AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **76(1)** (2020), Pages 24–40

Equitable induced decompositions of twin graphs

Peter Damaschke

Department of Computer Science and Engineering
Chalmers University, 41296 Göteborg
Sweden
ptr@chalmers.se

Abstract

Twins in a graph are vertices that are not adjacent, and which have exactly the same neighbors. In the twin graph of a graph G, every vertex represents an equivalence class of twins in G and is weighted by the size of this class. Motivated by (complex) group testing with two defectives and by fair division problems, we are interested in decomposing graphs with few twin classes into k induced subgraphs with nearly equal edge numbers, where every edge belongs to exactly one subgraph. Technically we consider the fractional version of the problem, where the vertices of a weighted twin graph can be split into arbitrary fractions, and the k induced subgraphs must receive exactly the same total edge weights. The results then apply to usual graphs, subject to a small discretization error. We show that such equitable induced decompositions are indeed possible for various twin graphs, including all bipartite graphs, cycles C_n and C_n-colorable graphs, and (C_3, C_5)-free graphs. We also pay attention to the necessary number of vertices (after the splittings) in the induced decompositions. Usually this number is bounded by $k + O(1)$, but for complete bipartite graphs, i.e., when the twin graph is a single edge, roughly $2\sqrt{k}$ vertices suffice, and their exact minimum number is easy to compute for many k.

1 Introduction

Complex group testing [5, 6] is a natural generalization of the well-studied group testing problem and is described as follows. We are given a set of items and an unknown family of nonempty subsets called complexes. We can choose an arbitrary subset, called a pool, and test it. The result of the test is positive if the pool contains at least one complex as a subset, and otherwise negative. The goal is to identify the complexes by a minimum number of tests. We know in advance, for given numbers d and r, that (at most or exactly) d complexes exist, and that each complex has (at most or exactly) r elements.

ISSN: 2202-3518

The case $r = 1$ is the ordinary group testing problem. The case $r = 1$, $d = 1$ is trivial in the sense that it can be optimally solved adaptively by binary search or non-adaptively by binary encoding. The case $r = 2$ with a general d has found special attention, e.g., in [7, 8].

The main motivation for complex group testing is the search for unknown factors that cause a certain effect if and only if all these factors are present, and where the candidate factors can be independently switched on and off in experiments. Examples include gene knockout experiments in biology, testing the properties of mixtures of chemicals, and fault diagnosis of technical installations or software.

In the following we only deal with the case of exactly one complex comprising exactly two items (or vice versa). This case should be quite common, as many phenomena are caused by the coincidence of two different factors, whereas none of the factors alone has an effect. (For instance, only the combination of two specific chemicals may be toxic.) Moreover, this case has a neat connection to graph theory, as explained below.

We assume that the reader is familiar with usual graph-theoretic concepts, like neighbor, clique, independent set, induced subgraph, bipartite graph, etc. A twin class is a maximal set I of vertices such that I is an independent set and all vertices in I have the same neighborhood. Let n and m denote the number of vertices and edges, respectively.

For $d = 1$ and $r = 2$ we may think of the items and candidate pairs as vertices and edges, respectively, of an undirected graph. When a pool is tested positively, the remaining graph is the subgraph induced by this pool. When a pool is tested negatively, the remaining graph is obtained by deleting the edges, but not the vertices, of this pool. Usual group testing with two defectives ($d = 2$ and $r = 1$) is equivalent: just take the complements of the pools.

If initially all pairs are candidates, the graph is a clique. Negative tests cut out the edge sets of some smaller cliques. Moreover, as soon as the negatively tested pools cover every vertex at least once, the vertex set of the remaining graph is partitioned into twin classes whose number depends only on the number of pools but not on the graph size. Since positive tests leave induced subgraphs, the statements remain true if we encounter some positive tests, too.

The candidate graph can naturally consist of a few twin classes also right from the beginning: The candidate factors may belong to different classes (e.g., genes with different roles, chemicals of different types), and we may know in advance that the two sought items can only belong to certain pairs of distinct classes. Altogether, graphs of arbitrary size but with a limited number of twin classes are of special interest in our context.

