Property Testing and the Container Method

Eric Blais   Cameron Seth

June 2024

UNIVERSITY OF WATERLOO

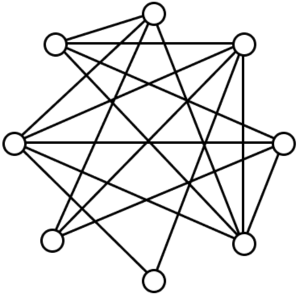
Outline

1. What is the Container Method?
2. How to use Container Method to analyze property testers.
3. New container lemmas for new property testing results.

What is the Container Method?

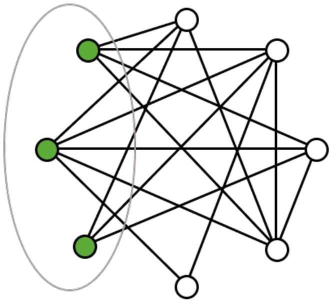
What is the Container Method?

Answer: a tool for characterizing independent sets in some graphs.



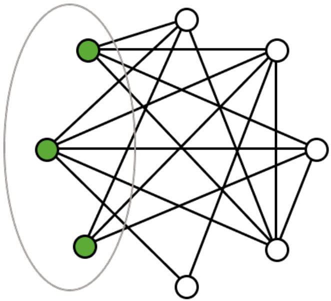
What is the Container Method?

Answer: a tool for characterizing independent sets in some graphs.



What is the Container Method?

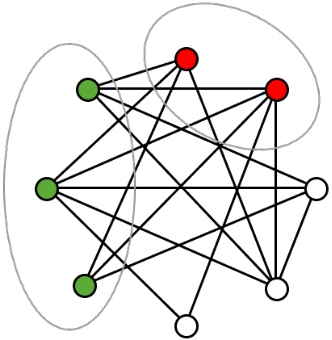
Answer: a tool for characterizing independent sets in some graphs.



Informal idea. For any graph satisfying some "nice" conditions, all independent sets in the graph are clustered together into a small number of containers.

What is the Container Method?

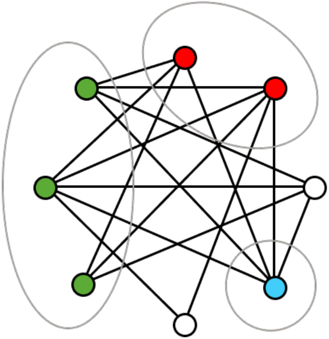
Answer: a tool for characterizing independent sets in some graphs.



Informal idea. For any graph satisfying some "nice" conditions, all independent sets in the graph are clustered together into a small number of containers.

What is the Container Method?

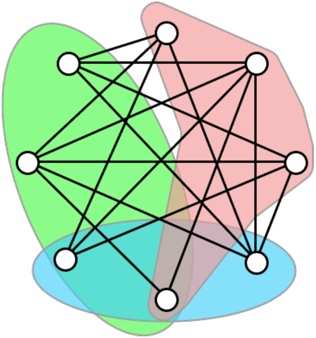
Answer: a tool for characterizing independent sets in some graphs.



Informal idea. For any graph satisfying some "nice" conditions, all independent sets in the graph are clustered together into a small number of containers.

What is the Container Method?

Answer: a tool for characterizing independent sets in some graphs.



Informal idea. For any graph satisfying some "nice" conditions, all independent sets in the graph are clustered together into a small number of containers.

An Initial Graph Container Lemma

An Initial Graph Container Lemma

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$. [Kleitman, Winston '82]

An Initial Graph Container Lemma

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

[Kleitman, Winston '82]

Examples:

- Random graphs
- d-regular graphs
- graphs that are ϵ-far from having a ρn independent set

An Initial Graph Container Lemma[1]

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

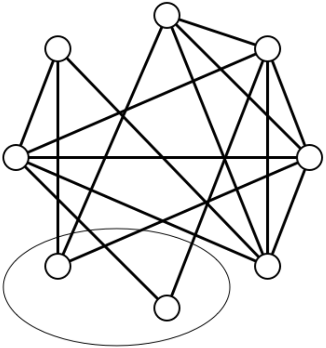
[Kleitman, Winston '82]

