Fully decomposable split graphs

Hajo Broersma[1], Dieter Kratsch[2], and Gerhard J. Woeginger[3⋆]

[1] Department of Computer Science, Durham University, Science Laboratories, South Road, Durham, DH1 3LE England, email: hajo.broersma@durham.ac.uk.
[2] LITA, Université de Metz, 57045 Metz Cedex 01, France, email: kratsch@lita.univ-metz.fr.
[3] Department of Mathematics and Computer Science, TU Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, email: gwoegi@win.tue.nl.

Abstract. We discuss various questions around partitioning a split graph into connected parts. Our main result is a polynomial time algorithm that decides whether a given split graph is fully decomposable.

Keywords: graph decomposition; integer partition; computational complexity.

1 Introduction

Throughout we only consider finite undirected graphs without loops or multiple edges. Let $G = (V, E)$ be a graph on n vertices, and let $\alpha = (\alpha_1, \ldots, \alpha_k)$ denote a partition of n, that is, a sequence of positive integers $\alpha_1, \ldots, \alpha_k$ with $\sum_{i=1}^{k} \alpha_i = n$. The graph G is called α-*decomposable*, if there exists a partition of V into disjoint subsets $A_1, \ldots, A_k$ of cardinality $|A_i| = \alpha_i$ for $1 \leq i \leq k$ such that every set A_i induces a connected subgraph of G. Such a partition is called an α-*decomposition* of G, and a (connected) subgraph induced by $|A_i| = \alpha_i$ vertices is also referred to as an α_i-*component* of the α-decomposition. A graph is called *fully decomposable* (or *arbitrarily vertex decomposable*) if it is α-decomposable for every partition α of n.

Fully decomposable graphs were introduced by Horňák & Woźniak [6]. There are two natural algorithmic questions centered around α-decompositions of graphs.

Q1: Decide whether a given graph G is α-decomposable for a given partition α.

⋆ This research has been supported by the Netherlands Organisation for Scientific Research (NWO), grant 639.033.403, and by BSIK grant 03018 (BRICKS: Basic Research in Informatics for Creating the Knowledge Society).

Q2: Decide whether a given graph G is fully decomposable.

Question Q1 is notoriously hard. For instance, Dyer & Frieze [3] proved that it is NP-hard to decide whether a planar graph is $(3, 3, \ldots, 3)$-decomposable. Barth & Fournier [2] showed that Q1 is NP-hard for trees. Generally speaking, Q1 seems to be NP-hard for every natural non-trivial class of specially structured graphs.

The computational complexity of question Q2 is not understood. We are aware of only a single result on question Q2 from the literature: Barth, Baudon & Puech [1] designed a polynomial time algorithm for deciding whether a given tripode (a tree with a single vertex of degree three, and all other vertices of degree one or two) is fully decomposable. Barth & Fournier [2] also proved that every fully decomposable tree has maximum vertex degree at most four. Determining the precise computational complexity of Q2 is an outstanding open problem: The problem is neither known to be NP-hard, nor is it known to be contained in the class NP.

2 Results of this paper

A graph $G = (V, E)$ is a *split graph* (see for instance Golumbic [5]) if its vertex set can be partitioned into an induced independent set I and a clique C. Often split graphs are specified in the form $G = (C, I, E)$.

In this paper, we will resolve the computational complexity of questions Q1 and Q2 for split graphs: The following two theorems show that for this graph class Q1 is hard, whereas Q2 is easy.

Theorem 1. *It is NP-hard to decide whether a given split graph with n vertices is α-decomposable for a given partition α of n.*

Theorem 2. *It can be decided in polynomial time whether a given split graph with n vertices is fully decomposable.*

Theorem 1 will be proved in the following Section 3, and Theorem 2 will be proved in the remaining part of this paper.

3 The hardness proof

In this section we will prove Theorem 1. The reduction is done from the following version of the NP-hard d-DIMENSIONAL MATCHING problem; see Garey & Johnson [4].

