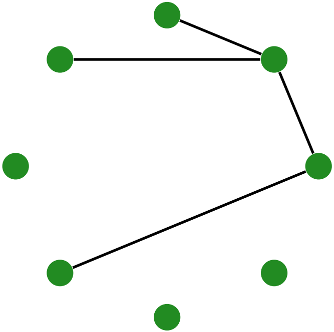
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



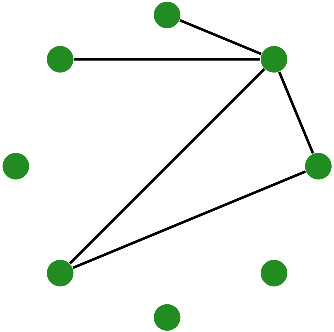
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



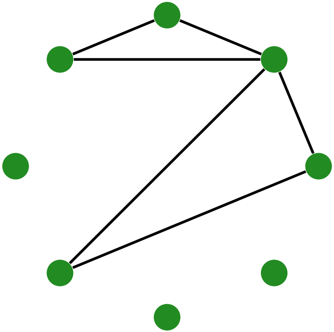
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



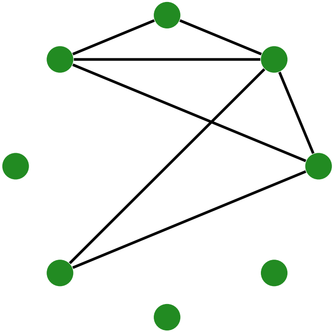
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



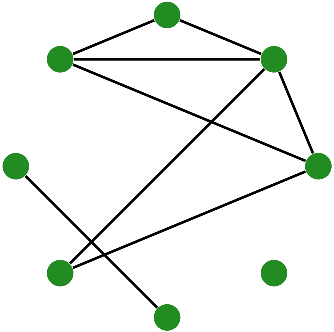
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



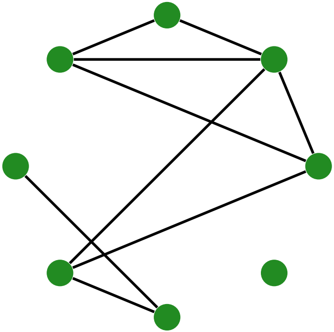
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



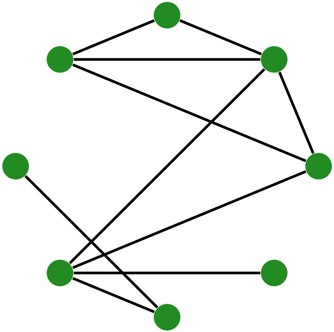
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



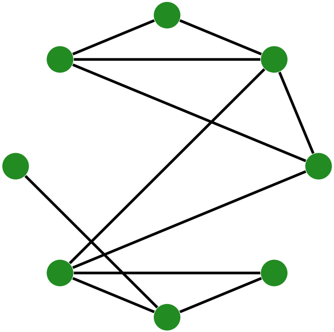
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



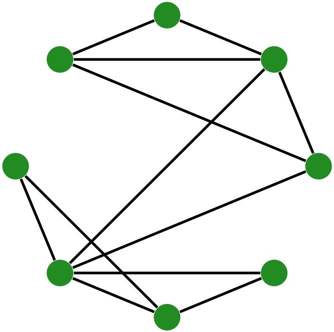
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



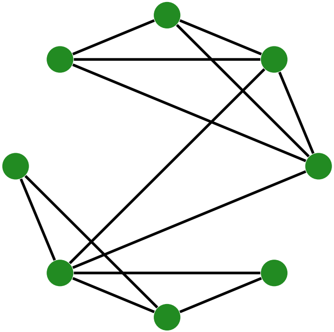
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



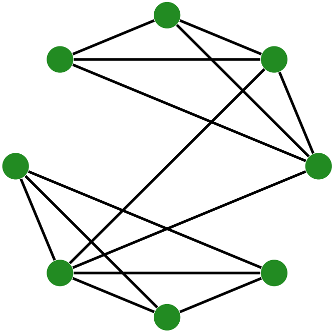
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



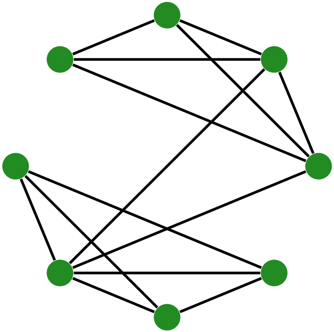
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



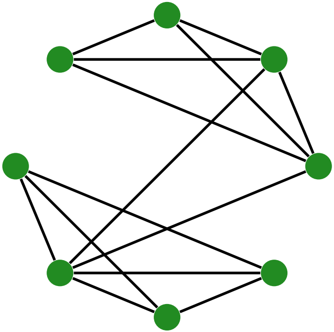
Streaming model

- edges of $G = (V, E)$ arrive in an arbitrary order in a stream; denote $|V| = n, |E| = m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream



Streaming model

- edges of $G=(V,E)$ arrive in an arbitrary order in a stream; denote $|V|=n, |E|=m$
- algorithm can only use $\widetilde{O}(n)$ space
- several passes over the stream **(ideally one pass)**



$\Omega(n)$ space is often needed:

- output size often $\Omega(n)$ (e.g., matching, sparsifier, spanner)
- even if output is a number (e.g. testing connectivity)

$\Omega(n)$ space is often needed:

- output size often $\Omega(n)$ (e.g., matching, sparsifier, spanner)
- even if output is a number (e.g. testing connectivity)

But not always:

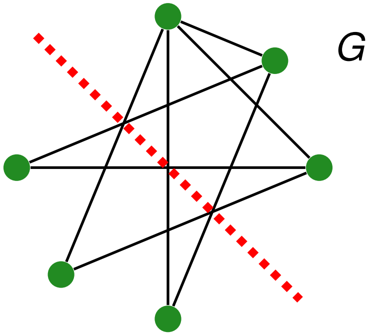
Kapralov-Khanna-Sudan'14 – can approximate matching size to $\text{poly}(\log n)$ factor using $\text{poly}(\log n)$ space in random streams.

Also, Efsaniari-Hajiaghayi-Liaghat-Monemizadeh-Onak'15, Bury-Schwiegelsohn'15, McGregor-Vorotnikova'16, Cormode-Jowhari-Monemizadeh-Muthukrishnan'16,...

Approximate solution cost for graph problems in $o(n)$ space?

MAX-CUT

Given a graph output value of maximum cut



- A random cut cuts half of the edges – trivial factor 2 approximation
- 1.318-approximation due to Goemans-Williamson'95 (best possible assuming UGC)
- 1.884 via spectral techniques Trevisan'09, Kale-Seshadhri'11

Streaming algorithms:

- factor 2 approximation: count the number of edges m and output $m/2$. Only $O(\log n)$ space.

- $(1+\varepsilon)$-approximation using $O(n/\varepsilon^2)$ space (keep a sample of the edge set)

Streaming algorithms:

- factor 2 approximation: count the number of edges m and output $m/2$. Only $O(\log n)$ space.

- $(1+\varepsilon)$-approximation using $O(n/\varepsilon^2)$ space (keep a sample of the edge set)

Better than factor 2 approximation in polylog(n) space?

Theorem (K.-Khanna-Sudan'15)

For any constant $\varepsilon > 0$ a single pass streaming algorithm for approximating MAX-CUT value to factor $2 - \varepsilon$ requires $\Omega(\sqrt{n})$ space, even in the random order model.

Theorem (K.-Khanna-Sudan'15)

For any constant $\varepsilon > 0$ a single pass streaming algorithm for approximating MAX-CUT value to factor $2 - \varepsilon$ requires $\Omega(\sqrt{n})$ space, even in the random order model.

Rules out $\text{poly}(\log n)$ space, suggests $\tilde{O}(\sqrt{n})$ space may be possible in some settings...

1. Hard input distribution

2. Boolean Hidden Partition Problem (BHP)

3. Analysis of BHP

1. Hard input distribution

2. Boolean Hidden Partition Problem (BHP)

3. Analysis of BHP

Hard distribution

We establish the main theorem using a hard distribution based on Erdős-Rényi graphs:

YES: random bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

NO: non-bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

Hard distribution

We establish the main theorem using a hard distribution based on Erdős-Rényi graphs:

YES: random bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

NO: non-bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

In the YES case MAX-CUT value is m, in the NO case MAX-CUT value is $(1+O(\varepsilon))m/2$.

Hard distribution

We establish the main theorem using a hard distribution based on Erdős-Rényi graphs:

YES: random bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

NO: non-bipartite (multi)graph with expected degree $\approx \frac{1}{\varepsilon^2}$

In the YES case MAX-CUT value is m, in the NO case MAX-CUT value is $(1 + O(\varepsilon))m/2$.

