Matroid-Constrained Maximum Vertex Cover: Approximate Kernels and Streaming Algorithms

Chien-Chung Huang[1] François Sellier*[2,3]

[1]CNRS, DI ENS, École normale supérieure, Université PSL, France

[2]Université Paris Cité, CNRS, IRIF, F-75013, Paris, France

[3]MINES ParisTech, Université PSL, F-75006, Paris, France

June 2022

Maximum k-vertex cover

Let $G = (V, E)$ be a graph.

- a non-negative *weight* $w(e)$ is associated to each edge $e \in E$
- an edge $e = (u, v)$ is called *covered* by a set $S \subseteq E$ if at least one of its endpoints is in S, *i.e* $u \in S$ or $v \in S$
- we will denote $E_G(S)$ the sum of the weights of the edges covered the set $S \subseteq V$, *i.e.*

$$E_G(S) = \sum_{e=(u,v)\in E, \{u,v\}\cap S \neq \emptyset} w(e)$$

- in the *maximum k-vertex cover* problem, we want a set containing at most k elements maximizing E_G, *i.e.*

$$\underset{S\subseteq V, |S|\leq k}{\text{argmax}}\ E_G(S)$$

Maximum k-vertex cover

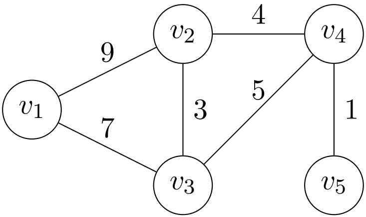

Figure: A maximum k-vertex cover problem with $k = 2$

Maximum k-vertex cover

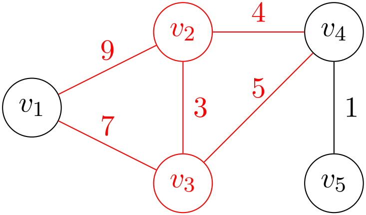

Figure: A maximum k-vertex cover problem with $k = 2$

Matroid-Constrained Maximum Vertex Cover

$\mathcal{M} = (V, \mathcal{I})$ is a *matroid* on the ground set V if these conditions hold for $\mathcal{I} \subseteq \mathcal{P}(V)$:

1. $\emptyset \in \mathcal{I}$,
2. if $X \subseteq Y \in \mathcal{I}$, then $X \in \mathcal{I}$,
3. if $X, Y \in \mathcal{I}, |Y| > |X|$, there exists an element $e \in Y \backslash X$ so that $X \cup \{e\} \in \mathcal{I}$,

the sets in $\mathcal{I}$ are the *independent sets* and the *rank* $r_{\mathcal{M}}$ of the matroid $\mathcal{M}$ is defined as $\max_{X \in \mathcal{I}} |X|$.

Matroid-Constrained Maximum Vertex Cover

Let $G = (V, \mathcal{I})$ be a graph and $\mathcal{M} = (V, \mathcal{I})$ a matroid:

- in the *matroid-constrained maximum vertex cover* problem, we want a set independent in $\mathcal{M}$ maximizing E_G, *i.e.*

$$\underset{S \subseteq V, S \in \mathcal{I}}{\operatorname{argmax}} E_G(S)$$

Matroid-Constrained Maximum Vertex Cover

Some examples of matroids:

- *uniform* matroid of rank k: $S \in \mathcal{I}$ iff $|S| \leq k$
- *partition* matroid defined by a partition $V_1, \ldots, V_r$ and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$
- *laminar* matroid defined by a laminar family $V_1, \ldots, V_r$ and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$ (for $V_i \cap V_j \neq \emptyset$ then either $V_i \subseteq V_j$ or $V_j \subseteq V_i$)
- *transversal* matroid defined by a family $V_1, \ldots, V_k$: $S \in \mathcal{I}$ iff there exists an injective function $\phi : S \rightarrow \{1, \ldots, k\}$ such that for all $v \in S$, $v \in V_{\phi(v)}$

