Partial kernelization of multiway cut: bounding the number of vertices with small excess

Igor Razgon *

Cork Constraint Computation Centre
Computer Science Department, University College Cork, Ireland
i.razgon@cs.ucc.ie

Abstract. We introduce the notion of partial kernelization where instead of bounding the size of the whole input we bound the size of the part of the input having particular property. We apply this notion to the multiway cut problem and show how to bound the number of vertices participating in small isolating cuts. We argue that the proposed results provide a considerable progress towards understanding the kernelizability of the multiway cut problem.

1 Introduction

Partial kernelization. Kernelization is probably the most practically applicable methodology of design of fixed-parameter algorithms. In a simplified form the kernelization can be defined as follows. Let $(I, k_1, \dots, k_r)$ be an instance of a decision problem **A** where I is the input and $k_1, \dots, k_r$ are parameters. Kernelization is an algorithm whose runtime polynomial in $|I|$ and the output is an instance $(I', k'_1, \dots, k'_r)$ of problem **A** equivalent to $(I, k_1, \dots, k_r)$ [1] and such that $k'_1 \leq k_1, \dots k'_r \leq k_r$ and (the most important condition!) $|I'|$ polynomially depends on $k'_1, \dots k'_r$. The above methodology is practically important because kernelization can be perceived as an algorithm of reducing the input size at the preprocessing stage providing (unlike many other preprocessing algorithms) a guaranteed upper bound on the size of the resulting instance.

Not all fixed-parameter tractable (FPT) problems admit kernelization. A recent research direction started from [1] has identified a number of problems that are not kernelizable unless some widely believed conjecture in the complexity theory fails. There are also problems whose kernelizability is a challenging open question, for example Directed Feedback Vertex Set (DFVS) [6]. For such problems it would be interesting to consider algorithms that allow significant reduction of a *considerable portion* of the input.

As an attempt to address the above question, we introduce the notion of *partial kernelization*. We define it w.r.t. graph-theoretic problems but this definition is easily extendable to other kinds of problems. Let **A** be a problem whose

* Supported by Science Foundation Ireland 05/IN/I886

[1] The equivalence is in the sense that the output of the former instance is 'YES' if and only if the output of the latter instance is 'YES'

input is a labeled graph $(G, V_1, \dots V_z)$, i.e. graph with specially identified subsets of its vertices. Let P be a predicate defining a property of vertices of G w.r.t. $V_1, \dots V_z$. Formally, for each $v \in V(G)$, $P(G, V_1, \dots V_z, v)$ is either *true* or *false* and the property is satisfied for those vertices where the predicate is *true*. Given parameters $k_1, \dots, k_r$, we say that **A** is *partially kernelizable* w.r.t. property P if there is a polynomial-time algorithm transforming the instance $(G, V_1, \dots V_z, k_1, \dots, k_r)$ into an equivalent instance $(G', V'_1, \dots, V'_z, k'_1, \dots, k'_r)$ such that $k'_1 \leq k_1, \dots k'_r \leq k_r$ and $|\{v \in V(G') | P(G'V'_1, \dots, V'_z, v) = \mathit{true}\}|$ is polynomial in $k'_1, \dots k'_r$. In other words, polynomial kernelization regarding P significantly reduces the number of vertices having property P. One also needs a criterion as to whether a particular property is sufficiently interesting to consider kernelization regarding it. To answer this question, we introduce the notion of *useful partial kernelization*. In particular, kernelization regarding property P is useful if problem **A** remains intractable even on graphs where all vertices have property P.

Technical Results. Let (G, T) be a pair where G is a graph and $T \subseteq V(G)$ is a subset of its vertices called *the terminals*. A subset C of $V(G) \setminus C$ is a multiway cut (MWC) of (G, T) if in $G \setminus C$ no two terminals belong to the same connected component. Given a parameter k, the MWC problem asks if (G, T) has an MWC of size at most k. In this paper we consider partial kernelization of the MWC problem regarding a property described below.

An *isolating cut* of $t \in T$ is a subset C of $V(G) \setminus T$ such that in $G \setminus C$ there is no path from t to $T \setminus \{t\}$. The *excess* of C w.r.t. t is the difference between C and the smallest possible size of an isolating cut of t. The excess of a vertex $v \in V(G) \setminus T$ with respect $t \in T$ is the smallest possible r such that there is a minimal isolating cut C of excess r w.r.t. t such that $v \in C$. The excess of v is the smallest possible excess w.r.t. a terminal of T. The property $P(G, T, v)$ considered in this paper is true if and only if $v \in T$ [2] or the excess of v is at most 1. We provide the following results.

- We show that the MWC problem parameterized by k and $|T|$ is partially kernelizable w.r.t. $P(G, T, v)$. More precisely, we show that there is a polynomial-time algorithm transforming (G, T, k) into an equivalent instance (G', T', k') where $k' \leq k$ and the number of vertices of $v \in V(G')$ such that $P(G', T', v)$ is true is at most $3(k'+1)|T'|$.
- We prove that the MWC problem remains NP-hard even for class instances of (G, T) such that $P(G, T, v)$ is true for all $v \in G$. Thus we demonstrate that the above kernelization is useful.
- We consider an instance of the multiway cut problem obtained as a result of the above kernelization. For this instance we show that if a question is to check existence of a multiway cut of size at most k being a subset of the vertices of excess 1, the instance can be further kernelized to become of a size polynomially dependent on k only.

[2] To make the description more intuitive, in the sequel of the paper we mainly refer to partial kernelization regarding vertices of excess 1 without explicit mentioning of terminal vertices.

Motivation. The MWC problem is a natural generalization of the polynomially solvable min-cut problem. The MWC problem is well known to be FPT [12, 5]. Understanding the kernelizability is a natural next step of investigation of this problem from the point of view of parameterized complexity. However, to the best of our knowledge, currently there are no kernelizability results even for special cases of the MWC problem.

