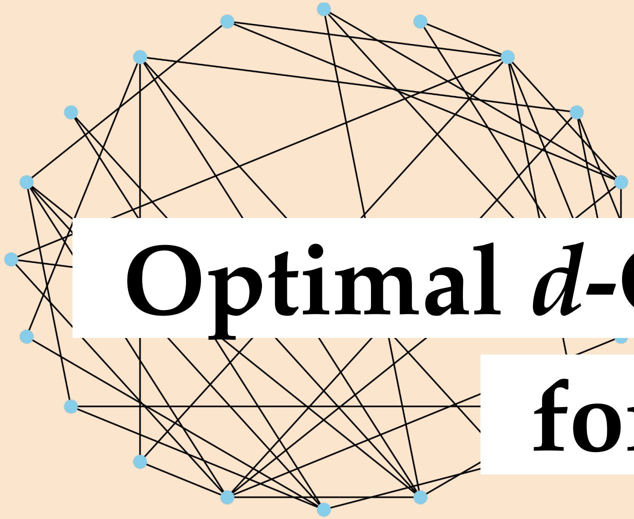
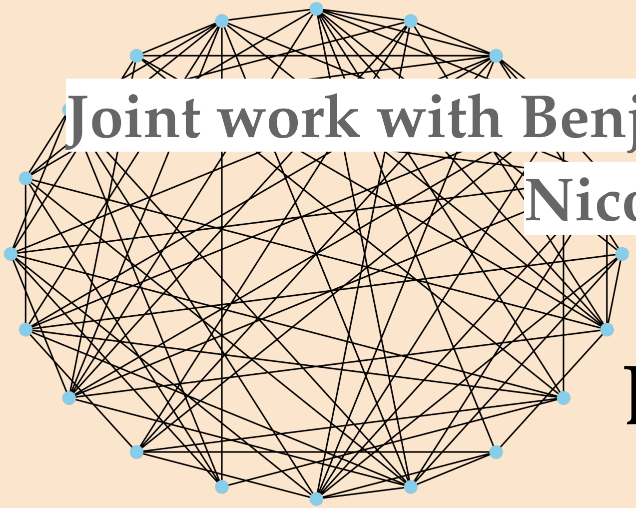
Optimal *d*-Clique Decompositions for Hypergraphs

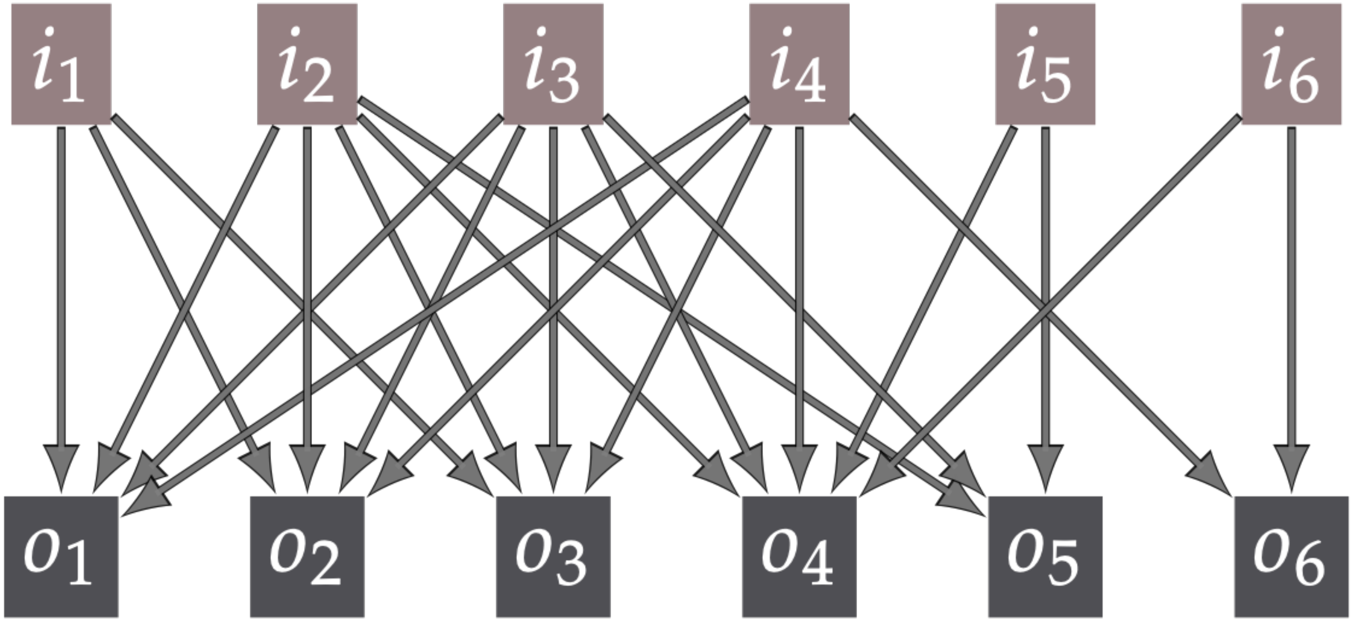




Joint work with Benjamin Przybocki, Andrew Krapiving, and Nicolás Sanhueza-Matamala

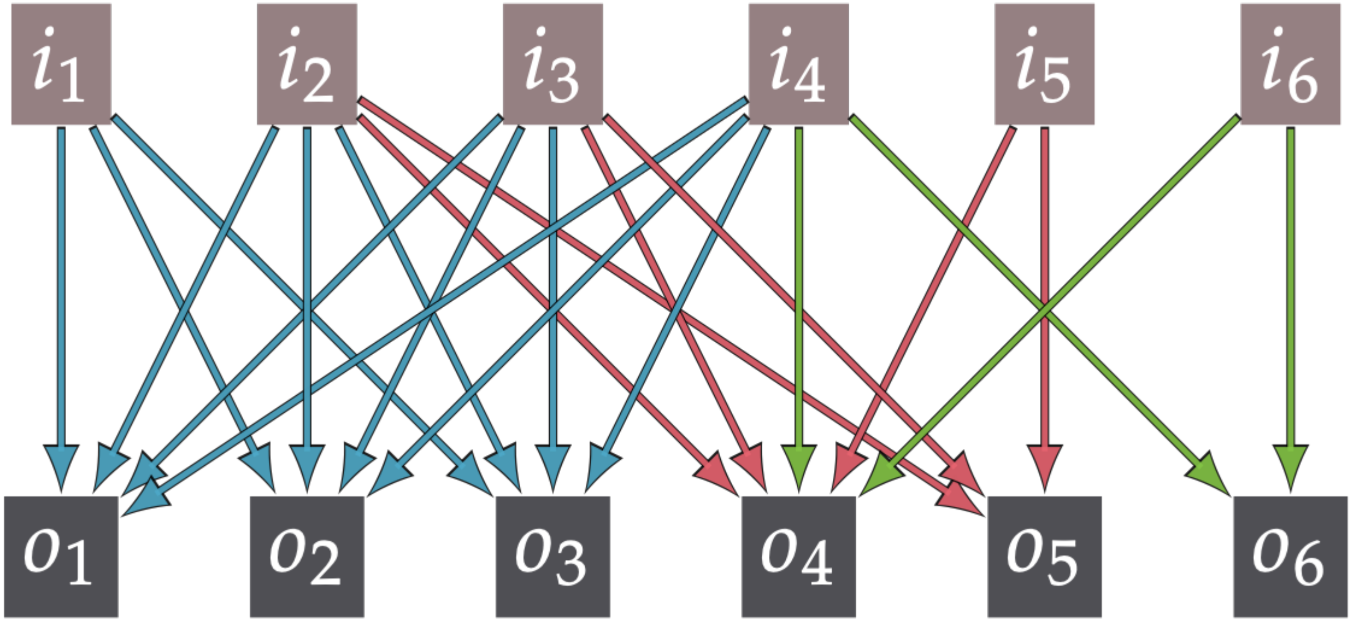
Bernardo Subercaseaux (CMU)