Examples:

- Random graphs
- d-regular graphs
- graphs that are ϵ-far from having a ρn independent set

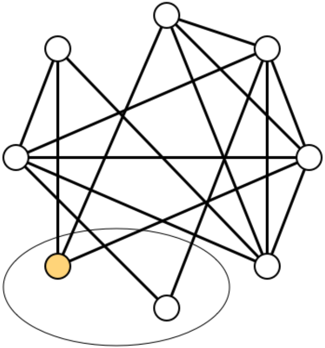
[1] For survey of combinatorial applications see "Counting Indepedent Sets in Graphs" by Samotij or "The method of hypergraph containers" by Balogh, Morris, and Samotij.

Graph Container Method – Encoding View



How can I give you the most information about an independent set I by just telling you about one of the vertices in I?

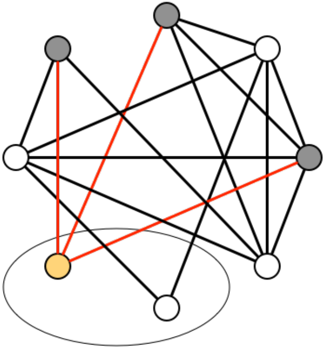
Graph Container Method – Encoding View



How can I give you the most information about an independent set I by just telling you about one of the vertices in I?

Answer: Send the vertex $v \in I$ with highest degree.

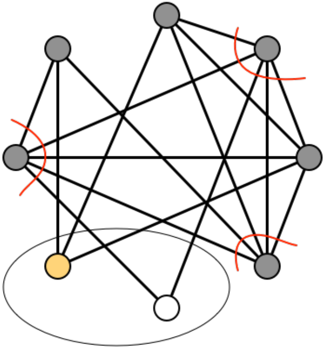
Graph Container Method – Encoding View



How can I give you the most information about an independent set I by just telling you about one of the vertices in I?

Answer: Send the vertex $v \in I$ with highest degree.

Graph Container Method – Encoding View



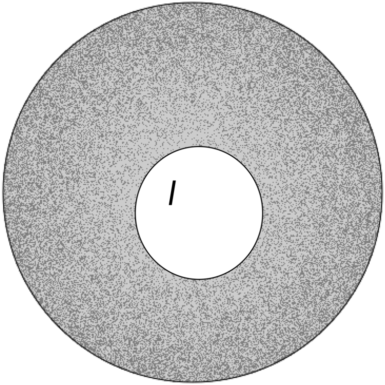
How can I give you the most information about an independent set I by just telling you about one of the vertices in I?

Answer: Send the vertex $v \in I$ with highest degree.

Container Generation Algorithm

Input: Graph G and an independent set I

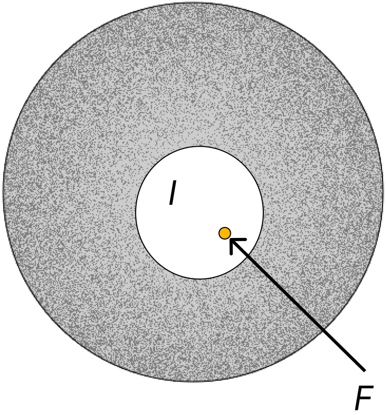
Output: Container $C(I)$ and Fingerprint $F(I)$