We believe that the results proposed in this paper provide an important insight into the kernelizability of MWC problem. Apart from being the first results in this direction, they create a *framework* posing a number of natural questions towards understanding the kernelizability of the MWC problem. These questions are:

– Is the MWC problem partially kernelizable w.r.t. vertices of excess at most 2? at most 3? any fixed excess s?
– Even if the answer to the previous question is affirmative it may happen that the resulting set of vertices of excess at most s is of size exponential in s. Is it possible to get rid of the exponential dependence? Observe that the affirmative answer to this question will imply kernelizability of the general MWC problem parameterized by k and $|T|$ because any (minimal) multiway cut C of size at most k is the union of minimal isolating cuts of size at most k for all the terminals of T. Therefore, all the vertices of C are of excess at most k.
– Assuming the affirmative answer to the previous question, is it possible to get rid of the dependence of $|T|$? We believe that the answer to this question is positive and the respective indication is provided by our last result in the list given in the previous subsection.

Related work Investigation of the methods of coping with NP-hardness of the MWC problem started from the seminal paper [7], where the NP-hardness of the problem has been proven and the first fixed-ratio approximation algorithm (for the edge version of problem) has been proposed. This result has been followed by a row of improved approximation algorithms (see e.g. [4] and [9]). The first fixed-parameter algorithm for the MWC problem has been proposed in [12] and a significantly improved algorithm has been proposed in [5]. The latter algorithm gave rise to the first fixed-parameter algorithms for the DFVS and min-2CNF deletion problems [6, 15]. Apart from the multiway cut, other special cases of the *multicut problem* have been investigated, see e.g. [12, 11, 10] and some of them have been found kernelizable [3, 2]. The result [2] addresses the correlation clustering problem whose unweighted version is equivalent to the edge multicut [8].

Organization of the paper. Section 2 provides additional terminology and a number of technical statements. Section 3 proves that the MWC problem is NP-hard even if all non-terminal vertices are of excess 0. Section 4 provides the partial kernelization (parameterized by k and T) for vertices of excess at most 1. Section 5 provides presents the last result in the list provided above in Technical Results subsection. The proofs omitted due to the space constraints are postponed to the appendix.

2 Preliminaries

We employ a standard notation related to graphs. In particular, given a graph G, let $C \subseteq V(G)$. Then $G[C]$ denotes the subgraph of G induced by C and $G \setminus C \equiv G[V(G) \setminus C]$. For $v \in V(G)$, $G \setminus v \equiv G[V(G) \setminus \{v\}]$ and $N(v)$ is the set of neighbors of v in G. Also, $N(C) \equiv (\bigcup_{v \in C} N(v)) \setminus C$.

Let X and Y be two sets of vertices of the given graph G. A set $K \subseteq V(G) \setminus (X \cup Y)$ is an $X - Y$ separator if in $G \setminus K$ there is no path from X to Y. Let A, B be two disjoint subsets of $V(G)$. We denote by $NR(G, A, B)$ the set of vertices that are not reachable from A in $G \setminus B$ Let K_1 and K_2 be two $X - Y$ separators. We say that $K_1 \geq K_2$ if $NR(G, Y, K_1) \supseteq NR(G, Y, K_2)$.

Proposition 1. *Let K_1 and K_2 be two minimal $X - Y$ separators. Then $K_1 \leq K_2$ if and only if $K_1 \setminus K_2 \subseteq NR(G, Y, K_2)$.*

Proposition 2. *Let K_1 and K_2 be two minimal $X - Y$ separators. Let $K_1^t = K_1 \cap NR(G, Y, K_2)$, $K_1^b = (K_1 \setminus K_1^t) \setminus (K_1 \cap K_2)$. Accordingly, let $K_2^t = K_2 \cap NR(G, Y, K_1)$ and $K_2^b = (K_2 \setminus K_2^t) \setminus (K_1 \cap K_2)$ (the superscripts 't' and 'b' correspond to the words 'top' and 'bottom'). Denote $K_1^t \cup K_2^t \cup (K_1 \cap K_2)$ and $K_1^b \cup K_2^b \cup (K_1 \cap K_2)$ by, respectively, KT and KB. Then both KT and KB are $X - Y$ separators. Moreover, $KB \geq K_1$ and $KB \geq K_2$.*

Let $\{u, v\} \in E(G)$. We define graph $G_{u \leftarrow v}$ as the graph obtained from $G \setminus v$ by making u adjacent to all the vertices of $N(v) \setminus (N(u) \cap \{u\})$. In other words, $G_{u \leftarrow v}$ is obtained by contraction of $\{u, v\}$ where the new vertex obtained as a result of contraction is identified with u.

Theorem 1. [**[5] (Theorem 3.2.)**] *Let $(G, T = \{t_1, \dots, t_m\})$ be an instance of the* MWC *problem. Let v be a non-terminal vertex adjacent to t_1 and assume that there is a minimum isolating cut of t_1 that does not contain v. Then the instances $(G_{t_1 \leftarrow v}, T)$ and (G, T) have the same size of a smallest* MWC.

3 Usefulness of partial kernelization w.r.t. vertices of small excess

Theorem 2. *The* MWC *problem is NP-hard even if all the non-terminal vertices of the graph in the considered* MWC *instance are of excess* 0.

Proof. We provide a polynomial time reduction from the NP-hard problem Vertex Cover (VC) for graph of Max-Degree 3 into the problem of computing a smallest MWC for a graph where all the non-terminal vertices of of excess 0.

So, let G be a graph of max-degree 3. Clearly, G can be colored in at most 4 colors and this coloring can be computed in a polynomial time. So, let A_1, A_2, A_3, A_4 be a partition of $V(G)$ into 4 independent sets. Introduce new vertices t_1, t_2, t_3, t_4 and make each t_i be adjacent to all the vertices of A_i. Let G' be the resulting graph and consider the instance $(G', \{t_1, \dots, t_4\})$ of the MWC problem. Observe that $C \subseteq V(G)$ is a VC of G if and only if it is a MWC of $(G', \{t_1, \ldots, t_4\})$. Indeed, assume that C is a VC of G. Then in $G' \setminus C$ the only remaining edges are incident to t_i. Since, by construction, no two distinct t_i and t_j have a common neighbor, it follows that no two terminals of $\{t_1, \ldots, t_2\}$ belong to the same connected component of $G' \setminus C$, i.e. C is a MWC of $(G', \{t_1, \ldots, t_4\})$. Conversely, assume that C is a MWC of $(G', \{t_1, \ldots, t_4\})$. If C is not a VC of G then there is $\{u, v\} \in E(G)$ disjoint with C. Clearly, u and v belong to distinct partition classes of $\{A_1, \ldots A_4\}$. Assume w.l.og. that $u \in A_1$ while $v \in A_2$. Then t_1, u, v, t_2 is a path between t_1 and t_2 in $G' \setminus C$ in contradiction to being C a MWC of $(G', \{t_1, \ldots, t_4\})$.