Sufficient to show $\Omega(\sqrt{n})$ space required to distinguish between the two cases.

Erdős-Rényi graphs

Sample $G = (V, E)$ from distribution $\mathcal{G}_{n,p}$

=

include each edge $(u, v) \in \binom{V}{2}$ independently with probability p

Erdős-Rényi graphs

Sample $G = (V, E)$ from distribution $\mathcal{G}_{n,p}$

$=$

include each edge $(u, v) \in \binom{V}{2}$ independently with probability p

If $p = \alpha/n$ for $\alpha < 1$, then G is a union of $O(\log n)$ size trees, with probability $1 - O(\alpha^3)$.

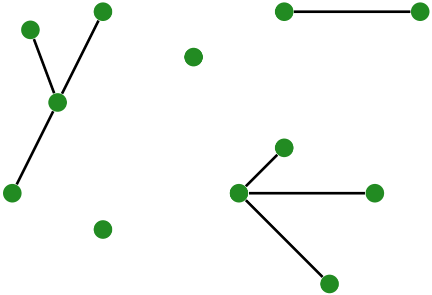
Erdős-Rényi graphs

Sample $G = (V, E)$ from distribution $\mathscr{G}_{n,p}$

=

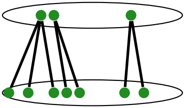
include each edge $(u, v) \in \binom{V}{2}$ independently with probability p

If $p = \alpha/n$ for $\alpha < 1$, then G is a union of $O(\log n)$ size trees, with probability $1 - O(\alpha^3)$.



Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

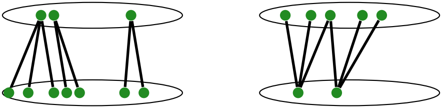
YES 

NO

Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

YES

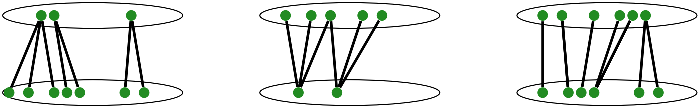


NO

Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

YES

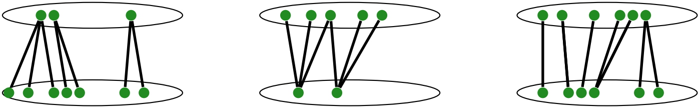


NO

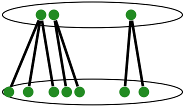
Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

YES



NO

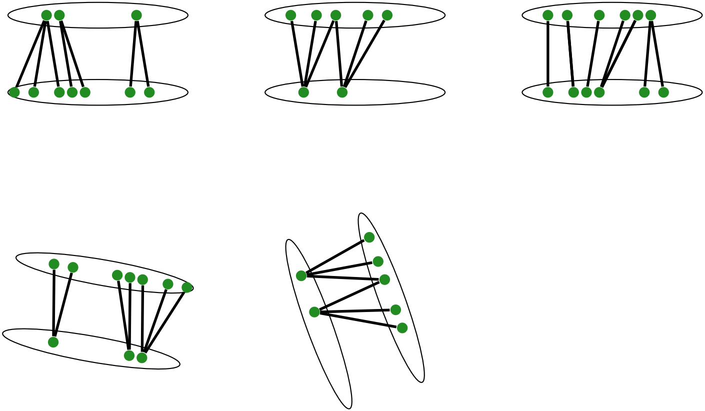


Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

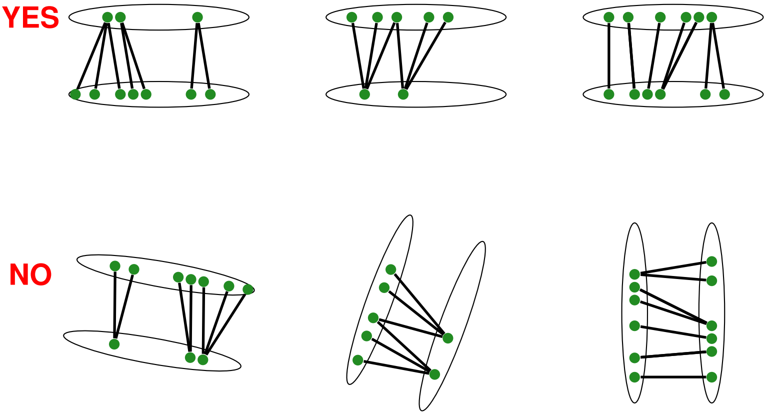
YES

NO



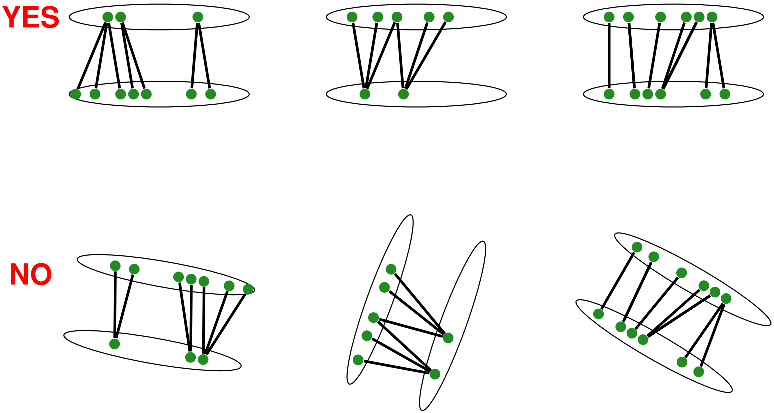
Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:



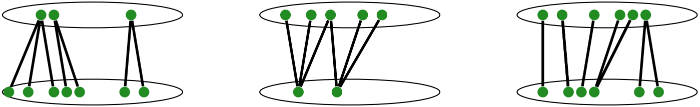
Hard input distribution

Partition the stream into $k \approx 1/\varepsilon^2$ phases:

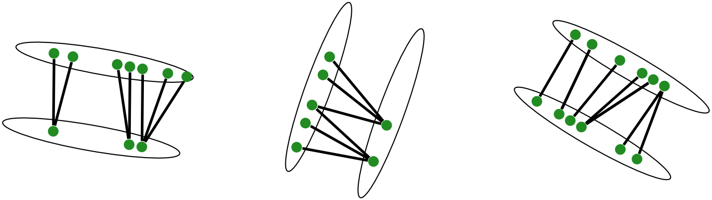


MAX-CUT value is m in **YES** case and $\leq (1+\varepsilon)m/2$ in **NO** case.