Matroid-Constrained Maximum Vertex Cover

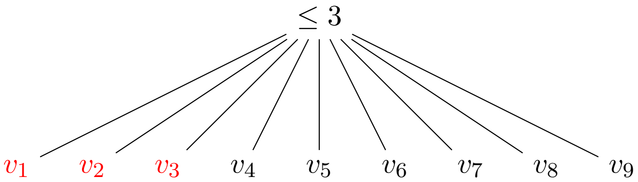

Figure: Representation of a uniform matroid

- *uniform* matroid of rank k: $S \in \mathcal{I}$ iff $|S| \leq k$

Matroid-Constrained Maximum Vertex Cover

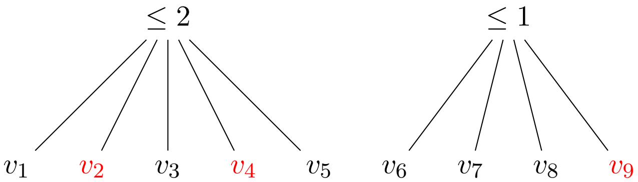

Figure: Representation of a partition matroid

- *partition* matroid defined by a partition $V_1, \ldots, V_r$ and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$

Matroid-Constrained Maximum Vertex Cover

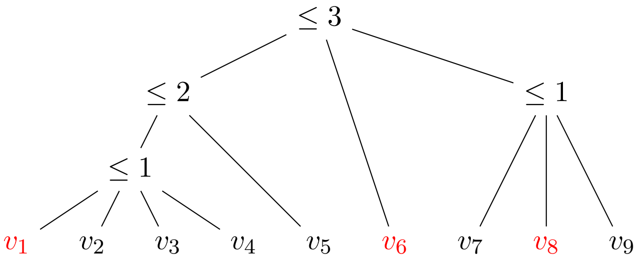

Figure: Representation of a laminar matroid

- *laminar* matroid defined by a laminar family $V_1, \ldots, V_r$ and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$ (for $V_i \cap V_j \neq \emptyset$ then either $V_i \subseteq V_j$ or $V_j \subseteq V_i$)

Matroid-Constrained Maximum Vertex Cover

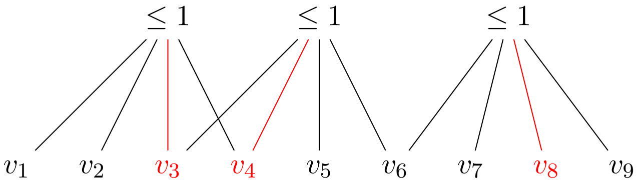

Figure: Representation of a transversal matroid

- *transversal* matroid defined by a family $V_1, \ldots, V_k$: $S \in \mathcal{I}$ iff there exists an injective function $\phi : S \to \{1, \ldots, k\}$ such that for all $v \in S$, $v \in V_{\phi(v)}$

Interest of the Problem

The problem of maximizing a submodular function under matroid constraint has been studied extensively

- the problem is easier when:
  - either the submodular function is linear
  - or the matroid is a simple uniform matroid (*i.e.*, cardinality constraint)
- in our work, we study cases that are slightly more complicated:
  - the function is a cover function of bounded frequency (but still we have decreasing marginal returns)
  - and the matroid is a partition, laminar, or transversal matroid (but still it is not a simple cardinality constraint)

Previous Results

For the maximum k-vertex cover problem:

- greedy provides a $1 - 1/e$ approximation [Hochbaum and Pathria, 1998]
- a LP-based approach and a technique of pipage rounding give a ratio of $3/4$ [Ageev and Sviridenko, 2000]
- the current best ratio is 0.92, attained using a kernelization method [Manurangsi, 2018]
- it is not possible to have a Polynomial Time Approximation Scheme [Guo, Niedermeier, and Wernicke, 2005]

Previous Results

Here we recall the definition of an FPT-AS [Marx, 2008]:

Definition

Given a parameter function κ associating a natural number to each instance $x \in I$ of a given problem, a *Fixed-Parameter Tractable Approximation Scheme* (FPT-AS) is an algorithm that can provide a $(1 - \varepsilon)$ approximation in $f(\varepsilon, \kappa(x)) \cdot |x|^{O(1)}$ time.