Suppose we have a dense electric circuit



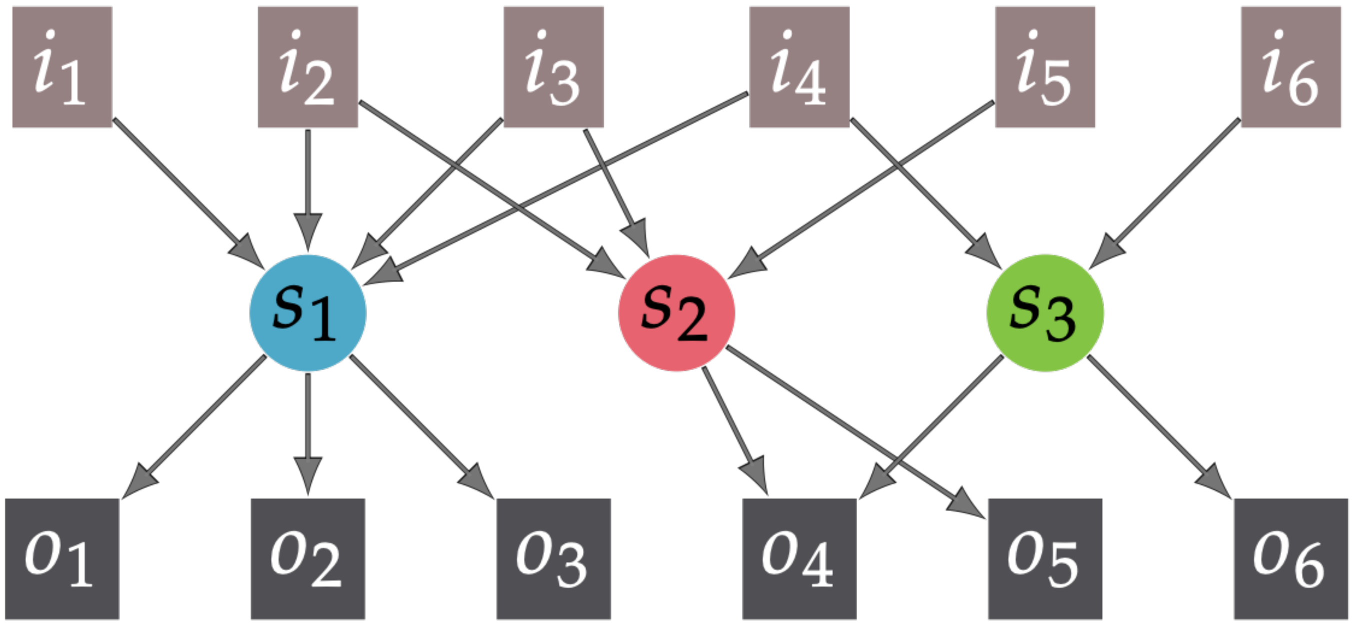
And we want to **reduce the number of wires**…

Lupanov (1956) noted that if we identify ***bicliques***



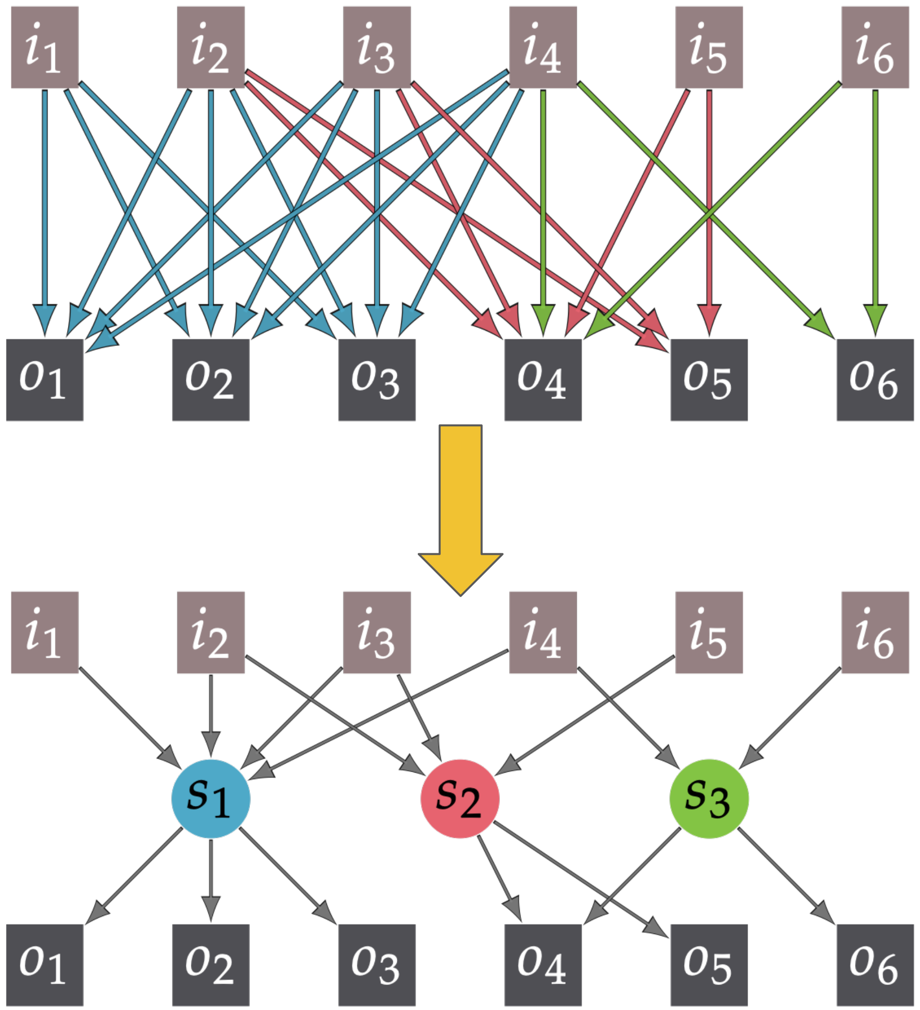
we can easily reduce each of them!

Lupanov (1956) noted that if we identify ***bicliques***

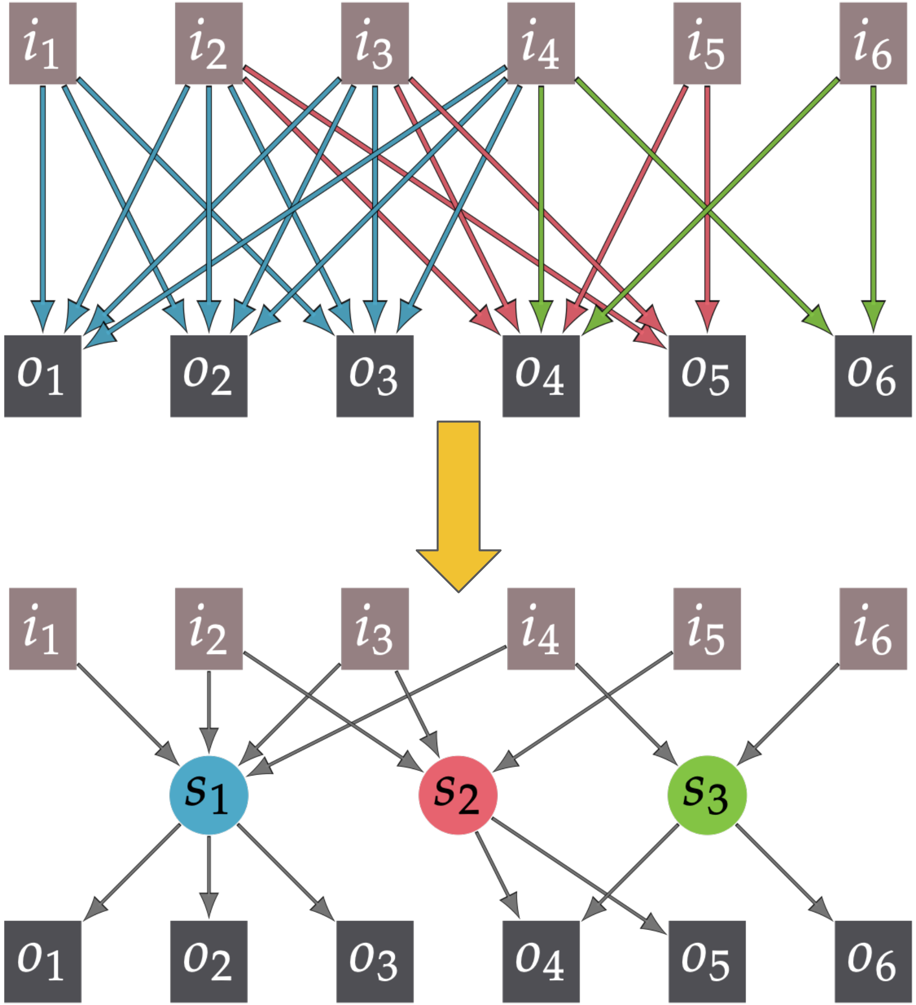


we can easily reduce each of them!

Introducing a *switch* we can reduce *ab* wires to $a + b$ if we identify a $K_{a,b}$



Introducing a *switch* we can reduce *ab* wires to $a + b$ if we identify a $K_{a,b}$



Reduction factor

$f(a, b) := ab/(a+b)$

Definition 1. *A biclique partition of a graph G is a set* $\mathcal{B} := \{B_1, \ldots, B_k\}$ *where each* B_i *is a complete bipartite graph, and such that* $E(G) = \bigsqcup_{i=1}^{k} E(B_i)$*. The* weight $w(\mathcal{B})$ *of such a partition is simply* $w(\mathcal{B}) := \sum_{i=1}^{k} |V(B_i)|$.