Container Generation Algorithm

Input: Graph G and an independent set I

Output: Container $C(I)$ and Fingerprint $F(I)$

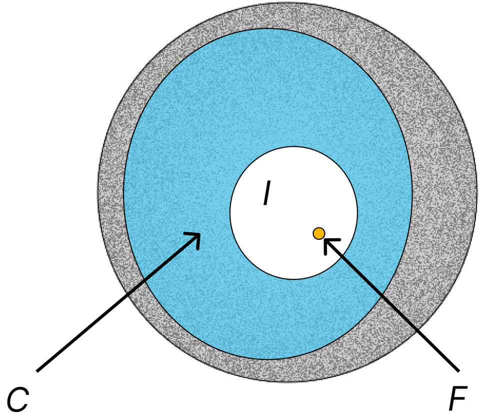


1st Iteration

Container Generation Algorithm

Input: Graph G and an independent set I

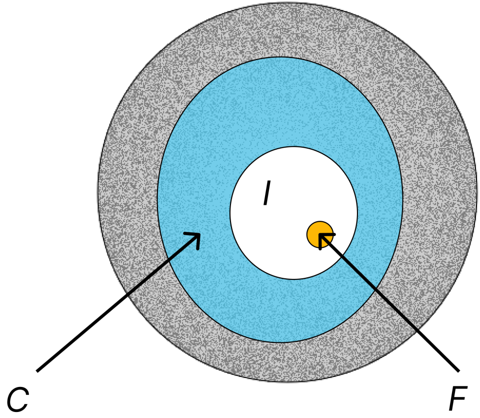
Output: Container $C(I)$ and Fingerprint $F(I)$



1st Iteration

Container Generation Algorithm

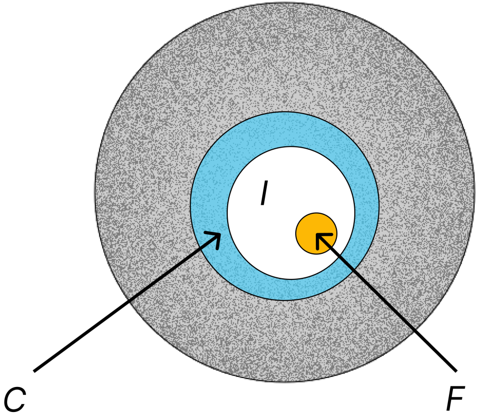
Input: Graph G and an independent set I
Output: Container $C(I)$ and Fingerprint $F(I)$



2nd Iteration

Container Generation Algorithm

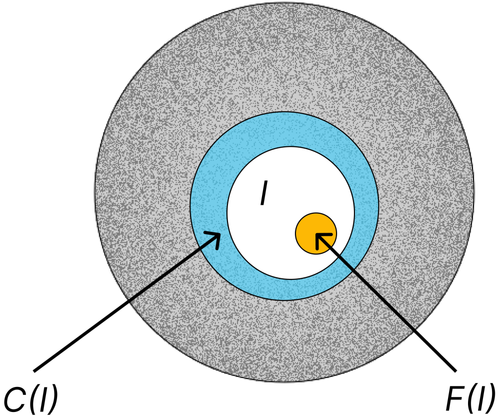
Input: Graph G and an independent set I
Output: Container $C(I)$ and Fingerprint $F(I)$



Final Iteration

Container Generation Algorithm

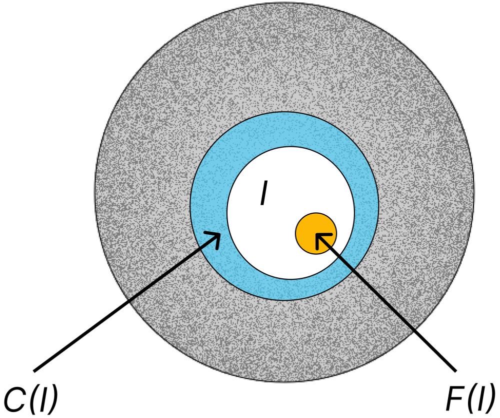
Input: Graph G and an independent set I
Output: Container $C(I)$ and Fingerprint $F(I)$



Final Iteration

Container Generation Algorithm

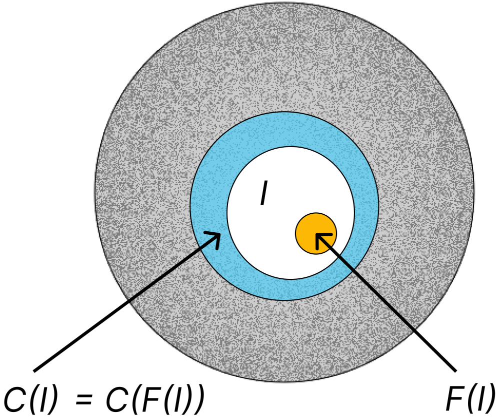
Input: Graph G and an independent set I
Output: Container $C(I)$ and Fingerprint $F(I)$



Key Observation: $C(I)$ can be constructed from $F(I)$.

