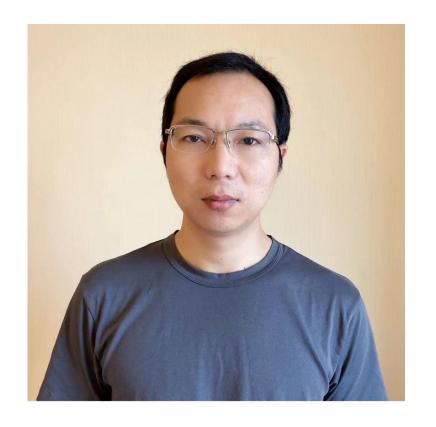
Learning-Augmented Streaming Algorithms for Approximating Max-Cut

Yinhao Dong (董寅灏)
University of Science and Technology of China (USTC)

Based on joint work with



Pan Peng USTC

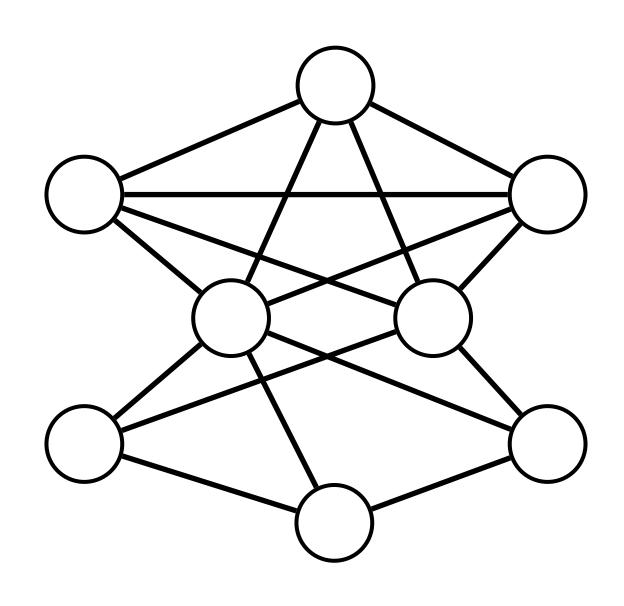


Ali Vakilian
TTIC -> Virginia Tech

第六届 CCF 理论计算机科学博士生论坛中南大学·2025 年 11 月

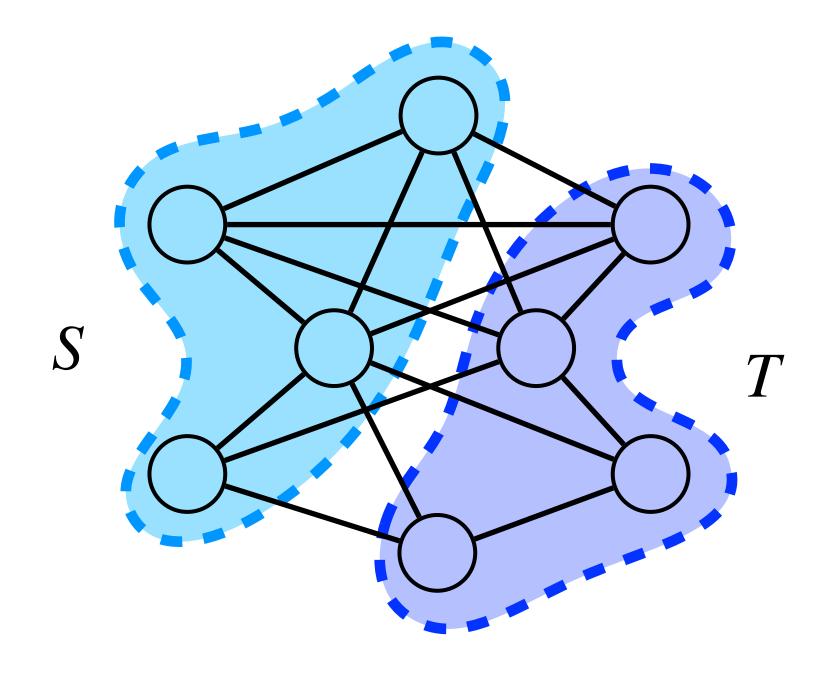
Input: An undirected, unweighted graph G = (V, E)

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$



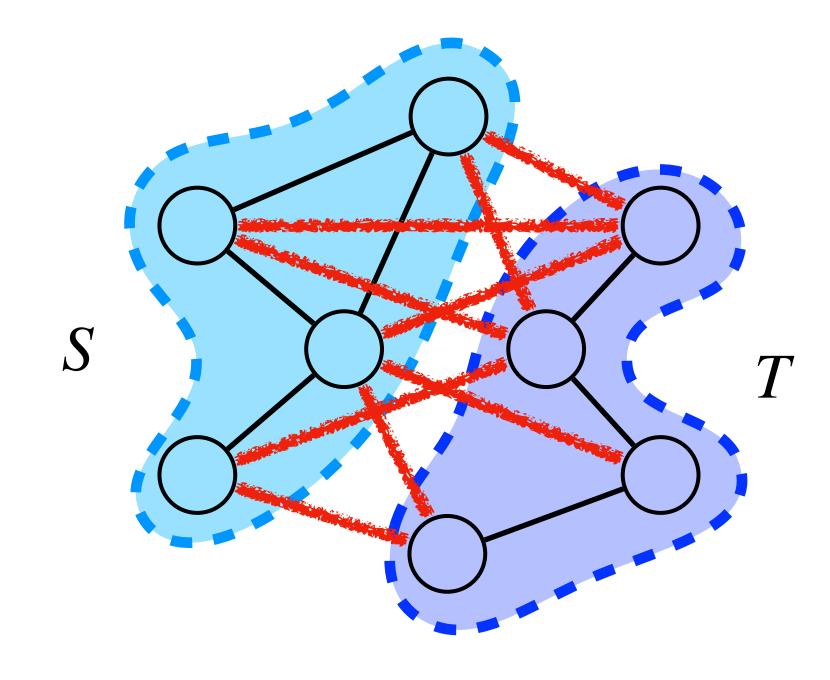
Input: An undirected, unweighted graph G = (V, E)

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$



Input: An undirected, unweighted graph G = (V, E)

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

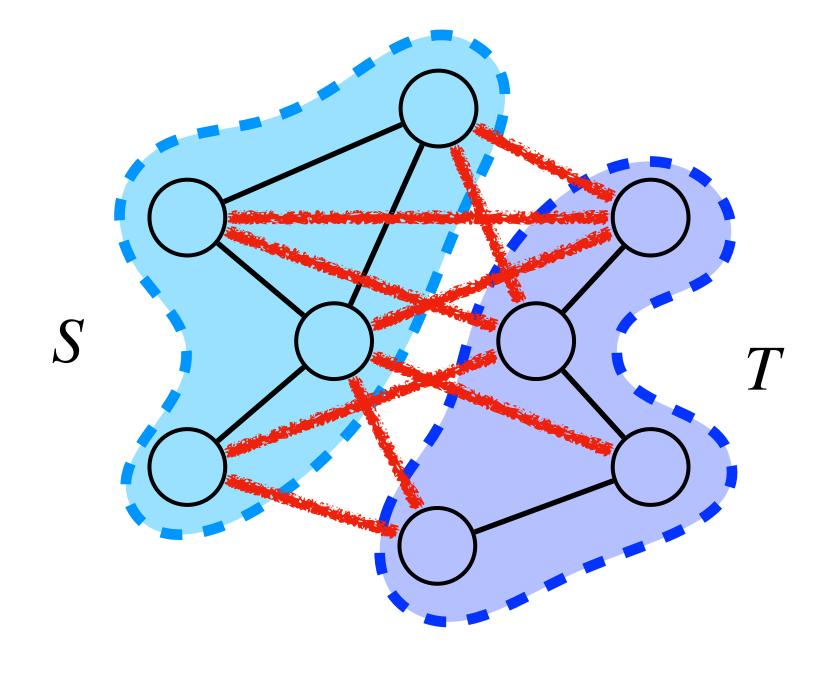


Input: An undirected, unweighted graph G = (V, E)

Goal: Find a bipartition of V into S and T that maximizes the cut value

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

NP-hard

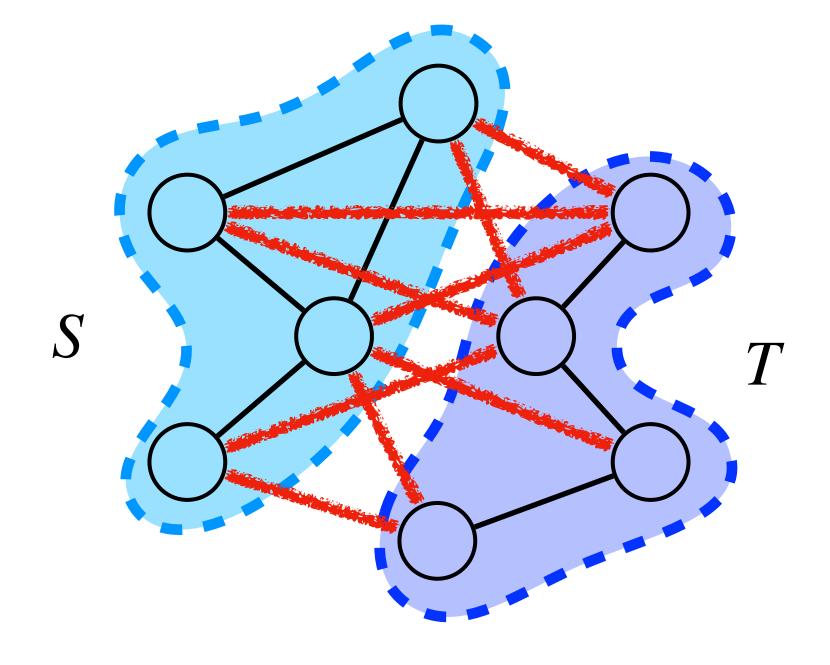


Input: An undirected, unweighted graph G = (V, E)

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

- NP-hard
- α -approximation: outputs a **cut** (S, T), s.t.

$$\alpha \cdot \mathsf{OPT} \leq |E(S,T)| \leq \mathsf{OPT}$$



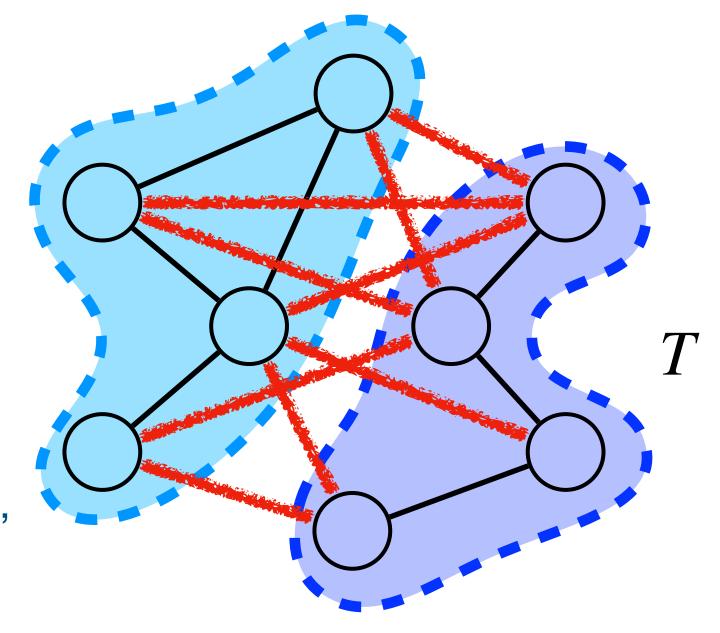
Input: An undirected, unweighted graph G = (V, E)

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

- NP-hard
- α -approximation: outputs a **cut** (S, T), s.t.

$$\alpha \cdot \mathsf{OPT} \leq |E(S,T)| \leq \mathsf{OPT}$$

- SDP rounding: 0.878-approx [Goemans, Williamson, JACM'95]
 - Unique Games Conjecture → best possible [Khot, Kindler, Mossel, O'Donnell, SICOMP'07]