Now, perform the following operation. Whenever G' has an edge $\{t_i, v\}$ such that v does not participate in a minimum isolating cut of t_i, perform the replacement operation $G' \leftarrow G'_{t_i \leftarrow v}$. Checking if there is required $\{t_i, v\}$ can be done in a polynomial time by network flow techniques and the number of iterations is $O(n)$ because each iteration but the final one decreases the number of vertices of the resulting graph. So, in a polynomial time we obtain a graph G^* for which the above condition is not satisfied for any edge $\{t_i, v\}$. Applying inductively Theorem 1, we observe that the size of a minimum MWC of the initial instance and of $(G^*, \{t_1, \ldots, t_4\})$ is the same. Taking into account the above discussion, it follows that the minimum size of a MWC of $(G', \{t_1, \ldots, t_4\})$ equal the minimum size of a VC of G. Finally, observe that all the non-terminal vertices of G^* are of excess 0. Indeed, assume that some $v \in V(G^*) \setminus \{t_1, \ldots t_4\}$ is of a higher excess. It is not hard to see that v is adjacent to some t_i (this property is true for the initial graph and it is preserved by each contraction). Since the excess of v is not zero, this means that v does not participate in any minimum isolating cut of t_i and the edge $\{t_i, v\}$ can be further contracted in contradiction to our assumption regarding G^*. It follows that indeed all the non-terminal vertices of G^* are of excess 0, completing the proof. ■

4 Bounding the number of vertices of excess at most 1

In this section we present a transformation that reduces the given instance (G, T) of the MWC problem into one where for each terminal $t \in T$ there are at most $3k$ vertices participating in a minimal isolating cut of t of excess at most 1. Thus the number of vertices of excess at most 1 will be at most $3k|T|$. We present two reduction rules, reason about instances where such reduction rules cannot be applied, and finish this section with showing that such an irreducible instance can be obtained in a polynomial time. We assume that all the terminals are mutually non-adjacent because otherwise 'NO' can be immediately returned.

Reduction Rule 1 *Whenever, there is $t \in T$ and $\{t, v\} \in E(G)$ such that v does not participate in some smallest isolating cut of t, replace of G with $G_{t \leftarrow v}$.*

Definition 1. *An instance (G, T) of the* MWC *problem is 1-irreducible if no edge $\{t, v\}$, where $t \in T$, can be contracted by reduction Rule 1 and no two terminals have a common neighbor.*

Lemma 1. *Let (G, T) be a 1-irreducible instance of the* MWC *problem. Then for each $t \in T$, $N(t)$ is the only smallest isolating cut of t.*

Proof. According to the assumption of the lemma, $N(t)$ is a subset of any smallest isolating cut of t. On the other hand, $N(t)$ itself is an isolating cut, implying the lemma. ■

Lemma 1 in fact says that the number of vertices of excess 0 in a 1-irreducible instance is at most $k|T|$ (unless a smallest isolating cut of some terminal is of size greater than k and 'NO' can be immediately returned). The rest of the section shows how to bound the number of vertices of excess 1.

Definition 2. *Let (G, T) be a 1-irreducible instance of the* MWC *problem and let $t \in T$. A subset S of $N(t)$ is a* coverable *set (of t) if there is a minimal isolating cut K of t of excess 1 such that $N(t) \setminus K \supseteq S$. We say that S is a* maximal coverable set *(of t) if S is not a proper subset of any coverable set of t.*

Our next reduction rule presented later in this section will reduce the number of vertices of excess 1 w.r.t. $t \in T$ *per* maximal coverable set of t. At the first glance such reduction looks of little use because the number of maximal coverable sets might be exponential in k. The following theorem justifies this reduction rule by showing that the number of maximal coverable sets of k is in fact linear in k.

Theorem 3. *Let (G, T) be a 1-irreducible instance of the* MWC *problem and let $t \in T$. Then maximal coverable sets of t are pairwise disjoint.*

Proof. Let S_1 and S_2 be two maximal coverable sets of t. It follows that there are two minimal isolating cuts K_1 and K_2 of t such that $N(t) \setminus K_1 = S_1$ and $N(t) \setminus K_2 = S_2$. Assume that $S_1 \cap S_2$ is non-empty. Let $K_1^t, K_1^b, K_2^t, K_2^b, KT, KB$ be as in the statement of Proposition 2. According to Proposition 2, KB is an isolating cut of t such that $K_1 \leq KB$ and $K_2 \leq KB$. It follows that $S_1 \subseteq NR(G, K_1, T \setminus \{t\}) \subseteq NR(G, KB, T \setminus \{t\})$. Consequently, $S_1 \subseteq N(t) \setminus KB$. Analogously, it can be shown that $S_2 \subseteq N(t) \setminus KB$. Since both $S_1 \setminus S_2$ and $S_2 \setminus S_1$ are non-empty due to their maximality, we conclude that $N(t) \setminus KB$ is a proper superset of both S_1 and S_2.

Next, observe that $|KB| = |N(t)| + 1$. Indeed, $2(|N(t)| + 1) = |K_1| + |K_2| = |K_1^t| + |K_1 \cap K_2| + |K_1^b| + |K_2^t| + |K_1 \cap K_2| + |K_2^b| = (|K_1^t| + |K_1 \cap K_2| + |K_2^t|) + (|K_1^b| + |K_1 \cap K_2| + |K_2^b|) = |KT| + |KB|$. According to our assumption $N(t) \setminus KB \neq \emptyset$, therefore $|KB| > |N(t)|$ by Lemma 1. The only possibility to avoid $|KB| = |N(t)| + 1$ is to assume that $|KB| = |N(t)| + 2$ and $|KT| = |N(t)|$. But then $KT = N(t)$ by Lemma 1. In this case $S_1 = K_2^t$ and $S_2 = K_1^t$. However, K_1^t and K_2^t are disjoint by construction, while S_1 and S_2 are not disjoint according to our assumption. This contradiction proves that $|KB| = |N(t)| + 1$.