PROBLEM: d-DIMENSIONAL MATCHING (d-DM)

INPUT: A ground set $X = \{x_1, \ldots, x_{qd}\}$ of qd elements; a family $\mathcal{S}$ of d-element subsets $S_1, \ldots, S_\ell$ of X.

QUESTION: Can set X be partitioned into q disjoint subsets from $S_1, \ldots, S_\ell$?

For an instance of this problem d-DM, we now construct the following corresponding split graph.

- For every element $x \in X$, the independent set I contains a corresponding vertex $i(x)$. Furthermore, in the independent set I there are $\ell - q$ groups $D_1, \ldots, D_{\ell-q}$ of dummy vertices; every such group consists of exactly $d-1$ vertices.
- For every set S in the family $\mathcal{S}$, the clique C contains a corresponding vertex $c(S)$. Furthermore, the clique C contains $\ell - q$ dummy vertices $c_1, \ldots, c_{\ell-q}$.
- There is an edge between any two vertices in the clique C.
- Whenever $x \in S$ for some $x \in X$ and some $S \in \mathcal{S}$, there is an edge between $i(x)$ and $c(S)$.
- Furthermore, for $k = 1, \ldots, \ell - q$ the dummy vertex c_k is joined to the $d-1$ dummy vertices in the group D_k.

The resulting split graph G has $2\ell - q$ vertices in C, has $(d-1)\ell + q$ vertices in I, and thus consists altogether of $(d+1)\ell$ vertices. Finally, we define the vector $\alpha = (d+1, d+1, \ldots, d+1)$ that consists of ℓ components of value $d+1$. We claim that the split graph G is α-decomposable, if and only if the instance of d-DM has answer YES.

First assume that the instance of d-DM has answer YES. Consider the partition of X into q subsets from $\mathcal{S}$. For every set S occurring in this partition, we put vertex $c(S)$ together with all vertices $i(x)$ with $x \in S$ into one connected component. For every set S not occurring in this partition, we put vertex $c(S)$ together with one of the dummy vertices c_k and the vertices in the group D_k into one connected component. This yields that G is α-decomposable.

Next assume that the graph G is α-decomposable. Every dummy vertex c_k must be in the same connected component with the vertices in group D_k, and with exactly one of the vertices $c(S)$. This leaves q of the vertices $c(S)$ unmatched, and each of them must be in one connected components with d vertices $i(x)$ of the independent set. This yields the desired partition of the set X.

This completes the NP-hardness argument, and the proof of Theorem 1. Since d-DM is NP-hard for every fixed $d \geq 3$, we have actually established the following stronger statement.

Corollary 1. *For every fixed integer $f \geq 4$, it is NP-hard to decide whether a given split graph on qf vertices is α-decomposable with respect to the vector $\alpha = (f, f, \ldots, f)$ consisting of q components of value f.* □

The following sections will show that the statement in Corollary 1 is essentially strongest possible: For $f \leq 3$, the corresponding decomposition problem allows a polynomial time solution.

4 Primitive partitions

For $n \geq 2$, a partition α of n is called *2-3-primitive*, if it is of one of the following forms.

- $\alpha = (1, 3, 3, \ldots, 3)$ consists of threes and a single one;
- $\alpha = (2, \ldots, 2, 3, 3, \ldots, 3)$ only consists of twos and threes.

The following lemma shows that for analyzing the full decomposability of a split graph, we can restrict our attention to 2-3-primitive partitions.