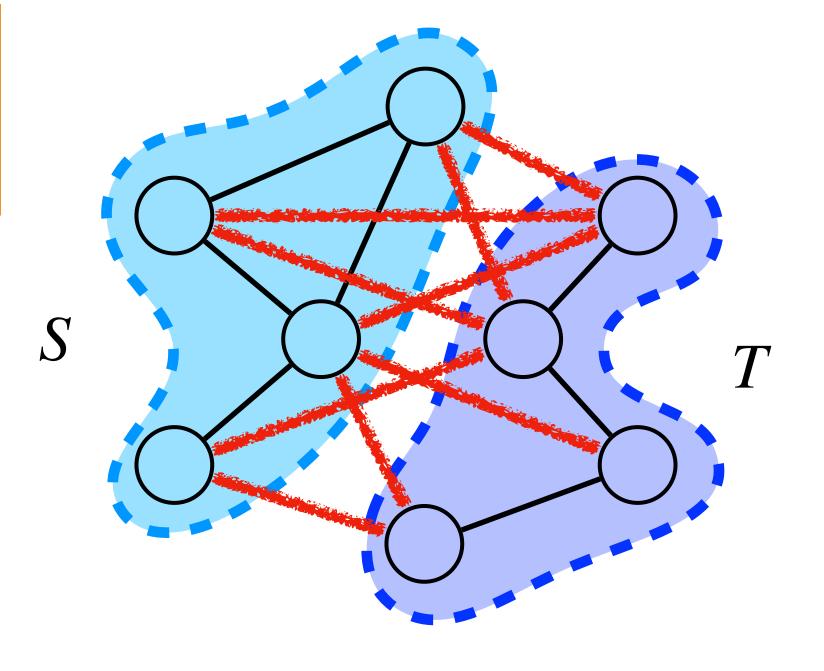
Input: An undirected, unweighted graph G = (V, E)

Goal: Find a bipartition of V into S and T that maximizes the cut value

$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

Estimation Task:

Compute an estimate v of the Max-Cut value



Input: An undirected, unweighted graph G = (V, E)

Goal: Find a *bipartition* of V into S and T that maximizes the cut value

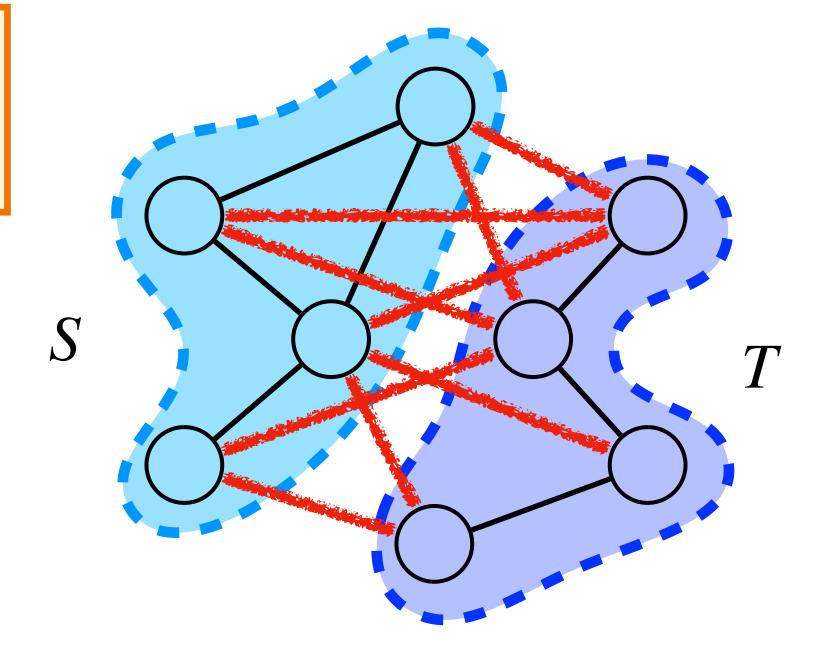
$$|E(S,T)| = |\{(u,v) \in E : u \in S \land v \in T\}|$$

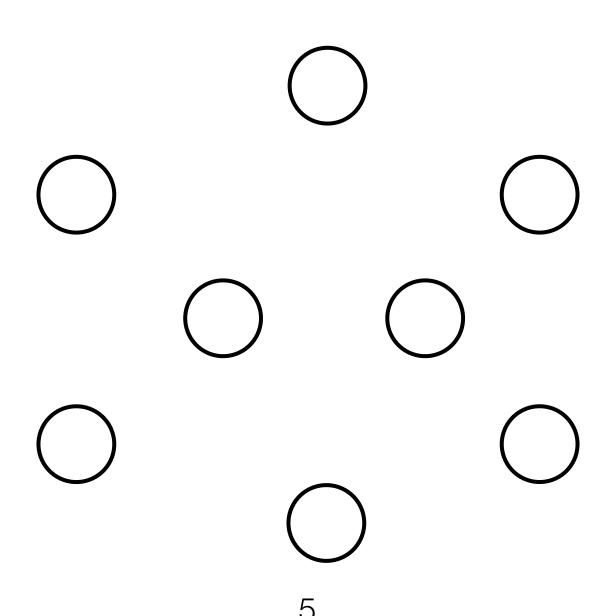
Estimation Task:

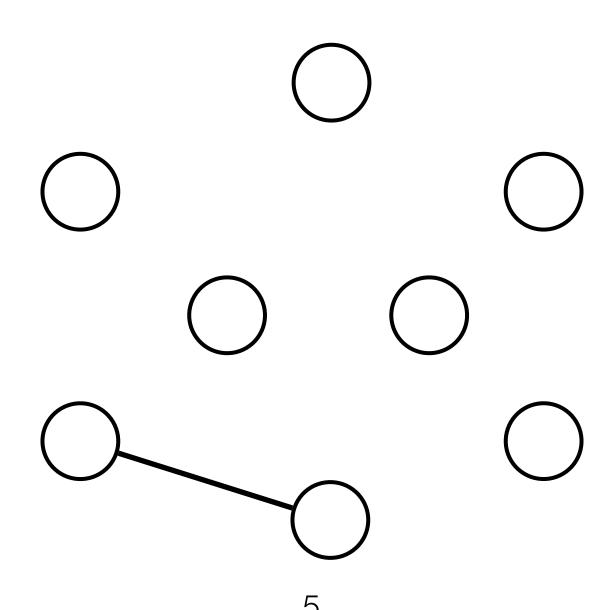
Compute an estimate v of the Max-Cut value

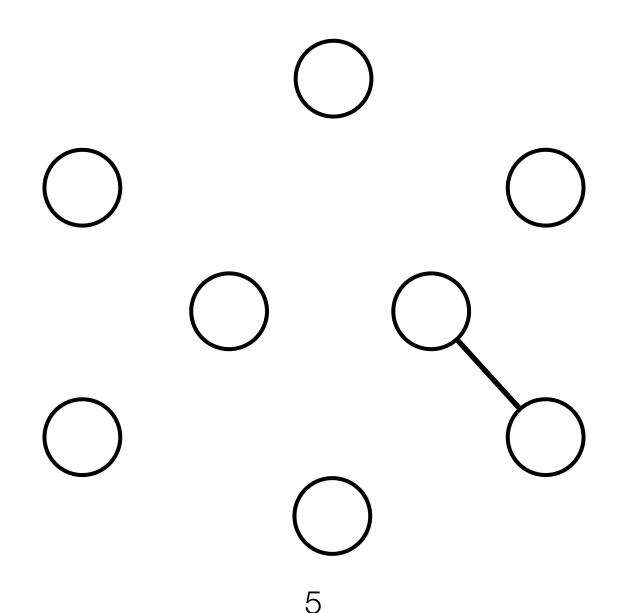
• α -approximation: outputs a value ν , s.t.

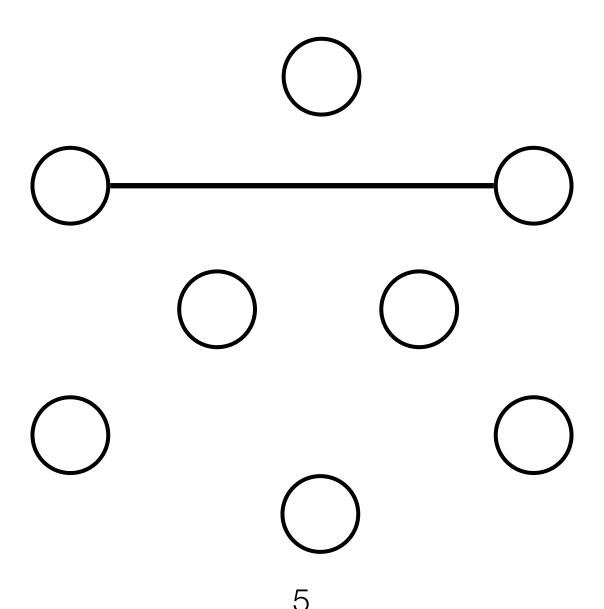
$$\alpha \cdot \mathsf{OPT} \leq v \leq \mathsf{OPT}$$

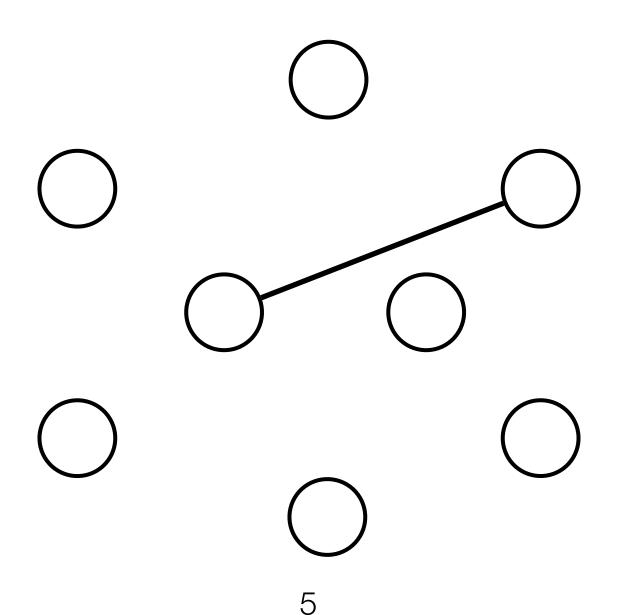


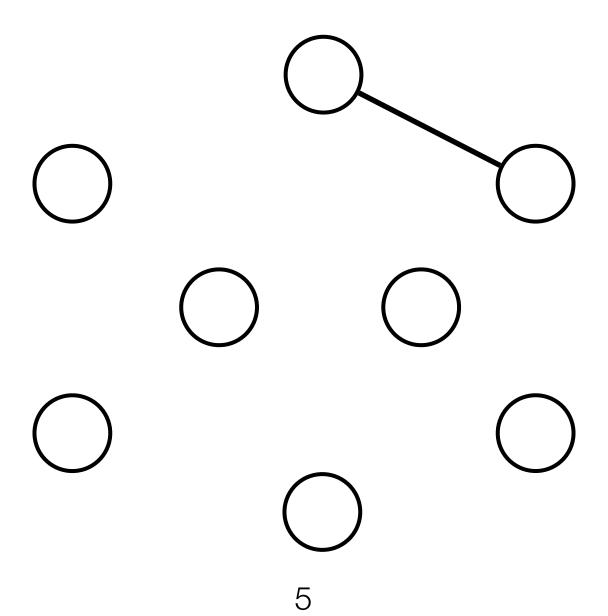


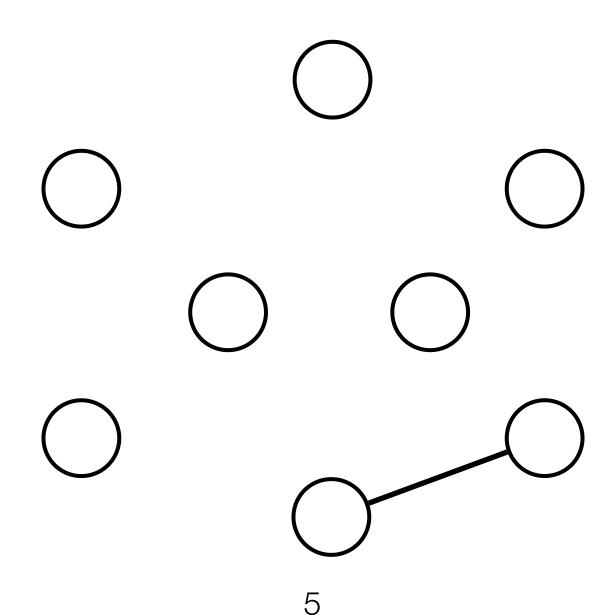


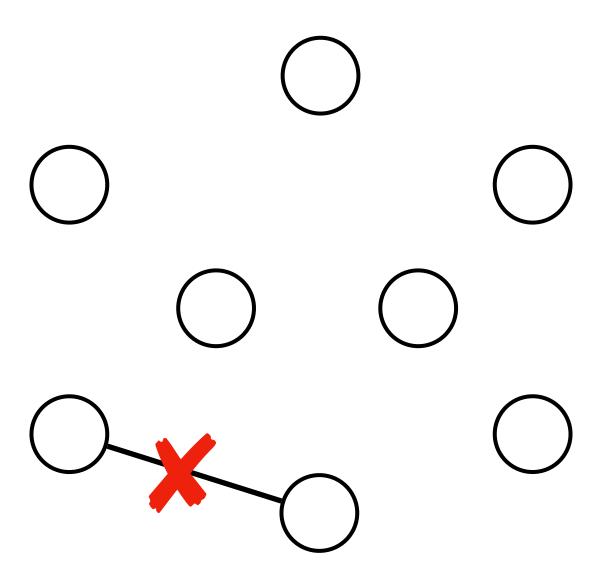


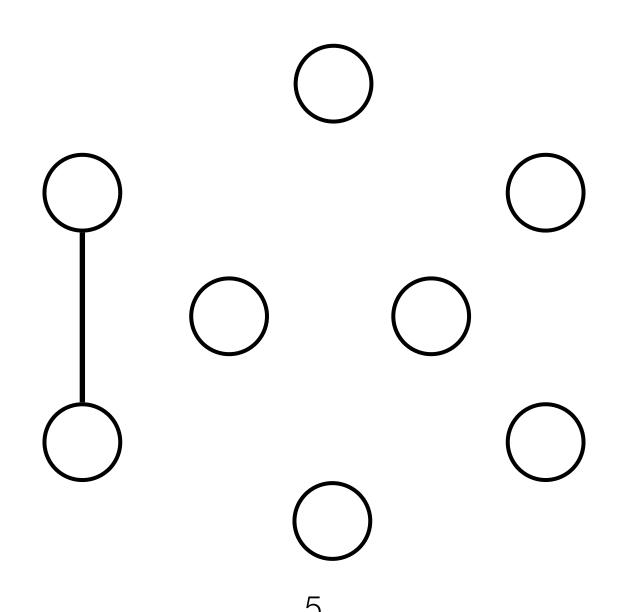












- Input graph is presented as a sequence of edge insertions and deletions
 - insertion-only streams: contains of edge insertions only
 - dynamic streams: has both insertions and deletions