Next, also the pool sizes may be restricted for practical reasons (e.g., one cannot inject too many faulty conditions in a system simultaneously, or mix too many different chemicals for a test). This imposes limits on the number of vertices or edges in each pool. In the following we briefly discuss a situation where, for some factor $\beta < 1$, at most βm edges are allowed in each pool. We use a known fact which is Proposition 1.25 in [1]:

Proposition 1.1. *For adaptive group testing with exactly one positive element ($d = 1$ and $r = 1$) and pools of size at most s, the following strategy minimizes the worst-case number of tests: First test pairwise disjoint sets with s elements until a positive result is obtained or less than $2s$ elements are left, then continue with bisection search.*

Suppose for a moment that we know a family of induced subgraphs, each with at most βm edges, that together cover all edges of a given graph. Then, Proposition 1.1 applied to the edge set yields that the following strategy minimizes the worst-case number of tests, up to an additive error of at most one test: First test a minimum number of the mentioned induced subgraphs. As soon as a positive test is found, the size limitation is trivially satisfied in the remaining graph. Henceforth do bisection search, that is, repeatedly test an induced subgraph with roughly half of the remaining edges. Bisection search is always possible, in the sense that there exists an efficient search strategy that misses the information-theoretic lower bound (binary logarithm of the number of edges) by at most one test; see [12] for an elegant proof of a more general result.

If $1/\beta$ happens to be an integer, an ideal covering would be a partitioning of the edge set into $1/\beta$ induced subgraphs with pairwise disjoint edge sets of equal, or nearly equal, sizes. In general graphs we cannot expect such partitionings to exist, and even if they do, their construction would involve the NP-hard problem of finding induced subgraphs with prescribed numbers of edges. However, in this paper we will show such constructions at least for certain graphs. These cases may appear very special, but they are twin graphs that arise from negative initial tests as described above.

The more general case when $1/\beta$ is fractional is less clear. We may still take the next smaller inverse of an integer. Some limited overlaps and size differences of the tested edge sets would not always increase the worst-case number of tests. If, regardless, we can efficiently construct equitable families of disjoint induced subgraphs, we might as well use them, thus performing the optimal number of tests in the first phase when the graph is larger then the allowed pool size. However, we will not prove formal results on that matter, and we must leave it to future research.

Finally, we have arrived at the following problem: *Partition a graph into induced subgraphs with pairwise disjoint edge sets of nearly equal sizes.*

Trivially, the same problem appears if complexes may be present or not, and we only want to check the absence of complexes by group tests where the number of edges in each pool is bounded. Moreover, the problem appears fascinating in its own right, and it may also find other applications. In general, problems of dividing objects equitably, where the parts must obey certain restrictions, have various applications in management and administration. Here we cannot survey the whole field, and we refer to some selected papers [2, 4, 9, 10].

2 Fractional Induced Decompositions

The notion of induced decompositions is already well known in graph theory [13, 3]. An *induced decomposition* of a graph G consists of induced subgraphs $H_1, \ldots, H_k$ of

G such that every edge of G belongs to exactly one subgraph H_i. In an induced H-decomposition, all H_i are isomorphic to some fixed graph H. But remember that the subject of the present paper is different: We are interested in induced decompositions where the graphs H_i are not necessarily isomorphic but have the same number of edges, as far as possible.

Two vertices u and v in a graph G are *twins* (sometimes called "false twins" in other literature), if uv is not an edge, and u and v are adjacent to exactly the same set of other vertices. The twin relation is an equivalence relation whose equivalence classes are the *twin classes* of G.

The *twin graph* $T(G)$ of G is defined as the graph whose vertices are the twin classes of G, and where two vertices are adjacent if and only if some (and hence all) vertices in the two corresponding twin classes are adjacent, i.e., if the two twin classes induce a complete bipartite graph.

Recall that we consider large graphs with a few twin classes. We can succinctly represent any such graph G by its twin graph $T(G)$, where every vertex of $T(G)$ has a weight indicating the number of vertices in the corresponding twin class in G. We also give every edge of $T(G)$ a weight indicating the number of edges between the corresponding twin class in G.