Lemma 1. *A split graph on n vertices is fully decomposable, if and only if it is α-decomposable for every 2-3-primitive partition α of n.*

Proof. The only-if-statement is implicit in the definition of a fully decomposable graph. For the if-statement, we recall that every integer $\ell \geq 2$ can be written in the form $\ell = 2a + 3b$ with non-negative integers a and b. Consider an arbitrary partition $\alpha = (\alpha_1, \ldots, \alpha_k)$ of n. Replace every $\alpha_i \geq 2$ in α by a partition of α_i into a_i twos and b_i threes. Let α_0 denote the number of 1s in the vector α. If $\alpha_0 \geq 2$, then replace the 1s in vector α by a partition of α_0 into a_0 twos and b_0 threes. If $\alpha_0 \leq 1$, then leave the 1s untouched. The resulting new partition $\alpha' = (\alpha'_1, \ldots, \alpha'_m)$ of n is of the form $(1, 3, 3, \ldots, 3)$ or $(2, \ldots, 2, 3, 3, \ldots, 3)$, and hence 2-3-primitive. By assumption the split graph G is α'-decomposable. We let $A'_1, \ldots, A'_m$ denote the corresponding connected vertex sets. Every set A'_j with $\alpha'_j = |A'_j| \geq 2$ contains at least one clique-vertex; therefore, the union of the a_i two-element sets and the b_i three-element sets corresponding to component α_i is a connected vertex set A_i with α_i elements. This yields that G is α-decomposable. □

We note that Lemma 1 already implies an NP-certificate for deciding whether an n-vertex split graph is fully decomposable: The certificate lists all 2-3-primitive partitions of n together with the corresponding decompositions into connected parts. The following sections prove even stronger results.

5 Canonical primitive partitions

Next let us introduce *canonical primitive partitions* as a crucial subfamily of the 2-3-primitive partitions. Let $n \geq 2$ be an integer.

- If $n = 2k$ is even, then the canonical 2-primitive partition of n consists of k twos.
  If $n = 2k + 1$ is odd, then the canonical 2-primitive partition of n consists of $k - 1$ twos and a single three.
- If $n = 3k$, then the canonical 3-primitive partition of n consists of k threes.
  If $n = 3k+1$, then the canonical 3-primitive partition of n consists of k threes and a single one.
  If $n = 3k+2$, then the canonical 3-primitive partition of n consists of k threes and a single two.

The following lemma strengthens the statement of Lemma 1.

Lemma 2. *A split graph with n vertices is fully decomposable, if and only if it is α-decomposable for the canonical 2-primitive partition α of n and for the canonical 3-primitive partition α of n.*

The rest of this section is dedicated to the proof of Lemma 2. We first introduce some additional notation and terminology.

We use 2^r3^s to denote a partition of $n = 2r + 3s$ into r (possibly $r = 0$) twos and s (possibly $s = 0$) threes. A partition of $n = 3k + 1$ into k threes and 1 one is denoted by 13^k.

Suppose $G = (C, I, E)$ is a split graph and H is a subgraph of G. Then a vertex of $V(H) \cap C$ or $V(H) \cap I$ is called a *C-vertex* or *I-vertex* of H, respectively. Analogously, we call a neighbor u of a vertex $v \in V(G)$ a *C-neighbor* or *I-neighbor* of v if $u \in C$ or $u \in I$, respectively. If $|V(H)| = 3$, we say that H is a T_i^c if $|V(H) \cap C| = c$ and $|V(H) \cap I| = i$; in the special case that $c = 2$ and $i = 1$ we add a bar (only) if T_1^2 is a triangle, so we use $\overline{T}_1^2$ instead of T_1^2 if and only if the three vertices induce a triangle in G.

For proving Lemma 2 it is sufficient to prove the following result.

Lemma 3. *If a split graph G with n vertices is α-decomposable for the canonical 2-primitive partition α of n and for the canonical 3-primitive partition α of n, then G is α-decomposable for every 2-3-primitive partition α of n.*

Proof. Let $G = (C, I, E)$ be a split graph on n vertices, and assume that G is α-decomposable for the canonical 2-primitive partition α of n and for the canonical 3-primitive partition α of n. First note that we may assume that $n \geq 10$; if $n < 10$ then the only possible 2-3-primitive partitions are the canonical 2-primitive and the canonical 3-primitive partitions. Secondly, note that G has a matching saturating at least $|I| - 1$ vertices of I (and all vertices of I if n is even); since I is an independent set, this follows immediately from the hypothesis that G is α-decomposable for the canonical 2-primitive partition α of n. This also implies that $|C| \geq |I| - 1$.