Container Generation Algorithm

Input: Graph G and an independent set I
Output: Container $C(I)$ and Fingerprint $F(I)$



Key Observation: $C(I)$ can be constructed from $F(I)$.

Graph Container Lemma - Restated

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

[Kleitman, Winston '82]

Graph Container Lemma - Restated

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

[Kleitman, Winston '82]

Question: How do we construct the *collection* of containers?

Graph Container Lemma - Restated

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

[Kleitman, Winston '82]

Question: How do we construct the *collection* of containers?

Answer:

Let $\mathscr{C} = \{C(I) : I \text{ is an independent set in } G\}$

Graph Container Lemma - Restated

Lemma. For any ϵ, ρ, let $G = (V, E)$ be a graph such that every induced subgraph on ρn vertices has at least ϵn^2 edges. Then, there exists a collection $\mathscr{C} \subseteq P(V)$ of containers that satisfies:

1. for every independent set I, there exists $C \in \mathscr{C}$ with $I \subseteq C$,
2. for every $C \in \mathscr{C}$, $|C| \lesssim (1 - \frac{\epsilon}{\rho^2})\rho n$, and
3. $|\mathscr{C}| \lesssim \binom{n}{1/\epsilon}$.

[Kleitman, Winston '82]

Question: How do we construct the *collection* of containers?

Answer:

$$\begin{aligned}\text{Let } \mathscr{C} &= \{C(I) : I \text{ is an independent set in } G\} \\ &= \{C(F) : F = F(I) \text{ for some independent set } I \text{ in } G\}\end{aligned}$$

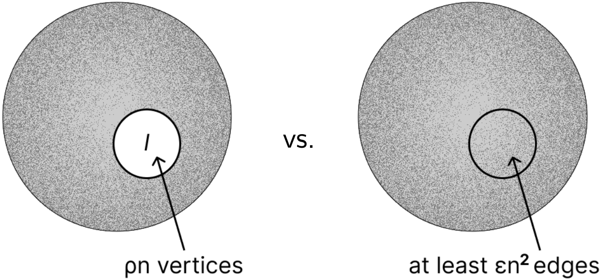
How to use the Container Method to Analyze Property Testers

Graph Property Testing - Example

Testing Independent Sets: Distinguish between the cases:

(i) G has a ρn independent set, and

(ii) every ρn-induced subgraph in G has at least ϵn^2 edges (ϵ-far).

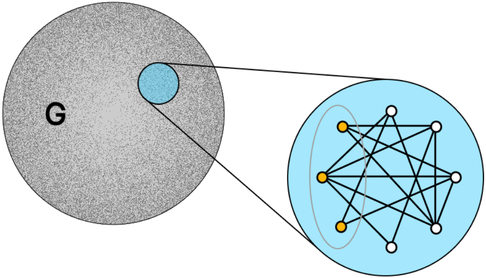


by inspecting a small fraction of G.

[Goldreich, Goldwasser, Ron '98]

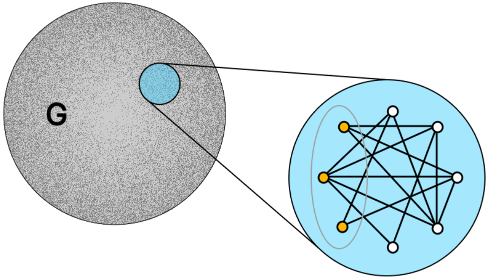
A Simple Tester for Independent Set Property

Testing Algorithm: Take a random sample S of s vertices, check if the induced subgraph $G[S]$ has a ρs independent set (for some small s, which we call the sample complexity).



A Simple Tester for Independent Set Property

Testing Algorithm: Take a random sample S of s vertices, check if the induced subgraph $G[S]$ has a ρs independent set (for some small s, which we call the sample complexity).