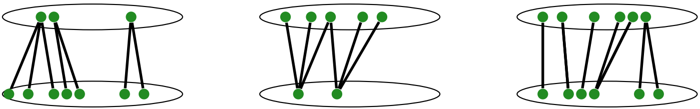
YES



NO

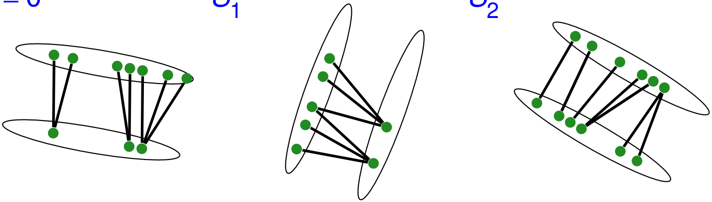


YES

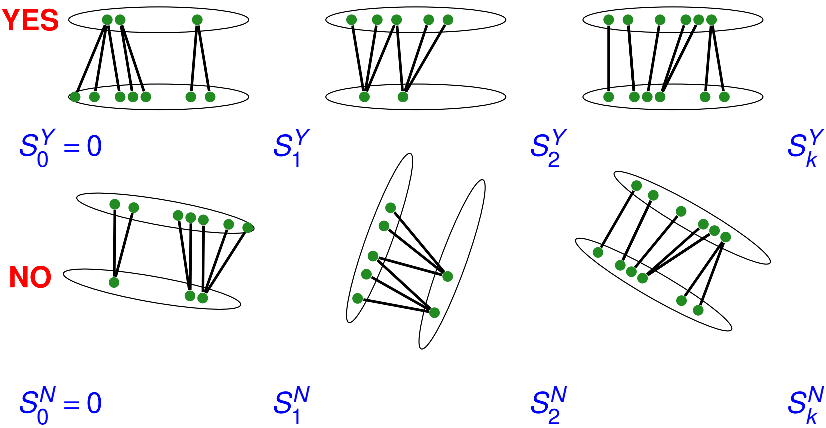


$S_0^Y = 0$ S_1^Y S_2^Y S_k^Y

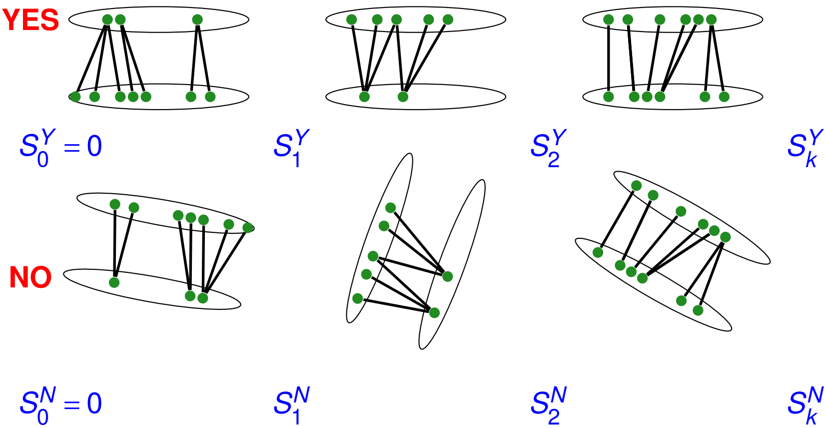
NO



$S_0^N = 0$ S_1^N S_2^N S_k^N



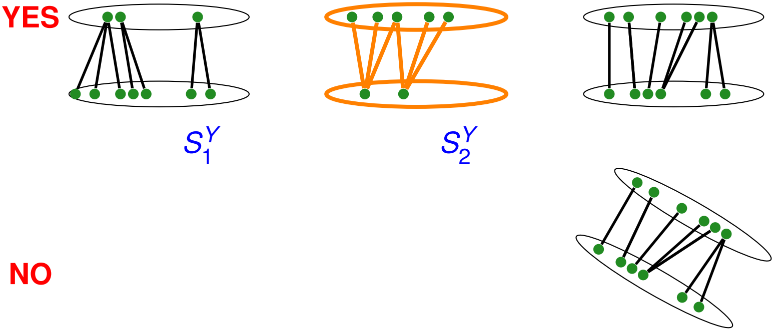
We have $S_0^N = S_0^Y = 0$ and $||S_k^Y - S_k^N||_{TV} = \Omega(1)$.



We have $S_0^N = S_0^Y = 0$ and $||S_k^Y - S_k^N||_{TV} = \Omega(1)$.

So there must exist j^* (informative index) such that

$$||S_{j^*+1}^Y - S_{j^*+1}^N||_{TV} \geq ||S_{j^*}^Y - S_{j^*}^N||_{TV} + \Omega(1/k)$$

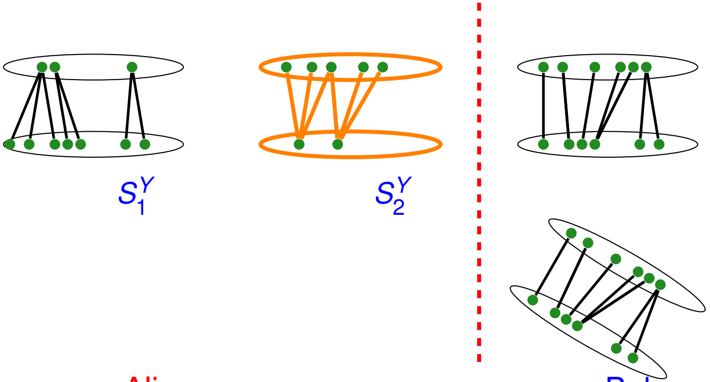


We have $S_0^N = S_0^Y = 0$ and $||S_k^Y - S_k^N||_{TV} = \Omega(1)$.
So there must exist j^* (informative index) such that

$$||S_{j^*+1}^Y - S_{j^*+1}^N||_{TV} \geq ||S_{j^*}^Y - S_{j^*}^N||_{TV} + \Omega(1/k)$$

YES

NO



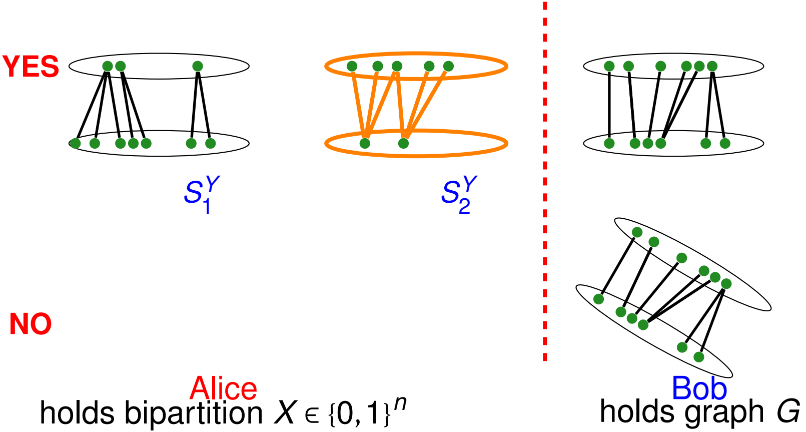
Alice
holds bipartition $X \in \{0,1\}^n$

Bob
holds graph G

We have $S_0^N = S_0^Y = 0$ and $||S_k^Y - S_k^N||_{TV} = \Omega(1)$.

So there must exist j^* (informative index) such that

$$||S_{j^*+1}^Y - S_{j^*+1}^N||_{TV} \geq ||S_{j^*}^Y - S_{j^*}^N||_{TV} + \Omega(1/k)$$



YES case: Bob's graph consistent with Alice's bipartition

NO case: Bob's graph inconsistent with Alice's bipartition

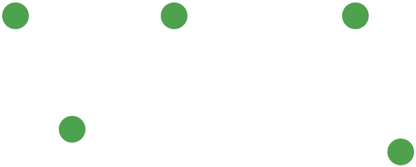
1. Hard input distribution
2. Boolean Hidden Partition Problem (BHP)
3. Analysis of BHP

1. Hard input distribution
2. Boolean Hidden Partition Problem (BHP)
3. Analysis of BHP

Boolean hidden partition problem (BHP)

Alice

binary string $x \in \{0,1\}^n$

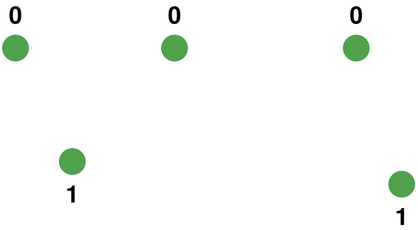


Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)

Alice

binary string $x \in \{0, 1\}^n$



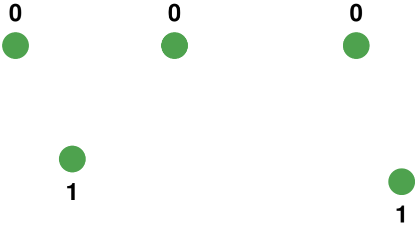
Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)

message m

Alice $\longrightarrow$ Bob

binary string $x \in \{0,1\}^n$

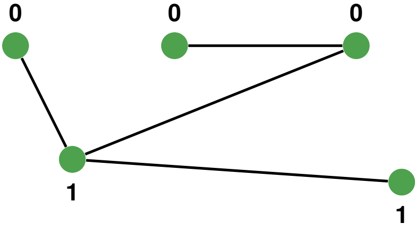


Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)

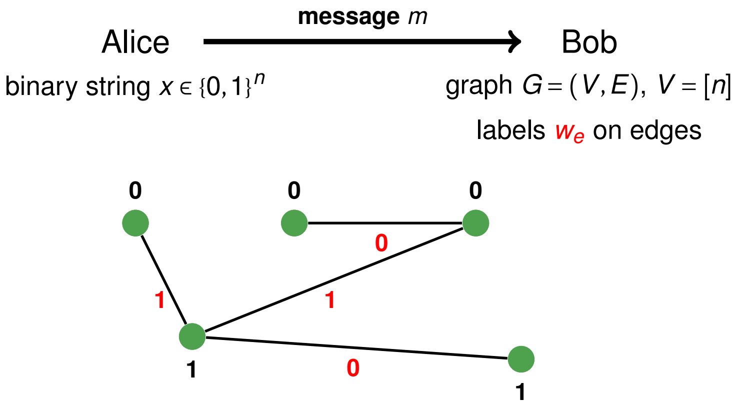
Alice —**message** m→ Bob

binary string $x \in \{0,1\}^n$ graph $G = (V,E)$, $V = [n]$



Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)



Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)