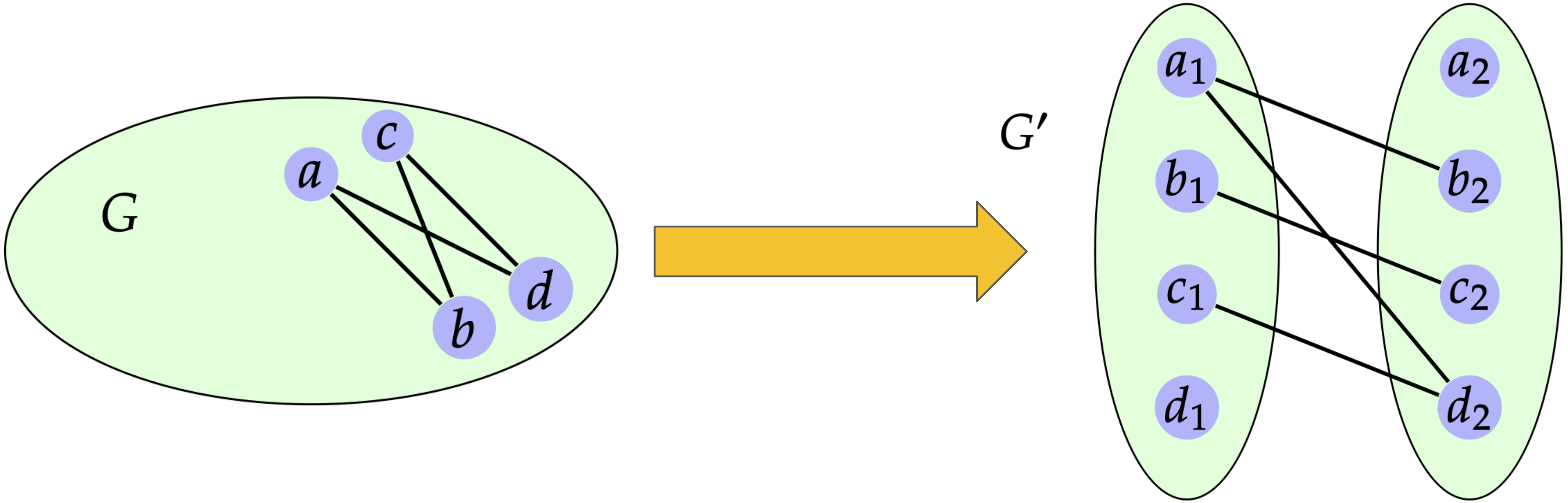
Definition 1. *A biclique partition of a graph G is a set* $\mathcal{B} := \{B_1, \ldots, B_k\}$ *where each* B_i *is a complete bipartite graph, and such that* $E(G) = \bigsqcup_{i=1}^{k} E(B_i)$*. The* weight $w(\mathcal{B})$ *of such a partition is simply* $w(\mathcal{B}) := \sum_{i=1}^{k} |V(B_i)|$.

Lupanov's Theorem (1956)

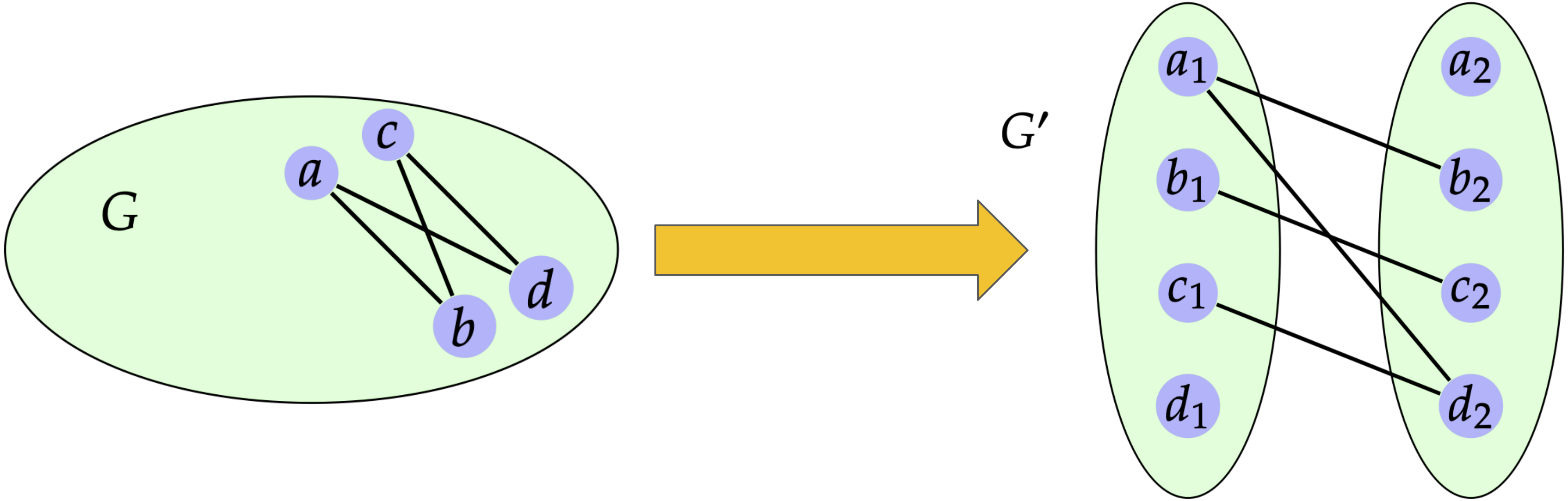
Let $G = (X \sqcup Y, E)$ be a bipartite graph with $|X| = |Y| = n$. Then, G admits a biclique partition of weight at most $(1 + o(1))\frac{n^2}{\lg n}$.

Moreover, there are some balanced bipartite graphs G for which every biclique partition has weight $(1 - o(1))\frac{n^2}{\lg n}$.

Lupanov for general graphs

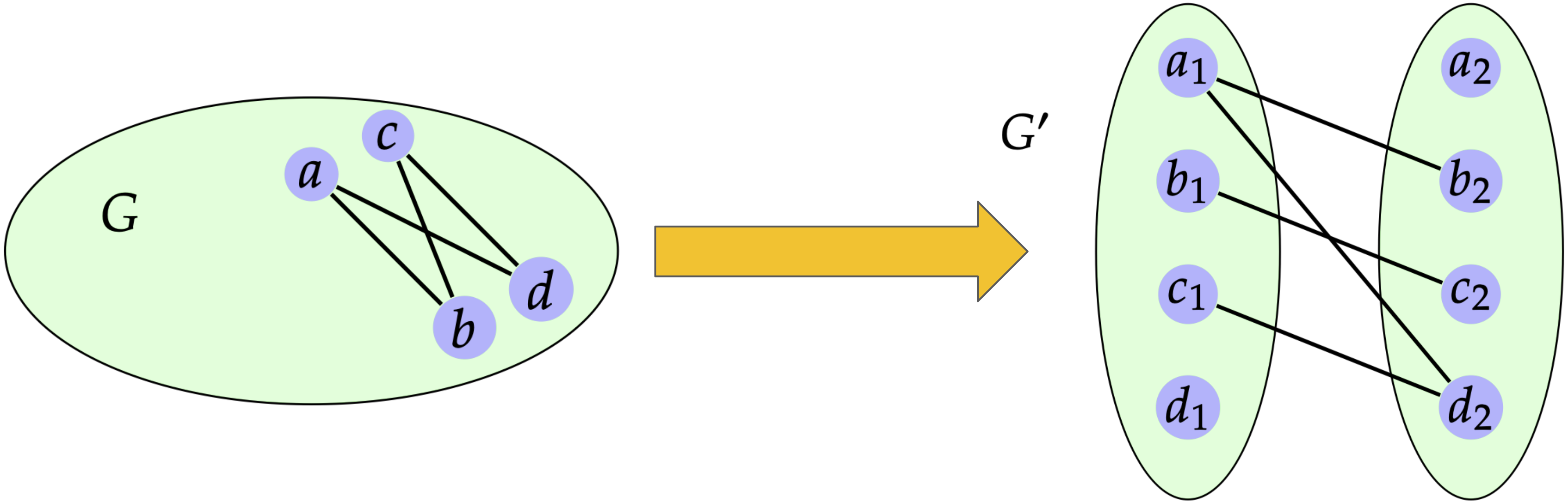


Lupanov for general graphs



Bicliques in G' correspond to bicliques in G (but not conversely)

Lupanov for general graphs



Bicliques in G' correspond to bicliques in G (but not conversely)

$$\left(\tfrac{1}{4} - o(1)\right)\frac{n^2}{\lg n} \leqslant BP(n) \leqslant (1 + o(1))\frac{n^2}{\lg n}.$$

Chung–Erdős–Spencer (1983)

$$\left(0.265 - o(1)\right)\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant \left(0.721 + o(1)\right)\frac{n^2}{\lg n}$$

Lower bound: random graph with edge probability *1/e*

$$\Pr[\text{contain } K_{a,b}] \leq \binom{n}{a}\binom{n}{b}e^{-ab} < e^{(a+b)\ln n - ab}$$

Chung–Erdős–Spencer (1983)

$$(0.265 - o(1))\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant (0.721 + o(1))\frac{n^2}{\lg n}$$

Lower bound: random graph with edge probability *1/e*

$$\Pr[\text{contain } K_{a,b}] \leq \binom{n}{a}\binom{n}{b}e^{-ab} < e^{(a+b)\ln n - ab}$$

Say a $K_{a,\,b}$ is "good" if ab/(a+b) ≥ ln n / (1-ε)

Chung–Erdős–Spencer (1983)

$$(0.265 - o(1))\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant (0.721 + o(1))\frac{n^2}{\lg n}$$

Lower bound: random graph with edge probability *1/e*

$$\Pr[\text{contain } K_{a,b}] \leq \binom{n}{a}\binom{n}{b}e^{-ab} < e^{(a+b)\ln n - ab}$$

Say a $K_{a,\, b}$ is "good" if ab/(a+b) ≥ ln n / (1-ε)

$$\Pr[\text{contain some good } K_{a,b}] \leq \sum_{(a,b)\text{ is good}} e^{-\varepsilon ab} \leq n^2 e^{-\varepsilon \ln^2 n} < 1$$

Chung–Erdős–Spencer (1983)

$$\left(0.265 - o(1)\right)\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant \left(0.721 + o(1)\right)\frac{n^2}{\lg n}$$

Upper bound: iteratively peel good bicliques

Theorem 1 (Baby KST). *Let* $E(G) = \gamma\binom{n}{2}$, *then G contains a* $K_{s,t}$ *such that the 'reduction factor'* $st/(s+t)$ *is* $(1-o(1))\lg n \cdot \lg(1/\gamma)$.