- Input graph is presented as a sequence of edge insertions and deletions
 - insertion-only streams: contains of edge insertions only
 - dynamic streams: has both insertions and deletions
- Goal: in one pass, using small space, compute the solution

Input: An undirected, unweighted graph G = (V, E) as a graph stream

Goal: Compute an estimate of the Max-Cut value using small space

Input: An undirected, unweighted graph G = (V, E) as a graph stream

Goal: Compute an estimate of the Max-Cut value using small space

• A trivial 1/2-approximation using $O(\log n)$ bits of space: count the number of edges m and output m/2

Input: An undirected, unweighted graph G = (V, E) as a graph stream

Goal: Compute an estimate of the Max-Cut value using small space

- A trivial 1/2-approximation using $O(\log n)$ bits of space: count the number of edges m and output m/2
- A series of works explored the approximation ratio vs space trade-offs [Kapralov, Khanna, Sudan, SODA'15] [Kogan, Krauthgamer, ITCS'15] [Kapralov, Khanna, Sudan, Velingker, SODA'17] [Kapralov, Krachun, STOC'19]

Input: An undirected, unweighted graph G = (V, E) as a graph stream

Goal: Compute an estimate of the Max-Cut value using small space

- A trivial 1/2-approximation using $O(\log n)$ bits of space: count the number of edges m and output m/2
- A series of works explored the approximation ratio vs space trade-offs [Kapralov, Khanna, Sudan, SODA'15] [Kogan, Krauthgamer, ITCS'15] [Kapralov, Khanna, Sudan, Velingker, SODA'17] [Kapralov, Krachun, STOC'19]
- Any (randomized) algorithm that achieves $(1/2+\epsilon)$ -approximation requires $\Omega_\epsilon(n)$ bits of space [Kapralov, Krachun, STOC'19]

Input: An undirected, unweighted graph G = (V, E) as a graph stream

Goal: Compute an estimate of the Max-Cut value using small space

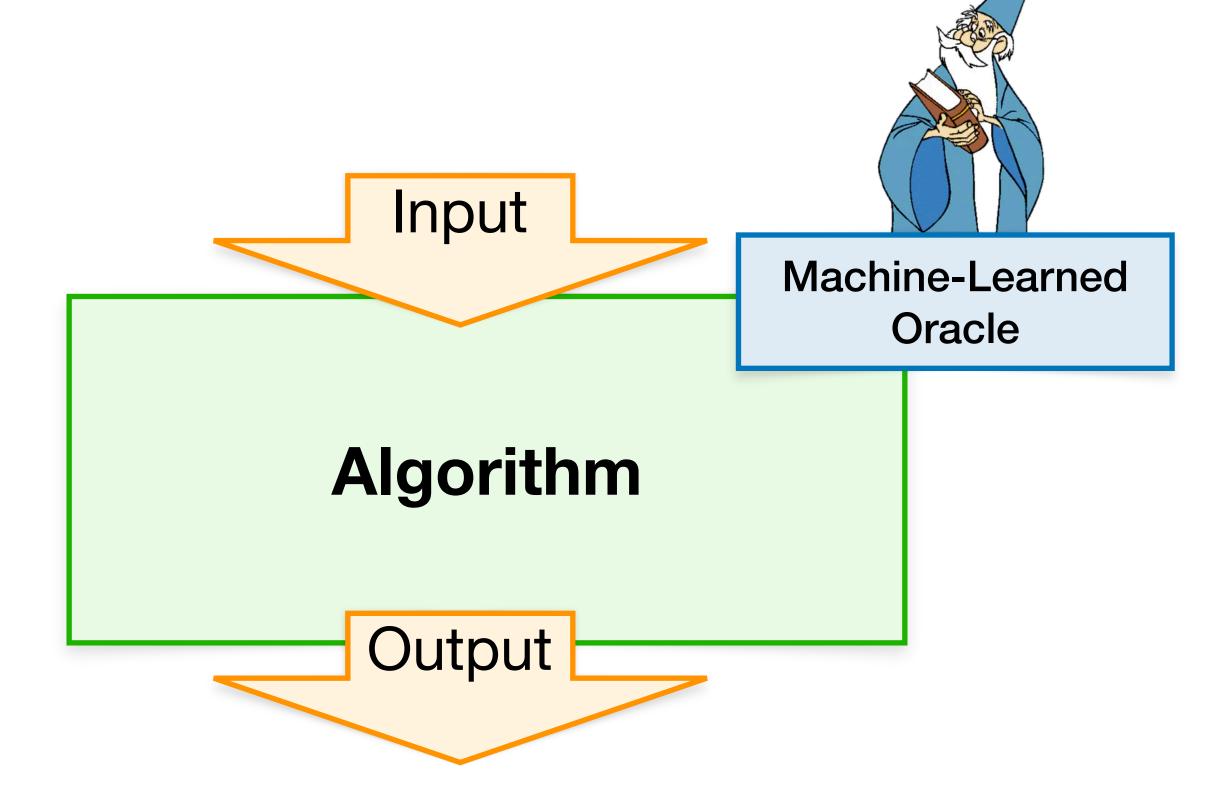
- A trivial 1/2-approximation using $O(\log n)$ bits of space: count the number of edges m and output m/2
- A series of works explored the approximation ratio vs space trade-offs [Kapralov, Khanna, Sudan, SODA'15] [Kogan, Krauthgamer, ITCS'15] [Kapralov, Khanna, Sudan, Velingker, SODA'17] [Kapralov, Krachun, STOC'19]
- Any (randomized) algorithm that achieves $(1/2+\epsilon)$ -approximation requires $\Omega_\epsilon(n)$ bits of space [Kapralov, Krachun, STOC'19]

What if we can obtain extra information about the input?

(a.k.a. Algorithms with Predictions)

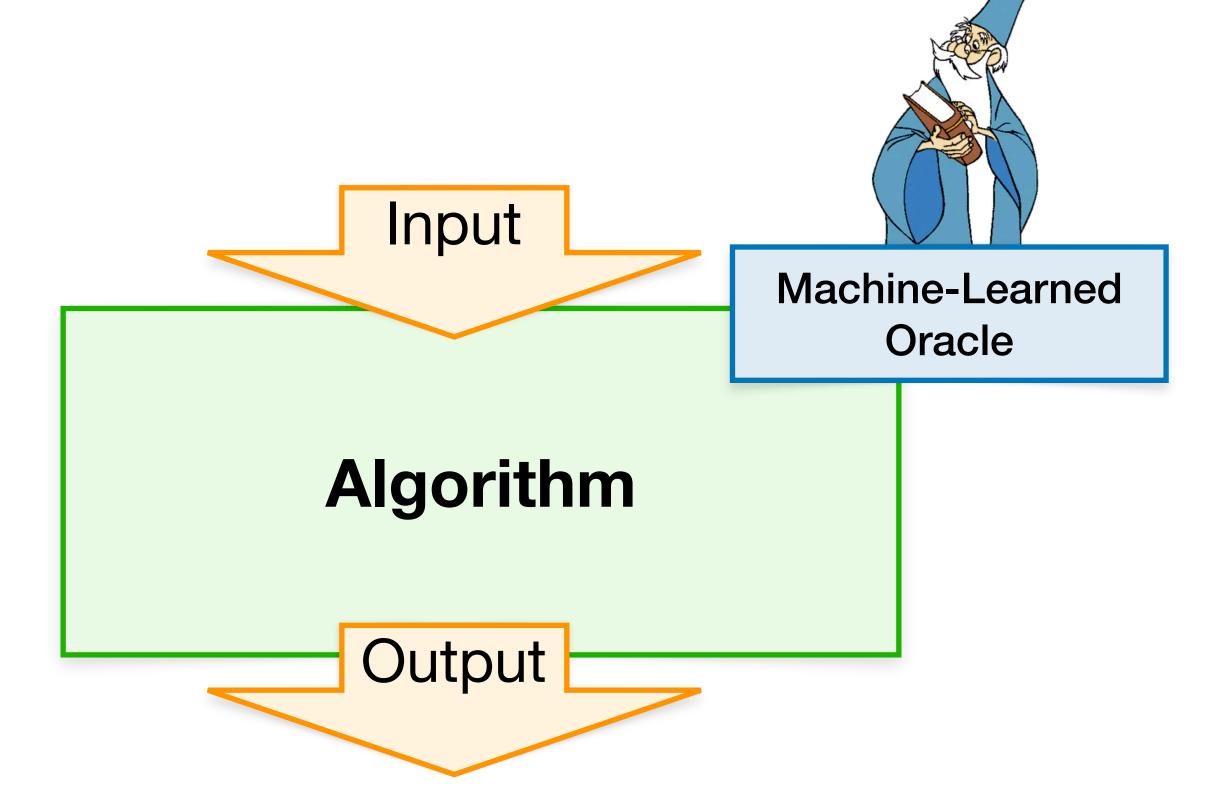
(a.k.a. Algorithms with Predictions)

 The algorithm has access to a learned oracle providing a certain type of predictions about the input instance



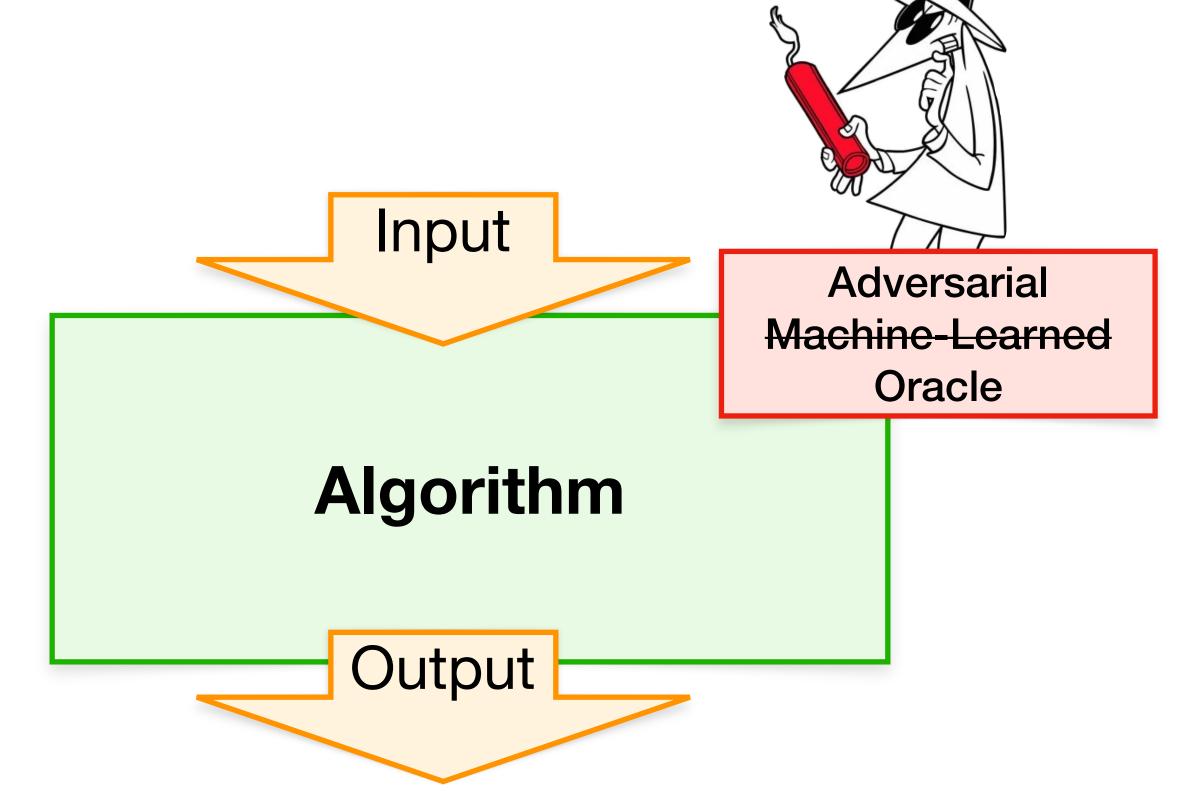
(a.k.a. Algorithms with Predictions)