Definition 1. *We say that G is* $(3, 3)$-reducible *if and only if it has the following property: If G is $2^r 3^s$-decomposable for some $r \geq 0$ and $s \geq 4$, then it is also $2^{r+3} 3^{s-2}$-decomposable.*

Similarly, we say that G is $(1, 3)$-reducible *if and only if G has the following property: If G is 13^k-decomposable for some $k \geq 3$, then it is also $2^2 3^{k-1}$-decomposable.*

Note that in the language of this definition, it is now sufficient to prove that G is both $(3, 3)$-reducible and $(1, 3)$-reducible. The following two claims establish these facts, and thus complete the proof of Lemma 3.

Claim. G is $(3, 3)$-reducible.

Proof. Suppose G has a $2^r 3^s$-decomposition α with $r \geq 0$ and $s \geq 4$. Then at least two of the 3-components in α have at least two C-vertices, since $|C| \geq |I| - 1$. It is obvious how to decompose the subgraph of G induced by the six vertices of two such 3-components into three 2-components.

Claim. G is $(1, 3)$-reducible.

Proof. Suppose G has a 13^k-decomposition α with $k \geq 3$. Then at least one of the 3-components in α has at least two C-vertices, since $|C| \geq |I| - 1$. Let H denote such a 3-component, and let v denote the vertex of the 1-component in α.

If $v \in C$ it is clear how to decompose the subgraph of G induced by $V(H) \cup \{v\}$ into two 2-components.

Next suppose $v \in I$. Clearly, v is not an isolated vertex since G is α-decomposable for the canonical 2-primitive partition α of n. Let u be a

C-neighbor of v. If u is in a $\overline{T}_1^2$ or T_0^3 of α, or if it is the vertex with degree 1 in a T_1^2 of α, then it is again clear how to decompose the subgraph of G induced by v and the vertices of the 3-component containing u into two 2-components.

If u is the vertex with degree 2 in a T_1^2 of α, we use that α contains at least one other 3-component H' with at least two C-vertices, since $|C| \geq |I| - 1$ and $v \in I$. In this case we can combine v with u and its I-neighbor in T_1^2 into a 3-component, and we can decompose the subgraph of G induced by the remaining vertex of this T_1^2 and the vertices of H' into two 2-components.

A similar transformation along a longer chain of 3-components can be used in the remaining case where u is the C-vertex of a T_2^1. In this case the existence of a matching that saturates at least $|I| - 1$ vertices of I implies there is an alternating path $P = v_1v_2 \dots v_{2t}$ starting at $v = v_1$ and terminating at a vertex $w = v_{2t}$ in a $\overline{T}_1^2$, T_1^2 or T_0^3, in which each v_{2j} with $1 \leq j < t$ is the C-vertex of a T_2^1 and each v_{2j+1} with $1 \leq j < t$ is an I-vertex adjacent to v_{2j} in the corresponding T_2^1 for $t - 1$ disjoint 3-components isomorphic to T_2^1. The chain of these $t - 1$ copies of a T_2^1 without the vertex v_{2t-1} together with the vertex v and the edges $v_{2j-1}v_{2j}$ with $1 \leq j < t$ can be transformed into $t - 1$ new T_2^1s by swapping the edges of P (meaning that we include all edges $v_{2j-1}v_{2j}$ with $1 \leq j < t$ and remove all edges $v_{2j}v_{2j+1}$ with $1 \leq j < t$). The remaining vertex v_{2t-1} and the 3-component H_w containing w can be treated as before, yielding a decomposition of the subgraph of G induced by $V(H_w) \cup \{v_{2t-1}\}$ into two 2-components in case w is not the vertex with degree 2 in a T_1^2; otherwise we use again that α contains at least one other 3-component H' with at least two C-vertices. In this case we can combine v_{2t-1} with w and its I-neighbor in T_1^2 into a 3-component, and we can decompose the subgraph of G induced by the remaining vertex of this T_1^2 and the vertices of H' into two 2-components.

