# Constant Approximating Parameterized *k*-SetCover is W[2]-hard

Bingkai Lin, Xuandi Ren, Yican Sun, Xiuhan Wang

Jan, 2023

### SetCover Problem

**Input:** (S, U), where S is a collection of n sets  $S_1$ , ...,  $S_n$  over the universe U. **Output:** find the smallest number of sets in S, whose union is U.

- Equivalent view:
  - Given a bipartite graph  $G = (S \cup U, E = \{(S_i, u) | u \in S_i\})$ , find the smallest number of left vertices, whose neighbors' union is U.
- Example:
  - $U = \{1,2,3,4,5\}, S = \{S_1 = \{1,2\}, S_2 = \{2,3,4\}, S_3 = \{1,4\}, S_4 = \{5\}\}$
  - Answer:  $3(S_1, S_2, S_4)$ .



### SetCover Problem

**Input:** (S, U), where S is a collection of n sets  $S_1$ , ...,  $S_n$  over the universe U. **Output:** find the smallest number of sets in S, whose union is U.

- Equivalent view:
  - Given a bipartite graph  $G = (S \cup U, E = \{(S_i, u) | u \in S_i\})$ , find the smallest number of left vertices, whose neighbors' union is U.
- Example:
  - $U = \{1,2,3,4,5\}, S = \{S_1 = \{1,2\}, S_2 = \{2,3,4\}, S_3 = \{1,4\}, S_4 = \{5\}\}$
  - Answer:  $3(S_1, S_2, S_4)$ .
- NP-complete [Kar72]



### Parameterized SetCover Problem

#### k-SetCover

**Input:** (S, U, k), where S is a collection of n sets  $S_1$ , ...,  $S_n$  over the universe U. **Output:** find the smallest number of sets in S, whose union is U. It's guaranteed that the optimum is k.

- W[2]-complete [DF95]
  - W[2] $\neq$ FPT  $\Rightarrow$  no algorithm in  $f(k)n^{O(1)}$  time
- W[1]-complete if  $|U| = \Theta(k^{O(1)} \log n)$ .
  - [KLM19,Lin19] + This work
- Naïve algorithm:
  - enumerating all *k*-tuple of sets in  $n^{O(k)}$  time
  - dynamic programming on subsets of *U* in  $2^{O(|U|)}$  time



- "The set cover problem plays the same role in approximation algorithms that the maximum matching problem played in exact algorithms - as a problem whose study led to the development of fundamental techniques for the entire field."
  - Approximation Algorithms, by Vijay V. Vazirani



• A simple greedy algorithm reaches  $(\ln n - \ln \ln n + \Theta(1))$  approximation ratio [Sla97].

Problem	Assumption	Hardness of Approx. Ratio		Running Time Bound	Reference & Comments	
SetCover	NP≠P	$(1-\varepsilon)\ln n$		<i>n</i> <sup>0(1)</sup>	[DS14]	
Parameterized SetCover	SETH	$(\log n)^{\frac{1}{k^{O(1)}}}$	$(1 - o(1)) \left(\frac{\log n}{\log \log n}\right)^{\frac{1}{k}}$	$f(k)n^{k-\varepsilon}$	[KLM19]	[Lin19]
	ETH			$f(k)n^{o(k)}$		
	<i>k-</i> SUM Hypothesis			$f(k)n^{\left\lfloor \frac{k}{2} \right floor - \varepsilon}$		
	W[1]≠FPT	$(\log n)^{\varepsilon(k)}$		$f(k)n^{O(1)}$		
	W[2]≠FPT	?		?		

• A simple greedy algorithm reaches  $(\ln n - \ln \ln n + \Theta(1))$  approximation ratio [Sla97].

Problem	Assumption	Hardness of Approx. Ratio		Running Time Bound	Reference & Comments	
SetCover	NP≠P	$(1-\varepsilon)\ln n$		$n^{0(1)}$	[DS14]	
Parameterized SetCover	SETH	$(\log n)^{\frac{1}{k^{O(1)}}}$	$(1 - o(1)) \left(\frac{\log n}{\log \log n}\right)^{\frac{1}{k}}$	$f(k)n^{k-\varepsilon}$	[KLM19]	[Lin19]
	ETH			$f(k)n^{o(k)}$		
	<i>k-</i> SUM Hypothesis			$f(k)n^{\left\lfloor\frac{k}{2}\right floor-\varepsilon}$		
	W[1]≠FPT	$(\log n)^{\varepsilon(k)}$		$f(k)n^{O(1)}$		
	W[2]≠FPT	Any constant		$f(k)n^{O(1)}$	This work	

### W[2]-hardness of Approx. *k*-SetCover



• A simple greedy algorithm reaches  $(\ln n - \ln \ln n + \Theta(1))$  approximation ratio [Sla97].