- The algorithm has access to a learned oracle providing a certain type of predictions about the input instance
- Goals:
 - Consistency: High prediction quality
 Better performance than the best-known classical algorithm



(a.k.a. Algorithms with Predictions)

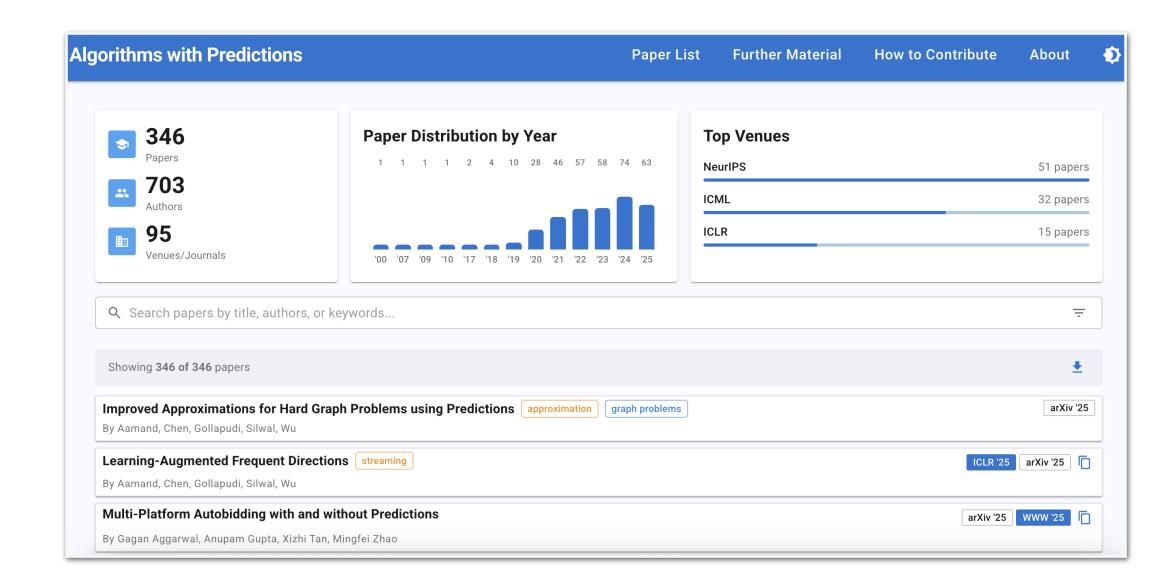
- The algorithm has access to a learned oracle providing a certain type of predictions about the input instance
- Goals:
 - Consistency: High prediction quality
 Better performance than the best-known classical algorithm
 - Robustness: Low prediction quality
 Performs no worse than the best-known classical algorithm



(a.k.a. Algorithms with Predictions)

- The algorithm has access to a learned oracle providing a certain type of predictions about the input instance
- Goals:
 - Consistency: High prediction quality
 Better performance than the best-known classical algorithm
 - Robustness: Low prediction quality

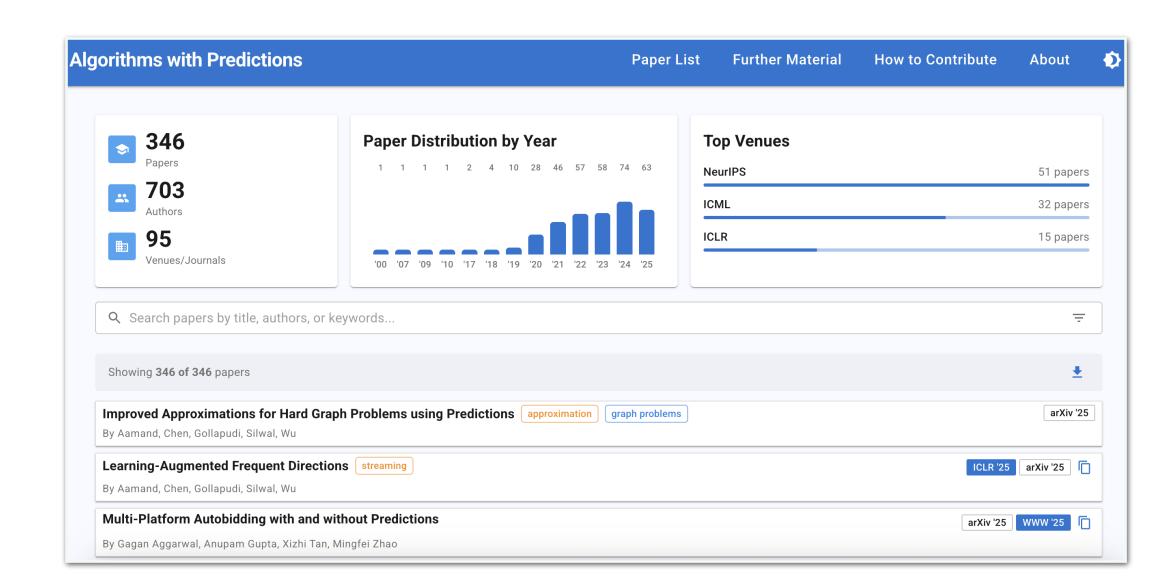
 ⇒ Performs no worse than the best-known classical algorithm



An up-to-date repository of publications (https://algorithms-with-predictions.github.io)

(a.k.a. Algorithms with Predictions)

- The algorithm has access to a learned oracle providing a certain type of predictions about the input instance
- Goals:
 - Consistency: High prediction quality
 Better performance than the best-known classical algorithm
 - Robustness: Low prediction quality
 Performs no worse than the best-known classical algorithm



An up-to-date repository of publications (https://algorithms-with-predictions.github.io)

How to define predictions?

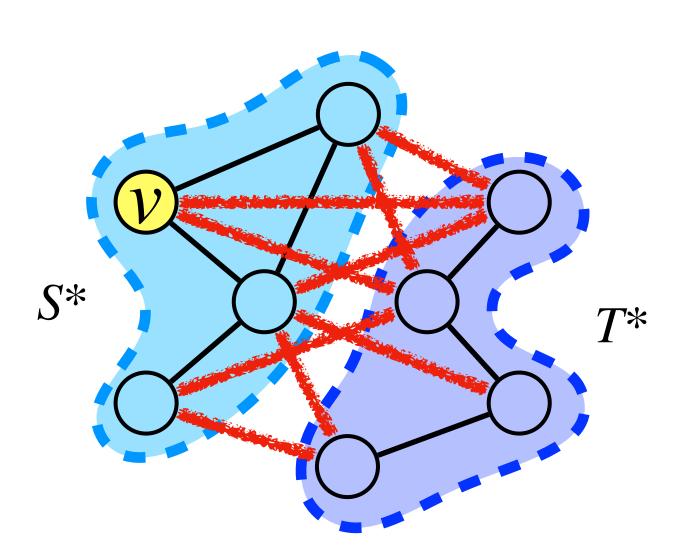
Our Prediction Model

• *C*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurIPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]

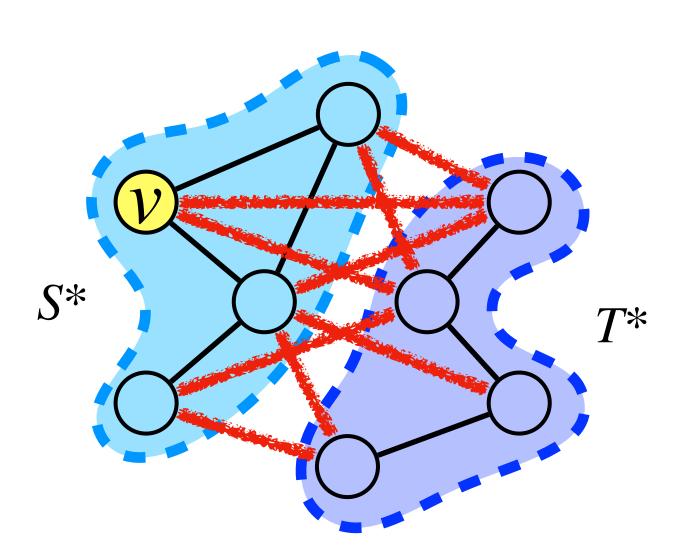
Our Prediction Model

- *E*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurlPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]
 - Let $x^* \in \{-1,1\}^n$ denote some fixed but unknown optimal solution

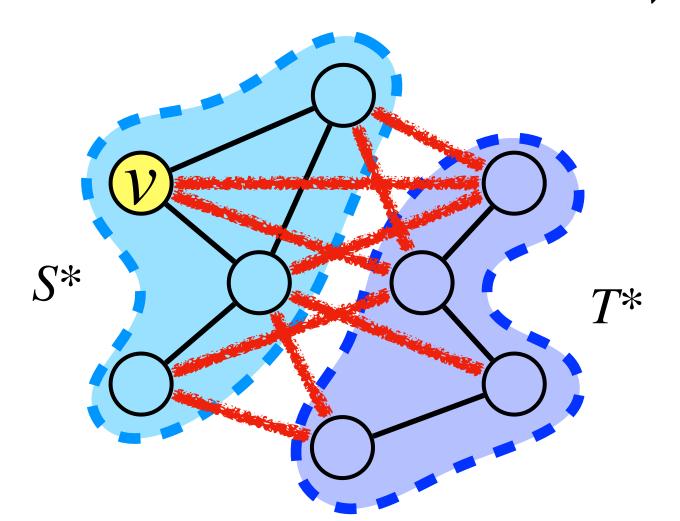
- *E*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurIPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]
 - Let $x^* \in \{-1,1\}^n$ denote some fixed but unknown optimal solution



- *E*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurlPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]
 - Let $x^* \in \{-1,1\}^n$ denote some fixed but unknown optimal solution
 - Oracle access to a prediction vector $Y \in \{-1,1\}^n$

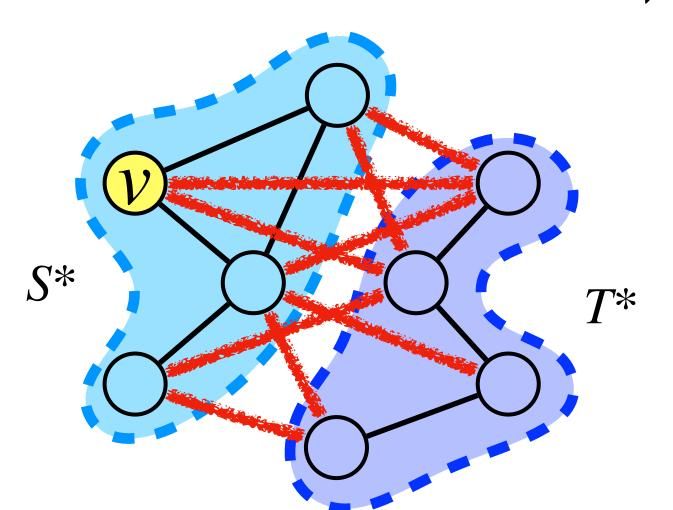


- *E*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurlPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]
 - Let $x^* \in \{-1,1\}^n$ denote some fixed but unknown optimal solution
 - Oracle access to a prediction vector $Y \in \{-1,1\}^n$
 - Each entry Y_{v} is independently correct with probability $1/2+\epsilon$



$$\forall v \in V, \ \Pr[Y_v = x_v^*] = \frac{1}{2} + \epsilon$$

- *E*-accurate predictions [Cohen-Addad, d'Orsi, Gupta, Lee, Panigrahi, NeurlPS'24] [Braverman, Dharangutte, Shah, Wang, APPROX'24] [Ghoshal, Makarychev, Makarychev, SODA'25]
 - Let $x^* \in \{-1,1\}^n$ denote some fixed but unknown optimal solution
 - Oracle access to a prediction vector $Y \in \{-1,1\}^n$
 - Each entry Y_{v} is independently correct with probability $1/2+\epsilon$



$$\forall v \in V, \ \Pr[Y_v = x_v^*] = \frac{1}{2} + \epsilon$$

Only slightly better than a random guess!

Our Results

Theorem [D., Peng, Vakilian, ITCS'25]:

Given ϵ -accurate predictions, there exists a single-pass streaming algorithm achieving a $(1/2 + \Omega(\epsilon^2))$ -approx with high probability, that uses

- $poly(1/\epsilon)$ words of space in **insertion-only streams**, and
- $poly(1/\epsilon, \log n)$ words of space in **dynamic streams**.