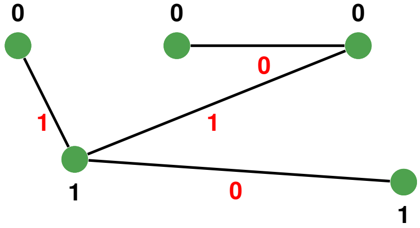
message m

Alice ⟶ Bob

binary string $x \in \{0,1\}^n$

graph $G = (V, E)$, $V = [n]$

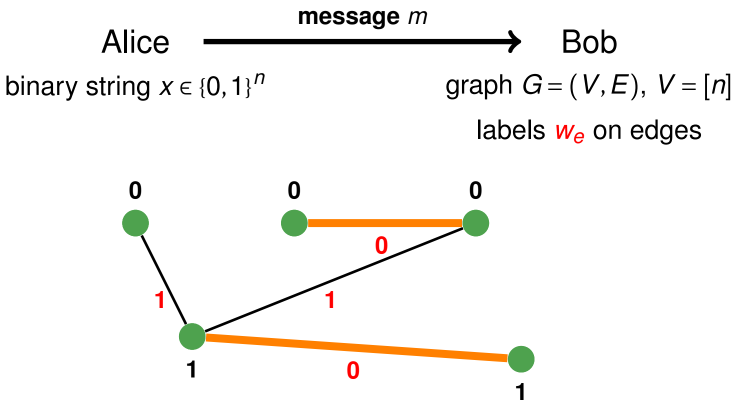
labels w_e on edges



YES: labels consistent with partition x: $w_{uv} = x_u + x_v$, i.e. $w = Mx$

Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)



YES: labels consistent with partition x: $w_{uv} = x_u + x_v$, i.e. $w = Mx$

Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

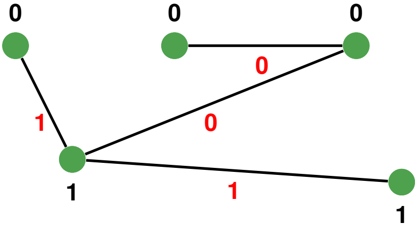
Boolean hidden partition problem (BHP)

Alice —— **message** m ——→ Bob

binary string $x \in \{0,1\}^n$

graph $G = (V, E)$, $V = [n]$

labels w_e on edges



YES: labels consistent with partition x: $w_{uv} = x_u + x_v$, i.e. $w = Mx$

NO: labels are uniformly random

Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Boolean hidden partition problem (BHP)

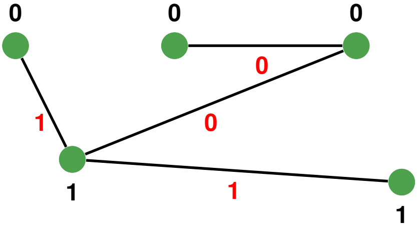
message m

Alice → Bob

binary string $x \in \{0,1\}^n$

graph $G = (V, E)$, $V = [n]$

labels w_e on edges



YES: labels consistent with partition x: $w_{uv} = x_u + x_v$, i.e. $w = Mx$

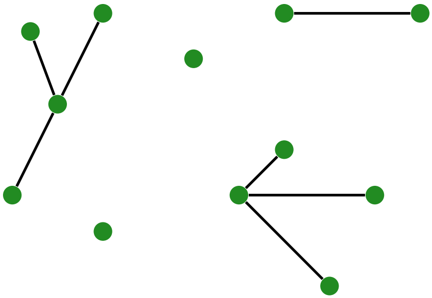
NO: labels are uniformly random

Extension of Gavinsky-Kempe-Kerenidis-Raz-de Wolf'07, Verbin-Yi'11

Distributional BHP (D-BHP)

Alice gets a uniformly random string $x \in \{0,1\}^n$

Bob gets graph G sampled from distribution $\mathcal{G}_{n,p}$ with $p = \alpha/n$, $\alpha \in (0,1)$ a small constant

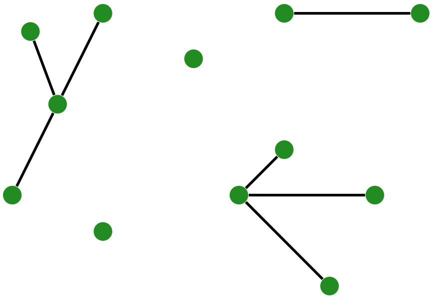


YES case independently with probability $1/2$, **NO** case otherwise.

Distributional BHP (D-BHP)

Alice gets a uniformly random string $x \in \{0,1\}^n$

Bob gets graph G sampled from distribution $\mathcal{G}_{n,p}$ with $p = \alpha/n$, $\alpha \in (0,1)$ a small constant



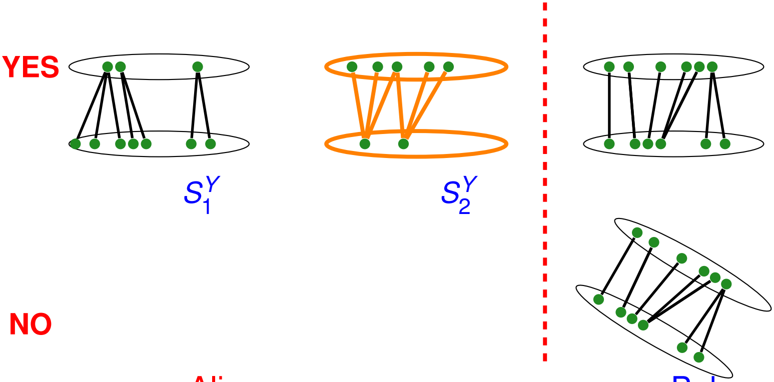
YES case independently with probability $1/2$, **NO** case otherwise.

$\sqrt{n}$ communication protocol by birthday paradox: Alice sends x_i for $\approx \sqrt{n}$ values of i!

Reduction from D-BHP to MAX-CUT

Lemma

A single-pass streaming algorithm **ALG** *that achieves $(2-\varepsilon)$-approximation to MAX-CUT with probability $\geq 99/100$ for our input distribution yields a protocol for D-BHP with advantage $\Omega(1/k)$ over random guessing.*



Alice
holds bipartition $X \in \{0,1\}^n$

Bob
holds graph G

Alice simulates $S_{j^*}^Y$ using bipartition X

Bob forms G' by including edges of G with $w_e = 1$

Communication complexity of D-BHP

Theorem

Let $G = (V, E)$ be sampled from $\mathcal{G}_{n,\alpha/n}$ for $\alpha \in (n^{-1/10}, 1/16)$. Then a one-way protocol with communication $\gamma\sqrt{n}, \gamma \in (n^{-1/10}, 1)$ achieves at most $O(\gamma + \alpha^{3/2})$ advantage over random guessing for D-BHP.

1. Hard input distribution
2. Boolean Hidden Partition Problem (BHP)
3. Analysis of BHP

1. Hard input distribution
2. Boolean Hidden Partition Problem (BHP)
3. Analysis of BHP

Show that distribution of MX in the YES case is close to uniform

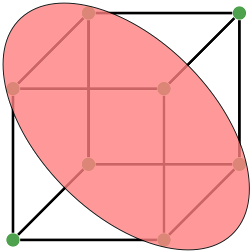
Show that distribution of MX in the YES case is close to uniform

Conditioned on Alice's message, is distribution of MX close to uniform?

Show that distribution of MX in the YES case is close to uniform

Conditioned on Alice's message, is distribution of MX close to uniform?

$X \sim UNIF(A)$ conditioned on m



$|A| \approx 2^{n-s}$

$f(x) :=$ indicator of A

Goal: show that

$$p_M(z) = \mathbf{Pr}[Mx = z | x \in A]$$

is close to uniform

Goal: show that

$$p_M(z) = \mathbf{Pr}[Mx = z | x \in A]$$

is close to uniform

Write $p_M(\cdot)$ in Fourier basis:

$$p_M(z) = \sum_{s \in \{0,1\}^E} \widehat{p}_M(s)(-1)^{s \cdot z}$$

Show that most Fourier mass is in the constant term, i.e. bound

$$\sum_{s \neq \emptyset} \widehat{p}_M(s)^2$$

Gavinsky et al'07:

$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Gavinsky et al'07:

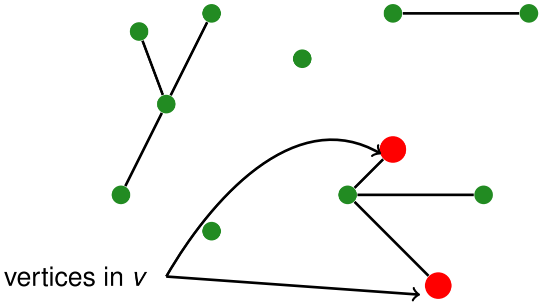
$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Given $v \in \{0,1\}^n$, when do we have $M^T s = v$ for some $s \in \{0,1\}^M$?

Gavinsky et al'07:

$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \hat{f}(M^T s)^2$$

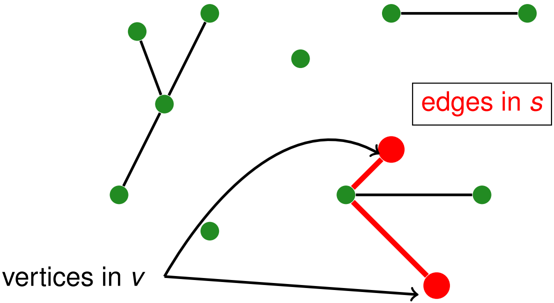
Given $v \in \{0,1\}^n$, when do we have $M^T s = v$ for some $s \in \{0,1\}^M$?