Problem	Assumption	Hardness of Approx. Ratio	Running Time Bound	Reference	
SetCover	NP≠P	$(1-\varepsilon)\ln n$	$n^{O(1)}$	[DS14]	
Parameterized SetCover	SETH		$f(k)n^{k-\varepsilon}$		
	ETH	$\left(\frac{\log n}{\log \log n}\right)^{1/k}$	$f(k)n^{o(k)}$		
	<i>k-</i> SUM Hypothesis		$f(k)n^{\left\lfloor\frac{k}{2}\right floor-arepsilon}$	[KLM19, Lin19]	
	W[1]≠FPT	$(\log n)^{\varepsilon(k)}$	$f(k)n^{O(1)}$		
	W[2]≠FPT	?	?		

# Approx. SetCover with Small OPT Size

Problem	Assumption	Hardness of Approx. Ratio	OPT Size	Technique	Reference
SetCover	NP≠P	$(1-\varepsilon)\ln n$	$\Omega(n)$	(Complex) PCP	[DS14]
	W[1]≠FPT	$o\left(\frac{\log n}{\log\log n}\right)$	O(poly log n)	New TGC	This work
	W[1]≠FPT	$(\log n)^{\varepsilon(k)}$	k	TGC / Distributed PCP	[KLM19, Lin19]

# Our Technique

- Threshold Graph Composition [Lin18, CL19, Lin19, BBE+21]
  - exploited new properties of threshold graphs
  - used the construction from error-correcting codes [KN21]
  - discovered new composition scheme

Constant Approximating *k*-SetCover is W[2]-hard!



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

#### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

#### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

#### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

#### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

#### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

**Completeness:** 

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i$ such that  $b_i$  has k + 1 neighbors in X, then |X| > h



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:

•  $\forall X \subseteq A \text{ if } \text{for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i$ such that  $b_i$  has k + 1 neighbors in X, then |X| > h



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:

•  $\forall X \subseteq A \text{ if } \text{for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i$ such that  $b_i$  has k + 1 neighbors in X, then |X| > h



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$ have a common neighbor in  $B_i$ 

#### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i$ such that  $b_i$  has k + 1 neighbors in X, then |X| > h



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying



• Threshold Graph  $G_T$ : a bipartite graph  $(A \cup B, E)$  with  $A = A_1 \cup \cdots \cup A_k$  and  $B = B_1 \cup \cdots \cup B_m$ , satisfying



### Threshold Graph Composition in [Lin19]

*k*-SetCover Instance  $\Gamma = (S, U)$ 

#### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A$ : If  $\forall i \in [m], \exists b_i \in B_i$  such that  $b_i$  has k + 1 neighbors in X, then |X| > h

*c*-gap *k*-SetCover Instance  $\Gamma' = (S', U')$ , where  $|U'| = |U|^{|B_i|}$ 

### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S'$  which can cover U'

### **NO Instance:**

• any covering of *U* has size  $\geq h$ 

### Threshold Graph Composition

*k*-SetCover Instance  $\Gamma = (\mathcal{S}, U)$ 

#### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i \text{ such that } b_i \text{ has } k + 1 \text{ neighbors in } X, \text{ then } |X| > ck'$ 

*c*-gap *k'*-SetCover Instance  $\Gamma' = (S', U')$ , where  $|U'| = (|U||B_i|)^{O(c)}$ 

### YES Instance:

•  $\exists S_1, \dots, S_k, \in S'$  which can cover U'

#### **NO Instance:**

any covering of *U* has size > *ck*'

### Threshold Graph Composition

*k*-SetCover Instance  $\Gamma = (S, U)$ 

### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i \text{ such that } b_i \text{ has } k + 1 \text{ neighbors in } X, \text{ then } |X| > ck'$ 

### **Our Reduction**

- Treat every  $A_i$  as a copy of S.
- The new sets  $S' = A \cup B$ .
- The new universe U' has m parts  $U'_1, ..., U'_m$ .

Desired Property

For any  $X \subseteq A$  and  $Y \subseteq B$  can cover  $U'_i$  iff (1) either  $\exists b_i \in Y \cap B_i$ , s.t.  $X \cap \mathcal{N}_{G_T}(b_i)$  cover U, (2) or  $|Y \cap B_i| \ge c + 1$ .

### Analysis of the YES Case

**YES Instance:** 

•  $\exists S_1, \dots, S_k, \in S'$  which can cover U'

*k*-SetCover Instance  $\Gamma = (S, U)$ 

**YES Instance:** 

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i \text{ such that } b_i \text{ has } k + 1 \text{ neighbors in } X, \text{ then } |X| > ck'$ 

### **Our Reduction**

- Treat every  $A_i$  as a copy of S.
- The new sets  $S' = A \cup B$ .
- The new universe U' has m parts  $U'_1, ..., U'_m$ .