We abstract away the size of G and keep only $T(G)$ and the proportions of twin class sizes. That is, we multiply all vertex weights and all edge weights with some scaling factor σ and σ^2, respectively, where σ is some free parameter. Note that the weight of every edge is the product of the weights of its end vertices. We call every weighted graph with this property a *weighted twin graph.* In particular, we may choose $\sigma = 1/\sqrt{m}$, such that the sum of edge weights is normalized to 1.

Next we introduce a fractional analogue to the notion of induced subgraphs, for these weighted twin graphs T. From every vertex, say with weight w, we take any "portion" of weight w', where $0 \leq w' \leq w$. We connect the vertices, with their chosen weights w', by the edges inherited from T to a graph T'. If T' contains twins (which can happen if some weights w' are 0), we may merge any twins to single vertices and sum up their weights. The edge weights in T' are again set to be the products of the vertex weights. We call every weighted twin graph T' obtained from T in this way an induced subgraph of T.

Obviously, any induced subgraph of G corresponds to an induced subgraph of $T = T(G)$, but the converse does not hold, as we cannot split vertices of G. However, for any fixed T and for large enough graphs G whose weighted twin graph is T (or approximately T) we can approximately realize any induced subgraph of T as an induced subgraph of G, by rounding fractional vertex numbers to integers. We do not further elaborate on this point in detail, however it is obvious that the relative error vanishes as G grows.

The removal of the edges of an induced subgraph T' from T is slightly more difficult to describe. We have to replace every vertex of T with two weighted copies: the portion belonging to and not belonging to T'. In an obvious way, every edge is replaced with four edges between the $2 + 2$ copies of its end vertices, and one of the four edges is deleted. In the worst case this can double the number of vertices (i.e., twin classes).

More generally, a family of k induced subgraphs $T_1, \ldots, T_k$ of T can be canonically represented as follows. Every vertex v of T is split into at most 2^k vertices v_I, one for each subset $I \subseteq \{1, \ldots, k\}$. The weight of every vertex v_I is the portion of v that belongs to the subgraphs T_i with indices $i \in I$, and does not belong to any subgraph T_i with an index $i \notin I$. Every edge uv is replaced with all possible edges $u_I v_J$, and the weight of every edge is again the product of the weights of its two end vertices. We keep only the vertices and edges with positive weights; note that many weights can be zero. We also refer to the resulting graph as the *generalized twin graph* of the family.

The family of induced subgraphs $T_1, \ldots, T_k$ is an induced decomposition of T if and only if $|I \cap J| = 1$ holds for every edge $u_I v_J$. Translated back to graphs G with (weighted) twin graph T, this condition just says that every edge of G belongs to exactly one of the induced subgraphs represented by the T_i. Note carefully that several T_i may share the same vertices.

Now we can formulate our main goal as follows: Given a weighted twin graph and an integer k, construct a family $\mathcal{F}$ of k induced subgraphs such that $\mathcal{F}$ is an induced decomposition, the total edge weights of the k induced subgraphs in $\mathcal{F}$ are equal, and the number of vertices in the generalized twin graph of $\mathcal{F}$ is kept to a minimum.

The latter condition is important for the approximate realization of the decompositions of T in actual graphs G with $T(G) = T$, since fractional numbers of vertices in G must be rounded to integers. The more vertices we produce in the twin graph, the more rounding operations we need. More informally, a decomposition with fewer vertices is just "simpler". Similar conditions appear in other equitable division problems where the fragmentation of the pieces, e.g., the number of cuts, shall be kept to a minimum.

In the following sections we study equitable fractional decompositions of various twin graphs. Let K_n and C_n denote the clique and the cycle, respectively, of n vertices.

3 The Twin Graph K_2

Consider $T = K_2$, which is the twin graph of complete bipartite graphs. It is trivial to construct an induced decomposition into k induced subgraphs with equal total edge weights, just by splitting one of the two vertices into k vertices with equal vertex weights. However, this creates a twin graph with $k + 1$ vertices. On second thought, $O(\sqrt{k})$ vertices should be enough, through a careful division of both initial vertices. Below we study the number of vertices needed. In fact we will show an upper bound $2\sqrt{k}(1 + o(1))$.