Gavinsky et al'07:

$$\|p_M - UNIF\|_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Given $v \in \{0,1\}^n$, when do we have $M^T s = v$ for some $s \in \{0,1\}^M$?



$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Each element of weight k appears with probability $\approx n^{-k/2}$.

$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Each element of weight k appears with probability $\approx n^{-k/2}$.

Lemma (Gavinsky et al'07; from KKL)

If $f : \{0,1\}^n \to \{0,1\}$ is the indicator function of a set $A \subset \{0,1\}^n$, $|A| \geq 2^{n-s}$, then for every $k \geq 1$,

$$\frac{2^{2n}}{|A|^2} \sum_{z \in \{0,1\}^n, |z|=2k} \widehat{f}(z)^2 \leq (O(s)/k)^{2k}$$

$$||p_M - UNIF||_{TVD} \leq \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Each element of weight k appears with probability $\approx n^{-k/2}$.

Lemma (Gavinsky et al'07; from KKL)

If $f : \{0,1\}^n \to \{0,1\}$ is the indicator function of a set $A \subset \{0,1\}^n$, $|A| \geq 2^{n-s}$, then for every $k \geq 1$,

$$\frac{2^{2n}}{|A|^2} \sum_{z \in \{0,1\}^n, |z|=2k} \widehat{f}(z)^2 \leq (O(s)/k)^{2k}$$

Plugging in $k = 1$, we get $\approx s^2/n$, so $s \ll \sqrt{n}$ suffices

$$\|p_M - UNIF\|_{TVD} \le \frac{2^{2n}}{|A|^2} \sum_{s \in \{0,1\}^M \setminus \{0\}} \widehat{f}(M^T s)^2$$

Each element of weight k appears with probability $\approx n^{-k/2}$.

Lemma (Gavinsky et al'07; from KKL)

If $f : \{0,1\}^n \to \{0,1\}$ is the indicator function of a set $A \subset \{0,1\}^n$, $|A| \ge 2^{n-s}$, then for every $k \ge 1$,

$$\frac{2^{2n}}{|A|^2} \sum_{z \in \{0,1\}^n, |z|=2k} \widehat{f}(z)^2 \le (O(s)/k)^{2k}$$

Plugging in $k = 1$, we get $\approx s^2/n$, so $s \ll \sqrt{n}$ suffices

Fourier mass bounds fairly tight for a coordinate subspace...

$(1+\Omega(1))$-Approximation to MAX-CUT Requires Linear Space

Main result

Theorem (K.-Khanna-Sudan-Velingker'17)

There exists a constant $\varepsilon_* > 0$ such that a single pass streaming algorithm for approximating MAX-CUT value to factor $1+\varepsilon_*$ requires $\Omega(n)$ space.

Q1: A poly($\log n$) space approximation scheme?

Q1: A poly$(\log n)$ space approximation scheme?

NO:

Better than factor 2 requires $\Omega(\sqrt{n})$ space K-Khanna-Sudan'14

$(1+\varepsilon)$-approximation requires $n^{1-O(\varepsilon)}$ space K-Khanna-Sudan'14, Kogan-Krauthgamer'14

Q1: A poly($\log n$) space approximation scheme?

NO:

Better than factor 2 requires $\Omega(\sqrt{n})$ space K-Khanna-Sudan'14

$(1+\varepsilon)$-approximation requires $n^{1-O(\varepsilon)}$ space K-Khanna-Sudan'14, Kogan-Krauthgamer'14

Q2: For every $1 < \alpha < 2$ there exists $0 \le \beta < 1$ such that α-approximation can be achieved in n^{β} space?

Q1: A poly($\log n$) space approximation scheme?

NO:
Better than factor 2 requires $\Omega(\sqrt{n})$ space K-Khanna-Sudan'14

$(1+\varepsilon)$-approximation requires $n^{1-O(\varepsilon)}$ space K-Khanna-Sudan'14, Kogan-Krauthgamer'14

Q2: For every $1 < \alpha < 2$ there exists $0 \leq \beta < 1$ such that α-approximation can be achieved in n^{β} space?

this result: NO

Q1: A poly($\log n$) space approximation scheme?

NO:
Better than factor 2 requires $\Omega(\sqrt{n})$ space K-Khanna-Sudan'14

$(1+\varepsilon)$-approximation requires $n^{1-O(\varepsilon)}$ space K-Khanna-Sudan'14, Kogan-Krauthgamer'14

Q2: For every $1 < \alpha < 2$ there exists $0 \le \beta < 1$ such that α-approximation can be achieved in n^{β} space?

this result: NO

Q3: There exist $1 < \alpha_* < 2$ and $0 \le \beta_* < 1$ such that α_*-approximation can be achieved in n^{β_*} space?

Q1: A poly($\log n$) space approximation scheme?

NO:
Better than factor 2 requires $\Omega(\sqrt{n})$ space K-Khanna-Sudan'14

$(1+\varepsilon)$-approximation requires $n^{1-O(\varepsilon)}$ space K-Khanna-Sudan'14, Kogan-Krauthgamer'14

Q2: For every $1 < \alpha < 2$ there exists $0 \leq \beta < 1$ such that α-approximation can be achieved in n^{β} space?

this result: NO

Q3: There exist $1 < \alpha_* < 2$ and $0 \leq \beta_* < 1$ such that α_*-approximation can be achieved in n^{β_*} space?

???

Hard distribution on MAX-CUT instances

YES: random bipartite graph with $\approx$ constant degrees

NO: non-bipartite graph with $\approx$ constant degrees

Hard distribution on MAX-CUT instances

YES: random bipartite graph with ≈ constant degrees

NO: non-bipartite graph with ≈ constant degrees

1. ensure MAX-CUT value gap between NO case and YES case
2. show $\Omega(n)$ space required to distinguish between the two cases

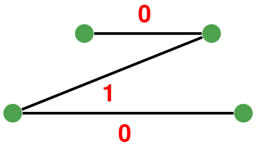
1. Implicit hidden partition problem
2. Reduction from MAX-CUT
3. Communication problem analysis via Fourier techniques

1. Implicit hidden partition problem

2. Reduction from MAX-CUT

3. Communication problem analysis via Fourier techniques

Implicit Hidden Partition Problem

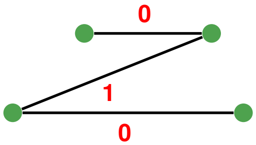


Player 1

graph G_1, labels

w^1 on edges

Implicit Hidden Partition Problem

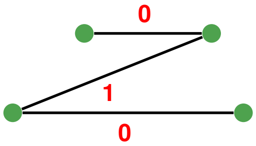


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

Implicit Hidden Partition Problem

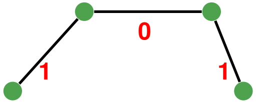


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮ ⋮

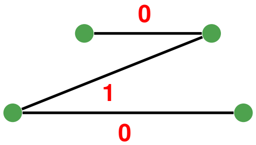


Player T

graph G_T, labels

w^T on edges

Implicit Hidden Partition Problem

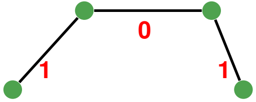


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

$\vdots$

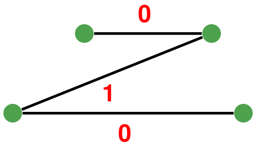


Player T $\longrightarrow$ m_T

graph G_T, labels

w^T on edges

Implicit Hidden Partition Problem

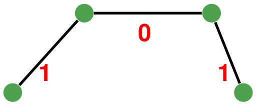


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮



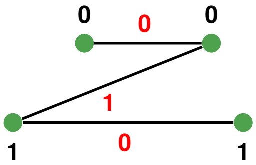
Player T $\longrightarrow$ m_T

graph G_T, labels

w^T on edges

YES case: $\exists$ partition $x \in \{0,1\}^n$ such that $w^t = M^t x$ for $1 \leq t \leq T$

Implicit Hidden Partition Problem

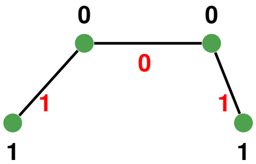


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮ ⋮



Player T $\longrightarrow$ m_T

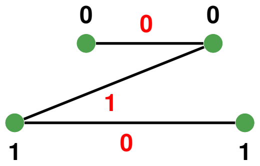
graph G_T, labels

w^T on edges

YES case: $\exists$ partition $x \in \{0,1\}^n$ such that $w^t = M^t x$ for $1 \le t \le T$

