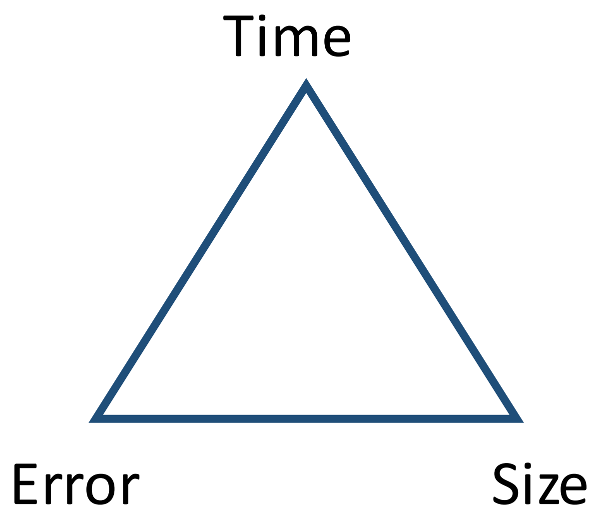
Computational Tradeoff in Network Problem

Advisor: I-Hsiang Wang

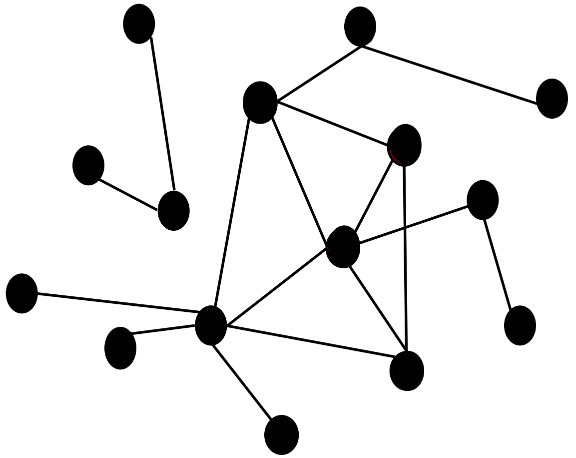
Student: Chi-Ning Chou

Overview

- Tradeoff among
  - Error
  - Time
  - Size



- Network problems:
  - Detecting/Locating hidden structures
  - Applications to other problems
  - Computational lower bound



Schedule For This Semester

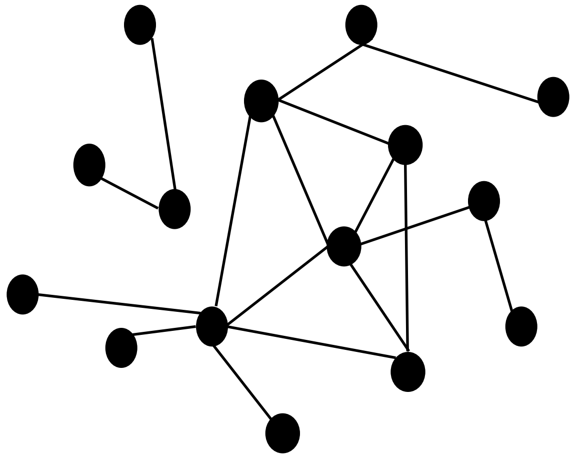
9/2015-10/2015	Survey on computational tradeoff in network problem.
10/2015-12/2015	Study SOS.
12/2015-1/2016	Working on information-theoretical bounds for several extension problems.