Our Results

Theorem [D., Peng, Vakilian, ITCS'25]:

Given ϵ -accurate predictions, there exists a single-pass streaming algorithm achieving a $(1/2 + \Omega(\epsilon^2))$ -approx with high probability, that uses

- $poly(1/\epsilon)$ words of space in insertion-only streams, and
- $poly(1/\epsilon, \log n)$ words of space in **dynamic streams**.
- Bypass the classical $(1/2+\epsilon)$ -approx vs $\Omega_\epsilon(n)$ space trade-off [Kapralov, Krachun, STOC'19]

Our Results

Theorem [D., Peng, Vakilian, ITCS'25]:

Given ϵ -accurate predictions, there exists a single-pass streaming algorithm achieving a $(1/2 + \Omega(\epsilon^2))$ -approx with high probability, that uses

- $poly(1/\epsilon)$ words of space in insertion-only streams, and
- $poly(1/\epsilon, \log n)$ words of space in **dynamic streams**.
- Bypass the classical (1/2 + ϵ)-approx vs $\Omega_\epsilon(n)$ space trade-off [Kapralov, Krachun, STOC'19]
- No need to store predictions (only oracle access suffices)

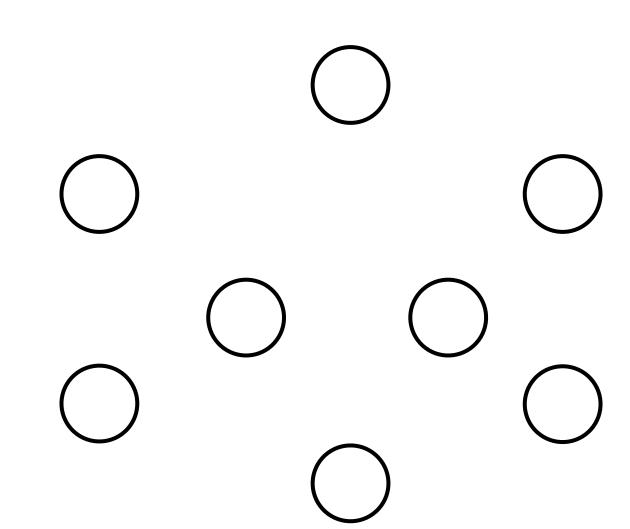
Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

- initially X = 0
- for each edge (u, v) in the stream do:
 - $Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$
 - if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X

Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

- initially X = 0
- for each edge (u, v) in the stream do:
 - $Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$
 - if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X

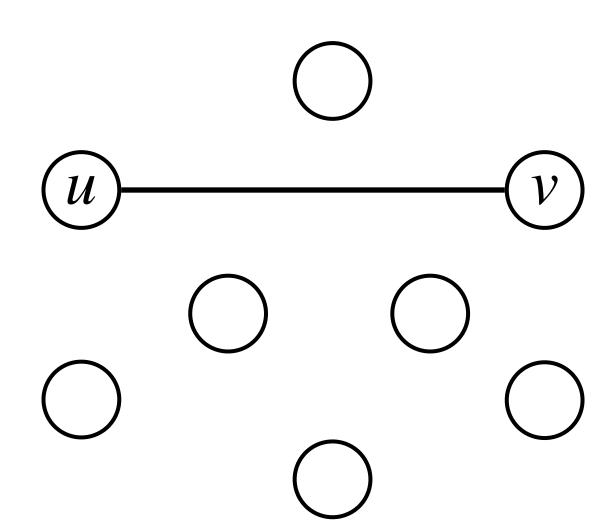


Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

- initially X = 0
- for each edge (u, v) in the stream do:

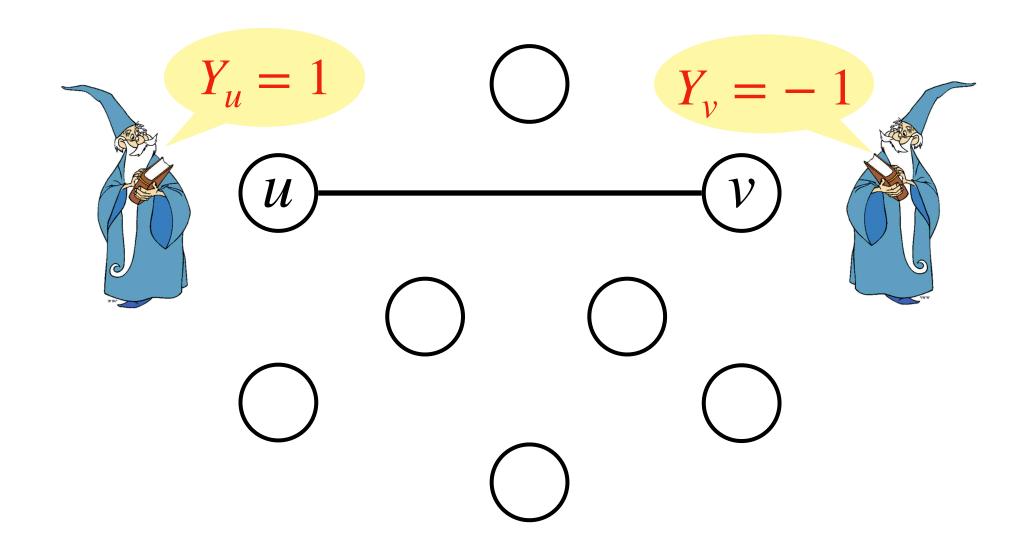
•
$$Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$$

- if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X



Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

- initially X = 0
- for each edge (u, v) in the stream do:
 - $Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$
 - if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X

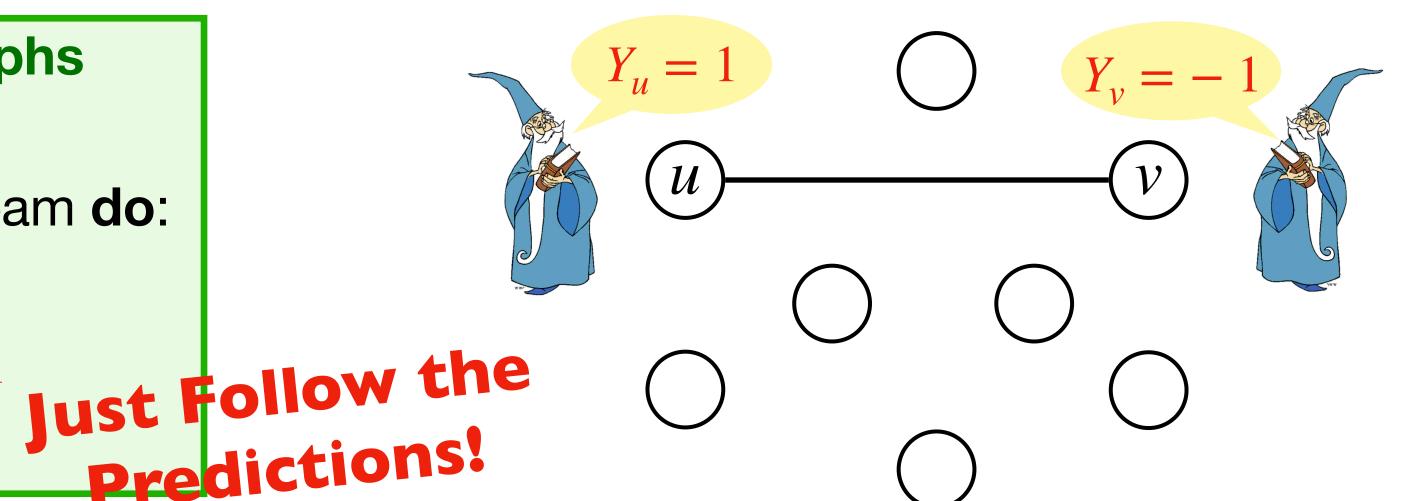


Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

- initially X = 0
- for each edge (u, v) in the stream do:

•
$$Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$$

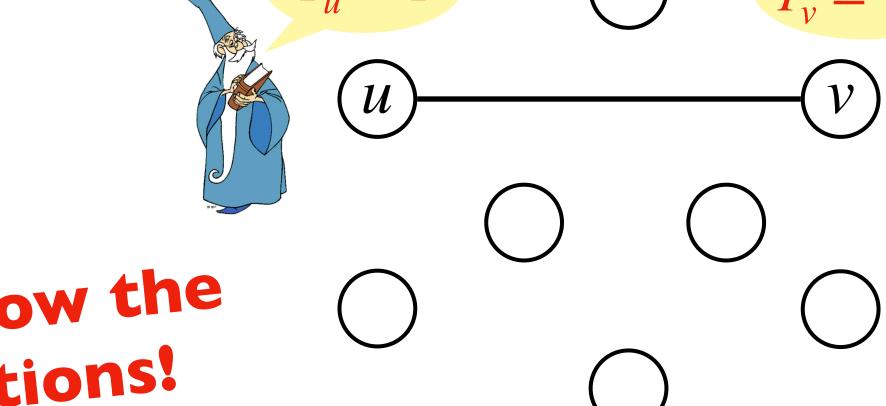
- if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X



Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

Algorithm for Low-Degree Graphs

- initially X = 0
- for each edge (u, v) in the stream do:
 - $Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$
 - if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X



Observation:

 ϵ -accurate predictions

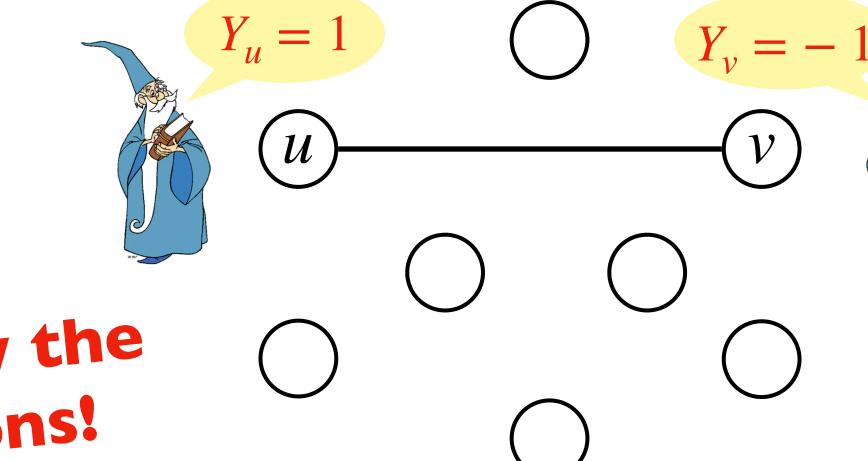
w.p.
$$\geq 2/3$$

• $(1/2 + \epsilon^2)$ -approx., O(1) words of space

Low-Degree Graphs: Graphs with maximum degree $\Delta = \Theta(\epsilon^2 m)$

Algorithm for Low-Degree Graphs

- initially X = 0
- for each edge (u, v) in the stream do:
 - $Y_u = \mathcal{O}(u), Y_v = \mathcal{O}(v)$
 - if $Y_u \neq Y_v$ then X = X + 1
- \cdot return X



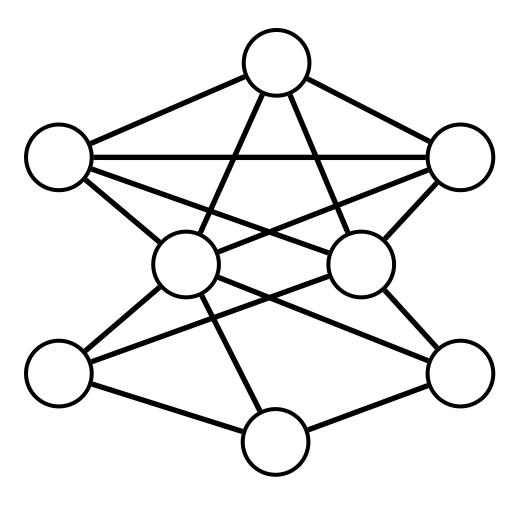
Observation:

 ϵ -accurate predictions

w.p.
$$\geq 2/3$$

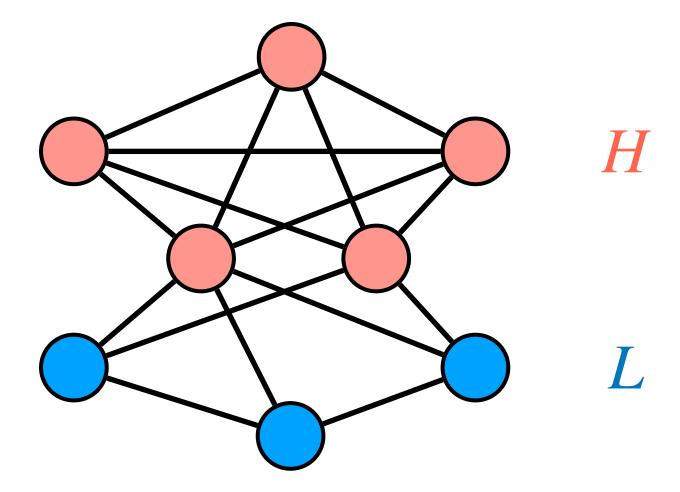
• $(1/2 + \epsilon^2)$ -approx., O(1) words of space

What to do in general graphs?