We can think of a weighted K_2 as a rectangle, where the side lengths and the area are the weights of the two vertices and of the edge, respectively. Any partitioning of the two vertices into, respectively, g and h smaller weighted vertices naturally defines a $g \times h$ grid of gh rectangles that we call *cells*. (A $g \times h$ grid is a grid with g rows and h columns.) Furthermore, any induced decomposition into k induced subgraphs is described by assigning a mark from $\{1, \dots, k\}$ to each cell, in such a way that, for every i, the union of all cells with mark i is a sub-grid. Here, the term *sub-grid* means a set of cells that can be turned into a rectangle by permutations of the rows and columns, respectively, of the entire grid. In particular, a usual rectangle of cells is also a sub-grid.

Definition 3.1. *For a given k we define $r := \lfloor \sqrt{k} \rfloor$, and we define $d := 2r$ if $r^2+1 \le k \le r^2+r$, and $d := 2r+1$ if $r^2+r+1 \le k \le r^2+2r+1$. If k is not clear from context, we may write $r(k)$ and $d(k)$.*

Note that $d+1 = \min\{g+h \mid gh \ge k\}$. It follows that every induced decomposition into k induced subgraphs needs at least $d+1$ vertices. Below we will give two contributions: we derive an upper bound on the number of vertices, and we characterize the numbers k for which $d+1$ vertices are sufficient.

Theorem 3.2. *For every k there exists an induced decomposition of the weighted twin graph K_2 (given with arbitrary vertex weights) into k induced subgraphs with equal total edge weights and with at most $d + \sqrt{d/2} + \log_2 d + O(1)$ vertices.*

Proof. First we describe the principal shape of our construction. We fix some number g of rows. For various divisors f of g we create columns for f marks, where always g/f consecutive cells get the same mark. The total number of marks must be k. Since, within every column, marks are assigned to the same number of cells, it is trivial to finally choose the breadths of the columns such that all k rectangles (unions of cells with the same mark) occupy equal areas.

We are left with a purely number-theoretic task: Write k as a sum of a minimum number h of divisors of g. (Note that a divisor may in general appear multiple times in this sum.) Finally we choose g so as to minimize $g+h$. Here we do not aim for the exact minimum $g+h$, but only for a simple general construction that yields some good upper bound.

The idea is to choose g both close to $\sqrt{k}$ and with a large number of prime factors 2. We work out the details now. Let e be the largest integer with $2^e \le \sqrt[4]{k}$, and let p be the largest integer with $g := 2^e \cdot p \le \sqrt{k}$ (which also defines our g). By maximality of p we have $\sqrt{k} - g < 2^e \le \sqrt[4]{k}$, thus $g \ge \sqrt{k} - \sqrt[4]{k}$.

First we generate a maximum number of columns with g marks, that is, we stop only when another column would raise the total number of marks above k. The number of such columns is at most $k/g \le k/(\sqrt{k} - \sqrt[4]{k}) = k(\sqrt{k} + \sqrt[4]{k})/(k - \sqrt{k}) = (\sqrt{k} + \sqrt[4]{k})/(1 - \sqrt{1/k}) = (\sqrt{k} + \sqrt[4]{k})(1 + O(\sqrt{1/k})) = \sqrt{k} + \sqrt[4]{k} + O(1)$.

There remain fewer than $g \le \sqrt{k}$ marks to assign. Since g has all divisors $g/2^i$ for $i \le e$, we can create a new column for at least half of the remaining marks, and continue doing so until fewer than p marks are left. (Note that the last such divisor is $g/2^e = p$.) Obviously, the number of these columns is at most e.

As said above, there remain fewer than p marks to assign. We observe that $2^{e+1} > \sqrt[4]{k}$ due to the maximality of e. Hence $p = g/2^e \le 2\sqrt{k}/2^{e+1} < 2\sqrt{k}/\sqrt[4]{k} = 2\sqrt[4]{k} \le 4 \cdot 2^e$. Since g has all divisors 2^i for $i \le e$, we conclude that at most $e+3$

further columns suffice to host all remaining marks, using the binary expansion of their number.