Key Challenge: If G is ϵ-far from having a ρn independent set, show that a random induced subgraph on s vertices has a ρs independent set with only small probability.

Application of Container Lemma

Key Challenge: If G is ϵ-far from having a ρn independent set, show that a random induced subgraph on s vertices has a ρs independent set with only small probability.

Application of Container Lemma

Key Challenge: If G is ϵ-far from having a ρn independent set, show that a random induced subgraph on s vertices has a ρs independent set with only small probability.

Use Container Lemma:

- Any independent set is contained in a container.
- Since each container is small, can upper bound the probability of getting ρs vertices from a specific container.

Application of Container Lemma

Key Challenge: If G is ϵ-far from having a ρn independent set, show that a random induced subgraph on s vertices has a ρs independent set with only small probability.

Use Container Lemma:

- Any independent set is contained in a container.
- Since each container is small, can upper bound the probability of getting ρs vertices from a specific container.
- Take union bound over all containers.

New Container Lemmas and Property Testing Results

Prior Testing Results using the Container Method

Theorem:
There is an ϵ-tester for ρ-IndepSet that only inspects a random induced subgraph on $\widetilde{O}(\rho^3/\epsilon^2)$ vertices.

There is an ϵ-tester for k-Color that only inspects a random induced subgraph on $\widetilde{O}(k/\epsilon)$ vertices.

[Blais, Seth FOCS '23]

Prior Testing Results using the Container Method

Theorem:
There is an ϵ-tester for ρ-IndepSet that only inspects a random induced subgraph on $\widetilde{O}(\rho^3/\epsilon^2)$ vertices.

There is an ϵ-tester for k-Color that only inspects a random induced subgraph on $\widetilde{O}(k/\epsilon)$ vertices.

[Blais, Seth FOCS '23]

- Both results use new graph container lemmas.
- Both results improve on the prior known bounds, and the tester for ρ-IndepSet nearly matches the best lower bound by Feige, Langberg, Schechtman ('04).

Prior Testing Results using the Container Method

Theorem:
There is an ϵ-tester for ρ-IndepSet that only inspects a random induced subgraph on $\widetilde{O}(\rho^3/\epsilon^2)$ vertices.

There is an ϵ-tester for k-Color that only inspects a random induced subgraph on $\widetilde{O}(k/\epsilon)$ vertices.

[Blais, Seth FOCS '23]

- Both results use new graph container lemmas.
- Both results improve on the prior known bounds, and the tester for ρ-IndepSet nearly matches the best lower bound by Feige, Langberg, Schechtman ('04).

Question: What else can the container method give in testing?

Result 1: Satisfiability Testing

Definition: An instance ϕ of (q, k)-$\textsc{Sat}$ is a constraint satisfaction problem with q variables per constraint and alphabet size k.

ϕ is ϵ-far from satisfiable if at least $\epsilon\binom{n}{q}$ constraints must be removed to make ϕ satisfiable.

Result 1: Satisfiability Testing

Definition: An instance ϕ of (q, k)-$\textsc{Sat}$ is a constraint satisfaction problem with q variables per constraint and alphabet size k.

ϕ is ϵ-far from satisfiable if at least $\epsilon\binom{n}{q}$ constraints must be removed to make ϕ satisfiable.

Theorem 1: There is an ϵ-tester for (q, k)-$\textsc{Sat}$ with sample complexity $\widetilde{O}(kq^3/\epsilon)$.

- Improving on prior bounds of $2^{k^{2q}}/\epsilon^2$ [Alon, Shapira '03] and k^{3q}/ϵ [Sohler '12].
- Via new **hypergraph** container lemma.

Corollaries of Result 1

Theorem 1: There is an ϵ-tester for (q, k)-$\textsc{Sat}$ with sample complexity $\widetilde{O}(kq^3/\epsilon)$.