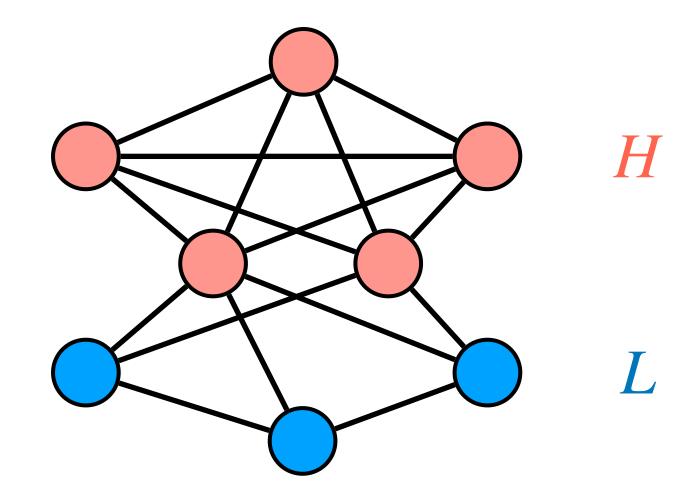
Offline Algorithm for General Graphs

• Divide V into H (high-degree) and L (low-degree)



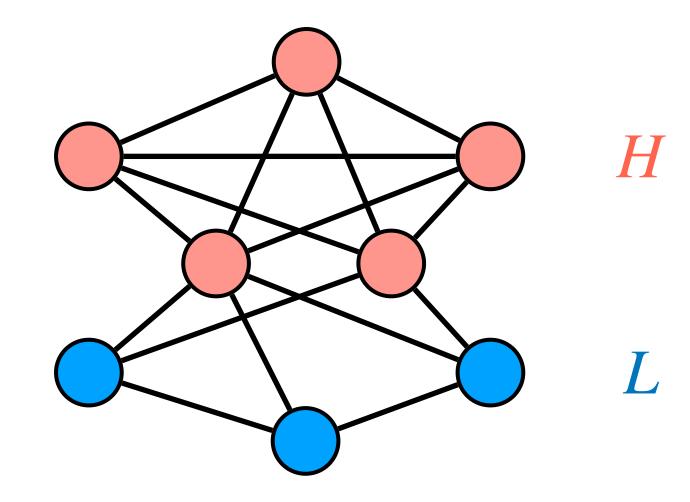
Offline Algorithm for General Graphs

• Divide V into H (high-degree) and L (low-degree)



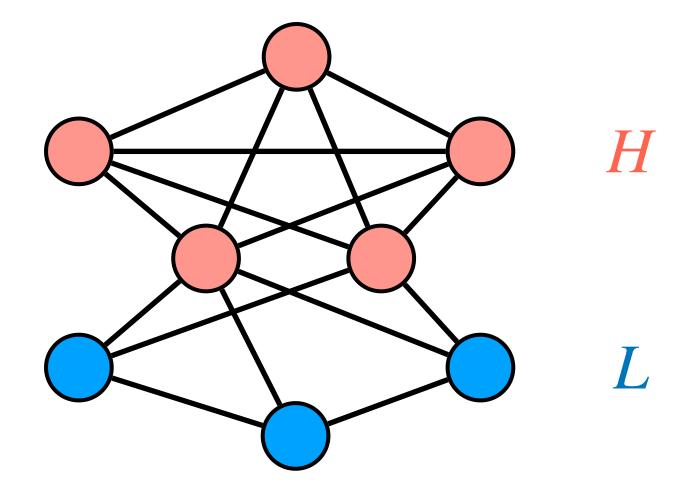
- Degree threshold is $\Theta(\epsilon^2 m)$, so
 - $|H| \le 1/\epsilon^2$, and $|E(H)| < |H|^2 \le 1/\epsilon^4$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:

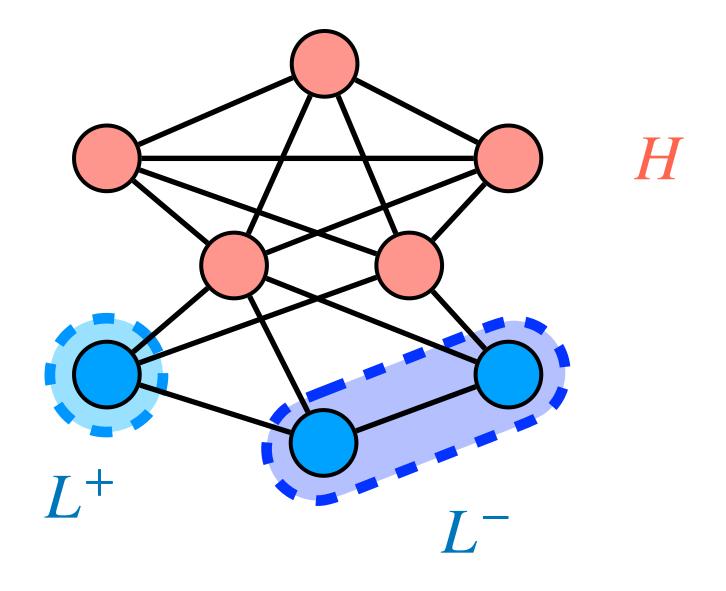


- Degree threshold is $\Theta(\epsilon^2 m)$, so
 - $|H| \le 1/\epsilon^2$, and $|E(H)| < |H|^2 \le 1/\epsilon^4$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)

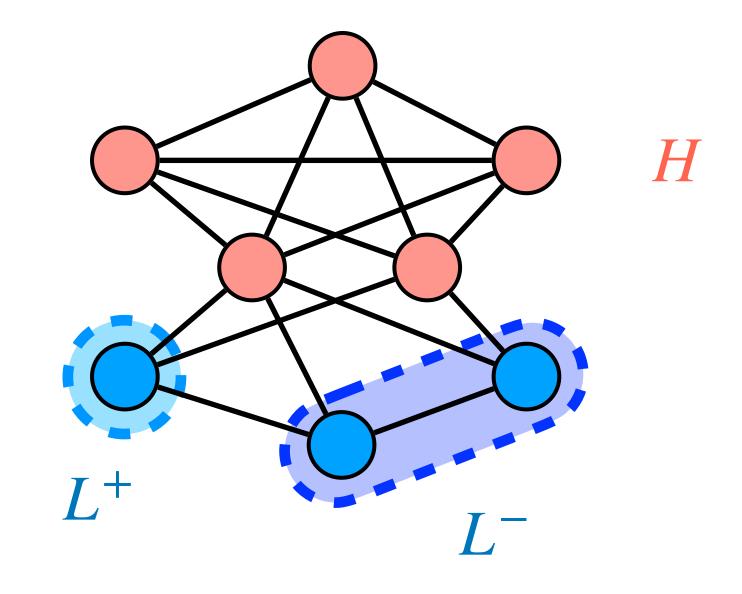


- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)



Offline Algorithm for General Graphs

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)



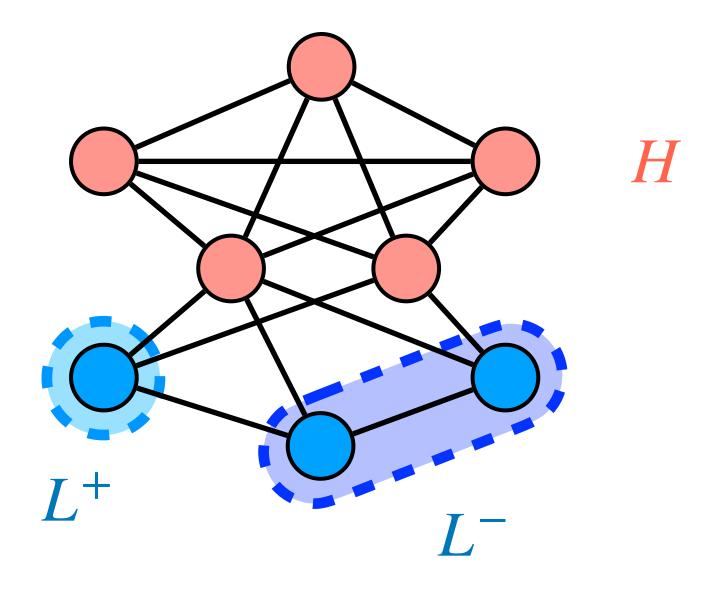
Observation on Low-Degree Graphs:

 ϵ -accurate predictions

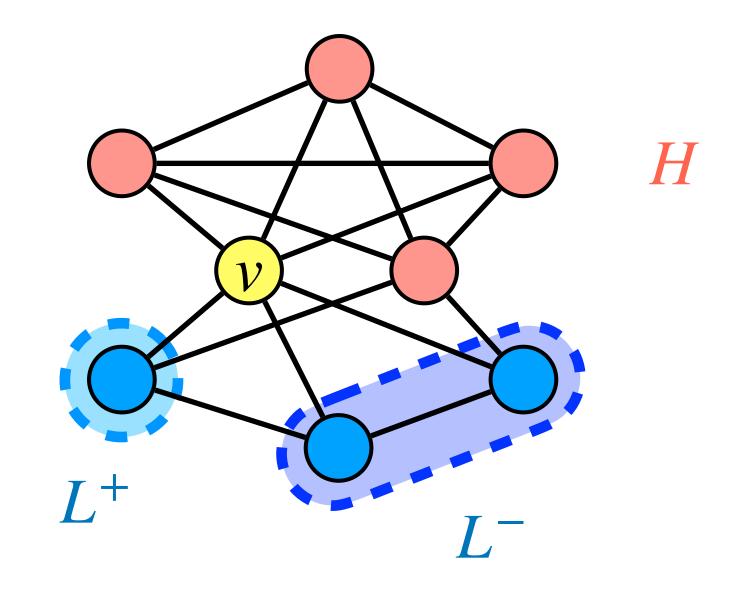
w.p.
$$\geq 2/3$$

 $(1/2 + \epsilon^2)$ -approx., O(1) words of space

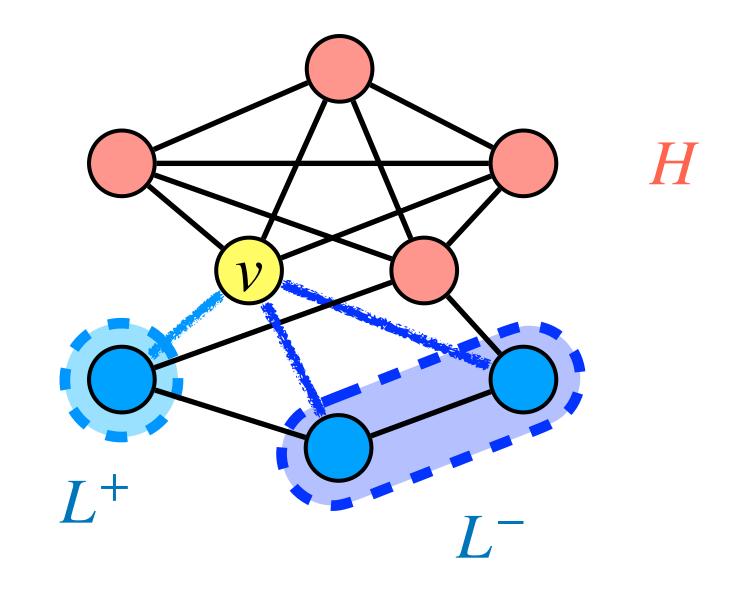
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



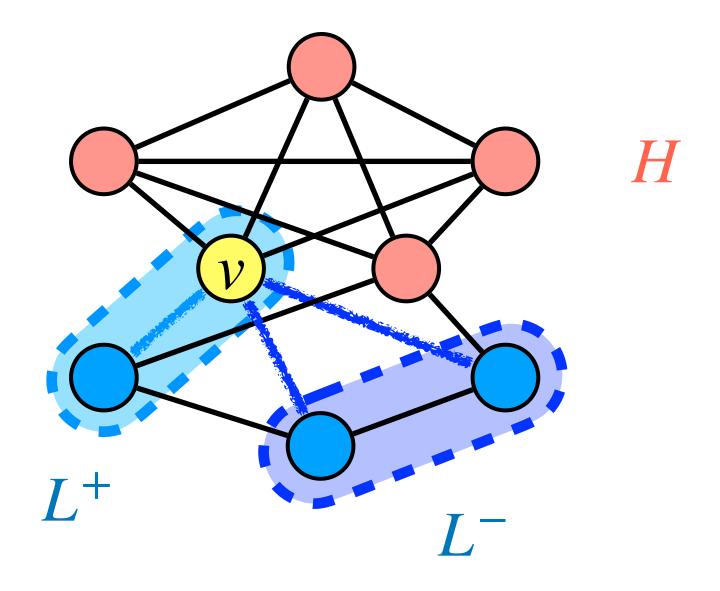
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