Since $e \leq \log_2 \sqrt[4]{k} = (1/4)\log_2 k$, the total number of columns needed is no larger than $\sqrt{k} + \sqrt[4]{k} + (1/2)\log_2 k + O(1)$. Together with the $g \leq \sqrt{k}$ rows these are at most $2\sqrt{k} + \sqrt[4]{k} + (1/2)\log_2 k + O(1)$ vertices. We plug in the definition of d to obtain the assertion. $\square$

Note that our upper and lower bound differ only by a factor $1 + O(1/\sqrt{d})$. An intriguing question is whether the dominating $\sqrt{d/2}$ term can be improved. Other numbers close to $\sqrt{k}$ may have even "nicer" divisors than in our construction based on powers of 2, but it seems to be a more intricate number-theoretic matter to derive a better bound.

In the second part of this section we follow a different route and characterize the numbers k for which $d(k) + 1$ vertices are sufficient. (Remember that this was the absolute minimum.) We call a cell *lonely* if no other cell has the same mark, that is, the cell is one of the k sub-grids of our induced decomposition.

Lemma 3.3. *In every induced decomposition of K_2 into induced subgraphs with equal total edge weights, the lonely cells form a sub-grid.*

Proof. We denote by $x(c)$ and $y(c)$ the horizontal and vertical side length, respectively, of a cell c. Note that all lonely cells must have equal areas, and the areas of other cells must be strictly smaller.

If two lonely cells a and b are in the same row, then trivially $y(a) = y(b)$, and since $x(a) \cdot y(a) = x(b) \cdot y(b)$, it also follows $x(a) = x(b)$. Similarly, any two cells in the same column are congruent as well. Now let a and b be lonely cells not being in the same row or column, and let c and e be the cells that form together with a and b a sub-grid, as displayed here:

$$\boxed{\begin{matrix} a & c \\ e & b \end{matrix}}$$

Since $x(a) \cdot y(a) \geq x(c) \cdot y(c) = x(b) \cdot y(a)$, we conclude $x(a) \geq x(b)$. By symmetry, and by involving the cell e, we also get the other inequalities $x(a) \leq x(b)$, $y(a) \geq y(b)$, $y(a) \leq y(b)$. Altogether this shows that all lonely cells are congruent.

Consider the sub-grid spanned by the lonely cells; it consists of all cells c such that both the row of c and the column of c contains some lonely cell. Since all lonely cells are congruent, it follows that all rows in this sub-grid have the same breadth, and so have the columns. Hence all cells in this sub-grid are congruent, which further implies that they are lonely. In other words, the lonely cells form a sub-grid. $\square$

Now we can show that induced decompositions with $d(k) + 1$ vertices necessarily have a rather strict structure.

Theorem 3.4. *Every induced decomposition of K_2 into k induced subgraphs with equal total edge weights and $d(k) + 1$ vertices is described by some $g \times h$ grid where all cells are lonely, except at most one row or column with f rectangles of h/f or g/f cells, where f is some divisor of h or g, respectively.*

Proof. Let $r^2+1 \le k \le r^2+r$. Consider any $g \times h$ grid with $g+h = d+1 = 2r+1$. Note that $gh \le r(r+1) = r^2+r$. The lonely cells form some $g' \times h'$ sub-grid. Assume that $g'+h' \le d-1 = 2r-1$. Then $g'h' \le (r-1)r = r^2-r$. Trivially, the number of rectangles beyond the lonely cells is at most half the number of further cells. This yields $k \le g'h' + (gh - g'h')/2 = (gh+g'h')/2 \le ((r^2+r)+(r^2-r))/2 = r^2$, which contradicts $r^2+1 \le k$. Thus $g'+h' \ge d = 2r$.

The argument in the other case is almost literally the same, but for easier reading we repeat it here, with the modifications:

Let $r^2+r+1 \le k \le r^2+2r+1$. Consider any $g \times h$ grid with $g+h = d+1 = 2r+2$. Note that $gh \le (r+1)^2 = r^2+2r+1$. The lonely cells form some $g' \times h'$ sub-grid. Assume that $g'+h' \le d-1 = 2r$. Then $g'h' \le r^2$. Trivially, the number of rectangles beyond the lonely cells is at most half the number of further cells. This yields $k \le g'h' + (gh - g'h')/2 = (gh+g'h')/2 \le ((r^2+2r+1)+r^2)/2 = r^2+r+1/2$, which contradicts $r^2+r+1 \le k$. Thus $g'+h' \ge d = 2r+1$.