Chung–Erdős–Spencer (1983)

$$(0.265 - o(1))\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant (0.721 + o(1))\frac{n^2}{\lg n}$$

Upper bound: iteratively peel good bicliques

Theorem 1 (Baby KST). *Let $E(G) = \gamma\binom{n}{2}$, then G contains a $K_{s,t}$ such that the 'reduction factor' $st/(s+t)$ is $(1 - o(1))\lg n \cdot \lg(1/\gamma)$.*

So if G has $O(n^2/\lg n)$ edges, then take each edge as a biclique.
Otherwise peel a biclique with reduction factor $\lg n$ and repeat!

Chung–Erdős–Spencer (1983)

$$(0.265 - o(1))\,\frac{n^2}{\lg n} \leqslant BC(n) \leqslant BP(n) \leqslant (0.721 + o(1))\,\frac{n^2}{\lg n}$$

In this paper, we have shown that $\alpha_{f_1}(n;\mathbf{B}) = O(n^2/\log n)$. However, we do not know the existence of

$$\lim_{n\to\infty} \frac{\alpha_{f_1}(n;\mathbf{B})}{n^2/\log n} \quad \text{or} \quad \lim_{n\to\infty} \frac{\beta_{f_1}(n;\mathbf{B})}{n^2/\log n},$$

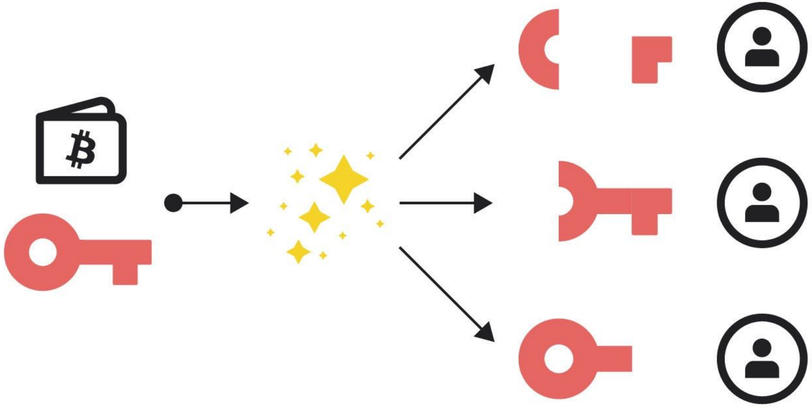
obviously.

Erdős–Pyber (1997)

Theorem 1 (Erdős–Pyber). *Every graph G admits a biclique partition in which every vertex v belongs to $O(n/\lg n)$ bicliques.*

Erdős–Pyber (1997)

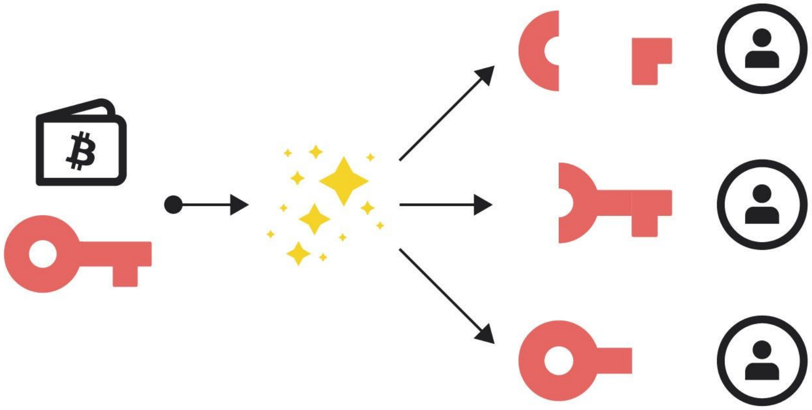
Theorem 1 (Erdős–Pyber). *Every graph G admits a biclique partition in which every vertex v belongs to $O(n/\lg n)$ bicliques.*

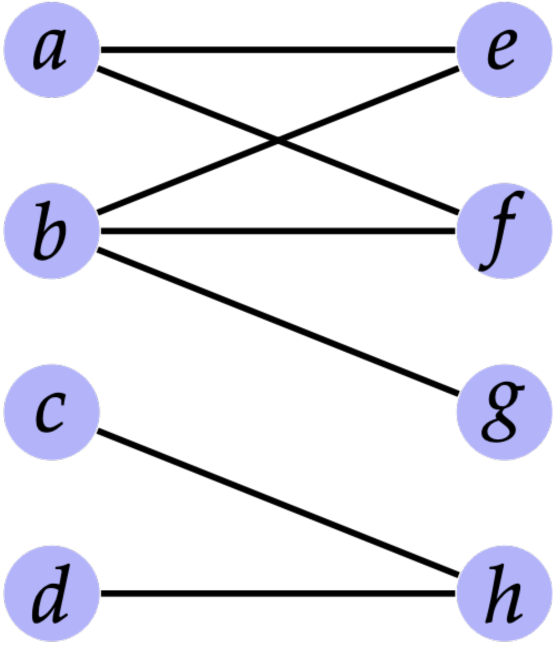


Important for secret sharing!

Erdős–Pyber (1997)

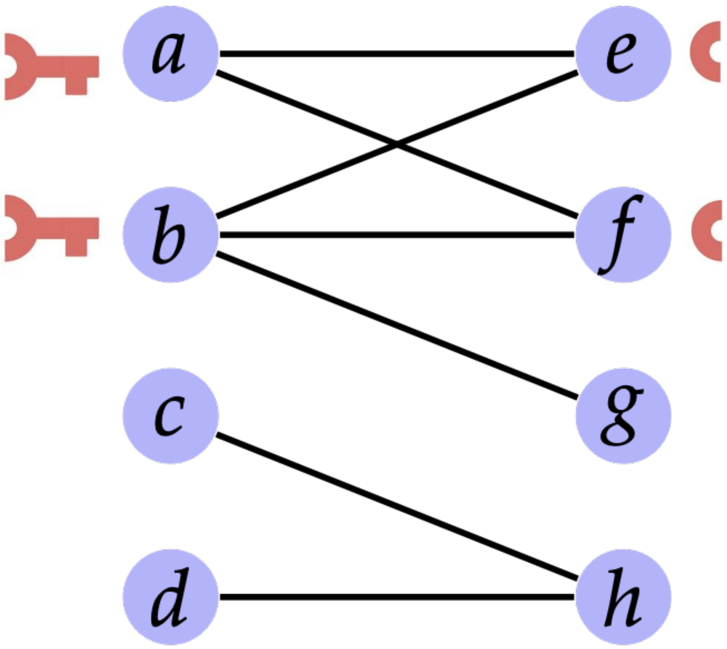
Theorem 1 (Erdős–Pyber). *Every graph G admits a biclique partition in which every vertex v belongs to $O(n/\lg n)$ bicliques.*





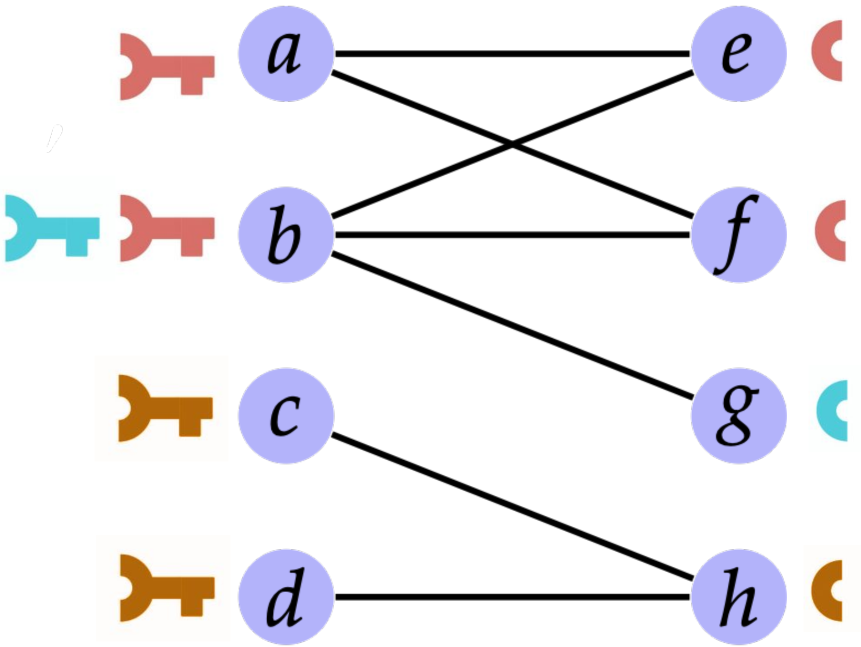
Erdős–Pyber (1997)

Theorem 1 (Erdős–Pyber). *Every graph G admits a biclique partition in which every vertex v belongs to $O(n/\lg n)$ bicliques.*



Erdős–Pyber (1997)

Theorem 1 (Erdős–Pyber). *Every graph G admits a biclique partition in which every vertex v belongs to $O(n/\lg n)$ bicliques.*



The number of bicliques per vertex determines the amount of secret they have to receive!