Previous Results

Here we recall the definition of an FPT-AS [Marx, 2008]:

Definition

Given a parameter function κ associating a natural number to each instance $x \in I$ of a given problem, a *Fixed-Parameter Tractable Approximation Scheme* (FPT-AS) is an algorithm that can provide a $(1-\varepsilon)$ approximation in $f(\varepsilon, \kappa(x)) \cdot |x|^{O(1)}$ time.

- in our case κ is the rank k of the matroid
- FTP-AS have been developed for maximum k-vertex cover in [Marx, 2008; Manurangsi, 2018]
- Munurangsi builds a kernel of size k/ε in [Manurangsi, 2018]

Previous Results

For the matroid-constrained maximum vertex cover problem:

- local search provides a $1 - 1/e$ approximation [Filmus and Ward, 2012]
- a LP-based approach and a technique of pipage rounding give a ratio of 3/4 [Cunningham, 1984; Ageev and Sviridenko, 2000; Calinescu, Chekuri, Pál, and Vondrák, 2011]

Our Contribution

Theorem

For every $\varepsilon > 0$, we can extract an approximate kernel $V' \subseteq V$ in polynomial time so that a $(1-\varepsilon)$-approximate solution is contained in V', such that: $\mathcal{M}$.

1. $|V'| \leq \frac{k}{\varepsilon}$ *when $\mathcal{M}$ is a partition matroid;*
2. $|V'| \leq \frac{2k}{\varepsilon}$ *when $\mathcal{M}$ is a laminar matroid;*
3. $|V'| \leq \frac{k}{\varepsilon} + k(k-1)$ *when $\mathcal{M}$ is a transversal matroid.*

Using a brute force enumeration, we can find the desired $1-\varepsilon$ approximation in $\left(\frac{1}{\varepsilon}\right)^{O(k)} n^{O(1)}$ time for partition and laminar matroids and $\left(\frac{1}{\varepsilon}+k\right)^{O(k)} n^{O(1)}$ time for transversal matroids.

Our Contribution

Corollary

Suppose that we are given a hypergraph $G = (V, E)$ *with edge size bounded by a constant* $\eta \geq 2$*. We can compute a* $(1 - (\eta - 1) \cdot \varepsilon)$ *approximation using* $\left(\frac{1}{\varepsilon}\right)^{O(k)} n^{O(1)}$ *time for partition and laminar matroids and* $\left(\frac{1}{\varepsilon} + k\right)^{O(k)} n^{O(1)}$ *time for transversal matroids.*

- if η is unbounded, one cannot obtain an approximation ratio better than $1 - 1/e + \varepsilon$, assuming GAP-ETH, in FPT time (the parameter being the rank k), even for the simplest uniform matroid [Manurangsi, 2020]

Kernelization Framework

The idea is to:

- build an *approximate kernel*, *i.e.*, a smaller graph containing a $(1-\varepsilon)$-approximation of the optimal solution
- find the optimal solution in that smaller graph using bruteforce

Previous Kernelization Technique

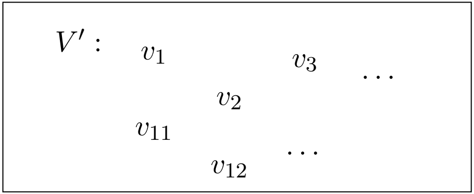

Figure: Kernelization technique developed in [Manurangsi, 2018] for the maximum k-vertex cover problem

- V' contains the k/ε vertices having the largest weighted degrees

Previous Kernelization Technique

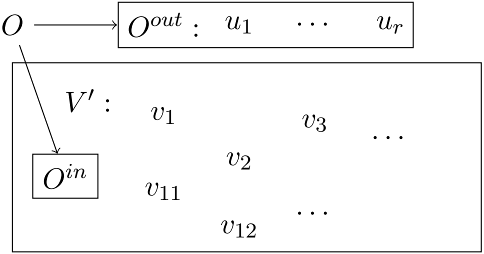

Figure: Kernelization technique developed in [Manurangsi, 2018] for the maximum k-vertex cover problem