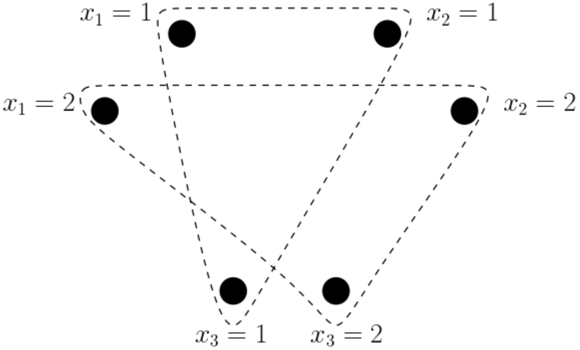
Corollaries:

- new bound for ϵ-testing k-colorability on q-uniform hypergraphs.
- new bound for ϵ-testing any $0 - 1$ graph partition property.

Hypergraph Encoding for Satisfiability

Idea: Encode (q, k)-SAT instance ϕ as a hypergraph H_ϕ. The vertex set is all possible single variable assignments. Edges correspond to falsified constraints.

Example: Two hyperedges formed by the constraint that variables x_1, x_2, x_3 must not all be equal.



New Hypergraph Container Lemma for Satisfiability

Lemma (informal). Let ϕ be an instance of (q, k)-$\textsc{Sat}$. If ϕ is ϵ-far from satisfiable then there exists a collection of containers $\mathscr{C} \subseteq P(V(H_\phi))$ that satisfies:

New Hypergraph Container Lemma for Satisfiability

Lemma (informal). Let ϕ be an instance of (q, k)-$\textsc{Sat}$. If ϕ is ϵ-far from satisfiable then there exists a collection of containers $\mathscr{C} \subseteq P(V(H_\phi))$ that satisfies:

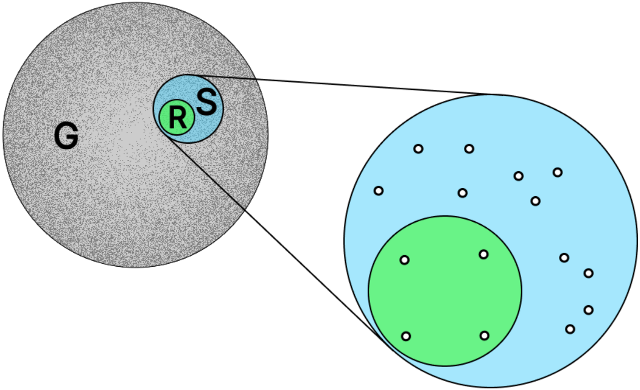
1. for every independent set I in H_ϕ, there exists $C \in \mathscr{C}$ such that $I \subseteq C$,
2. each $C \in \mathscr{C}$ is small, and
3. $|\mathscr{C}|$ is small.

Result 2: Query Complexity of Independent Set Testing

Theorem 2: There is an ϵ-tester for ρ-IndepSet with query complexity $\widetilde{O}(\rho^5/\epsilon^{7/2})$.

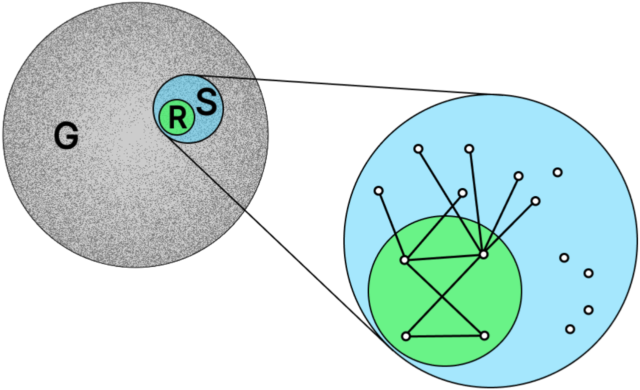
- Demonstrates gap between query complexity of optimal sample based tester of $\widetilde{O}(\rho^6/\epsilon^4)$ by Blais, Seth ('23).
- Via new graph container lemma for an object we call independent set stars.