Erdős-Pyber theorem for hypergraphs and secret sharing

László Csirmaz · Péter Ligeti · Gábor Tardos

Theorem 1 (Czirmaz, Ligetti, Tardos (2014)). *We have*

$$\mathsf{LBP}(n) \le (1+o(1))\frac{n}{\lg n}, \quad \textit{and} \quad \left(\frac{0.53}{2}+o(1)\right)\frac{n}{\lg n} \le \mathsf{LBP}^*(n) \le \left(\frac{1}{2}+o(1)\right)\frac{n}{\lg n}$$

Erdős-Pyber theorem for hypergraphs and secret sharing

László Csirmaz · Péter Ligeti · Gábor Tardos

Theorem 1 (Czirmaz, Ligetti, Tardos (2014)). *We have*

$$\mathsf{LBP}(n) \leq (1+o(1))\frac{n}{\lg n}, \quad \textit{and} \quad \left(\frac{0.53}{2}+o(1)\right)\frac{n}{\lg n} \leq \mathsf{LBP}^*(n) \leq \left(\frac{1}{2}+o(1)\right)\frac{n}{\lg n}$$

Theorem 2 (Czirmaz, Ligetti, Tardos (2014)). *For d-uniform hypergraphs we have*

$$\left(\frac{0.53}{d!}+o(1)\right)\frac{n^{d-1}}{\lg n} \leq \mathsf{LBP}^*_d(n) \leq \mathsf{LBP}_d(n) \leq \left(\frac{1}{(d-2)!}+o(1)\right)\frac{n^{d-1}}{\lg n}$$

Our main result

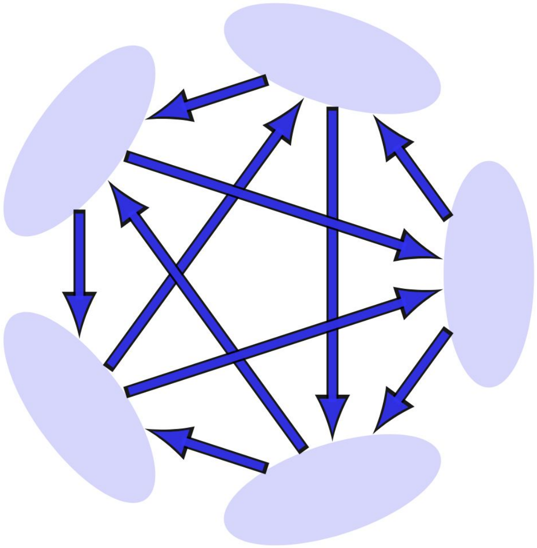
Theorem 1. *For d-uniform hypergraphs we have*

$$\left(\frac{1}{d!} - o(1)\right)\frac{n^{d-1}}{\lg n} \le \mathsf{LBP}_d(n) \le \left(\frac{1}{d!} + o(1)\right)\frac{n^{d-1}}{\lg n}.$$

Graph Upper bound

Step 1: Partition vertices with part size $\boxed{r := \lfloor \lg n - 2 \lg \lg n \rfloor}$

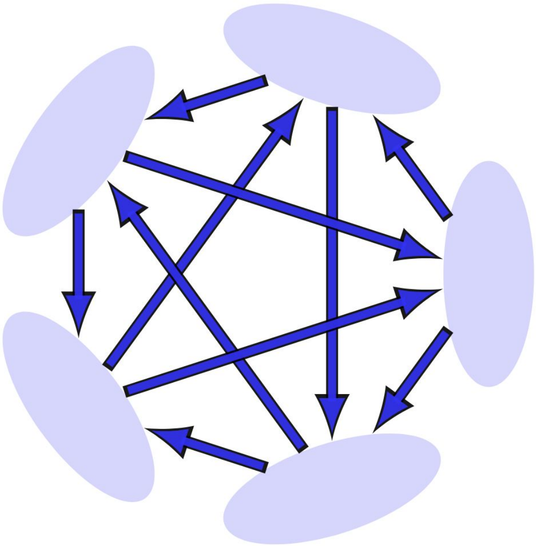
and set up an almost-regular tournament between parts



Graph Upper bound

Step 1: Partition vertices with part-size $r := \lfloor \lg n - 2 \lg \lg n \rfloor$

and set up an almost-regular tournament between parts

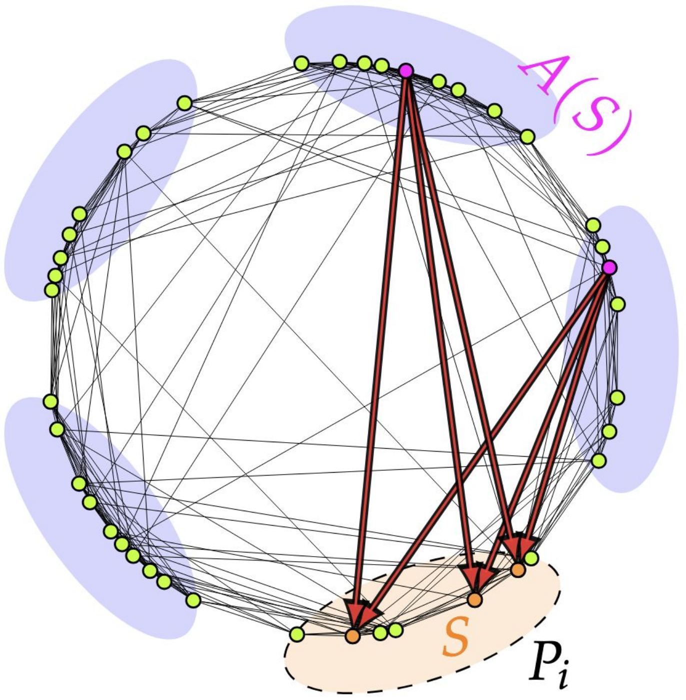


Step 2: Take each edge within a part as its own biclique.
Each vertex participates in at most *r - 1* = *o*(*n*/ lg *n*) of these.

Graph Upper bound

$r := \lfloor \lg n - 2 \lg \lg n \rfloor$

Step 3: For each part P_i, consider each of its $2^r \approx n / \lg^2 n$ subsets S, and emit biclique $(S, A(S))$ with $A(S)$ the vertices whose directed neighborhood in P_i is exactly S.

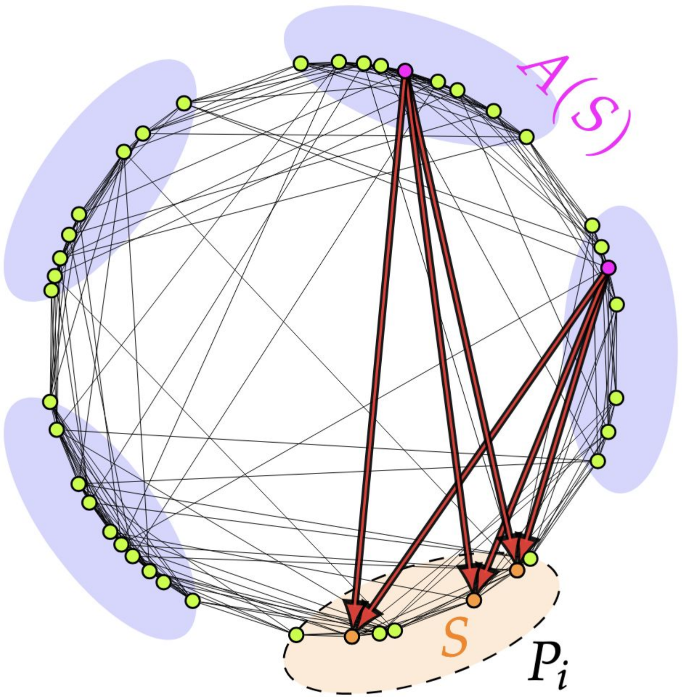


Graph Upper bound

$r := \lfloor \lg n - 2 \lg \lg n \rfloor$

Step 3: For each part P_i, consider each of its $2^r \approx n / \lg^2 n$ subsets S, and emit biclique $(S, A(S))$ with $A(S)$ the vertices whose directed neighborhood in P_i is exactly S.

Q: In how many such bicliques does each vertex v participate?