Finally, observe that KB is a minimal isolating cut of t indeed, otherwise it follows that there is an isolating cut of size at most $|N(t)|$ which does not coincide with $N(t)$ in contradiction to Lemma 1.

Thus, in contradiction to the maximality of S_1 as a coverable set of t, KB is a minimal isolating cut of t with excess 1 such that $N(t) \setminus KB$ is a proper superset of S_1. This contradiction shows that our initial assumption regarding non-disjointness of two maximal coverable sets of t is false and any two maximal coverable sets are indeed disjoint. ■

Now we are ready to provide the reduction rule that bounds the number of vertices of excess 1.

Reduction Rule 2 *Let (G, T) be a 1-irreducible instance of the* MWC *problem. Replace G by $G_{u \leftarrow v}$ whenever there is an edge $\{u, v\}$ and a terminal $t \in T$ satisfying the following conditions.*

- *$u \in N(t)$, while $v \notin N(t) \cup \{t\}$;*
- *u belongs to a maximal coverable set S of t;*
- *there is a minimal isolating cut of t of excess 1 such that $N(t) \setminus K = S$ and $v \notin K$.*

The following theorem shows that Reduction Rule 2 is correct.

Theorem 4. *Let (G, T) be a 1-irreducible instance of the* MWC *problem. Let $\{u, v\}$ be an edge of G satisfying the conditions of Reduction Rule 2 w.r.t. a terminal t. Then the instances (G, T) and $(G_{u \leftarrow v})$ have the same smallest size of the multiway cut.*

Proof. Let K be a smallest multiway cut of (G, T) and let C be a minimal isolating cut of t of excess 1 such that $v \notin C$. If $v \notin K$ then K remains a multiway cut of $(G_{u \leftarrow v}, T)$ and there is nothing to prove. Assume that $v \in K$. Denote $K \cap NR(G, C, T \setminus \{t\})$ by K'. Denote by C' the subset of $C \setminus K$ consisting of all the vertices w such that in $G \setminus K'$ any path from t to w meets at least one vertex of C other than w. Finally, denote $(K \setminus K') \cup C'$ by K^*. Arguing analogously to Theorem 3.2. in [5], we can observe that K^* is a multiway cut [3]

Observe that $|K^*| \leq |K|$. By definition of K' and C', $C^* = (C \setminus C') \cup K'$ is an isolating cut of t, $|C^*| = |C| - |C'| + |K'|$, and $|K^*| = |K| - |K'| + |C'|$. Therefore, it is sufficient to ensure that $|K'| \geq |C'|$. Assume the opposite. In this case $|C^*| < |C|$, i.e. C^* is a smallest isolating cut of t. Since, according to Lemma 1, $N(t)$ is the only such isolating cut, it follows that $C^* = N(t)$. We claim that this causes a contradiction. Indeed, no vertex of $v \in N(t)$ can belong to C' due to the adjacency of v to t. Consequently, $(C \setminus C') \cap N(t) = C \cap N(t)$. To ensure that $(C \setminus C') \cup K' = N(t)$ it is necessary that $C \setminus C' \subseteq N(t)$. Taking into account the disjointness of C' and $N(t)$, it is necessary that $C \setminus C' = C \cap N(t)$. But then $K' = N(t) \setminus C$, a contradiction since $v \in K' \setminus N(t)$. Thus we have verified that indeed $K^* \geq K$. To complete the theorem it remains to add that by construction $v \notin K^*$ and it is not hard to verify that K^* is a MWC of $(G_{u \leftarrow v}, T)$. ■

Definition 3. *An instance (G, T) of the* MWC *problem is 2-irreducible if it is 1-irreducible, for each $t \in T$, $|N(t)| \leq k$, and Reduction Rule 2 cannot contract any edge of G.*

[3] In the proof in [5] the isolating cut is a smallest one but this fact is not used for the proof that K^* is a multiway cut.

Let (G, T) be a 2-irreducible instance of the MWC problem and let $t \in T$. Denote the union of all maximal coverable set of t by $CS(t)$. Denote $N(CS(t)) \setminus (N(t) \cup \{t\})$ by $CS_2(t)$. The following theorem shows that in a 2-irreducible instance of the MWC problem the number of vertices of excess 1 is linearly bounded by k.

Theorem 5. *Let (G, T) be a 2-irreducible instance of the* MWC *problem. For each $t \in T$, $|N(t) \cup CS_2(t)| \leq 3k$ and all vertices of excess at most 1 w.r.t. t belong to $N(t) \cup CS_2(t)$.*

Proof. Fix a terminal t and let S be a maximal coverable set of t. Denote $N(S) \setminus (N(t) \cup \{t\})$ by NS. Due to the impossibility of applying Reduction Rule 2, it follows that NS is a subset of any minimal isolating cut K of t of excess 1 such that $N(t) \setminus K = S$. In fact, by definition of K, $NS \cup (N(t) \setminus S)$ is a subset of any such cut. However, $NS \cup (N(t) \setminus S)$ is an isolating cut of t. It follows that $|N(S) \cup (N(t) \setminus S)| = |N(t)| + 1$. Indeed, otherwise, it will follow that there is no minimal isolating cut K of excess 1 such that $N(t) \setminus K = S$ in contradiction to being S a maximal coverable set. It follows that $|NS| = |S| + 1$. Taking into account Theorem 3 and that the set $CS_2(t)$ is just the union of sets NS over all maximal coverable sets S of t, $|CS_2(t)| \leq 2k$. Thus $|N(t) \cup CS_2(t)| \leq 3k$ as required.