Desired PropertyFor any  $X \subseteq A$  and  $Y \subseteq B$  can cover  $U'_i$  iff(1) either  $\exists b_i \in Y \cap B_i$ , s.t.  $X \cap \mathcal{N}_{G_T}(b_i)$  cover U,(2) or  $|Y \cap B_i| \ge c + 1$ .

In the YES case, it suffices to pick  $S_1, ..., S_k$  and their common neighbors in every  $B_i$ .

**In the NO case**, one of the following holds:

- $(\# \text{ of } 1 \ge \varepsilon m) \Rightarrow |X| > ck'.$
- $(\# \text{ of } (2) \ge (1-\varepsilon)m) \Rightarrow |Y| \ge (c+1)(1-\varepsilon)m$ ,

### Analysis of the NO case

**NO Instance:** 

• any covering of *U* has size > *ck*'

*k*-SetCover Instance  $\Gamma = (S, U)$ 

### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i \text{ such that } b_i \text{ has } k + 1 \text{ neighbors in } X, \text{ then } |X| > ck'$ 

### **Our Reduction**

- Treat every  $A_i$  as a copy of S.
- The new sets  $S' = A \cup B$ .
- The new universe U' has m parts  $U'_1, \ldots, U'_m$ .

**Desired Property** For any  $X \subseteq A$  and  $Y \subseteq B$  can cover  $U'_i$  iff (1) either  $\exists b_i \in Y \cap B_i$ , s.t.  $|X \cap \mathcal{N}_{G_T}(b_i)| \ge k + 1$ , (2) or  $|Y \cap B_i| \ge c + 1$ .

In the YES case, it suffices to pick  $S_1, ..., S_k$  and their common neighbors in every  $B_i$ .

### **In the NO case**, one of the following holds:

- $(\# \text{ of } \underline{1} \geq \varepsilon m) \Rightarrow |X| > ck'.$
- $(\# \text{ of } 2 \ge (1 \varepsilon)m) \Rightarrow |Y| \ge (c + 1)(1 \varepsilon)m$ ,

### Analysis of the NO case

**NO Instance:** 

• any covering of *U* has size > *ck*'

*k*-SetCover Instance  $\Gamma = (S, U)$ 

### **YES Instance:**

•  $\exists S_1, \dots, S_k \in S$  which can cover U

### **NO Instance:**

• any covering of *U* has size  $\geq k + 1$ 

Threshold Graph  $G_T = (A \cup B, E)$ 

### **Completeness:**

•  $\forall a_1 \in A_1, ..., a_k \in A_k$  and  $i \in [m], a_1, ..., a_k$  have a common neighbor in  $B_i$ 

### Soundness:

•  $\forall X \subseteq A \text{ if for } \varepsilon \text{ fraction of } i \in [m], \exists b_i \in B_i \text{ such that } b_i \text{ has } k + 1 \text{ neighbors in } X, \text{ then } |X| > ck'$ 

### **Our Reduction**

- Treat every  $A_i$  as a copy of S.
- The new sets  $S' = A \cup B$ .
- The new universe U' has m parts  $U'_1, \ldots, U'_m$ .

**Desired Property** A and  $V \subseteq B$  can cover  $U'_{i}$  iff

For any  $X \subseteq A$  and  $Y \subseteq B$  can cover  $U'_i$  iff (1) either  $\exists b_i \in Y \cap B_i$ , s.t.  $|X \cap \mathcal{N}_{G_T}(b_i)| \ge k + 1$ , (2) or  $|Y \cap B_i| \ge c + 1$ .

**In the YES case**, it suffices to pick  $S_1, ..., S_k$  and their common neighbors in every  $B_i$ .

### **In the NO case**, one of the following holds:

- $(\# \text{ of } (1) \ge \varepsilon m) \Rightarrow |X| > ck'.$
- $(\# \text{ of } (2) \ge (1-\varepsilon)m) \Rightarrow |Y| \ge (c+1)(1-\varepsilon)m,$

### Summary

**Theorem 1.** Assuming W[2] $\neq$ FPT, there is no FPT algorithm which can approximate *k*-SetCover within any constant ratio.

**Theorem 2.** Assuming W[1]≠FPT, there is no polynomial time algorithm which can approximate *k*-SetCover within  $o\left(\frac{\log n}{\log \log n}\right)$  ratio, even if *k* is as small as  $O\left(\frac{\log n}{\log \log n}\right)^3$ .

### **Open Questions**

**Open Question 1.** *Is it* W[2]*-hard to approximate* k*-SetCover within*  $\omega(1)$  *ratio?* 

• Our reduction has running time  $\Omega(|U|^c)$ , thus can not get super-constant inapproximability.

**Open Question 2.** *Is there any FPT algorithm which can approximate k-SetCover with approximation ratio*  $o(\log n)$ ?

• Current best lower bound is  $(\log n)^{\varepsilon(k)}$  for any function  $\varepsilon(k) = o(1)$ .



• Thanks!