Graph Upper bound

$r := \lfloor \lg n - 2 \lg \lg n \rfloor$

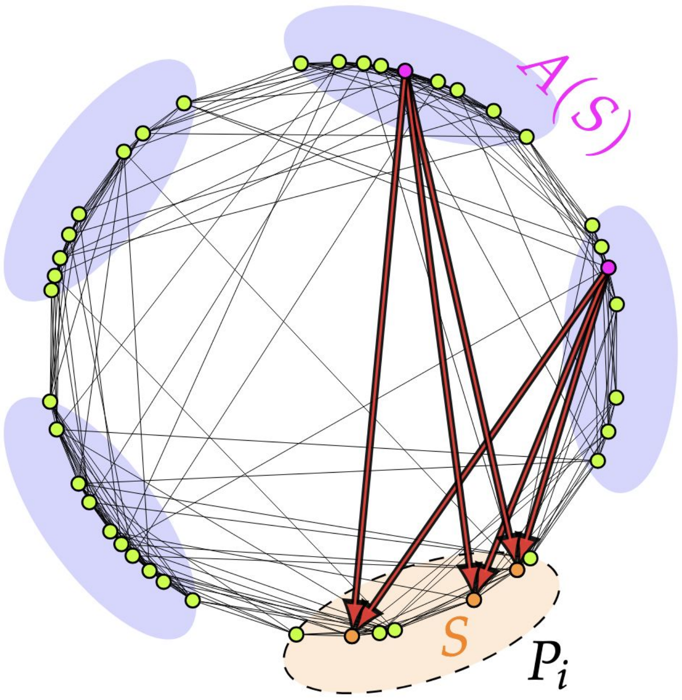
Q: In how many such bicliques does each vertex v participate?

Answer:

On the S-side, to at most $O(n / \lg^2 n)$

On the $A(S)$-side, by almost regularity to at most

$$\left\lceil \frac{n}{2r} \right\rceil \sim \frac{n}{2 \lg n}$$



Graph Upper bound

Efficient $O(n^2)$ implementation!

Algorithm 1 Biclique Partition

Require: A graph $G = (V, E)$ with $V = [n]$.

$r \leftarrow \lfloor \lg n - 2 \lg \lg n \rfloor$
function $R(i, j)$ $\triangleright$ *R*: $[\lceil n/r \rceil]^2 \rightarrow$ *Bool*
    return $(i < j$ and $d_n(i,j) \equiv 0 \pmod 2)$ or $(i > j$ and $d_n(i,j) \equiv 1 \pmod 2)$
for all $i \in [\lceil n/r \rceil]$ **do**
    $P_i \leftarrow \{(i-1) \cdot r + 1, \ldots, i \cdot r\}$
    $A \leftarrow \varnothing$ $\triangleright$ $A \colon \mathcal{P}(P_i) \rightarrow \mathcal{P}(V)$, *implemented as an array of linked lists of size* $2^{|P_i|} = O\left(\frac{n}{\lg^2 n}\right)$
    for all $v \in V$ **do**
        if $R(\lceil v/r \rceil, i)$ **then**
            $S \leftarrow \varnothing$ $\triangleright$ $S \colon \mathcal{P}(P_i)$
            for all $u \in P_i$ **do**
                if $\{u, v\} \in E$ **then**
                    $S \leftarrow S \cup \{u\}$
            $A(S) \leftarrow A(S) \cup \{v\}$ $\triangleright$ *We think of S as a binary string that indexes A*
    for all $S \in \mathcal{P}(P_i)$ **do**
        if $|S| > 0$ and $|A(S)| > 0$ **then**
            Emit $(S, A(S))$
    for all $\{u, v\} \in \binom{P_i}{2}$ **do**
        if $\{u, v\} \in E$ **then**
            Emit $(\{u\}, \{v\})$

Information theory lower bound

There are $2^{\binom{n}{2}}$ labeled graphs on n vertices.
So any binary encoding for them

must use $\lg(2^{\binom{n}{2}}) \sim \frac{n^2}{2}$ bits on some graph.

But a biclique covering of a graph fully identifies the graph, and as each vertex can be encoded with lg n bits, coverings of weight asymptotically less than $n^2/(2\lg n)$ would break the information barrier!

Optimal CES for d-Hypergraphs

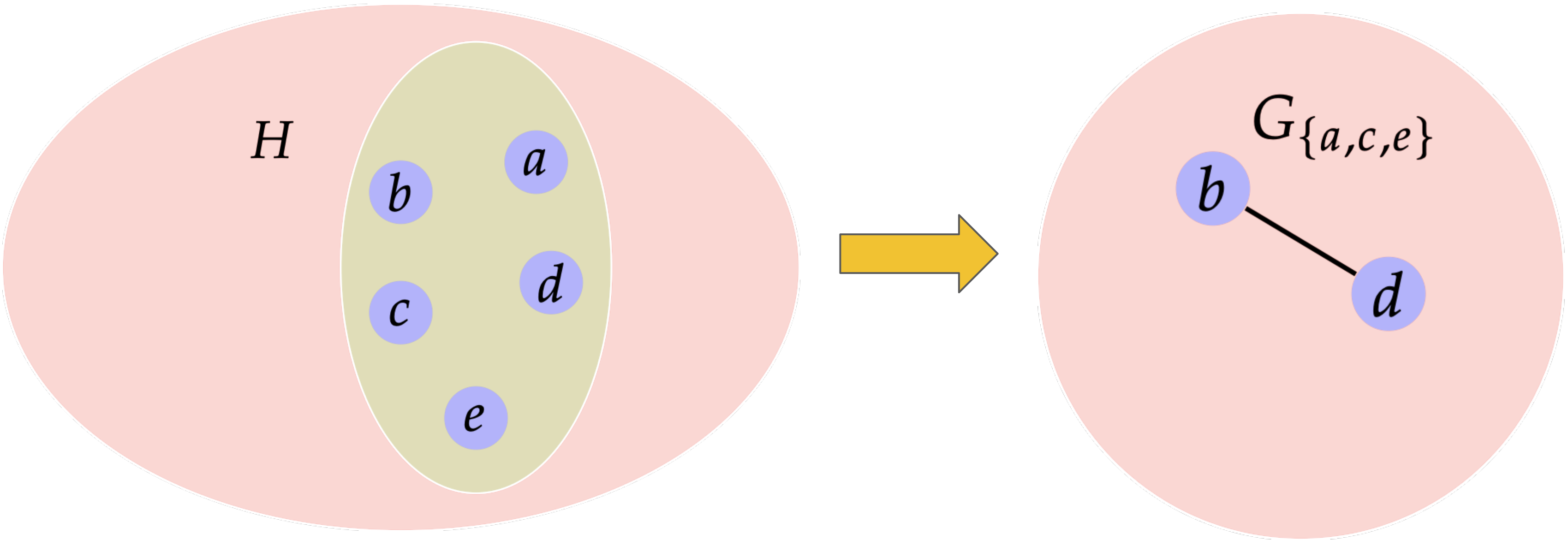
Here we want a d-clique partition of weight at most

$$\left(\frac{1}{d!} + o(1)\right) \frac{n^d}{\lg n}$$

And it turns out this is easy to get after having solved the graph case!

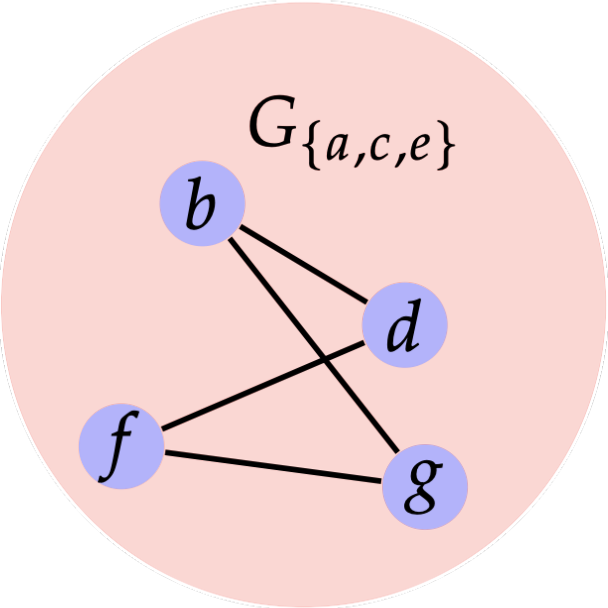
Optimal CES for *d*-Hypergraphs

We will reduce to d=2 through ***link graphs***



Optimal CES for *d*-Hypergraphs

We will reduce to d=2 through ***link graphs***



Bicliques in these link graphs are d-cliques in the original graph!

$$(\{b, f\}, \{d, g\}, \{a\}, \{c\}, \{e\})$$

Optimal CES for d-Hypergraphs

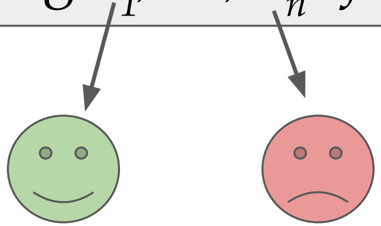
Step 1: Fix an ordering $v_1, \ldots, v_n$ *of the vertices*

Step 2: Separate each d-edge $e := \{x_1, \ldots, x_d\}$ into $S :=$ first d-2 vertices in e according to ordering, and $e \setminus S$, and place edge $e \setminus S$ in link graph G_S