Outline

- Motivations
- Historical Works
- Problem Formulations
- My Current Works

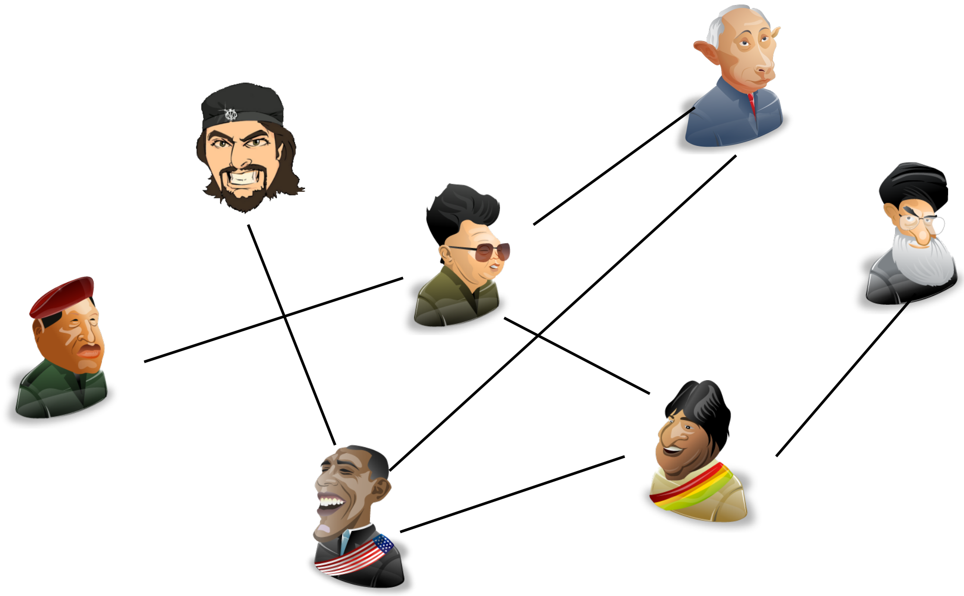
Motivations

Detecting/Locating Hidden Structures



- Detection: Is there a 5 clique?
- Location: Can you find a 5 clique?

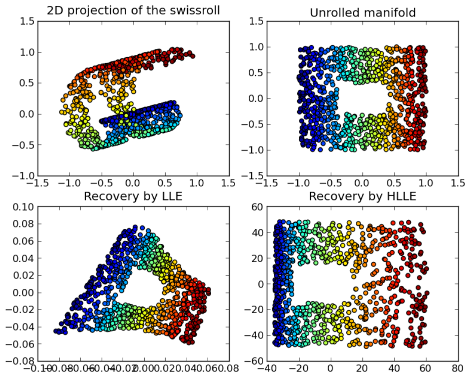
- Network problems



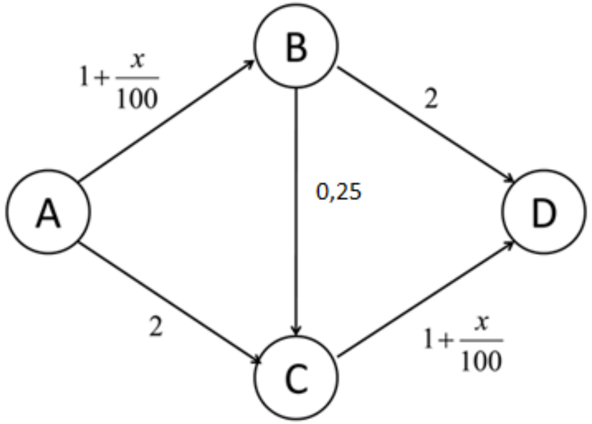
- Average-case complexity issues

Applications

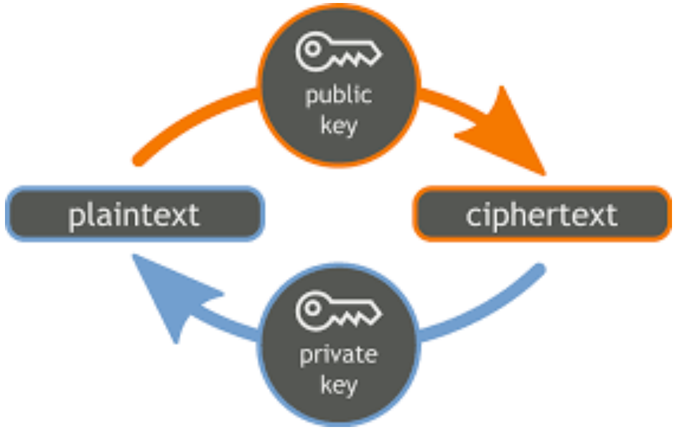
Biology: DNA



Statistics: Sparse PCA



Economy: Nash equilibrium



Cryptography: PKE

- Biology: Finding subtle signal in DNA sequences [PS00]
- Statistics: Lower bound for sparse PCA [BR13]
- Economy: Approximation hardness of Nash equilibrium [HK11]
- Cryptography: Public-key cryptography [ABW10]
- Others: Certification of the restricted isometry property [KZ14], etc.