This shows that, in a $g \times h$ grid with $d+1$ vertices, at most one row or column consists of cells that are not lonely. Since all rectangles must have equal areas, this row (column) must consist of f rectangles with h/f (g/f) cells each, where f is some divisor of h (g). $\square$

Using Theorem 3.3 we can now very easily enumerate all numbers $k = 1, 2, 3, \ldots$ for which induced decompositions with $d(k)+1$ vertices exist: Depending on whether $d(k) = 2r(k)$ or $d(k) = 2r(k)+1$ we start from an $r \times r$ grid or an $r \times (r+1)$ grid, then we decrement the number of rows and increment the number of columns, and we try all divisors to add one row or column to the sub-grid of lonely cells. If we cannot hit exactly the number k of marks, then we know from Theorem 3.3 that $d(k)+1$ vertices are not enough. In this case we try and add another divisor (that is, either two rows or two columns), in order to obtain $d(k)+2$ vertices.

An interesting observation is that, for growing k, the cases $d(k)+1$ and $d(k)+2$ take turns quite unsystematically, but $d(k)+2$ vertices are sufficient for all $k \le 166$. The smallest k for which the question cannot be settled in this way is 167.

Problem A. For which k do we need $d(k)+2$ vertices, $d(k)+3$ vertices, and so on?

4 The Twin Graph C_3

The next larger twin graph is $T = C_3$ ($= K_3$). Unlike the case of K_2, the existence of equitable induced decompositions is no longer obvious. But it is again helpful to look at a suitable geometric model.

Lemma 4.1. *Let us be given a twin graph C_3 with arbitrary positive vertex weights, and 4 positive numbers $x_1, \ldots, x_4$ such that $\sum_{i=1}^{4} x_i$ equals the total edge weight. Then there exists an induced decomposition into 4 induced subgraphs, which are again weighted C_3, with total edge weights $x_1, \ldots, x_4$, and with 6 vertices.*

Proof. We can naturally think of a weighted twin graph C_3 as an axis-parallel cuboid in an x, y, z-coordinate system: The 3 vertices of C_3 are represented by half-open segments on the 3 coordinate axes, whose lengths are proportional to the vertex weights. For instance, if the first vertex has weight w, we represent it by the segment $[0, w)$ on the x-axis. The cuboid is the Cartesian product of the 3 segments. The 3 edges of C_3 are represented by rectangles on the planes with $x = 0$, $y = 0$, and $z = 0$, respectively, which are the Cartesian products of the mentioned segments, i.e., 3 faces of the cuboid that meet in one corner.

For any point (r, s, t) in the cuboid we divide the twin graph as follows. First we divide the space into the 8 axis-parallel octants that have their origins at (r, s, t). More formally, each point (x, y, z) belongs to an octant depending on whether $x < r$, $y < s$, $z < t$ or not. The intersections of the cuboid with these octants define a partitioning into 8 smaller cuboids. We select 4 of them, that pairwise intersect in a line segment. (They would be the black "fields" in a $2 \times 2 \times 2$ "checkerboard", and the other 4 would be the white "fields".)

According to this partitioning we split each vertex of C_3 in 2 vertices whose weights are again proportional to the segment lengths. Each of the 4 chosen cuboids represents an induced subgraph, in the obvious sense. We also claim that every edge of the twin graph, after the splitting, is contained in exactly one of these 4 induced subgraphs. To see that this claim is true, just consider the projections of the 4 selected cuboids to each of the 3 coordinate planes. Next we adjust the weights.

Let $y_1, \ldots, y_4$ be the edge weights of our 4 induced subgraphs, where the indexing is arbitrary. Define $d_i := y_i - x_i$. We will show that the point (r, s, t) can be chosen such that $y_i = x_i$ for $i = 1, \ldots, 4$. Note that $\sum_{i=1}^4 y_i = \sum_{i=1}^4 x_i$, hence $\sum_{i=1}^4 d_i = 0$, and that we want $\sum_{i=1}^4 |d_i| = 0$.