Assume now that there G has a vertex u of excess at most 1 w.r.t. t which is not a subset of $N(t) \cup CS_2(t)$. Since, according to Lemma 1, any vertex of excess 0 w.r.t. t is a subset of $N(t)$, the excess u w.r.t. t is 1. Let K_1 be a minimal isolating cut of t of excess 1 such that $u \in K_1$. Let $S_1 = N(t) \setminus K_1$. Let S be the maximal coverable set of t such that $S_1 \subseteq S$. According to Theorem refcoverdisjoint, S is unique. Let K_2 be a maximal isolating cut of t of excess 1 such that $N(t) \setminus K_2 = S$. Let KB be as in the proof of Theorem 3 w.r.t. K_1 and K_2. Applying the analogous argumentation, we conclude that KB is a minimal isolating cut of excess 1 and that $N(t) \setminus KB \supseteq S_1 \cup S$. Due to the maximality of S, $N(t) \setminus KB = S$. Applying the argumentation of the previous paragraph, we obtain that $KB = NS \cup (N(t) \setminus S)$. We derive a contradiction by showing that at least one vertex of NS belongs to $NR(G, K_1, T \setminus \{t\})$ and hence cannot belong to KB by construction. This contradiction will show the vertex u cannot exist and complete the proof of the present theorem.

Clearly, $t \in NR(G, K_1, T \setminus \{t\})$. It follows that $S_1 \subseteq NR(G, K_1, T \setminus \{t\})$. Consequently, each vertex of $NS_1 = N(S_1) \setminus (N(t) \cup \{t\})$ belongs either to $NR(G, K_1, T \setminus \{t\})$ or to K_1. Assume that $NS_1 \subseteq K_1$. It follows that $NS_1 \cup (N(t) \setminus S_1) \subseteq K_1$. Observe that $NS_1 \cup (N(t) \setminus S_1)$ is an isolating cut of t. Due to the minimality of K_1, it follows that $NS_1 \cup (N(t) \setminus S_1) = K_1$, i.e. $K_1 \subseteq N(t) \cup CS_2(t)$ in contradiction to $u \in K_1$. ■

Theorem 5 shows that the desired bound of vertices of excess at most 1 is achieved for 2-irreducible instances of the MWC problem. Thus, to claim the desired partial kernelizability of the MWC problem, it is enough to show that any instance of the MWC problem can be transformed into 2-irreducible one in a polynomial time. This is done in the following theorem preceded by two auxiliary lemmas.

Lemma 2. *Let (G,T) be a 2-irreducible instance of the* MWC *problem, $t \in T$, and $S \subseteq N(t)$. For $S \subseteq N(t)$, it can be checked in a polynomial time whether S is coverable.*

Lemma 3. *Let (G,T) be a 2-irreducible instance of the* MWC *problem and $t \in T$. There is a polynomial algorithm that outputs the collection of maximal coverable sets of t.*

Proof. Let H be the graph whose set of vertices is the subset of $N(t)$ consisting of all vertices v such that $\{v\}$ is a coverable set of t. Two vertices u and v are adjacent in H if and only if $\{u,v\}$ is a coverable set of t. According to Lemma 2, H can be computed in a polynomial time. We claim that the sets of vertices of connected components of H are the maximal coverable sets of t. Clearly, the theorem immediately follows from this claim.

To prove the above claim, let U be a set of vertices of a connected component of H and let $u_1,\dots,u_r$ be the vertices of U ordered in such a way that for each i from 1 to r, $H[U_i]$ is connected, where $U_i = \{u_1,\dots,u_i\}$. We first show that each U_i is a coverable set of t The proof is by induction on i. U_1 and U_2 are clearly coverable sets by construction of H. So, assume that $i \geq 3$ and assume w.l.o.g. that $\{u_{i-1}, u_i\} \in E(H)$. By the induction assumption and the construction of H, there are sets S_1 and S_2 that are supersets of, respectively, U_{i-1} and $\{u_{i-1}, u_i\}$ and such that there are minimal isolating cuts K_1 and K_2 of t, both having excess 1 and such that $S_1 = N(t) \setminus K_1$ and $S_2 = N(t) \setminus K_2$. Arguing analogously to the proof of Theorem 3, we conclude that U_i is a coverable set. Consequently, $U = U_r$ is also a coverable set. To show that U is a maximal coverable set, assume that there is a vertex $v \in N(t) \setminus U$ such that $U \cup \{v\}$ is a coverable set. But then, both $\{v\}$ and $\{u_r, v\}$ are coverable sets. It follows that v is a vertex of H adjacent to u_r. Consequently, v belongs to the set of vertices of the connected component of u_r, namely to U, a contradiction completing the proof of the required claim and of the present theorem. ∎

Theorem 6. *The* MWC *problem is partially kernelizable w.r.t. to vertices of excess at most* 1.

Proof sketch According to Theorem 5, it is sufficient to show that that the given instance (G,T) of the MWC problem can be transformed in a polynomial time into an equivalent 2-irreducible instance without increasing the parameter and the number of terminals. This can be done by $O(n)$ removal of common neighbors between the terminals and application of Reduction Rules 1 and 2. The only part of this algorithm whose polynomial time computability is non-trivial is computing the maximal coverable sets of terminals in order to check applicability of Reduction Rule 2. That this part can be efficiently computed follows from Lemma 3. ∎

5 A tighter kernelization under an additional assumption

Let (G,T) be a 2-irreducible instance of the MWC problem ad k be a parameter. The 1-MWC problem asks if (G,T) has an MWC consisting of vertices of excess at most 1. In this section we show that the 1-MWC problem is kernelizable being parameterized by k only. The motivation behind considering the 1-MWC is that the approach to its kernelization may be useful for getting rid of the parameter $|T|$ in case we know how to kernelize the MWC parameterized by k and $|T|$.

The algorithm described in this section is an adaptation of a simple kernelization of the VC problem [14]. The subtle point of this adaptation is that it is applied to *sets* of vertices rather than to *single* vertices as in the case of VC.

Denote $\bigcup_{t \in T}(\{t\} \cup N(t) \cup CS_2(t))$ by V'. Since we are interested to find a smallest multicut which is a subset of V', we get rid of all the vertices that are not in V'. In particular, we replace G with a graph obtained from $G[V']$ by adding edges between any two nonadjacent vertices u and v that are adjacent to the same connected component of $G \setminus V'$. It is not hard to verify that all the multiway cuts that are subsets of V' are preserved by this transformation (see e.g. Proposition 2.5. in [13]).