The New Tester



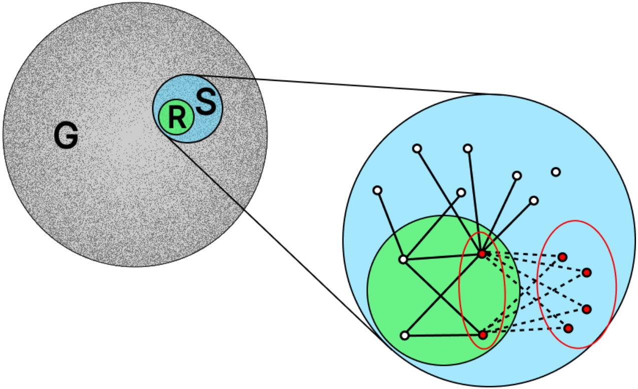
- Take vertex sample S and smaller sample $R \subset S$.

The New Tester



- Take vertex sample S and smaller sample $R \subset S$.
- Query all edges in R, and all edges between R and S.

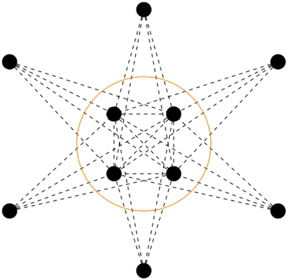
The New Tester



- Take vertex sample S and smaller sample $R \subset S$.
- Query all edges in R, and all edges between R and S.
- Accept if and only if R contains a $\rho|R|$ independent set I and S contains a $\rho|S|$ set J such that $E(I, J) = 0$.

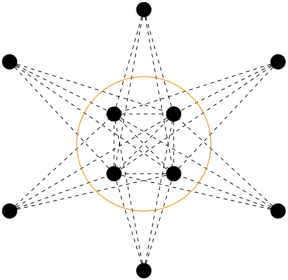
Independent Set Stars and New Container Lemma

Definition: Let $G = (V, E)$ be a graph. A pair of subsets (I, J) of V form an independent set star if I is an independent set and $|E(I, J)| = 0$.



Independent Set Stars and New Container Lemma

Definition: Let $G = (V, E)$ be a graph. A pair of subsets (I, J) of V form an independent set star if I is an independent set and $|E(I, J)| = 0$.



Lemma (informal). Let G be ϵ-far from having a ρn independent set. Then, there exists a collection of containers that cover all independent set stars.

Summary: Property Testing and the Container Method

New graph and hypergraph container lemmas can be used to show:

- ϵ-tester for ρ-IndepSet with sample complexity $\widetilde{\Theta}(\rho^3/\epsilon^2)$
- ϵ-tester for k-Color with sample complexity $\widetilde{O}(k/\epsilon)$

[Blais, Seth '23]

- ϵ-tester for (q, k)-Sat with sample complexity $\widetilde{O}(kq^3/\epsilon)$
- ϵ-tester for ρ-IndepSet with query complexity $\widetilde{O}(\rho^5/\epsilon^{7/2})$

[this work]

Summary: Property Testing and the Container Method

New graph and hypergraph container lemmas can be used to show:

- ϵ-tester for ρ-IndepSet with sample complexity $\widetilde{\Theta}(\rho^3/\epsilon^2)$
- ϵ-tester for k-Color with sample complexity $\widetilde{O}(k/\epsilon)$

[Blais, Seth '23]

- ϵ-tester for (q, k)-Sat with sample complexity $\widetilde{O}(kq^3/\epsilon)$
- ϵ-tester for ρ-IndepSet with query complexity $\widetilde{O}(\rho^5/\epsilon^{7/2})$

[this work]

Thank you!

Summary: Property Testing and the Container Method

New graph and hypergraph container lemmas can be used to show:

- ϵ-tester for ρ-$\textsc{IndepSet}$ with sample complexity $\widetilde{\Theta}(\rho^3/\epsilon^2)$
- ϵ-tester for k-$\textsc{Color}$ with sample complexity $\widetilde{O}(k/\epsilon)$

[Blais, Seth '23]

- ϵ-tester for (q, k)-$\textsc{Sat}$ with sample complexity $\widetilde{O}(kq^3/\epsilon)$
- ϵ-tester for ρ-$\textsc{IndepSet}$ with query complexity $\widetilde{O}(\rho^5/\epsilon^{7/2})$

[this work]

Thank you!