6 The polynomial time result

Lemma 4. *Let $G = (V, E)$ be a split graph on n vertices, and let α be the canonical 2-primitive partition of n. Then it can be decided in polynomial time whether G is α-decomposable.*

Proof. This boils down to a bipartite matching problem. If n is even, we need to find a matching from the independent set I into the clique C. If n

is odd, then we check all possibilities for the extra component with three vertices. □

Lemma 5. *Let $G = (V, E)$ be a split graph on n vertices, and let α be the canonical 3-primitive partition of n. Then it can be decided in polynomial time whether G is α-decomposable.*

Proof. Our main tool is the following result from matching theory; see for instance Lovász & Plummer [7]: Let $G' = (V', E')$ be an edge-weighted graph, and for every vertex $v \in V'$ let $d(v)$ be a non-negative integer. Then we can determine in polynomial time a maximum-weight subset $F' \subseteq E'$ of the edges, such that in the graph (V', F') every vertex v has degree $d(v)$, or find out that no such set F' exists.

Consider a split graph $G = (C, I, E)$. We only discuss the case where the number of vertices is of the form $n = 3k$; the other cases can be handled by checking all possibilities for the extra component with one or two vertices. We construct an auxiliary graph $G' = (V', E')$.

- The graph G' contains all vertices in $C \cup I$, together with all edges in E between C and I. All these edges have weight 0.
- For every vertex $v \in C$, the graph G' contains two additional vertices v' and v''. The three vertices v, v', v'' form a triangle. The weight of the edge $v'v''$ is 1, and the weight of the other two edges is 0.
- There is a special vertex v^* that is adjacent to all vertices in C. All edges between v^* and C have weight -1.

For $0 \leq s \leq k$, we define an instance G'_s of the above matching problem. The underlying edge-weighted graph is G', and the values $d(v)$ are defined as follows.

- For every vertex $v \in I$, we set $d(v) = 1$. For every vertex $v \in C$, we set $d(v) = 2$ and $d(v') = d(v'') = 1$. Finally, we set $d(v^*) = s$.

We claim that the considered split graph $G = (C, I, E)$ is α-decomposable for $\alpha = (3, 3, \ldots, 3)$ if and only if at least one of these graphs G'_s (with $0 \leq s \leq k$) possesses a subgraph (V', F') that satisfies all degree constraints and that has $w(F') \leq |C| - 2s$.

First, consider a subset $F' \subseteq E'$ of the edges in some graph G'_s such that in (V', F') every vertex v has degree $d(v)$, and such that $w(F') \leq |C| - 2s$. If an edge $v'v''$ is in F', then the corresponding vertex $v \in C$ must have one or two I-neighbors. If the edge $v'v''$ is not in F', then the edges vv' and vv'' are both in F' and vertex $v \in C$ has no I-neighbors. Denote the sets of vertices $v \in C$ that have zero, one, two I-neighbors, respectively by C_0, C_1, C_2. Note that $|C| = |C_0|+|C_1|+|C_2|$, that $|C_1| = s$, and that the total weight $w(F')$ of the edge set F' equals $(|C_1|+|C_2|)-|C_1| = |C_2|$. The condition $w(F') \leq |C|-2s$ can be equivalently written as $|C_0| \geq |C_1|$. We group every vertex in C_1 together with its I-neighbor in F' and together with an arbitrary vertex from C_0 into a connected triple. Furthermore, we group every vertex in C_2 with its two I-neighbors into a triple, and finally we group the remaining unused vertices in C_0 into triples. The resulting triples form an α-decomposition of the split graph G for $\alpha = (3, 3, \ldots, 3)$.