- V' contains the k/ε vertices having the largest weighted degrees
- the elements in O^{out} are replaced by random elements drawn from V'

Previous Kernelization Technique

The idea in random sampling is that:

- a vertex in O is replaced by a vertex in V' having a larger weighted degree
- however, even if the overall sum of the weighted degrees of the selected vertices is always it least equal to that of the vertices in O, some edges may have been double-counted
- using random sampling, in expectation the proportion of the double-counted edges is $1/\varepsilon$

Kernelization Framework

Definition

Let $\mathcal{M} = (V, \mathcal{I})$ be a matroid with weights $\omega : V \to \mathbb{R}^+$. We say $V' \subseteq V$ is t-robust if given any base $O \in \mathcal{I}$, there is a bijection from the elements $u_1, \cdots, u_t \in O \backslash V'$ to subsets $U_{u_1}, \cdots, U_{u_r} \subseteq V' \backslash O$ so that

1. the U_{u_i}s are mutually disjoint and $|U_{u_i}| = t$,
2. all elements in U_{u_i} have weights no less than u_i,
3. by taking an arbitrary element $u'_i \in U_{u_i}$ for all i, $(V' \cap O) \cup \{u'_i\}_{i=1}^{r}$ is a base in $\mathcal{M}$.

Kernelization Framework

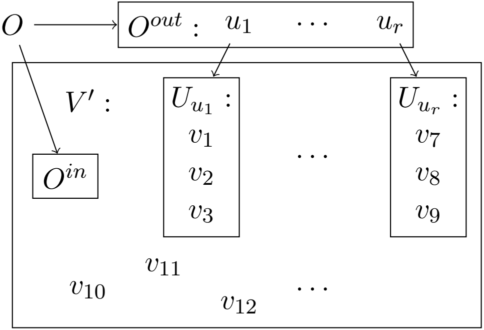

Figure: Representation a robust decomposition, for $t = 3$

- $\forall\, 1 \leq i \leq r, \forall\, v \in U_{u_i}, \omega(u_i) \leq \omega(v)$
- by taking an arbitrary element $u_i' \in U_{u_i}$ for all i, $(V' \cap O) \cup \{u_i'\}_{i=1}^{r}$ is independent in $\mathcal{M}$

Kernelization Framework

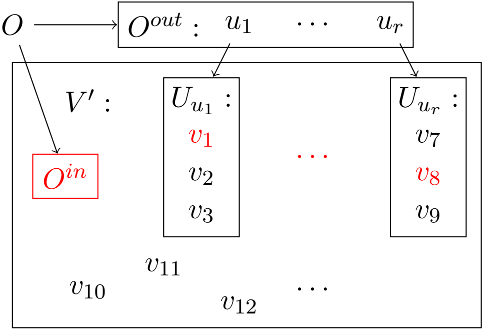

Figure: Representation a robust decomposition, for $t = 3$

- $\forall\, 1 \leq i \leq r, \forall\, v \in U_{u_i}, \omega(u_i) \leq \omega(v)$
- by taking an arbitrary element $u_i' \in U_{u_i}$ for all i, $(V' \cap O) \cup \{u_i'\}_{i=1}^{r}$ is independent in $\mathcal{M}$

Kernelization Framework

Definition

Suppose that $\mathcal{M} = (V, \mathcal{I})$ is a matroid. Then we can define $\tau\mathcal{M} = (V, \mathcal{I}_\tau)$ as the union of τ matroids $\mathcal{M}$, as follows: $S \in \mathcal{I}_\tau$ if S can be partitioned into $S_1 \cup \cdots \cup S_\tau$ so that each $S_i \in \mathcal{I}$.