Step 3: Decompose each link graph with an optimal clique partition for graphs

Optimal Erdős–Pyber for d-Hypergraphs

Step 1: Fix an ordering $v_1, \ldots, v_n$ *of the vertices*



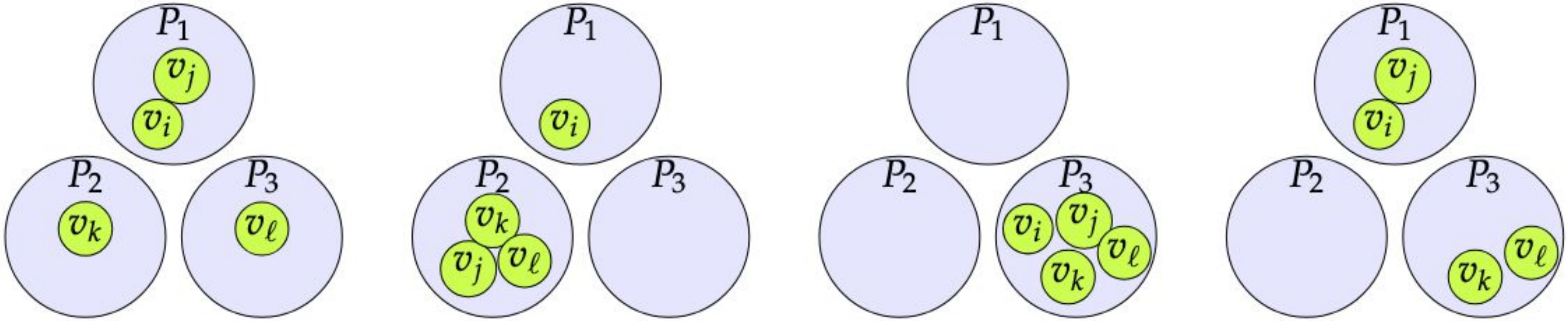
The issue with the previous proof is that any fixed ordering puts much more weight on some vertices than in others!

We need some way to assign edges that is **equitable**!

Optimal Erdős–Pyber for d-Hypergraphs

Step 1: Partition *V(H)* into *d-1* parts

Step 2: Use an equitable selection strategy for assigning edges to link graphs!



(a) $(2,1,1)$ (b) $(1,3,0)$ (c) $(0,0,4)$ (d) $(2,0,2)$

Optimal Erdős–Pyber for *d*-Hypergraphs

Step 1: Partition $V(H)$ into d-1 parts

Step 2: Use an equitable selection strategy for assigning edges to link graphs!

How to assign shapes like (2,2,3,3,0,0,0,0,0) or (2,2,2,2,1,1,0,0,0)?

Optimal Erdős–Pyber for d-Hypergraphs

Given a uniformly random potential hypergraph edge e, we want the probability that part i pays for it to be $1/(d-1)$

Optimal Erdős–Pyber for d-Hypergraphs

Given a uniformly random potential hypergraph edge e, we want the probability that part i pays for it to be $1/(d-1)$

Note that this depends only on the part-distribution of e, not on its individual vertices. Therefore, if f is the function that assigns part-distributions to parts, we want

Optimal Erdős–Pyber for *d*-Hypergraphs

Given a uniformly random potential hypergraph edge *e*, we want the probability that part *i* pays for it to be *1/(d-1)*

Note that this depends only on the part-distribution of *e*, not on its individual vertices. Therefore, if *f* is the function that assigns part-distributions to parts, we want

$$\Pr_e[f(\text{distribution of } e) = i] = \sum_{(x_1,\ldots,x_{d-1}) \in f^{-1}(i)} \frac{\binom{d}{x_1,\ldots,x_{d-1}}}{(d-1)^d} = \frac{1}{d-1}$$

Optimal Erdős–Pyber for d-Hypergraphs

Step 2: Use an equitable selection strategy for assigning edges to link graphs!

Definition 16. Let $d \geqslant 2$ be an integer. A *d-distribution* is a tuple $(x_1, \ldots, x_{d-1}) \in \mathbb{N}^{d-1}$ such that $x_1 + \ldots + x_{d-1} = d$. Let $\mathcal{S}_d$ be the set of d-distributions. A *selection strategy* is a function $f : \mathcal{S}_d \to [d-1]$ such that $f(x_1, \ldots, x_{d-1}) = i$ for some i such that $x_i \geqslant 2$. Then, we say a selection strategy is *equitable* if for all $i, j \in [d-1]$,

$$\sum_{(x_1,\ldots,x_{d-1}) \in f^{-1}(i)} \binom{d}{x_1, \ldots, x_{d-1}} = \sum_{(x_1,\ldots,x_{d-1}) \in f^{-1}(j)} \binom{d}{x_1, \ldots, x_{d-1}}.$$

Lemma 17. *For every $d \geqslant 2$, there exists an equitable selection strategy.*

Optimal Erdős–Pyber for *d*-Hypergraphs

How to assign shapes like (2,2,3,3,0,0,0,0,0) ?

It suffices to prove that their sets of cyclic rotations

{ (0,2,2,3,3,0,0,0,0), … (3,3,0,0,0,0,0,2,2),…}

always have size *d-1 :)*

That way, from each equivalence class under rotation we assign one element to each part.

Optimal Erdős–Pyber for *d*-Hypergraphs

How to assign shapes like (2,2,3,3,0,0,0,0,0) ?

It suffices to prove that their sets of cyclic rotations

{ (0,2,2,3,3,0,0,0,0), … (3,3,0,0,0,0,0,2,2),…}

always have size *d-1 :)*

The claim follows from word combinatorics, or the orbit-stabilizer theorem!

Density-aware partitions

Theorem 1. *Let γ such that* $\max\{\gamma^{-1}, (1-\gamma)^{-1}\} = n^{o(1)}$. *Then,*

$$\mathsf{BP}(n,\gamma) \sim \lg \binom{\binom{n}{2}}{\gamma\binom{n}{2}} / \lg n \sim \frac{h_2(\gamma)}{2} \frac{n^2}{\lg n},$$

where $h_2(x) := -x\lg(x) - (1-x)\lg(1-x)$ *is the binary entropy function.*

Based on Stirling: $$\binom{n}{n\gamma_n} \sim \frac{2^{h_2(\gamma_n)n}}{\sqrt{2\pi\gamma_n(1-\gamma_n)n}},$$

Density-aware partitions

Theorem 1. *Let* γ *such that* $\max\{\gamma^{-1}, (1-\gamma)^{-1}\} = n^{o(1)}$. *Then,*

$$\mathsf{BP}(n,\gamma) \sim \lg \binom{\binom{n}{2}}{\gamma\binom{n}{2}} / \lg n \sim \frac{h_2(\gamma)}{2} \frac{n^2}{\lg n},$$

where $h_2(x) := -x\lg(x) - (1-x)\lg(1-x)$ *is the binary entropy function.*

So biclique partitions are an ***information-theoretically optimal representation*** at almost every density!

Density-aware partitions (***Proof Idea***)

On top of splitting into equal-sized parts (now bigger), and setting up the almost-regular tournament, we view the adjacencies of vertex v into part P_i as a binary string like

00111010100000100000001111111100000100000100000100000001

Density-aware partitions (***Proof Idea***)

On top of splitting into equal-sized parts (now bigger), and setting up the almost-regular tournament, we view the adjacencies of vertex v into part P_i as a binary string like

00111010100000010000000111111110000010000100001000000001

And split it into slices of roughly equal ***entropy*** $:= \lg \binom{\ell}{w}$

Density-aware partitions (***Proof Idea***)

001110101000001000000011111111000001000010000100000001

And split it into slices of roughly equal ***entropy*** $:= \lg \binom{\ell}{w}$

Now bicliques will be within slices!

Recall: $\prod_{i=1}^{k} \binom{a_i}{b_i} \leqslant \binom{a_1 + \cdots + a_k}{b_1 + \cdots + b_k}$

Density-aware partitions (***Proof Idea***)

0011101010000010000000111111110000001000010000100000001

And split it into slices of roughly equal ***entropy*** $:= \lg \binom{\ell}{w}$

Now bicliques will be within slices!

Recall: $\prod_{i=1}^{k} \binom{a_i}{b_i} \leqslant \binom{a_1 + \cdots + a_k}{b_1 + \cdots + b_k}$

$$\sum_{v \in V} \sum_{\text{slice index } i} \lg \binom{\ell_i}{w_i} = \lg \left(\prod_{v \in V} \prod_{\text{slice index } i} \binom{\ell_i}{w_i} \right) \leqslant^{\text{Rem. 24}} (1 + o(1)) \lg \binom{n^2/2}{|E|} \sim^{\text{Rem. 23}} h_2(\gamma) \frac{n^2}{2},$$

Algorithmic applications

The previous result implies that biclique partitions make for a ***Succinct data structure!***

(*and* ***compact*** if we also store the vertex-biclique incidence list)

Algorithmic applications

The previous result implies that biclique partitions make for a ***Succinct data structure!***