Next, assume that the split graph is α-decomposable where $\alpha = (3, 3, \ldots, 3)$ is the canonical 3-primitive partition of n. The triples in this decomposition can be classified into three types: T_2^1-triples have one C-vertex and two I-vertices; we mark the corresponding two edges between C and I. T_1^2 or $\overline{T}_1^2$-triples have two C-vertices and one I-vertex; we mark one corresponding edge between C and I (if the I-vertex is adjacent to both C-vertices, then choose the marked edge arbitrarily). T_0^3-triples have three C-vertices; we mark no edges for them. Let x, y, z, respectively denote the number of triples of these three types. Note that $x+2y+3z = |C|$. If a vertex $v \in C$ is incident to one or two marked edges, then we also mark the edge $v'v''$. If a vertex $v \in C$ is not incident to any marked edges, then we also mark the two edges vv' and vv''. Finally, if a vertex $v \in C$ is incident to exactly one marked edge, then we also mark the edge vv^*. It can be verified that for the set F' of marked edges, the subgraph (V', F') satisfies all degree constraints in the graph G_y. The total weight of F' equals $w(F') = x + y - y = x \leq |C| - 2y$, as desired. □

7 Conclusions

We have settled the complexity of recognizing fully decomposable split graphs. We feel that it might be very difficult to come up with other graph classes for which this problem is tractable. The algorithm of Barth, Baudon & Puech [1] for recognizing fully decomposable tripodes (trees with a single vertex of degree three, and all other vertices of degree one or two) is highly non-trivial. Unfortunately, many other graph classes contain graphs with a similar connectivity structure as tripodes (with respect to full decomposability); hence settling the problem for these classes would amount to generalizing the proof of [1].

Let us illustrate this claim for the class of co-graphs. Consider a tripode T that consists of a root and three paths with ℓ_1, ℓ_2, and ℓ_3 vertices, respectively. We define a corresponding co-graph $G(T)$ that consists of three independent cliques with ℓ_1, ℓ_2, and ℓ_3 vertices, and a single vertex that is connected to all vertices in the cliques. It can be seen that the tripode T is fully decomposable if and only if the co-graph $G(T)$ is fully decomposable. We pose the computational complexity of recognizing fully decomposable co-graphs as an open problem.

Furthermore, we are not aware of any natural NP-certificates or coNP-certificates for deciding full decomposability of general graphs. In fact, this problem might be located in one of the complexity classes above NP (see for instance Chapter 17 in Papadimitriou's book [8]). If the problem is hard, then the complexity class DP=BH_2, the second level of the Boolean Hierarchy, might perhaps be a reasonable guess.

Finally, we will formulate a conjecture that would imply that the problem is easy. Let us call a vector α with positive integer components *balanced*, if $k-1$ of these components are equal to each other, and the last component does not exceed the other components. We did not manage to construct a counter-example to the following bold conjecture.

Conjecture 1. An n-vertex graph G is fully decomposable, if and only if G is α-decomposable for every balanced vector α whose components add up to n.

If this conjecture turns out to be true (for which admittedly we do not have the slightest evidence), then this would yield an NP-certificate for fully decomposable graphs: There are only $O(n)$ many balanced vectors α whose components add up to the number n of vertices in a graph. The α-decompositions for these $O(n)$ vectors form a certificate of polynomial length that can easily be verified in polynomial time.

References

1. D. Barth, O. Baudon, and J. Puech (2002). Decomposable trees: A polynomial algorithm for tripodes. *Discrete Applied Mathematics 119*, 205–216.
2. D. Barth and H. Fournier (2006). A degree bound on decomposable trees. *Discrete Mathematics 306*, 469–477.
3. M.E. Dyer and A.M. Frieze (1985). On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics 10*, 139–153.
4. M.R. Garey and D.S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, San Francisco.
5. M.C. Golumbic (1980). *Algorithmic graph theory and perfect graphs.* Academic Press, New York.
6. M. Horňák and M. Woźniak (2003). Arbitrarily vertex decomposable trees are of maximum degree at most six. *Opuscula Mathematica 23*, 49–62.
7. L. Lovász and M.D. Plummer (1986). Matching Theory. *Annals of Discrete Mathematics 29*, North-Holland.
8. C.H. Papadimitriou (1994). *Computational Complexity.* Addison-Wesley.