Implicit Hidden Partition Problem

$w_{uv} = x_u + x_v$

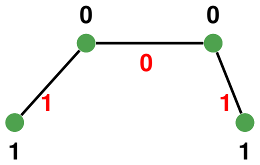


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

$\vdots$



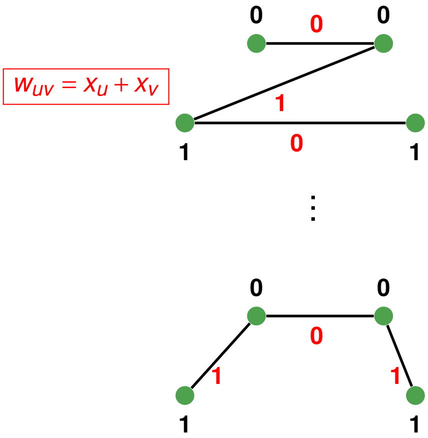
Player T $\longrightarrow$ m_T

graph G_T, labels

w^T on edges

YES case: $\exists$ partition $x \in \{0,1\}^n$ such that $w^t = M^t x$ for $1 \leq t \leq T$

Implicit Hidden Partition Problem



Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮

Player T $\longrightarrow$ m_T

graph G_T, labels

w^T on edges

YES case: $\exists$ partition $x \in \{0,1\}^n$ such that $w^t = M^t x$ for $1 \le t \le T$

NO case: no such partition exists

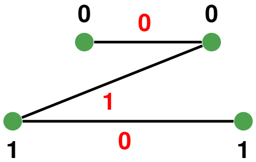
Distributional communication problem

Choose a hidden partition $X \in UNIF(\{0,1\}^n)$

Distributional communication problem

Choose a hidden partition $X \in UNIF(\{0,1\}^n)$

$w_{uv} = x_u + x_v$

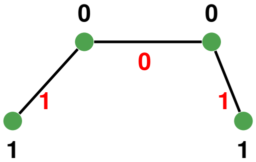


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮



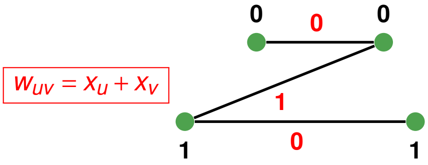
Player T $\longrightarrow$ m_T

graph G_T, labels

w^T on edges

Distributional communication problem

Choose a hidden partition $X \in UNIF(\{0,1\}^n)$

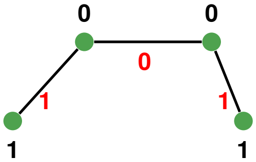


Player 1 $\longrightarrow$ m_1

graph G_1, labels

w^1 on edges

⋮



Player T $\longrightarrow$ m_T

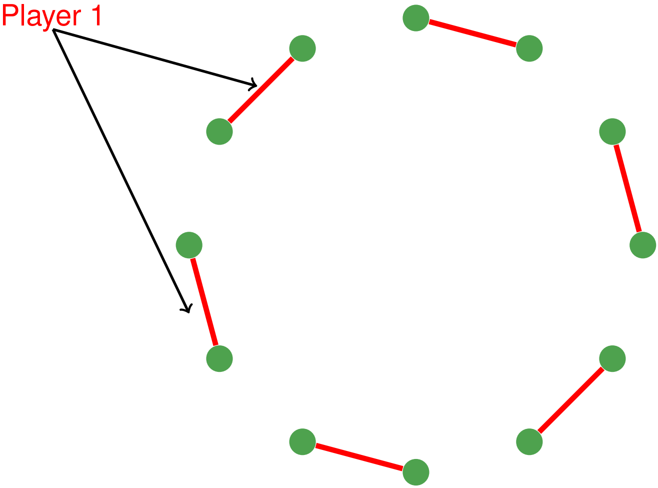
graph G_T, labels

w^T on edges

YES case: labels satisfy $w^t = M^t X$ for $1 \leq t \leq T$

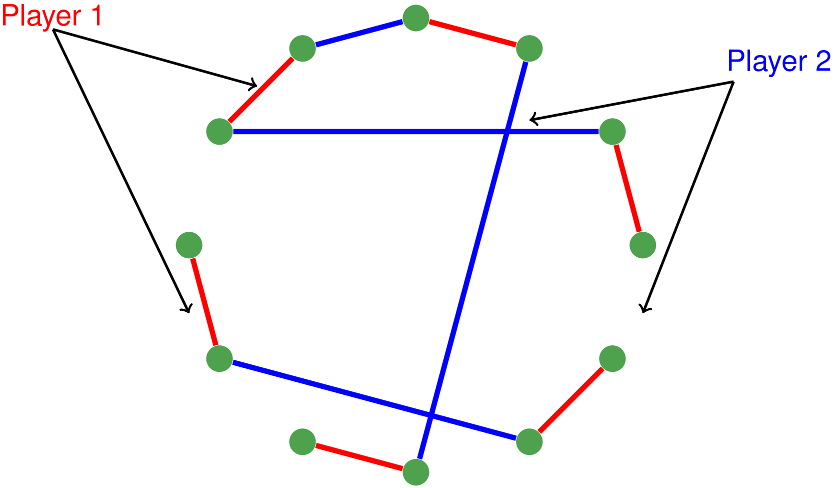
NO case: labels are random: $w^t \sim UNIF$

Distribution on players' graphs



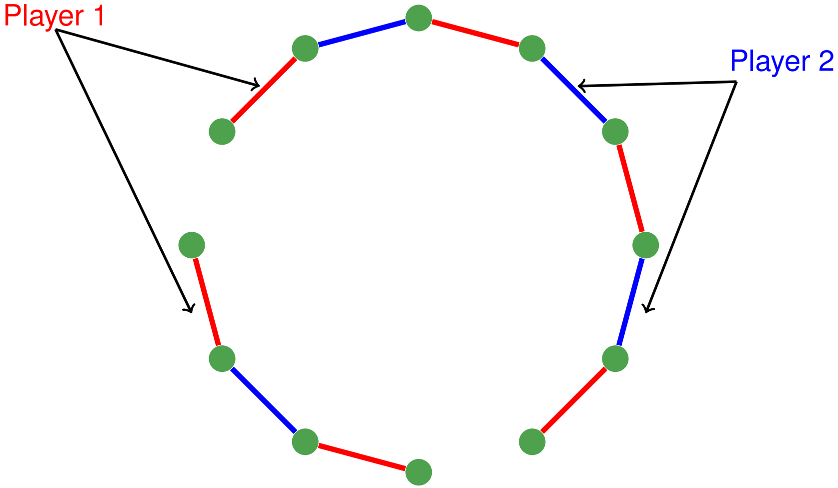
G_1 a perfect matching

Distribution on players' graphs



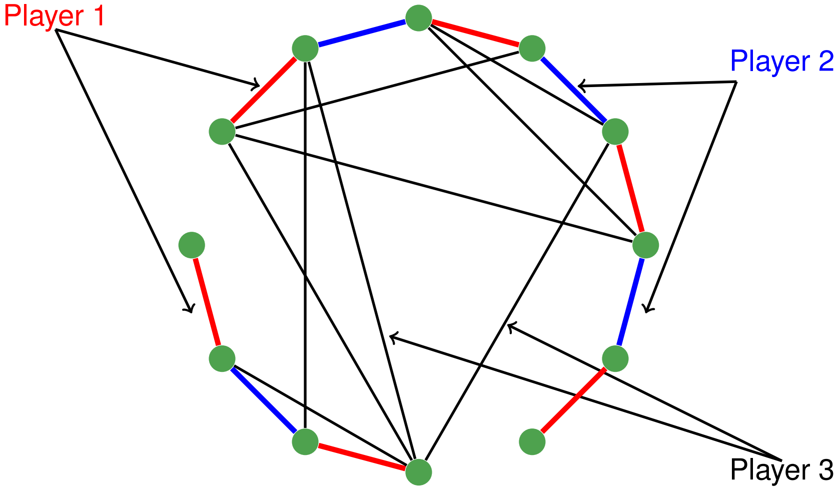
G_1 a perfect matching, G_2 a (random) near perfect matching

Distribution on players' graphs



G_1 a perfect matching, G_2 a (random) near perfect matching

Distribution on players' graphs



G_1 a perfect matching, G_2 a (random) near perfect matching, G_3 an Erdős-Rényi graph

1. Implicit hidden partition problem
2. Reduction from MAX-CUT
3. Communication problem analysis via Fourier techniques

1. Implicit hidden partition problem
2. Reduction from MAX-CUT
3. Communication problem analysis via Fourier techniques

Reduction from MAX-CUT

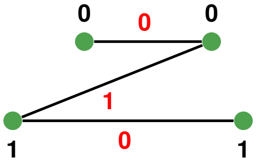
YES: random bipartite graph with ≈ constant degrees