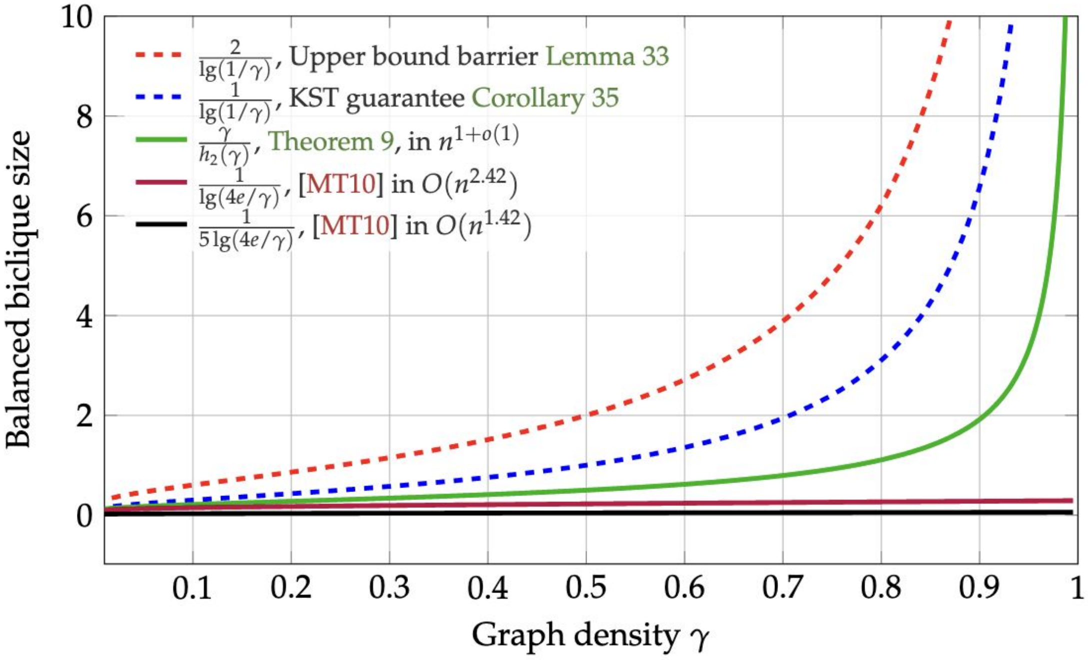
(*and* ***compact*** if we also store the vertex-biclique incidence list)

Theorem 1. *Using a biclique partition of weight $O(n^2/\lg n)$ as a succinct data structure we can:*

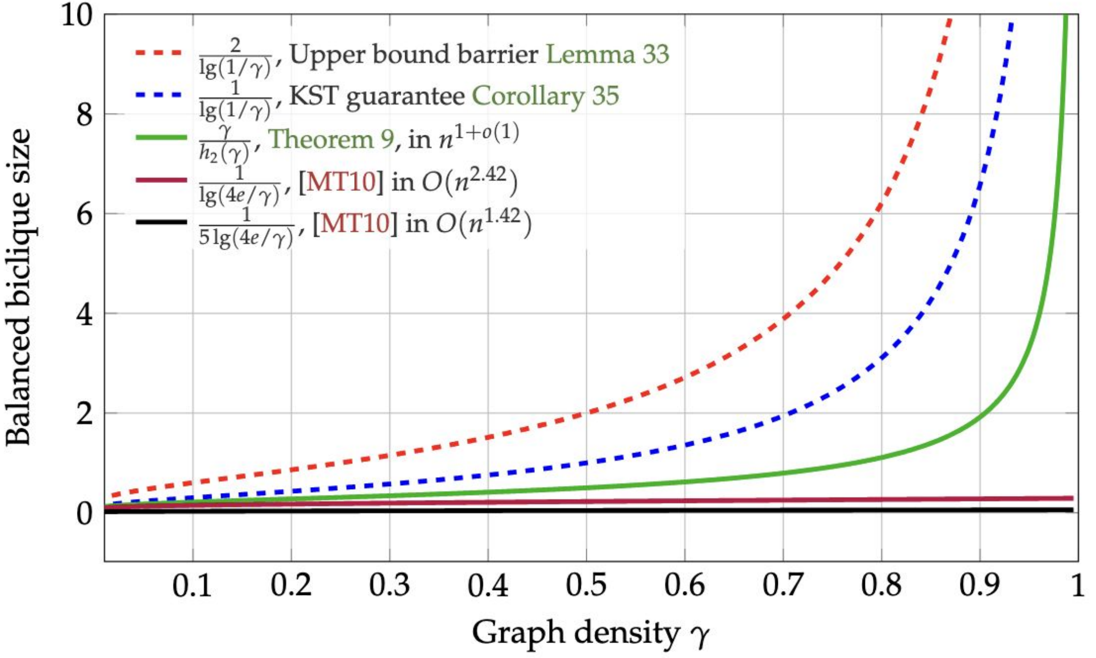
1. *Given $S \subseteq V$, determine if S is independent in $O(n^2/\lg n)$,*
2. *Given disjoint $S, T \subseteq V$, determine $e(S, T)$ in $O(n^2/\lg n)$.*

Moreover, as a compact data structure, for any $\alpha > 1$, we can compute a 2α-approximation for densest subgraph in time $O(n^2/\lg \alpha)$.

Faster algorithm for finding density-guaranteed balanced bicliques



Theorem 9. *Given a graph with edge density γ such that $\max\{\gamma^{-1}, (1-\gamma)^{-1}\} = n^{o(1)}$, there is a randomized $O(\frac{n \lg n}{h_2(\gamma)}) = n^{1+o(1)}$ time algorithm that returns a biclique $K_{t,t}$ with $t = (1-o(1))\frac{\gamma \lg n}{h_2(\gamma)}$ with high probability. Furthermore, if we are allowed to query the degree of a vertex in $O(1)$ time, the algorithm can be made deterministic.*



Optimal d-Clique Decompositions for Hypergraphs

Joint work with Benjamin Przybocki, Andrew Krapiving, and Nicolás Sanhueza-Matamala

Bernardo Subercaseaux (CMU)

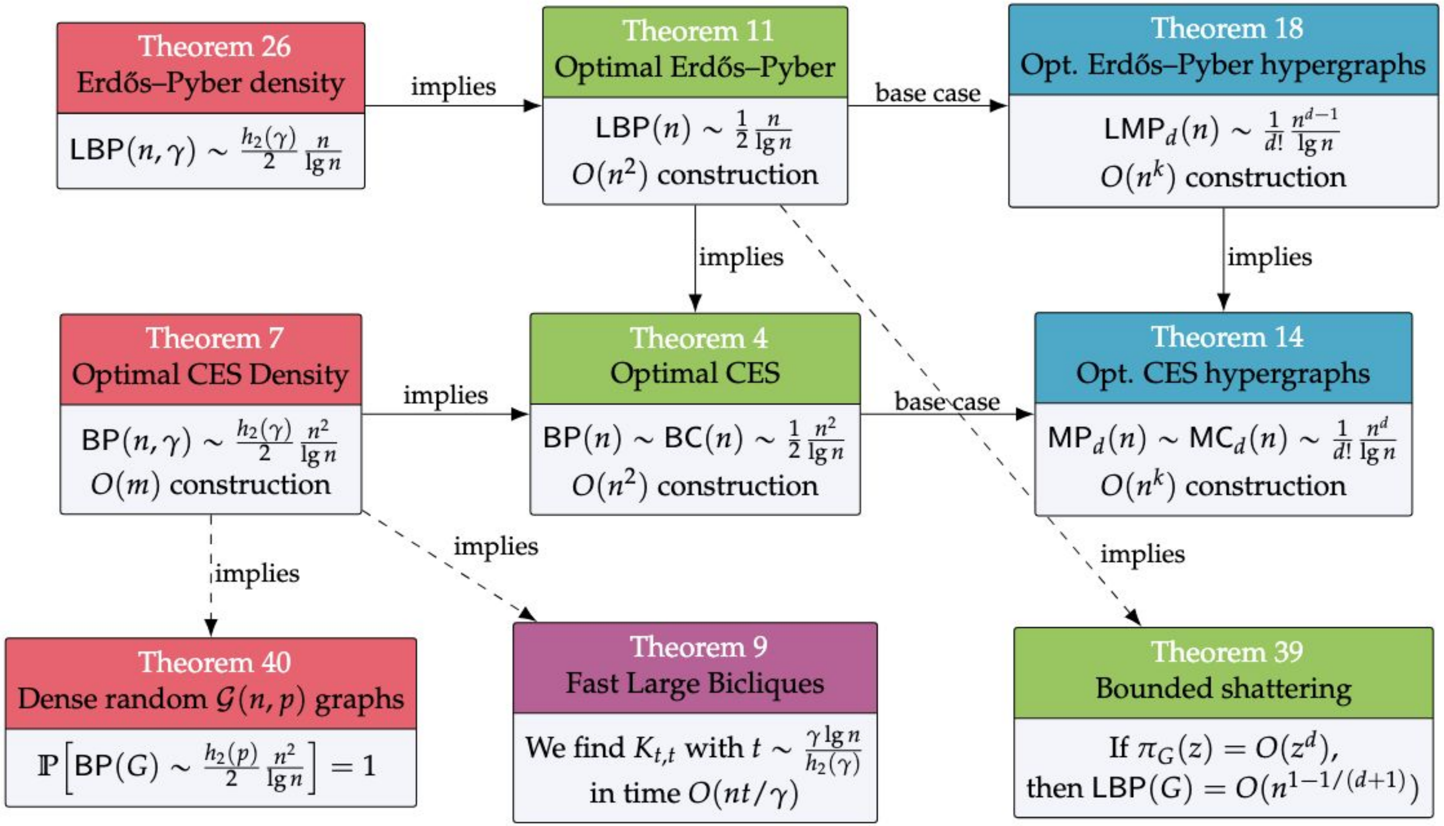


Figure 2: Summary of our results. Dashed implications represent that other tools are also required. (■ := density-aware, ■ := graph partitions, ■ := hypergraph partitions, ■ := finding bicliques)