Kernelization Framework

Let $\mathcal{M} = (V, \mathcal{I})$ be a matroid with weights $\omega : V \to \mathbb{R}^+$ and rank k. Consider the following greedy procedure on $\tau\mathcal{M} = (V, \mathcal{I}_\tau)$ to construct V':

- initially $V' = \emptyset$
- process the elements in V by non-increasing weights ω
- for each element u, if $V' \cup \{u\} \in \mathcal{I}_\tau$ (independent in matroid union), add u into V', otherwise, ignore it.

Kernelization Framework

Let $\mathcal{M} = (V, \mathcal{I})$ be a matroid with weights $\omega : V \to \mathbb{R}^+$ and rank k. Consider the following greedy procedure on $\tau\mathcal{M} = (V, \mathcal{I}_\tau)$ to construct V':

- initially $V' = \emptyset$
- process the elements in V by non-increasing weights ω
- for each element u, if $V' \cup \{u\} \in \mathcal{I}_\tau$ (independent in matroid union), add u into V', otherwise, ignore it.

Theorem

The final V' is t-robust

1. *if $\mathcal{M}$ is a partition matroid and $\tau \geq t$,*
2. *if $\mathcal{M}$ is a laminar matroid and $\tau \geq 2t$,*
3. *if $\mathcal{M}$ is a transversal matroid and $\tau \geq t + k - 1$.*

Kernelization Framework

Assuming the previous theorem, we proceed as follows:

- let the weight $\omega : V \to \mathbb{R}^+$ be the weighted degrees in the graph $G = (V, E)$, that is, $\omega(u) = \deg_w(u)$
- apply the previous theorem by setting $t = \frac{1}{\varepsilon}$
- then V' is $\frac{1}{\varepsilon}$-robust

Lemma

V' contains a set S such that $S \in \mathcal{I}$ and $E_G(S) \geq (1 - \varepsilon) E_G(O)$ where O denotes an optimal solution of the problem.

Kernelization Framework

Let $O \in \mathcal{I}$ be an optimal solution:

- we denote $O^{in} = O \cap V'$, $O^{out} = O \backslash O^{in}$
- by $\frac{1}{\varepsilon}$-robustness, we have mutually disjoint sets $U_v \subseteq V' \backslash O$ for each $v \in O^{out}$, each of size $\frac{1}{\varepsilon}$
- set $U = \cup_{v \in O^{out}} U_v$

Kernelization Framework

Let $O \in \mathcal{I}$ be an optimal solution:

- we denote $O^{in} = O \cap V'$, $O^{out} = O \backslash O^{in}$
- by $\frac{1}{\varepsilon}$-robustness, we have mutually disjoint sets $U_v \subseteq V' \backslash O$ for each $v \in O^{out}$, each of size $\frac{1}{\varepsilon}$
- set $U = \cup_{v \in O^{out}} U_v$

We construct a set $S \subseteq V'$ as follows:

- S is initialized as O^{in}
- from each set U_v, for all $v \in O^{out}$, pick an element at random and add it into S
- hence, by definition of $\frac{1}{\varepsilon}$-robustness, $S \in \mathcal{I}$.

Kernelization Framework

Let $O \in \mathcal{I}$ be an optimal solution:

- we denote $O^{in} = O \cap V'$, $O^{out} = O \backslash O^{in}$
- by $\frac{1}{\varepsilon}$-robustness, we have mutually disjoint sets $U_v \subseteq V' \backslash O$ for each $v \in O^{out}$, each of size $\frac{1}{\varepsilon}$
- set $U = \cup_{v \in O^{out}} U_v$

We construct a set $S \subseteq V'$ as follows:

- S is initialized as O^{in}
- from each set U_v, for all $v \in O^{out}$, pick an element at random and add it into S
- hence, by definition of $\frac{1}{\varepsilon}$-robustness, $S \in \mathcal{I}$.