Trivially, we can select any $j \in \{1, \ldots, 4\}$ and move the point (r, s, t) so as to increase or decrease y_j by some desired small amount (without caring how this affects the other y_i). Also observe that moving the point (r, s, t) parallel to some axis, i.e., changing exactly one of the three coordinates, increases two of the variables y_i and decreases the two others. Conversely, for each of the six partitionings of $\{1, \ldots, , 4\}$ in two sets I and J with $|I| = |J| = 2$, there exists such a moving direction that increases the y_i with $i \in I$ and decreases the y_j with $j \in J$. These manipulations enable the following existence proof, which can also be turned into an algorithm to find the said point.

Consider a point (p, q, r) that minimizes $\max_i |d_i|$. Such a point exists, since a continuous function on a compact set attains its minimum. We call every index $j \in \{1, \ldots, 4\}$ with $|d_j| = \max_i |d_i|$ a critical index. Assume that $\max_i |d_i| > 0$.

If exactly one critical index j exists, then we can decrease its $|d_j|$, yet keeping j critical, and thus decrease $\max_i |d_i|$.

Assume that exactly 2 indices j are critical, without loss of generality $j = 1$ and $j = 2$. If both $d_1 > 0$ and $d_2 > 0$, then they cannot be the only critical indices (contrary to the assumption), because of $\sum_{i=1}^4 d_i = 0$. The same reasoning applies if both $d_1 < 0$ and $d_2 < 0$. Thus we can assume $d_1 > 0$ and $d_2 < 0$; the opposite case is symmetric. Now we can decrease d_1 and another difference, say d_3, and increase d_2 and d_4, thus making $\max_i |d_i|$ smaller.

Assume that exactly 3 indices i are critical. Similarly as before, no matter which of the corresponding d_i are assumed to be positive or negative, $\sum_{i=1}^4 d_i = 0$ implies that this case is impossible.

Finally assume that all 4 indices i are critical. Again, due to $\sum_{i=1}^{4} d_i = 0$, two of them must be positive and negative, respectively. Hence we can decrease the former and increase the latter, thereby making $\max_i |d_i|$ smaller.

For every case we have seen that either the case is impossible or $\max_i |d_i|$ is not minimal. Hence the assumption was wrong, and $\max_i |d_i| = 0$ holds true. □

Lemma 4.2. *Every weighted twin graph C_3 with arbitrary positive vertex weights has an induced decomposition into 9 induced subgraphs with equal total edge weights and with 9 vertices.*

Proof. We use again the cuboid and partition the segment on the x-axis into 3 segments X_1, X_2, X_3 of equal lengths. Similarly we define Y_1, Y_2, Y_3 and Z_1, Z_2, Z_3 on the y- and z-axis, respectively. This splits every vertex of C_3 into 3 new vertices of equal weights. We choose the $3 + 6 = 9$ cuboids of the form $X_i \times Y_j \times Z_k$, where either $i = j = k$ or $\{i, j, k\} = \{1, 2, 3\}$. Indeed, every edge belongs to exactly one of the induced subgraphs represented by these cuboids, and their total edge weights are trivially equal. □

Theorem 4.3. *Every weighted twin graph C_3 with arbitrary positive vertex weights possesses, for $k \in \{4, 7, 9, 10, 13\}$ and for every $k \geq 16$, an induced decomposition into k induced subgraphs with equal edge weights and with at most $k + 2$ vertices.*

Proof. First we show the following claim: For $k = 3j+1$ ($j \geq 0$ integer) and arbitrary positive numbers x_i with $\sum_{i=1}^{k} x_i = 1$, there exists an induced decomposition into k induced subgraphs C_3 with total edge weights $x_1, \ldots, x_k$.

