# 2a. Kernelization

## COMP6741: Parameterized and Exact Computation

Serge Gaspers

Semester 2, 2018

## Contents



# 1 Reminder

### Kernelization

**Definition 1.** A *kernelization* (*kernel*) for a parameterized problem $\Pi$ is a **polynomial time** algorithm, which, for any instance $I$ of $\Pi$ with parameter $k$, produces an **equivalent** instance $I'$ of $\Pi$ with parameter $k'$ such that $|I'| \leq f(k)$ and $k' \leq f(k)$ for a computable function $f$. We refer to the function $f$ as the *size* of the kernel.

### Fixed-parameter tractability

**Definition 2.** A parameterized problem $\Pi$ is *fixed-parameter tractable* (FPT) if there is an algorithm solving $\Pi$ in time $f(k) \cdot \mathsf{poly}(n)$, where $n$ is the instance size, $k$ is the parameter, $\mathsf{poly}$ is a polynomial function, and $f$ is a computable function.

**Theorem 3.** *Let* $\Pi$ *be a decidable parameterized problem.* $\Pi$ *has a kernelization* $\Leftrightarrow$ $\Pi$ *is* FPT.

# 2 Kernel for Hamiltonian Cycle

A *Hamiltonian cycle* of $G$ is a subgraph of $G$ that is a cycle on $|V(G)|$ vertices.

| vc-Hamiltonian Cycle | |
|---|---|
| Input: | A graph $G = (V, E)$. |
| Parameter: | $k = vc(G)$, the size of a smallest vertex cover of $G$. |
| Question: | Does $G$ have a Hamiltonian cycle? |

**Thought experiment**: Imagine a very large instance where the parameter is tiny. How can you simplify such an instance?
**Issue**: We do not actually know a vertex cover of size $k$.

- Obtain a vertex cover of size $\leq 2k$ by applying Vertex Cover-kernelizations to $(G, 0), (G, 1), \ldots$ until the first instance where no trivial No-instance is returned.
- If $C$ is a vertex cover of size $\leq 2k$, then $I = V \setminus C$ is an independent set of size $\geq |V| - 2k$.
- No two consecutive vertices in the Hamiltonian Cycle can be in $I$.

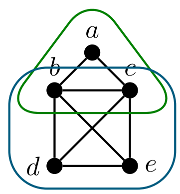
- A kernel with $\leq 4k$ vertices can now be obtained with the following simplification rule.

**(Too-large)**
Compute a vertex cover $C$ of size $\leq 2k$ in polynomial time. If $2|C| < |V|$, then return No

## 3 Kernel for Edge Clique Cover

**Definition 4.** An *edge clique cover* of a graph $G = (V, E)$ is a set of cliques in $G$ covering all its edges. In other words, if $\mathcal{C} \subseteq 2^V$ is an edge clique cover then each $S \in \mathcal{C}$ is a clique in $G$ and for each $\{u, v\} \in E$ there exists an $S \in \mathcal{C}$ such that $u, v \in S$.

Example: $\{\{a, b, c\}, \{b, c, d, e\}\}$ is an edge clique cover for this graph.




| Edge Clique Cover | |
| --- | --- |
| Input: | A graph $G = (V, E)$ and an integer $k$ |
| Parameter: | $k$ |
| Question: | Does $G$ have an edge clique cover of size at most $k$? |

The *size* of an edge clique cover $\mathcal{C}$ is the number of cliques contained in $\mathcal{C}$ and is denoted $|\mathcal{C}|$.

**Helpful properties**

**Definition 5.** A clique $S$ in a graph $G$ is a *maximal* clique if there is no other clique $S'$ in $G$ with $S \subset S'$.

**Lemma 6.** *A graph $G$ has an edge clique cover $\mathcal{C}$ of size at most $k$ if and only if $G$ has an edge clique cover $\mathcal{C}'$ of size at most $k$ such that each $S \in \mathcal{C}'$ is a maximal clique.*

*Proof sketch.* $(\Rightarrow)$: Replace each clique $S \in \mathcal{C}$ by a maximal clique $S'$ with $S \subseteq S'$.
$(\Leftarrow)$: Trivial, since $\mathcal{C}'$ is an edge clique cover of size at most $k$. □

**Simplification rules for Edge Clique Cover**
**Thought experiment**: Imagine a very large instance where the parameter is tiny. How can you simplify such an instance?
The instance could have many degree-0 vertices.

**(Isolated)**
If there exists a vertex $v \in V$ with $d_G(v) = 0$, then set $G \leftarrow G - v$.

**Lemma 7.** *(Isolated) is sound.*

*Proof sketch.* Since no edge is incident to $v$, a smallest edge clique cover for $G - v$ is a smallest edge clique cover for $G$, and vice-versa. □

**(Isolated-Edge)**
If $\exists uv \in E$ such that $d_G(u) = d_G(v) = 1$, then set $G \leftarrow G - \{u, v\}$ and $k \leftarrow k - 1$.