Then we can show that:

$$\mathbb{E}[E_G(S)] \geq (1 - \varepsilon) \cdot \mathbb{E}[E_G(O)].$$

Proof of Robustness

For partition matroids, $\tau = t$, and the proof is easy:

- recall that a *partition* matroid is defined by a partition V_1, …, V_r and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$
- if there are elements in $(O \cap V_i) \backslash V'$, then it means that $|V_i \cap V'| = \tau \cdot k_i$ (some elements were ignored)

Proof of Robustness

For laminar matroids, the situation is a bit more complex:

- recall that a *laminar* matroid is defined by a laminar family $V_1, \ldots, V_r$ and bounds $k_1, \ldots, k_r$: $S \in \mathcal{I}$ iff $\forall\, 1 \leq i \leq r$, $|S \cap V_i| \leq k_i$ (for $V_i \cap V_j \neq \emptyset$ then either $V_i \subseteq V_j$ or $V_j \subseteq V_i$)

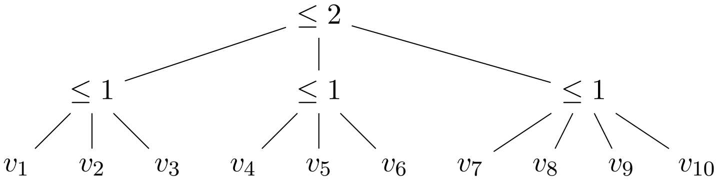

Figure: Example of set V' extracted from a laminar matroid of rank k (extracted with parameter $\tau = 5$): V' is not 5-robust.

Proof of Robustness

Key ideas:

- set $\tau = 2 \cdot t$
- the laminar tree can be binarized
- an element $u \in O \backslash V'$ can be associated to a *blocking node*: the node in the laminar tree that would have violated its cardinality constraint if u was added to V' at that time

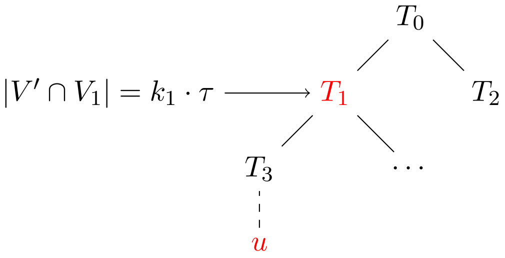

Figure: A blocking node

Proof of Robustness

Key ideas:

- allocate the sets from bottom blocking nodes to top ones

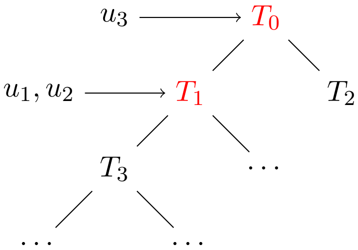

Figure: Building the associated sets (example, might not reflect a real situation)

Proof of Robustness

Key ideas:

- allocate the sets from bottom blocking nodes to top ones

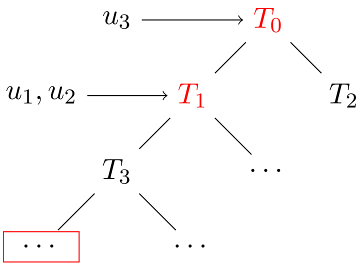

Figure: Building the associated sets (example, might not reflect a real situation)

Proof of Robustness

Key ideas:

- allocate the sets from bottom blocking nodes to top ones

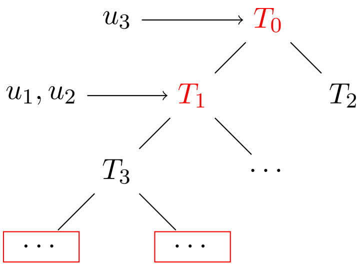

Figure: Building the associated sets (example, might not reflect a real situation)

Proof of Robustness

Key ideas:

- allocate the sets from bottom blocking nodes to top ones

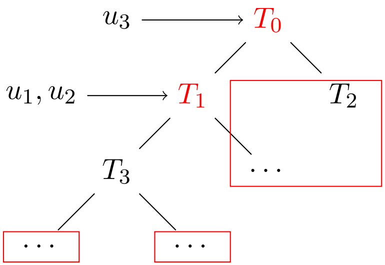

Figure: Building the associated sets (example, might not reflect a real situation)

Conclusion

- generalization of cardinality constraint to a larger set of constraints
- possible to generalize this framework to other kinds of matroids?