Let $T = \{t_1, \ldots, t_m\}$ and assume that the vertices of G are partitioned into sets $C_1, \ldots C_m$ such that for each C_i, $t_i \in C_i$ and $G[C_i]$ is connected. We say that such partition is *nice*. In addition, we say that a non-terminal vertex $v \in C_i$ is *removable* v is adjacent to at least $k+2$ partition classes other than C_i and $G[C_i] \setminus v$ is connected.

Lemma 4. *Let $C_1, \ldots C_m$ be a nice partition and let $v \in C_i$ be a removable vertex. Then $(G \setminus v, T)$ has a multiway cut of size $k-1$ if and only if (G, T) has a multiway cut of size k.*

Lemma 4 inspires the following reduction rule.

Reduction Rule 3 *Let $C_1, \ldots, C_m$ be a nice partition. Then while $k \geq 0$ and there is a removable vertex v of some C_i, replace G with $G \setminus v$, C_i with $C_i \setminus \{v\}$ and k with $k-1$. If there are no removable vertices then perform the following 'cleaning' operations. If $k = 0$ and at least two terminals are connected then return 'NO'. If all the terminals have been separated then return 'YES'. If G has a connected component G' containing at most one terminal exclude this component from G, replace G with $G \setminus V(G')$, T with $T \setminus V(G')$ and each C_i with $C_i \setminus V(G')$.*

Corollary 1. *Reduction Rule 3 is correct.*

In order to apply Reduction Rule 3, we introduce the following nice partition. For each $t_i \in T$, $N(t_i) \cup \{t_i\} \subseteq C_i$; for each remaining vertex v, choose an arbitrary t_i such that $v \in CS_2(t_i)$ and introduce v into C_i. Note that by construction, each vertex of G belongs to $\{t_i\} \cup N(t_i) \cup CS_2(t_i)$ of some t_i, so this rule is well defined.

Observe that the above partition is nice. Indeed, for each t_i, the vertices of $N(t_i)$ are adjacent to t_i and each vertex of $CS_2(t_i)$ is adjacent to some vertex of $N(t_i)$. It follows that all the vertices of each $G[C_i]$ belong to the same connected component with t_i. Consequently, $G[C_i]$ is connected.

To understand the effect Reduction Rule 3, we specify a class of vertices of G and prove two its properties. In particular, we say that the set $\{t_i\} \cup CS(t_i)$ are *internal* vertices of t_i and a vertex is *internal* if its an internal vertex of some t_i.

Lemma 5. *Let v be an internal vertex and let C_i be the partition class containing v. Then v is adjacent to at most $k+1$ partition classes other than C_i.*

Proof. If $v = t_i$ then v is not adjacent to any partition class besides C_i. Otherwise, $v \in S$ where S is a maximal coverable set of t_i. The only neighbors of v that can belong to other partition classes are the vertices of $NS = N(S) \setminus (N(t) \cap \{t\})$. As argued in the proof of Theorem 5, NS is a subset of a minimal isolating cut of t_i of excess at most 1. Consequently, $|NS| \leq k+1$. Consequently, the vertices of NS cannot belong to more than $k+1$ partition classes. ■

Lemma 6. *Let C_i be a partition class and let $V' \subseteq C_i$ be an arbitrary subset of non-internal vertices. Then $G[C_i] \setminus V'$ is connected.*

Proof. By definition $t_i \in C_i \setminus V'$. Each vertex of $N(t_i) \setminus V'$ is adjacent to t_i. Each vertex of $CS_2(t_i) \setminus V'$ is, by definition, adjacent to some *internal* vertex that belongs to $N(t_i)$, i.e. to some vertex of $C_i \setminus V'$ adjacent to t_i. It follows that all the vertices of $G[C_i] \setminus V'$ are connected to t_i. ■

Now we are ready to state the effect Reduction Rule 3.

Theorem 7. *Let (G, T, k) be a 2-irreducible instance of the* MWC *problem such that $V(G) = \bigcup_{t \in T}(\{t\} \cup N(t) \cup CS_2(t))$ and let (G^*, T^*, k^*) be the instance obtained from (G, T, k) as a result of application of Reduction Rule 3 (on the assumption that no stopping condition occurred). Denote $|T^*|$ by m_1. Let $C_1 \dots C_m$ be the initial nice partition of $V(G)$ and let $C^*_1, \dots C^*_{m_1}$ be the partition of $V(G^*)$ resulting from application of Reduction Rule 3. Then each C^*_i is adjacent to at most $3k(k+1)$ classes C^*_j other than C^*_i.*

Proof. We are going to show that for each C^*_i and for each $v \in C^*_i$, v is adjacent to at most $k+1$ classes C^*_j other than C^*_i. Taking into account that according to Theorem 5, C_i has at most $3k+1$ vertices, the theorem will follow.

Assume w.l.o.g. that C^*_i is a subset C_i, i.e. elimination of some terminals has not changed their enumeration. If there is some $v \in C^*_i$ adjacent to at least $k+2$ classes C^*_j different from C^*_i then v is clearly adjacent to at least $k+2$ classes C_j different from C_i. Consequently, v is a non-internal vertex according to Lemma 5. Moreover, according to the same lemma, all the vertices in $C_i \setminus C^*_i$ are non-internal ones. Consequently, $C^*_i \setminus \{v\} = C_i \setminus ((C_i \setminus C^*_i) \cap \{v\})$ induces a connected subgraph of G and hence of G^* (the latter is an induced subgraph of G). It follows that in this case v is a removable vertex in contradiction to the fact that Reduction Rule 3 has terminated on (G^*, T^*, k^*) due to the absence of removable vertices. It follows that v cannot be adjacent to more than $k+1$ partition classes C^*_j other than C^*_i. ■

Corollary 2. *The* 1-MWC *problem is polynomially kernelizable.*

Proof sketch. Theorem 7 implies that if graph H has too many edges then it has a large matching and hence the instance (G^*, T^*, k^*) cannot have a multiway cut of size at most k. This in turn causes the kernelizability of the 1-MWC problem.