Planted clique/subgraph problem

$\approx$

- Network problem
- Average-case complexity

This is a fundamental problem!

Computational Tradeoff

- Biology: Finding subtle signal in DNA sequences [PS00]
- Statistics: Lower bound for sparse PCA [BR13]
- Economy: Approximation hardness of Nash equilibrium [HK11]
- Cryptography: Public-key cryptography [ABW10]
- Others: Certification of the restricted isometry property [KZ14], etc.

Planted clique/subgraph problem

This is hard

Historical Works

History

1984: Karp suggested that finding clique with size $\mathrm{k} = (1+\epsilon)\log n$ is hard.

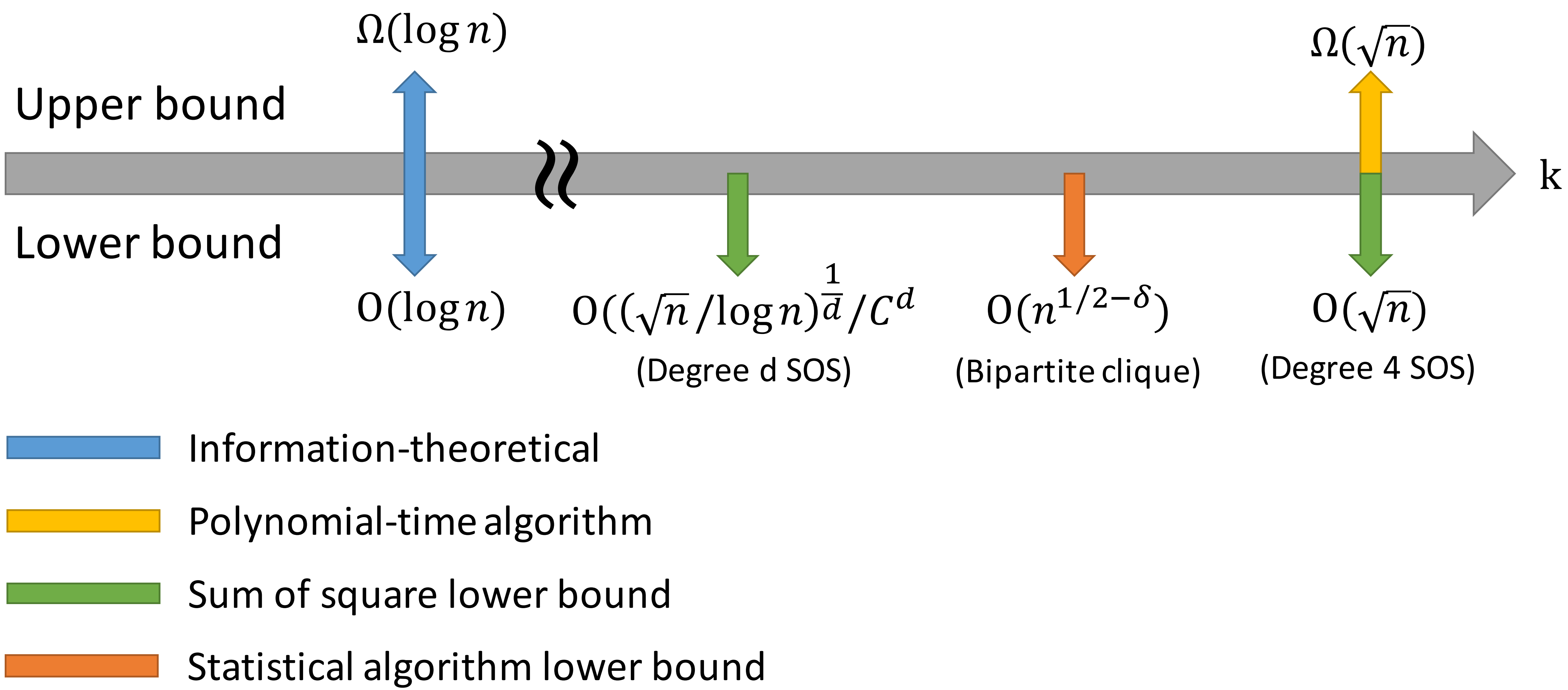
1995: Kucera formulated the planted clique problem.