NO: non-bipartite graph with ≈ constant degrees

Reduction from MAX-CUT

YES: random bipartite graph with ≈ constant degrees

NO: non-bipartite graph with ≈ constant degrees



Player t $\longrightarrow$ m_t

graph G_t, labels

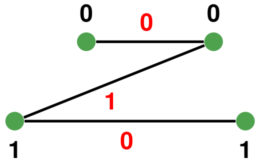
w^t on edges

t-th player generates graph G'_t by including edges $e \in G_t$ with $w^t_e = 1$

Reduction from MAX-CUT

YES: random bipartite graph with ≈ constant degrees

NO: non-bipartite graph with ≈ constant degrees



Player t $\longrightarrow$ m_t

graph G_t, labels

w^t on edges

t-th player generates graph G'_t by including edges $e \in G_t$ with $w^t_e = 1$

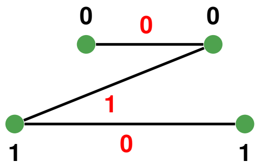
YES case: labels satisfy $w^t = M^t X$ for $1 \leq t \leq T$

$\cup_t G'_t$ is bipartite

Reduction from MAX-CUT

YES: random bipartite graph with ≈ constant degrees

NO: non-bipartite graph with ≈ constant degrees



Player t $\longrightarrow$ m_t

graph G_t, labels

w^t on edges

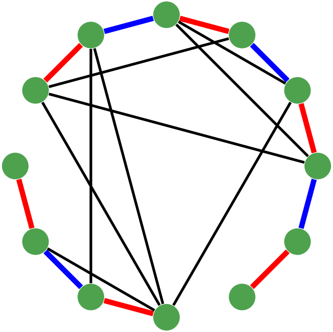
t-th player generates graph G'_t by including edges $e \in G_t$ with $w^t_e = 1$

YES case: labels satisfy $w^t = M^t X$ for $1 \leq t \leq T$

$\cup_t G'_t$ is bipartite

NO case: labels are random: $w^t \sim UNIF$

$\cup_t G'_t$ is a sample of $\cup_t G_t$ at rate $1/2$



Distributional Implicit Hidden Partition Problem (DIHP): G_1 a perfect matching, G_2 a (random) near perfect matching, G_3 an Erdős-Rényi graph close to the giant component threshold

Theorem

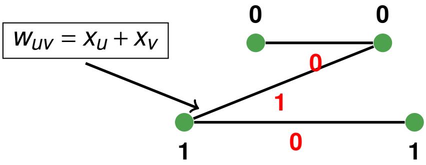
If $G_i(1/2), i = 1,2,3$ is G_i subsampled at rate $1/2$, then $G_1(1/2) \cup G_2(1/2) \cup G_3(1/2)$ is $\Omega(1)$-far from bipartite with high probability.

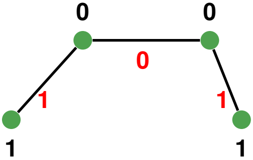
1. Implicit hidden partition problem

2. Reduction from MAX-CUT

3. Communication problem analysis via Fourier techniques

1. Implicit hidden partition problem
2. Reduction from MAX-CUT
3. Communication problem analysis via Fourier techniques





Player 0

bipartition $X \in \{0,1\}^n$ $\rightarrow$ m_0

Player 1 $\longrightarrow$ m_1

graph G_1, labels

$w^1 = M^1 X$ on edges

$\vdots$

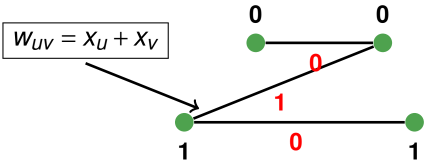
Player 3 $\longrightarrow$ m_3

graph G_3, labels

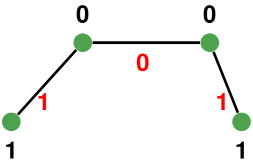
$w^3 = M^3 X$ on edges

player 0 dominates communication!

K.-Khanna-Sudan'15



⋮



Player 0

bipartition $X \in \{0,1\}^n$ → m_0

Player 1 ⟶ m_1

graph G_1, labels

$w^1 = M^1 X$ on edges

⋮

Player 3 ⟶ m_3

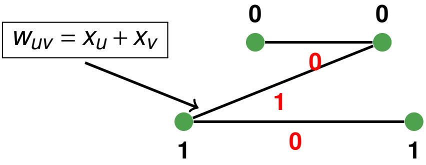
graph G_3, labels

$w^3 = M^3 X$ on edges

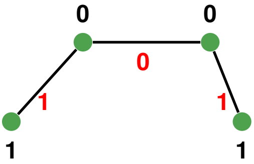
Our approach: Implicit Hidden Partition Problem

player 0 dominates communication!

K.-Khanna-Sudan'15



⋮



Player 0

bipartition $X \in \{0,1\}^n$ $\rightarrow$ m_0

Player 1 $\longrightarrow$ m_1

graph G_1, labels

$w^1 = M^1 X$ on edges

⋮

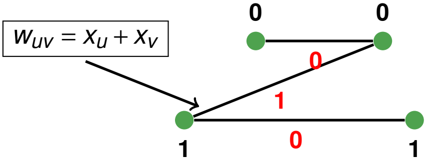
Player 3 $\longrightarrow$ m_3

graph G_3, labels

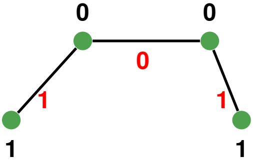
$w^3 = M^3 X$ on edges

Our approach: Implicit Hidden Partition Problem

information about X revealed implicitly!



⋮



~~Player 0~~
~~bipartition $X \in \{0,1\}^n$~~ $\rightarrow m_0$

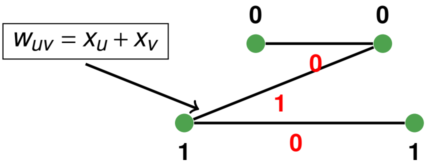
Player 1 $\longrightarrow m_1$
graph G_1, labels
$w^1 = M^1 X$ on edges

⋮

Player 3 $\longrightarrow m_3$
graph G_3, labels
$w^3 = M^3 X$ on edges

Our approach: Implicit Hidden Partition Problem

information about X revealed implicitly!

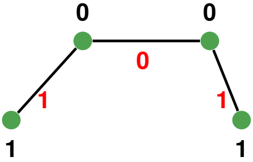


Player 1 $\longrightarrow$ m_1

graph G_1, labels

$w^1 = M^1 X$ on edges

⋮

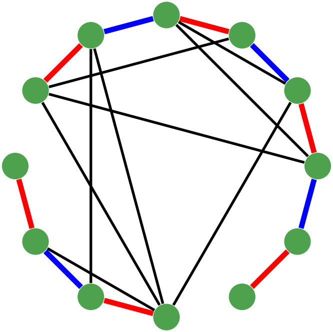


Player 3 $\longrightarrow$ m_3

graph G_3, labels

$w^3 = M^3 X$ on edges

Communication complexity of D-IHP



Theorem

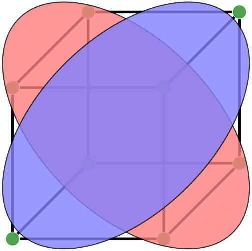
Any one-way protocol with communication $o(n)$ achieves at most $o(1)$ advantage over random guessing for D-IHP.

Fourier analysis (convolution theorem) and graph theoretic considerations.

Conditioned on messages of player 1 and player 2, is distribution of $M_3 X$ close to uniform?

Conditioned on messages of player 1 and player 2, is distribution of $M_3 X$ close to uniform?

$X \sim UNIF(A_1 \cap A_2)$
conditioned on (m_1, m_2)



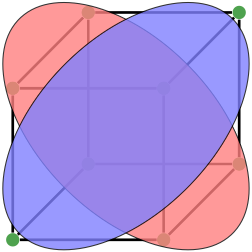
$|A_1| \approx 2^{n-s}, |A_2| \approx 2^{n-s}$

$f_1(x)$:=indicator of A_1
$f_2(x)$:=indicator of A_2

The indicator of $A_1 \cap A_2$ is $f_1 \cdot f_2$.

Conditioned on messages of player 1 and player 2, is distribution of $M_3 X$ close to uniform?