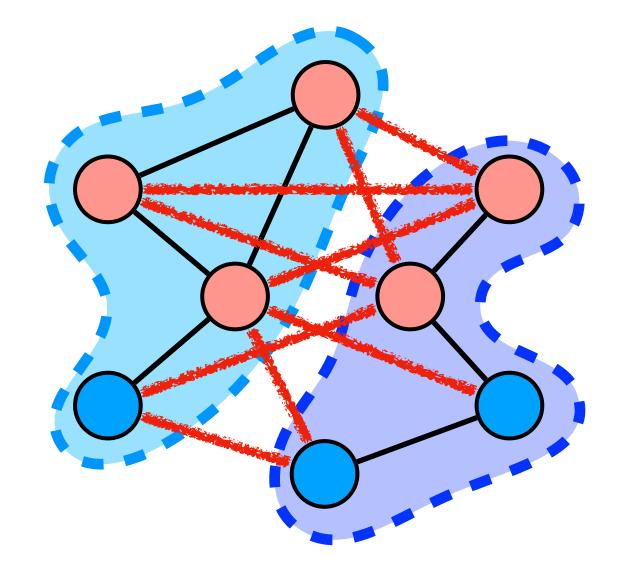
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



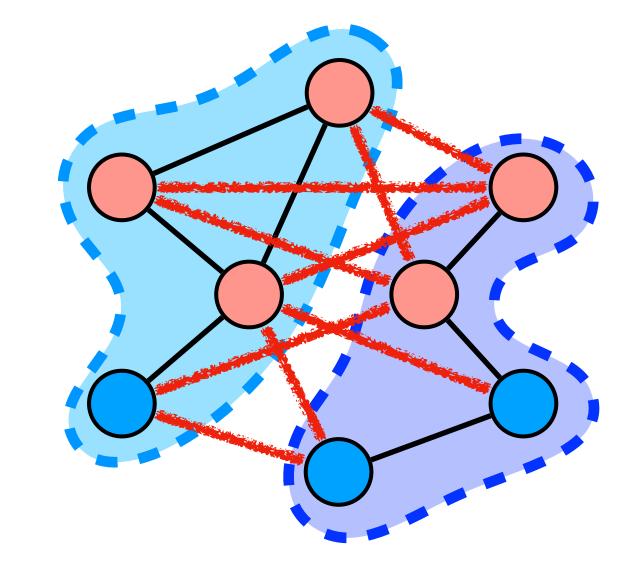
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



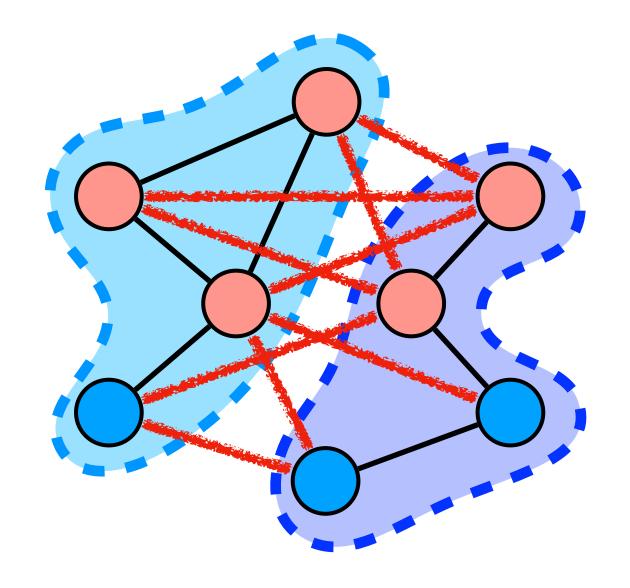
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach



$$Z_1 = |E(L^+, L^-)| + \sum_{v \in H} \max\{|E(v, L^+)|, |E(v, L^-)|\}$$

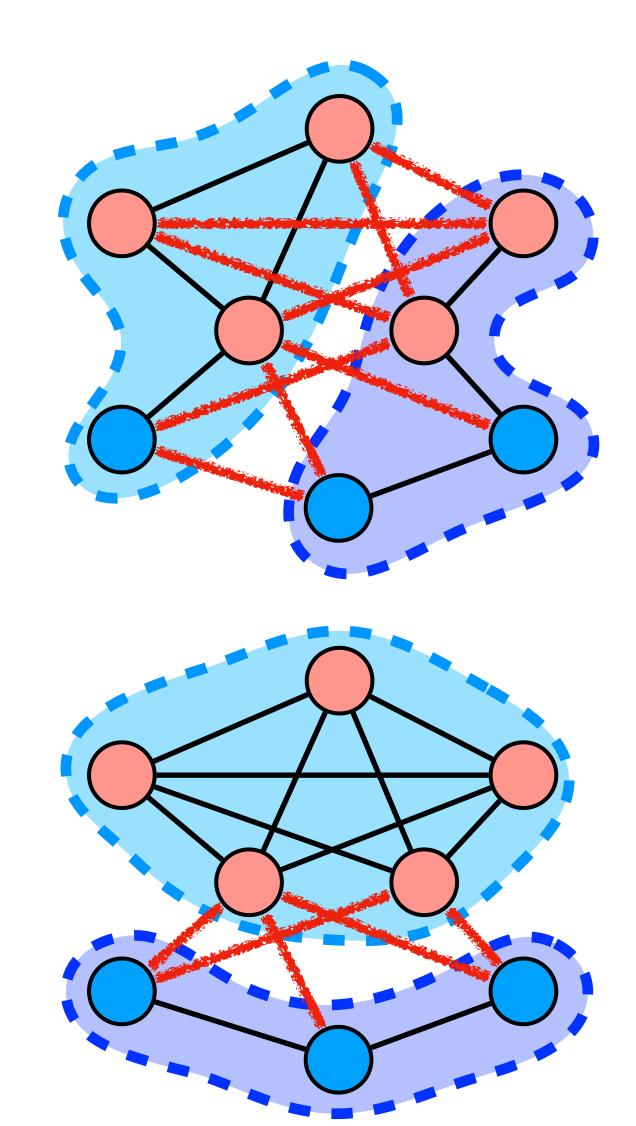
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

$$Z_1 = |E(L^+, L^-)| + \sum_{v \in H} \max\{|E(v, L^+)|, |E(v, L^-)|\}$$



- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

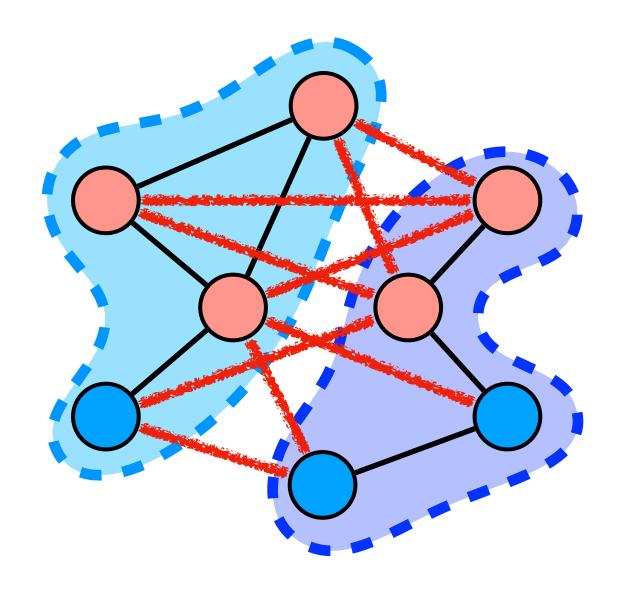
$$Z_1 = |E(L^+, L^-)| + \sum_{v \in H} \max\{|E(v, L^+)|, |E(v, L^-)|\}$$

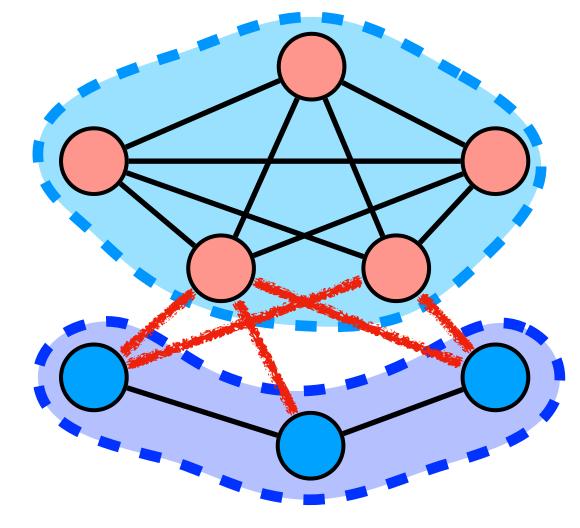


- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

$$Z_1 = |E(L^+, L^-)| + \sum_{v \in H} \max\{|E(v, L^+)|, |E(v, L^-)|\}$$

•
$$Z_2 = |E(H, L)|$$





Approximation Analysis of Our Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

Approximation Analysis of Our Algorithm

Offline Algorithm for General Graphs

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

• Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$

Approximation Analysis of Our Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$Z_1 = |E(L^+, L^-)| + \sum_{v \in H} \max\{|E(v, L^+)|, |E(v, L^-)|\}$$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$Z_{1} = |E(L^{+}, L^{-})| + \sum_{v \in H} \max\{|E(v, L^{+})|, |E(v, L^{-})|\}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \cdot \mathsf{OPT}_{L} + \frac{|E(H, L)|}{2}$$

Offline Algorithm for General Graphs

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$Z_{1} = |E(L^{+}, L^{-})| + \sum_{v \in H} \max\{|E(v, L^{+})|, |E(v, L^{-})|\}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \cdot \mathsf{OPT}_{L} + \frac{|E(H, L)|}{2}$$

Observation: $|E(L^+, L^-)| \ge (1/2 + \epsilon^2) \cdot \mathsf{OPT}_L$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$Z_{1} = |E(L^{+}, L^{-})| + \sum_{v \in H} \max\{|E(v, L^{+})|, |E(v, L^{-})|\}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \cdot \mathsf{OPT}_{L} + \frac{|E(H, L)|}{2}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \left(1 - \alpha - \epsilon^{4}\right) \cdot \mathsf{OPT} + \frac{\alpha \cdot \mathsf{OPT}}{2}$$

Offline Algorithm for General Graphs

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

Lemma: Suppose $|E(H,L)| = \alpha \cdot \text{OPT}$ where $\alpha < 1/2 + \epsilon^2$, then $\text{OPT}_L > (1 - \alpha - \epsilon^4) \cdot \text{OPT}$

- OPT \leq OPT_L + OPT_H + |E(H, L)|
- OPT_H $\leq |E(H)| \leq 1/\epsilon^4$

• Goal: To prove
$$\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$$

• If
$$Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$$
, DONE

• Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$Z_{1} = |E(L^{+}, L^{-})| + \sum_{v \in H} \max\{|E(v, L^{+})|, |E(v, L^{-})|\}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \cdot \mathsf{OPT}_{L} + \frac{|E(H, L)|}{2}$$

$$\geq \left(\frac{1}{2} + \epsilon^{2}\right) \left(1 - \alpha - \epsilon^{4}\right) \cdot \mathsf{OPT} + \frac{\alpha \cdot \mathsf{OPT}}{2}$$

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

- Goal: To prove $\max\{Z_1, Z_2\} \ge (1/2 + \Omega(\epsilon^2)) \cdot \mathsf{OPT}$
- If $Z_2 = |E(H, L)| \ge (1/2 + \epsilon^2) \cdot \text{OPT}$, DONE
- Otherwise, $|E(H,L)| < (1/2 + \epsilon^2) \cdot \text{OPT}$, it suffices to prove $Z_1 \geq (1/2 + \Omega(\epsilon^2)) \cdot \text{OPT}$

$$\begin{split} Z_1 &= |E(L^+, L^-)| + \sum_{v \in H} \max\{ |E(v, L^+)|, |E(v, L^-)| \} \\ &\geq \left(\frac{1}{2} + \epsilon^2 \right) \cdot \mathsf{OPT}_L + \frac{|E(H, L)|}{2} \\ &\geq \left(\frac{1}{2} + \epsilon^2 \right) \left(1 - \alpha - \epsilon^4 \right) \cdot \mathsf{OPT} + \frac{\alpha \cdot \mathsf{OPT}}{2} \\ &\geq \left(\frac{1}{2} + \Omega(\epsilon^2) \right) \cdot \mathsf{OPT} \end{split}$$

Implementation of Our Algorithm in Streaming

Implementation of Our Algorithm in Streaming

Random Order Streams (insertion-only)

Arbitrary Order Streams (insertion-only)