1998: Alon, Krivelevich, and Sudakov presented a poly-time algorithm for $\mathrm{k} = \Omega(\sqrt{n})$

2002: Feige, and Krauthgamer showed Lov́ asz-Schrijver lwer bound for $\mathrm{k} = O(\sqrt{n/2^d})$

2015:Meka, Potechin, and Wigderson showed SOS lower bound for $\mathrm{k} = O((\sqrt{n}/\log n)^{\frac{1}{d}}/C^d)$

Current Status: Planted Clique

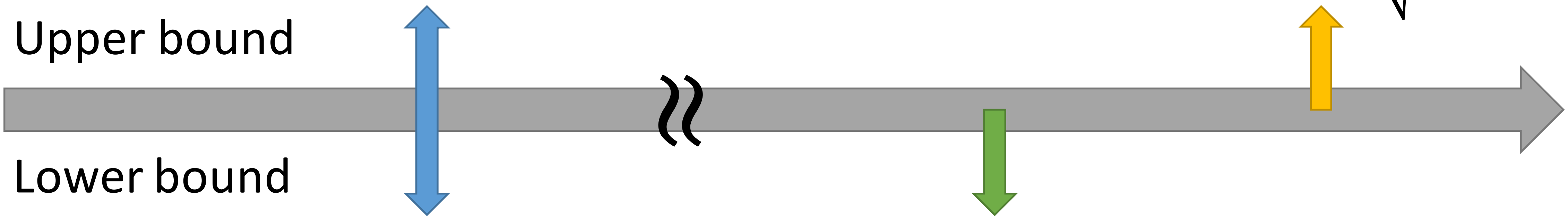


Current Status: Detecting Hidden Subgraph

$$\frac{d(H_n)}{\log n} > \frac{1}{\log 1/q_0}$$

$$\frac{\lambda_1\left(\boldsymbol{A}_{H_n}\right)}{\sqrt{n}} > 3\sqrt{\frac{q_o}{1-q_0}}$$

Upper bound



Lower bound

$$\frac{d(H_n)}{\log n} < \frac{1}{\log 1/q_0}$$

$$\frac{\lambda_1\left(\boldsymbol{A}_{H_n}\right)}{\sqrt{n}} < \frac{1}{4}\sqrt{\frac{q_o}{1-q_0}}$$

Information-theoretical

Spectral algorithm

A direct SDP relaxation

Problem Formulations

Two Problems

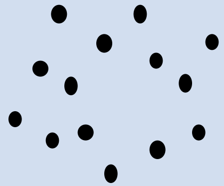
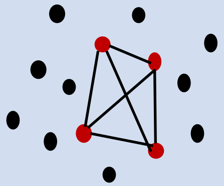
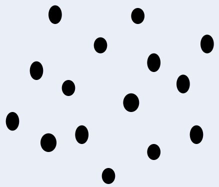
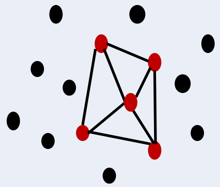
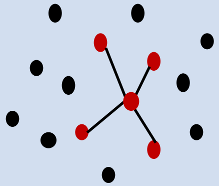
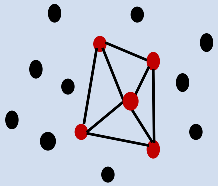
- Detecting problem & locating problem

$\underset{?}{\cancel{\leq}}$

- Detecting is no harder than locating
- Is there a gap between these two problems?
- Extend to general settings.
  - E.g. regular subgraph, two distinct graphs, etc.

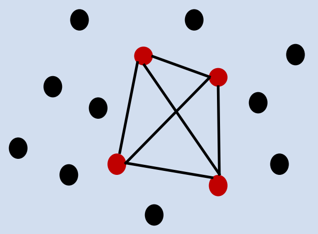
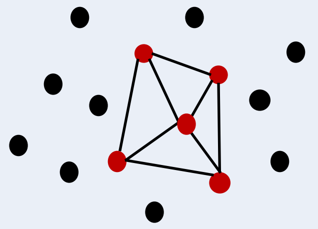
Detecting Problem

- Given a graph G, can you determine which distribution does G follow?