References

1. Hans L. Bodlaender, Rodney G. Downey, Michael R. Fellows, and Danny Hermelin. On problems without polynomial kernels. *J. Comput. Syst. Sci.*, 75(8):423–434, 2009.
2. Hans L. Bodlaender, Michael R. Fellows, Pinar Heggernes, Federico Mancini, Charis Papadopoulos, and Frances A. Rosamond. Clustering with partial information. In *MFCS*, pages 144–155, 2008.
3. Nicolas Bousquet, Jean Daligault, Stéphan Thomassé, and Anders Yeo. A polynomial kernel for multicut in trees. In *STACS*, pages 183–194, 2009.
4. Gruia Calinescu, Howard J. Karloff, and Yuval Rabani. An improved approximation algorithm for multiway cut. *Journal of Computer and System Sciences*, 60(3):564–574, 2000.
5. Jianer Chen, Yang Liu, and Songjian Lu. An improved parameterized algorithm for the minimum node multiway cut problem. In *WADS*, pages 495–506, 2007.
6. Jianer Chen, Yang Liu, Songjian Lu, Barry O'Sullivan, and Igor Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. *Journal of the ACM*, 55(5), 2008.
7. Elias Dahlhaus, David S. Johnson, Christos H. Papadimitriou, Paul D. Seymour, and Mihalis Yannakakis. The complexity of multiterminal cuts. *SIAM J. Comput.*, 23(4):864–894, 1994.
8. Dotan Emanuel and Amos Fiat. Correlation clustering - minimizing disagreements on arbitrary weighted graphs. In *ESA*, pages 208–220, 2003.
9. Naveen Garg, Vijay V. Vazirani, and Mihalis Yannakakis. Multiway cuts in node weighted graphs. *Journal of Algorithms*, 50(1):49–61, 2004.
10. Georg Gottlob and Stephanie Tien Lee. A logical approach to multicut problems. *Inf. Process. Lett.*, 103(4):136–141, 2007.
11. Jiong Guo, Falk Hüffner, Erhan Kenar, Rolf Niedermeier, and Johannes Uhlmann. Complexity and exact algorithms for vertex multicut in interval and bounded treewidth graphs. *European Journal of Operational Research*, 186(2):542–553, 2008.
12. Dániel Marx. Parameterized graph separation problems. *Theor. Comput. Sci.*, 351(3):394–406, 2006.
13. Dániel Marx, Barry O'Sullivan, and Igor Razgon. Treewidth reduction for constrained separation and bipartization problems. In *STACS*, pages 561–572, 2010.
14. Rolf Niedermeier. *Invitation to Fixed-Parameter Algorithms*, volume 31. Oxford Lecture Series in Mathematics and Its Applications, 2006.
15. Igor Razgon and Barry O'Sullivan. Almost 2-SAT is fixed-parameter tractable. *Journal of Computer and System Sciences*, 75:435–450, 2009.

A Proofs omitted from the main body of the paper

Proof of Proposition 1 Assume first that $K_1 \leq K_2$. Assume that there is $v \in K_1 \setminus K_2$ such that $v \notin NR(G, Y, K_2)$. Due to the minimality of K_1 there is a path p from X to Y that includes v and does not include any other vertex of K_1. Let pv be the prefix of p ending at v. Since $v \notin NR(G,Y,K_2)$, there is path p' from v to Y that does not pass through K_2. Let p'' be the $X-Y$ walk obtained by appending p' to the end of pv. Since K_2 is a $X-Y$ separator, p'' intersects with K_2. Since p' does not intersect with K_2, pv intersects with K_2. Let w be a vertex of K_2 in pv and let pw be the prefix of pv and hence of p. By our assumption about p, pw does not intersect with K_1. Also, there is a path from w to Y that does not intersect with K_1 because otherwise $w \in NR(G,Y,K_1)$ in contradiction to $K_1 \leq K_2$. It follows that there is a path from X to Y in $G \setminus K_1$, a contradiction, confirming such v does not exist.

Conversely, assume that $K_1 \setminus K_2 \subseteq NR(G,Y,K_2)$. Assume by contradiction that $NR(G,Y,K_1) \nsubseteq NR(G,Y,K_2)$. Let $v \in NR(G,Y,K_1) \setminus NR(G,Y,K_2)$. Observe that there is a path pv from v to Y whose intermediate vertices do not intersect with K_2. If $v \notin K_2$ such path exist by selection of v, otherwise such path exists by the minimality of K_2. Since $v \in NR(G,Y,K_1)$, pv contains a vertex $w \in K_1 \setminus K_2$. But this means that w is followed in pv by a vertex of K_2, a contradiction. ■

Proof of Proposition 2. We prove that KT is an $X-Y$ separator, for KB the argument is similar. Assume that KT does not separate X from Y and let p be a $X-Y$ path in $G \setminus KT$. By definition of K_1 and K_2, p necessarily intersects with $K^b_1 \cup K^b_2$. Let v be the first vertex of $K^b_1 \cup K^b_2$ occurring in p as it is traversed from X to Y. Let pv be the prefix of p ending at v. Assume w.l.o.g. that $v \in K^b_1$. By definition of K^b_1, there is path from v to Y that does not intersect with K_2. Also observe that pv does not intersect with K_2. Indeed, it does not intersect with K^b_2 because otherwise, v is not the first vertex of $K^b_1 \cup K^b_2$ occurring in p in contradiction to our assumption and pv does not intersect with $K_2 \setminus K^b_2$ as being a subset of KT. It follows that Y is reachable from X in $G \setminus K_2$, a contradiction proving that KT indeed separates X from Y.

Now, observe that $KB \geq K_1$ and $KB \geq K_2$. We prove only the former, the reasoning for the latter is similar. Assume that $KB \ngeq K_1$. It follows that there is a vertex $v \notin K_1$ which is not reachable from Y in $G \setminus K_1$ and regarding which either $v \in KB$ or v is reachable from Y in $G \setminus KB$.