$X \sim UNIF(A_1 \cap A_2)$
conditioned on (m_1, m_2)



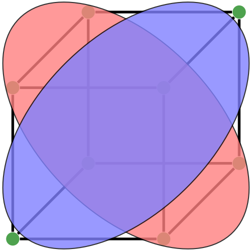
$|A_1| \approx 2^{n-s}, |A_2| \approx 2^{n-s}$

$f_1(x) :=$indicator of A_1
$f_2(x) :=$indicator of A_2

The indicator of $A_1 \cap A_2$ is $f_1 \cdot f_2$. Will prove that for $k \geq 1$

$$\frac{2^{2n}}{|A_1 \cap A_2|^2} \sum_{\substack{v \in \{0,1\}^n \\ |v| = \mathbf{2k}}} \widehat{f_1 \cdot f_2}(v)^2 \leq (O(s)/k)^{\mathbf{k}}$$

$X \sim UNIF(A_1 \cap A_2)$
conditioned on (m_1, m_2)



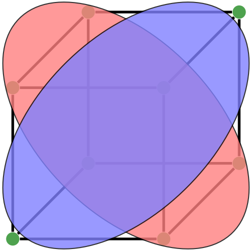
$|A_1| \approx 2^{n-s}, |A_2| \approx 2^{n-s}$

$f_1(x)$:=indicator of A_1
$f_2(x)$:=indicator of A_2

Players only access X via $M_i X$, so $\widehat{f_i}$ is **supported on edges** and has strong spectral properties:

$$2^{2s} \sum_{|v|=\mathbf{2k}} \widehat{f_i}(v)^2 \leq (O(s)/k)^{\mathbf{k}}$$

$X \sim UNIF(A_1 \cap A_2)$
conditioned on (m_1, m_2)



$|A_1| \approx 2^{n-s}, |A_2| \approx 2^{n-s}$

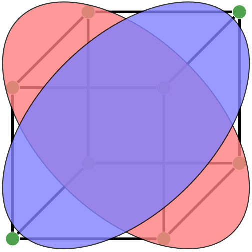
$f_1(x)$:=indicator of A_1
$f_2(x)$:=indicator of A_2

Players only access X via $M_i X$, so $\widehat{f_i}$ is **supported on edges** and has strong spectral properties:

$$2^{2s} \sum_{|v|=\mathbf{2k}} \widehat{f_i}(v)^2 \leq (O(s)/k)^{\mathbf{k}}$$

Intuition: with s space can only learn about $\approx s$ pairs
Prior work, with player 0: with s space can only learn about $\approx s^2$ pairs

$X \sim UNIF(A_1 \cap A_2)$
conditioned on (m_1, m_2)



$|A_1| \approx 2^{n-s}, |A_2| \approx 2^{n-s}$

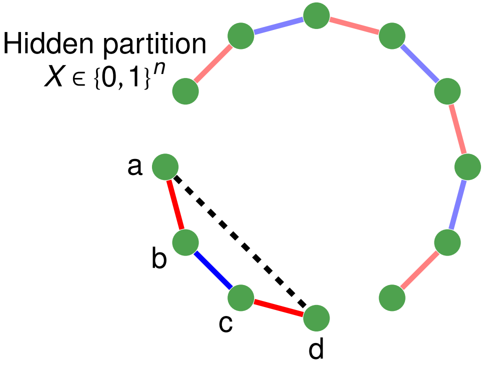
$f_1(x) :=$indicator of A_1
$f_2(x) :=$indicator of A_2

Players only access X via $M_i X$, so $\widehat{f}_i$ is **supported on edges** and has strong spectral properties:

$$2^{2s} \sum_{|v|=\mathbf{2k}} \widehat{f}_i(v)^2 \leq (O(s)/k)^k$$

The indicator of $A_1 \cap A_2$ is $f_1 \cdot f_2$, so by the convolution theorem

$$\widehat{f_1 \cdot f_2} = \widehat{f_1} * \widehat{f_2}$$



Intuition: $\widehat{f_1}(a,b,c,d)^2 \approx$ how much information player 1 transmits about parity $X_a + X_b + X_c + X_d$

$\widehat{f_2}(b,c)^2 \approx$ how much information player 2 transmits about parity $X_b + X_c$

$\widehat{f_1 \cdot f_2}(a,d)^2 = \widehat{f_1}(a,b,c,d)^2 \cdot \widehat{f_2}(b,c)^2 \approx$ how much information players 1 and 2 transmit about parity $X_a + X_d$

For any $\ell \geq 0$,

$$\sum_{\substack{v\in\{0,1\}^n,\\ |v|=2\ell}} \widehat{f_1 \cdot f_2}(v)^2 = \sum_{k\geq 0} \underbrace{\sum_{w\in\{0,1\}^n,|w|=2k} \widehat{f}_1(w)^2}_{\text{large for } k\gg\ell!} \cdot \underbrace{\left(\sum_{v\in\{0,1\}^n,|v|=2\ell} \widehat{f}_2(w+v)^2\right)}_{\text{small for } k\gg l?}$$

For any $\ell \geq 0$,

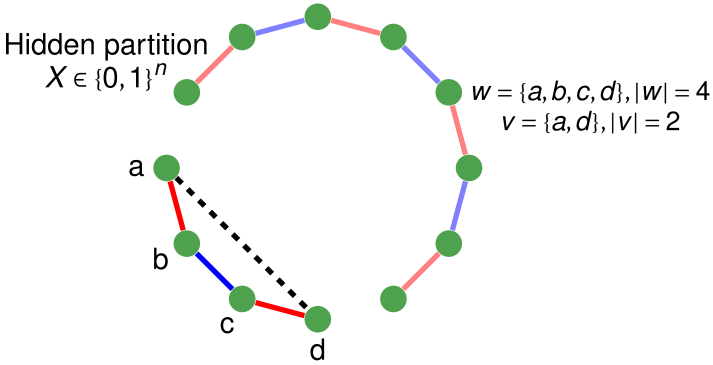
$$\sum_{\substack{v\in\{0,1\}^n,\\|v|=2\ell}} \widehat{f_1 \cdot f_2}(v)^2 = \sum_{k\geq 0} \underbrace{\sum_{w\in\{0,1\}^n,|w|=2k} \widehat{f}_1(w)^2}_{\text{large for } k\gg\ell!} \cdot \underbrace{\left(\sum_{v\in\{0,1\}^n,|v|=2\ell} \widehat{f}_2(w+v)^2 \right)}_{\text{small for } k\gg l?}$$

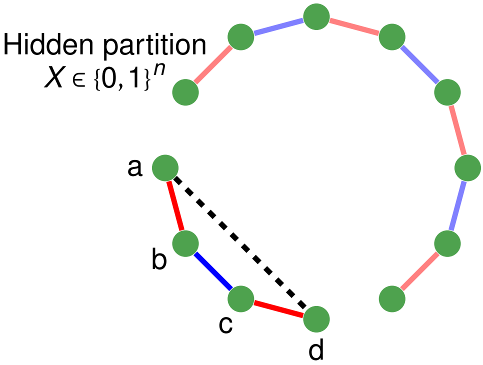
Show that the last term decays for $k > l$?

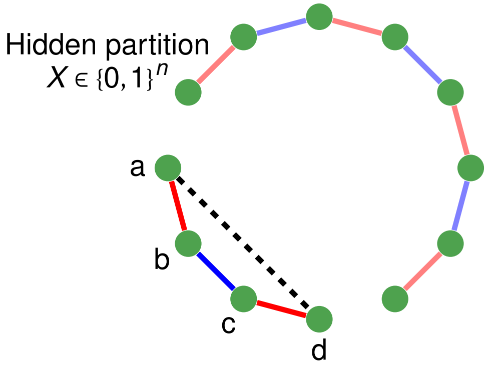
For any $\ell \geq 0$,

$$\sum_{\substack{v \in \{0,1\}^n, \\ |v|=2\ell}} \widehat{f_1 \cdot f_2}(v)^2 = \sum_{k \geq 0} \underbrace{\sum_{w \in \{0,1\}^n, |w|=2k} \widehat{f_1}(w)^2}_{\text{large for } k \gg \ell!} \cdot \underbrace{\left(\sum_{v \in \{0,1\}^n, |v|=2\ell} \widehat{f_2}(w+v)^2 \right)}_{\text{small for } k \gg l?}$$

Show that the last term decays for $k > l$?




Hidden partition
$X \in \{0,1\}^n$
a
b
c
d

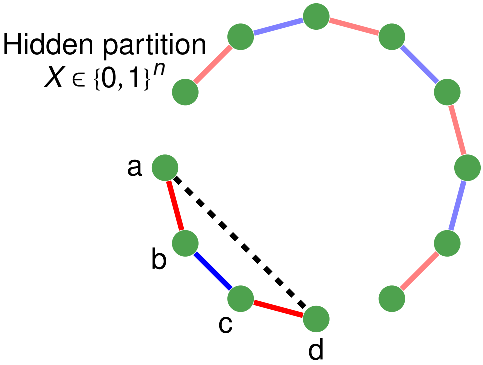


Open problems

Any improvement over factor 2 requires $\Omega(n)$ space?

$(2 - \varepsilon_*)$-approximation in $n^{1-\delta}$ space?

Analyze $\widehat{f}_1 * \widehat{f}_2 * \cdots * \widehat{f}_T$ for large T?



Open problems

Any improvement over factor 2 requires $\Omega(n)$ space?

$(2-\varepsilon_*)$-approximation in $n^{1-\delta}$ space?

Analyze $\widehat{f}_1 * \widehat{f}_2 * \cdots * \widehat{f}_T$ for large T?

Thank you!