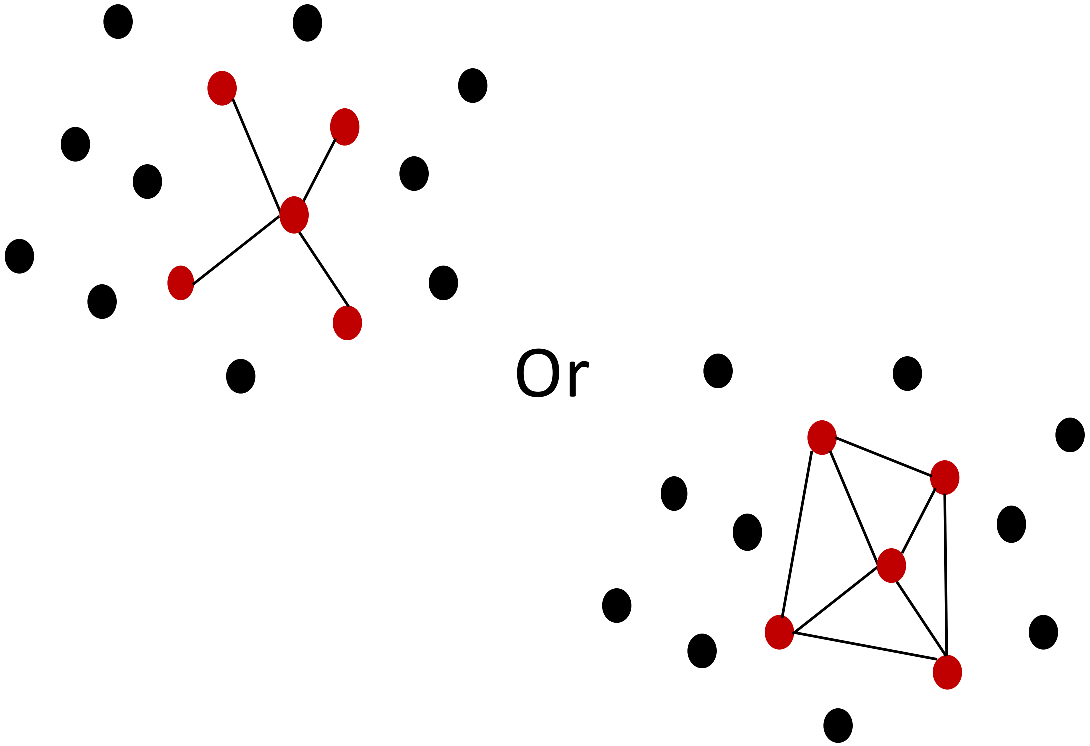
	$\mathcal{H}_0$	$\mathcal{H}_1$
Planted clique		
Hidden structure		
Two distinct structures		

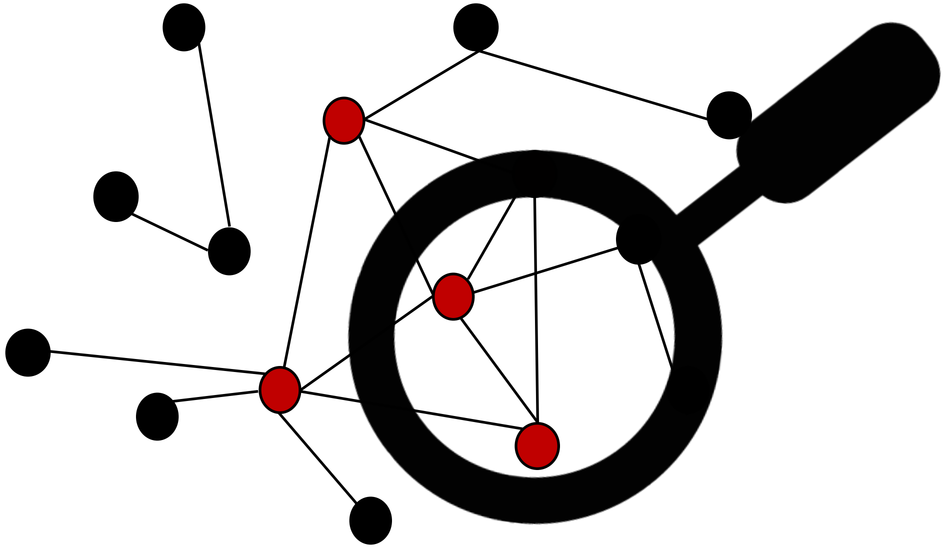
Locating Problem

- Given a graph G, can you find the planted subgraph?