**(Twins)**
If $\exists u, v \in V$, $u \neq v$, such that $N_G[u] = N_G[v]$, then set $G \leftarrow G - v$.

**Lemma 8.** *(Twins) is sound.*

*Proof.* We need to show that $G$ has an edge clique cover of size at most $k$ if and only if $G - v$ has an edge clique cover of size at most $k$.

($\Rightarrow$): If $\mathcal{C}$ is an edge clique cover of $G$ of size at most $k$, then $\{S \setminus \{v\} : S \in \mathcal{C}\}$ is an edge clique cover of $G - v$ of size at most $k$.

($\Leftarrow$): Let $\mathcal{C}'$ be an edge clique cover of $G - v$ of size at most $k$. Partition $\mathcal{C}$ into $\mathcal{C}_u = \{S \in \mathcal{C} : u \in S\}$ and $\mathcal{C}_{\neg u} = \mathcal{C} \setminus \mathcal{C}_u$. Note that each set in $\mathcal{C}'_u = \{S \cup \{v\} : S \in \mathcal{C}_u\}$ is a clique since $N_G[u] = N_G[v]$ and that each edge incident to $v$ is contained in at least one of these cliques. Now, $\mathcal{C}'_u \cup \mathcal{C}_{\neg u}$ is an edge clique cover of $G$ of size at most $k$. □

**(Size-V)**
If the previous simplification rules do not apply and $|V| > 2^k$, then return No.

**Lemma 9.** *(Size-V) is sound.*

*Proof.* For the sake of contradiction, assume neither (Isolated) nor (Twins) are applicable, $|V| > 2^k$, and $G$ has an edge clique cover $\mathcal{C}$ of size at most $k$. Since $2^{\mathcal{C}}$ (the set of all subsets of $\mathcal{C}$) has size at most $2^k$, and every vertex belongs to at least one clique in $\mathcal{C}$ by (Isolated), we have that there exists two vertices $u, v \in V$ such that $\{S \in \mathcal{C} : u \in S\} = \{S \in \mathcal{C} : v \in S\}$. But then, $N_G[u] = \bigcup_{S \in \mathcal{C} : u \in S} S = \bigcup_{S \in \mathcal{C} : v \in S} S = N_G[v]$, contradicting that (Twin) is not applicable. □

**Kernel for Edge Clique Cover**

**Theorem 10.** Edge Clique Cover *has a kernel with $O(2^k)$ vertices and $O(4^k)$ edges.*

**Corollary 11.** Edge Clique Cover *is* FPT.

## 4 Frequently Arising Issues

**Issue 1**: A kernelization needs to produce an instance of the same problem.

How could we turn the following lemma into a simplification rule?

**Lemma 12.** *If there is an edge $\{u, v\} \in E$ such that $S = N_G[u] \cap N_G[v]$ is a clique, then there is a smallest edge clique cover $\mathcal{C}$ with $S \in \mathcal{C}$.*

*Proof.* By Lemma 6, we may assume the clique covering the edge $\{u, v\}$ is a maximal clique. But, $S$ is the unique maximal clique covering $\{u, v\}$. □

**(Neighborhood-Clique)**
If there exists $\{u, v\} \in E$ such that $S = N_G[u] \cap N_G[v]$ is a clique, then ...???

Edges with both endpoints in $S \setminus \{u, v\}$ are covered by $S$ but might still be needed in other cliques.

We could design a kernelization for a more general problem.

| Generalized Edge Clique Cover | |
|---|---|
| Input: | A graph $G = (V, E)$, a set of edges $R \subseteq E$, and an integer $k$ |
| Parameter: | $k$ |
| Question: | Is there a set $\mathcal{C}$ of at most $k$ cliques in $G$ such that each $e \in R$ is contained in at least one of these cliques? |

**(Neighborhood-Clique)**
If there exists $\{u, v\} \in R$ such that $S = N_G[u] \cap N_G[v]$ is a clique, then set $G \leftarrow (V, E \setminus \{u, v\})$, $R \leftarrow R \setminus \{\{x, y\} : x, y \in S\}$, and $k \leftarrow k - 1$.

**Issue 2**: A proposed simplification rule might not be sound.

Consider the following simplification rule for Vertex Cover.

**(Degk)**
If $\exists v \in V$ such that $d_G(v) \geq k$, then set $G \leftarrow G - v$ and $k \leftarrow k - 1$.

To show that a simplification rule is not sound, we exhibit a counter-example.
**Issue 3**: A problem might be FPT, but only an exponential kernel might be known / possible to achieve.