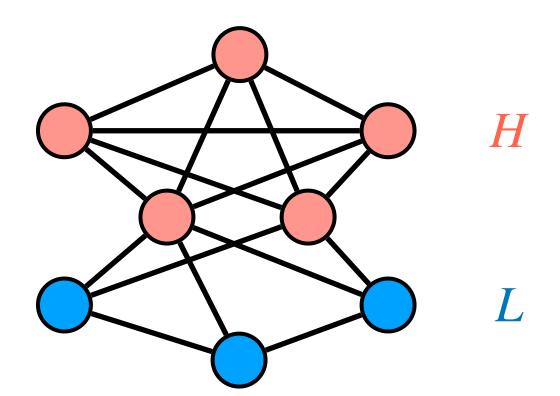
Dynamic Streams

Random Order Stream: Edges of the input graph arrive in a uniformly random order

Random Order Stream: Edges of the input graph arrive in a uniformly random order

Offline Algorithm

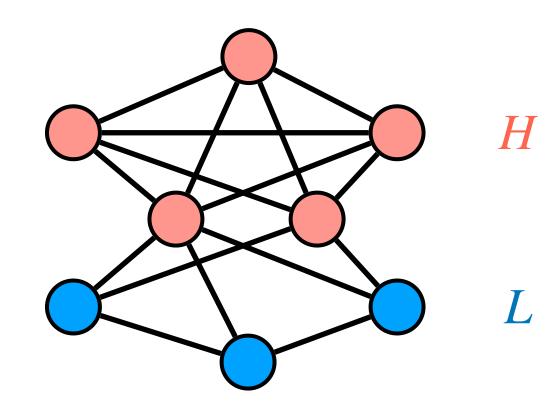
- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)



Random Order Stream: Edges of the input graph arrive in a uniformly random order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)



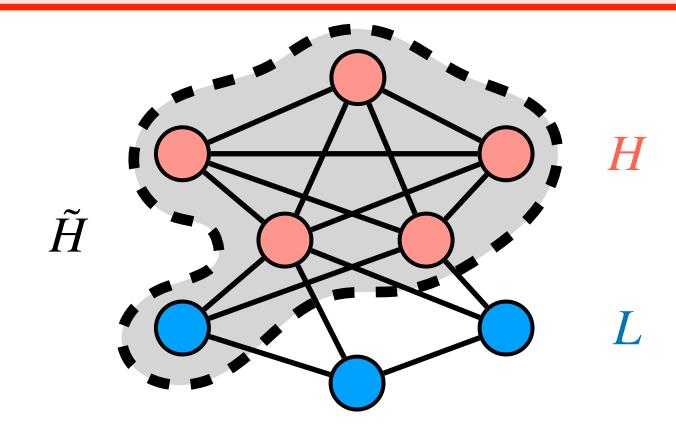
Algorithm in Random Order Streams

- Streaming Phase:
 - Store the first $\operatorname{poly}(1/\epsilon)$ edges, $\det \tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$
 - Compute Estimate (1) and Estimate (2) w.r.t. \tilde{H} and \tilde{L}
 - Store the degree info of \tilde{H} and $E(\tilde{H})$
- Post-Processing Phase:
 - Correct H, L and the estimates using degree info of H

Random Order Stream: Edges of the input graph arrive in a uniformly random order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)



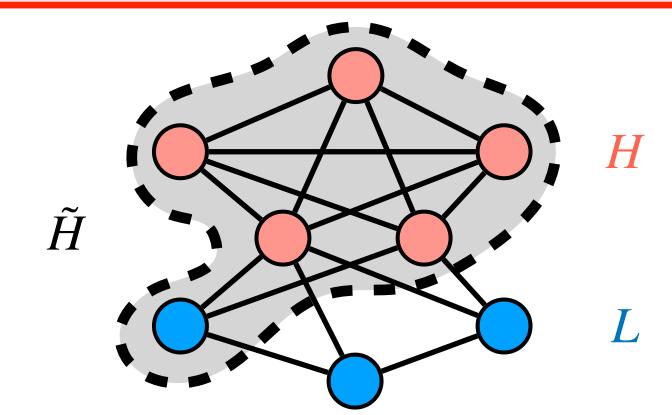
Algorithm in Random Order Streams

- Streaming Phase:
 - Store the first $\operatorname{poly}(1/\epsilon)$ edges, $\det \tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$
 - Compute Estimate (1) and Estimate (2) w.r.t. \tilde{H} and \tilde{L}
 - Store the degree info of \tilde{H} and $E(\tilde{H})$
- Post-Processing Phase:
 - Correct H, L and the estimates using degree info of H

Random Order Stream: Edges of the input graph arrive in a uniformly random order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)
- Approx: $1/2 + \Omega(\epsilon^2)$
- Space: $poly(1/\epsilon)$ words



Algorithm in Random Order Streams

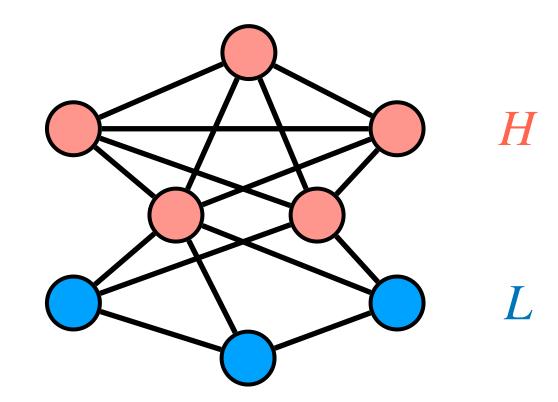
- Streaming Phase:
 - Store the first $\mathrm{poly}(1/\epsilon)$ edges, $\det \tilde{H} \supseteq H$ and $\tilde{L} \subseteq L$
 - Compute Estimate (1) and Estimate (2) w.r.t. \tilde{H} and \tilde{L}
 - Store the degree info of \tilde{H} and $E(\tilde{H})$
- Post-Processing Phase:
 - Correct H, L and the estimates using degree info of H

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

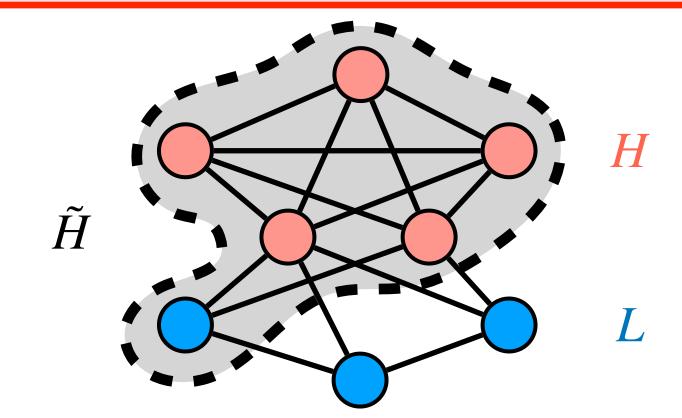


- Streaming Phase:
 - Perform **reservoir sampling** to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch (or Misra-Gries)
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

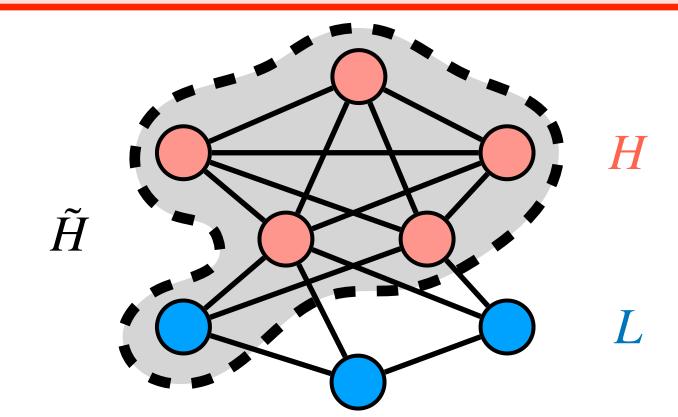


- Streaming Phase:
 - Perform **reservoir sampling** to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch (or Misra-Gries)
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)
- Since $|\tilde{H}|, |E(\tilde{H})| = \text{poly}(1/\epsilon),$ we can safely drop them in the calculation

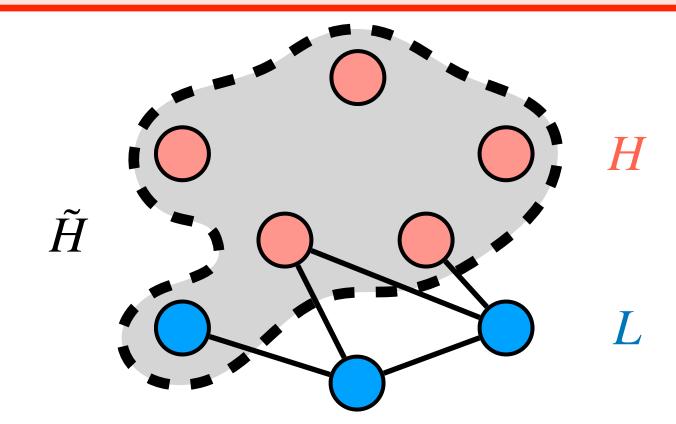


- Streaming Phase:
 - Perform **reservoir sampling** to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch (or Misra-Gries)
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)
- Since $|\tilde{H}|, |E(\tilde{H})| = \text{poly}(1/\epsilon),$ we can safely drop them in the calculation

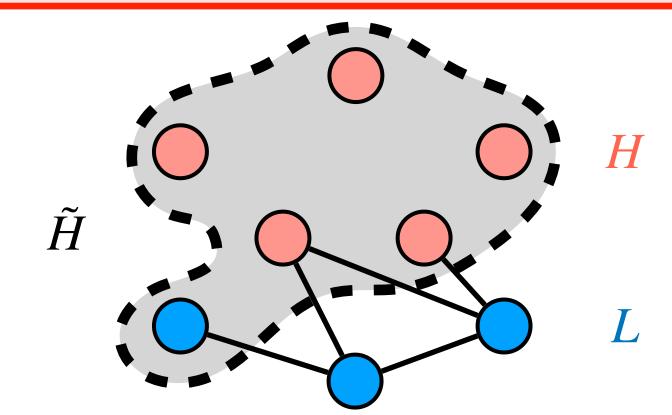


- Streaming Phase:
 - Perform **reservoir sampling** to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch (or Misra-Gries)
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Arbitrary Order Stream: Edges of the input graph arrive in an arbitrary/adversarial order

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)
- Approx: $1/2 + \Omega(\epsilon^2)$
- Space: $poly(1/\epsilon)$ words



- Streaming Phase:
 - Perform **reservoir sampling** to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch (or Misra-Gries)
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Dynamic Stream: The stream contains both edge insertions and deletions

Dynamic Stream: The stream contains both edge insertions and deletions

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)

Algorithm in Dynamic Streams

- Streaming Phase:
 - Perform ℓ_0 -sampling to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

Dynamic Stream: The stream contains both edge insertions and deletions

Offline Algorithm

- Divide V into H (high-degree) and L (low-degree)
- Output the maximum of the following Max-Cut estimates:
 - Candidate Estimate (1):
 - Follow predictions on G[L], compute the cut value of (L^+, L^-)
 - Extend (L^+, L^-) with vertices in H via a greedy approach
 - Candidate Estimate (2):
 - Compute the cut value of (H, L)
- Approx: $1/2 + \Omega(\epsilon^2)$
- Space: $poly(1/\epsilon, \log n)$ words

Algorithm in Dynamic Streams

- Streaming Phase:
 - Perform ℓ_0 -sampling to obtain $poly(1/\epsilon)$ edges
 - Store the degree info of all vertices to V^+ and V^- using CountMin Sketch
- Post-Processing Phase:
 - Detect $\tilde{H}\supseteq H$ and $\tilde{L}\subseteq L$ from the sampled edges
 - Compute Estimate (1) and Estimate (2) approximately

- Streaming Max-Cut with predictions
 - Takeaway: ϵ -accurate predictions can help bypass the classical $(1/2 + \epsilon)$ -approx vs $\Omega_{\epsilon}(n)$ space trade-off!

Streaming Max-Cut with predictions

• Takeaway: ϵ -accurate predictions can help bypass the classical $(1/2+\epsilon)$ -approx vs $\Omega_{\epsilon}(n)$ space trade-off!

Open Problems

- Extension to weighted graphs?
- Other graph problems in streaming?
- Other prediction model? E.g., *€*-accurate edge predictions [Aamand, Chen, Gollapudi, Silwal, Wu, ICML'25]

Streaming Max-Cut with predictions

• Takeaway: ϵ -accurate predictions can help bypass the classical $(1/2+\epsilon)$ -approx vs $\Omega_{\epsilon}(n)$ space trade-off!

Open Problems

- Extension to weighted graphs?
- Other graph problems in streaming?
- Other prediction model? E.g., ϵ -accurate edge predictions [Aamand, Chen, Gollapudi, Silwal, Wu, ICML'25]