	$\boldsymbol{H_n}$
Planted clique	
Hidden structure	

Current Works





Distinguishing two distinct subgraphs

Locating hidden subgraph

Distinguishing Two Distinct Subgraphs

- Motivations
  - Generalization of detecting problem.
  - We might want to know which is the actual underlying substructure.

- Difficulties
  - How to measure the difference of two graphs?
  - Need to divide the possible error cases in to more categories for the tightness of bounds.

Locating Hidden Subgraph

- Motivations
  - We might want to know where is the actual underlying substructure.
  - Find out whether there is a gap between detecting and locating.

- Difficulties
  - Union bound cannot work for constructing lower bound.
  - There might be no unified bound for all kinds of graphs.

Central Ideas

Identifying critical graph property

Union bound for upper bound

Partition error cases or Janson inequality for lower bound

Identifying Critical Graph Properties

- The size of the hidden clique can affect the computation!

Hard ⟵ Clique size k ⟶ Easy

- What about hidden subgraph?
  - We need to identify some nice properties.
    - Density: $d(H) = \max_{F \subseteq H} \frac{e(F)}{d(F)}$
    - Edge expansion: $\alpha(H) = \min_{S \subseteq V(H)} \frac{|\partial S|}{|S|}$
    - Maximum degree: $\mathrm{d}_{max}(H) = \max_{v \in V(H)} d(v)$
    - Graph distortion: $D(H_1, H_0)$

Upper Bound: Union Bound

Exact k vertices to k vertices error

Union bound

Exact k vertices error

Union bound

k vertices error

Union bound

Detecting error

Lower Bound: Janson Inequality

- Goal: Lower bound the error probability

$$A_i$$

Events
(Single error)

$$E = A_1^C \cap A_2^C \cap \cdots \cap A_n^C$$

Intersection of events
(No error)

$$\mathbb{P}[E] \leq \exp -\mu + \frac{\Delta}{2}$$

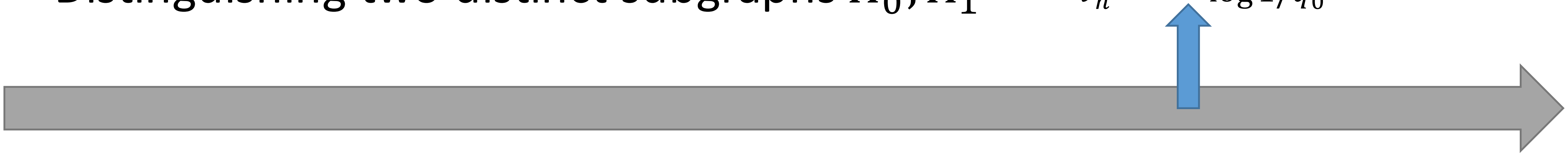
Upper bound
(Probability of no error)

Lower bound for the error probability!

Current Results

- Distinguishing two distinct subgraphs H_0, H_1

$$\frac{D(H_1, H_0)}{v_n} \geq \frac{\log n}{\log 1/q_0}$$



- Locating hidden subgraph H_n

$$\alpha(H_n) = \omega(\frac{\log n}{\log 1/q_0})$$

$$d_{max}(H_n) = o(\frac{\log n}{\log 1/q_0})$$

Future Works

Information-theoretical bounds for every possible extensions.

SOS lower bound.

Statistical algorithm lower bound.

References: Motivations

[ABW10] Public-key cryptography from different assumptions.

[BR13] Complexity Theoretic Lower Bounds for Sparse Principal Component Detection.

[HK11] How hard is it to approximate the best Nash equilibrium?

[KZ14] Hidden Cliques and the Certification of the Restricted Isometry Property.

[PS00] Combinatorial approaches to finding subtle signal in DNA sequences.

References: Historical Works

[Kar76] Probabilistic analysis of some combinatorial search problems.

[Kuc95] Expected complexity of graph partitioning problems.

[AKS98] Finding a large hidden clique in a random graph.

[FK02] The probable value of the Lovász-Schrijver relaxations for maximum independent set.

[MPW15] Sum-of-Squares Lower Bounds for Planted Clique.

Q & A