If $v \in KB$ then $v \in K^b_2$. By definition of K^b_2, all of its vertices are reachable from Y in $G \setminus K_1$, a contradiction. It remains to assume that v is reachable from Y in $G \setminus KB$. Let p' be a path from v to Y in $G \setminus KB$. By our assumption p' intersects with $K_1 \setminus KB = K^t_1 \subseteq KT$. Let w be the last vertex of KT occurring in p' being traversed from v to Y. Since $K_1 \cup K_2 \subseteq KB$, $w \in K^t_1 \cup K^t_2$. Assume w.l.o.g. that $w \in K^t_1$. Let pw be the suffix of p' starting at w. Observe that pw does not intersect with K_2. Indeed, pw does not intersect with K^t_2 by our assumption about w and it does not intersect with $K_2 \setminus K^t_2$ as being a subset of KB. However, this is a contradiction since w is not reachable from y in $G \setminus K_2$. Thus we have shown that $KB \geq K_1$ and $KB \geq K_2$. ■

Proof of Lemma 2. Split each vertex of S into many copies, say $(10n + 100)$. Then compute a minimum isolating cut K of t. If $|K| = |N(t)| + 1$ then return 'YES'. Otherwise, return 'NO'. Assume that this algorithm returns 'YES'.

Observe that the isolating cut K witnessing the 'YES' answer does not contain any of the new copies of the vertices of S (otherwise, its size would be much larger than $|N(t)|+1$). Due to the minimality of K, it follows that K is a minimal isolating cut of t of excess 1 such that $S \subseteq N(t) \setminus K$. Thus S is indeed a coverable set.

Conversely, assume that there is a minimal isolating cut K of t with excess 1 such that $S \subset N(t) \setminus K$. Then K remains an isolating cut of t after splitting of the vertices of S. Moreover, since $N(t)$ is the only isolating cut of t of size $|N(t)|$, K is the smallest isolating cut of t. Thus in this case the algorithm cannot returns 'YES'. Thus we have verified correctness of the above algorithm. ■

Proof of Theorem 6. According to Theorem 5, it is sufficient to show that that (G,T) can be transformed in a polynomial time into an equivalent 2-irreducible instance without increasing the parameter and the number of terminals. This can be done by an algorithm that iteratively performs as follows. If there are two terminals having a common neighbor v, replace G by $G \setminus v$ and reduce parameter by 1. If all the terminals get separated then return 'YES'. If $k=0$ and some terminals are not separated then return 'NO'. If the condition of Reduction Rule 1 is satisfied then apply Reduction Rule 1. If the instance is 1-irreducible the condition of Reduction Rule 2 is satisfied then apply Reduction Rule 2. The algorithm finishes when the last iteration has not provided any reduction of the graph. Since each iteration decreases the number of vertices of the graph, there are $O(n)$ iterations. Applying inductively Theorems 1 and 4, we observe that the algorithm correctly returns 'YES' or correctly returns 'NO or returns a 2-irreducible instance of the MWC problem equivalent to the original one. Therefore it only remains to verify that the applicability of reduction rules can be checked in a poly-time. For Reduction Rule 1, this can be easily done by network flow techniques. The same can be said about Reduction Rule 2 provided we know the maximal coverable sets of each terminal of T. Therefore the theorem follows from Lemma 3. ■

Proof of Lemma 4. Assume that (G,T) has a multiway cut of size k. Then this cut necessarily contains v. Indeed, consider $k+2$ arbitrary partition classes other than C_i adjacent to v. Assume w.l.o.g. that they are $C_1, \dots C_{k+2}$. If v does not belong to some multiway cut C then, in order to separate $t_1, \dots t_{k+2}$ at least $k+1$ of $C_1, \dots C_{k+2}$ must contribute a vertex to C, i.e. $|C| \geq k+1$, a contradiction. Thus, in the considered case, $(G \setminus v, T)$ indeed has a multiway cut of size at most $k+1$.

Conversely, if (G,T) does not have a multiway cut of size at most k then, clearly $(G \setminus v, T)$ does not have a multiway cut of size at most $k-1$. ■

Proof of Corollary 1. The correctness of the iterative removal of vertices follows from inductive application of Lemma 4 to each iteration and from observing that after each removal, the resulting partition remains nice. The only non-trivial part of the 'cleaning' algorithm is the removal of connected components containing at most 1 terminal. The vertices of these components are redundant because they do not participate in any path between distinct two terminals. ■

Proof of Corollary 2. As specified by the reasoning above, an instance of the 1-MWC problem can be transformed in a polynomial time into the instance (G^*, T^*, k^*) of the MWC problem as appears in the statement of Theorem 7. Let H be a graph on $C_1^*, \dots C_{m_1}^*$ from the statement of Theorem 7 such that two classes C_i and C_j are adjacent in H if and only if they are adjacent in G, i.e. if there is $\{u, v\} \in E(G^*)$ such that $u \in C_i^*$ while $v \in C_j^*$. According to Theorem 7, the degree of each vertex of H is at most $3k(k+1)$. We claim that if H has more than $6k^2(k+1)$ edges then G has a matching of size at least $k+1$. Indeed, let M be the largest matching of H and assume that $|M| \leq k$. It follows that each edge of H is incident to at least one vertex of M. Since $|V(M)| \leq 2k$ and the degree of each vertex is at most $3k(k+1)$, the number of available edges is at most $6k^2(k+1)$ as required.

Observe that if H has a matching of size at least $k+1$ then (G^*, T^*) does not have a multiway cut of size at most k. Indeed, let M be a matching of H of size at least $k+1$ and let $\{C_i^*, C_j^*\}$ be an edge of M. It follows that $G^*[C_i^* \cup C_j^*]$ is connected and therefore, to separate C_i^* and C_j^*, at least one vertex of $C_i^* \cup C_j^*$ has to be contributed. Let $\{C_{i_1}^*, C_{j_1}^*\}, \dots, \{C_{i_r}^*, C_{j_r}^*\}$ be the edges of M. Taking into account that for any distinct x, y, $C_{i_x}^* \cup C_{j_x}^*$, is disjoint with $C_{i_y}^* \cup C_{j_y}^*$ at least $k+1$ vertices have to be contributed to separate all the terminals of T^*. It follows that if H has more than $6k^2(k+1)$ edges, 'NO' can be returned immediately. If 'NO' is not returned then H ahs at most $12k^2(k+1)$ vertices. Taking into account that each C_i^* has at most $3k+1$ vertices (the additional 1 is on the account of t_i), G^* has at most $36k^2(k+1)^2$ vertices in case 'NO' is not returned. Thus we have established a polynomial kernelizability of the 1-MWC problem with $O(k^4)$ of the resulting kernel size. ■