We prove the claim by induction on j. The induction base $j = 1$ is stated by Lemma 4.1. Suppose that the claim is true for $j - 1$. We replace the desired edge weights $x_1, \ldots, x_{3j+1}$ with $x_1, \ldots, x_{3(j-1)}, x_{3j-2} + x_{3j-1} + x_{3j} + x_{3j+1}$ and apply the induction hypothesis to these $3(j - 1) + 1$ edge weights. This yields $3(j - 1) + 1$ induced subgraphs which are C_3 and have the indicated edge weights. We apply Lemma 4.1 once more to the last subgraph and split its total edge weight into the desired proportions. This yields $3j + 1$ induced subgraphs with total edge weights $x_1, \ldots, x_{3j+1}$. The number of vertices increases by 3 in the inductive step.

If $k = 3j + 1$ for some integer $j \geq 0$, we choose all $x_i = 1/k$, and we are done. If $k = 3j$ for some integer $j \geq 4$, we choose $x_1 = 9x_2 = \ldots = 9x_{k-8}$ and divide the first subgraph into 9 smaller subgraphs using Lemma 4.2. This last step increases the number of subgraphs by 8 but the number of vertices by only 6. If $k = 3j - 1$ for some integer $j \geq 6$, we choose $x_1 = x_2 = 9x_3 = \ldots = 9x_{k-16}$ and divide the first two subgraphs into $9 + 9$ smaller subgraphs using Lemma 4.2. This last step increases the number of subgraphs by 16 but the number of vertices by only 12. These cases modulo 3 cover all numbers k in the assertion. □

Problem B. Do there exist induced decompositions of the weighted K_3 with equal edge weights also for $k \in \{2, 3, 5, 6, 8, 11, 12, 14, 15\}$?

We conjecture that they do not exist, not even with unequal weights. One can in principle solve this problem by naive exhaustive search, however the number of cases to consider would be prohibitive. It may be interesting to close these remaining gaps without excessive case distinctions.

5 Independent Separators

For any graph $G = (V, E)$ and any subset $U \subseteq V$ let $G[U]$ denote the subgraph induced by U. When we give the set U explicitly by its list of elements $U = \{u_1, u_2, u_3, \ldots, \}$, we omit the curly brackets and write $G[u_1, u_2, u_3, \ldots]$.

A set $N \subset V$ is a *separator* if $G[V \setminus N]$ has at least two connected components. For any such connected component C, we refer to $C \cup N$ or $G[C \cup N]$ as a *wing*.

First notice an equivalent characterization of induced decompositions that follows instantly from the definition: A family of induced subgraphs $H_1, \ldots, H_k$ of a graph G forms an induced decomposition of G if and only if the H_i cover together all edges of G, and the vertex sets of any two graphs H_i intersect in some independent set in G.

Lemma 5.1. *Let $G = (V, E)$ be a graph, and let $N \subset V$ be both an independent set and a separator. Then, any induced decompositions of its wings form together an induced decomposition of G.*

Proof. Consider any two induced subgraphs from the given induced decompositions. If they are in the same wing, they intersect in an independent set, due to the characteristic property of induced decompositions. If they are in different wings, they intersect in an independent set as well, since N was assumed to be an independent set. $\square$

We remark that Lemma 5.1 straightforwardly extends to twin graphs. Next we turn to special twin graphs.

Lemma 5.2. *Consider any bipartite weighted twin graph whose vertex set consists of two independent sets S and N, and let y denote its total edge weight. Let y' and y'' be any positive numbers with $y' + y'' = y$. Then there exists an induced decomposition into two induced subgraphs with total edge weights y' and y'', where at most one vertex of S (and no vertex of N) is split in two vertices.*

Proof. We may go through the vertices of S in an arbitrary order, until the total weight of the edges incident to them reaches y'. At this moment we split the current vertex $u \in S$ in two vertices u' and u'', which have the same neighbors in N as u had. We transfer some portion of the weight of u to u', such that the total weight of the edges incident to u and all previous vertices of S is exactly y'. We transfer the remaining weight of u to u''. Let S' to be the vertex set containing u' and all previous vertices of S, and let S'' be the vertex set containing u'' and all subsequent vertices of S. The two subgraphs induced by $N \cup S'$ and $N \cup S''$ cover together all edges, they intersect at most in the independent set N, and they have total edge weights y' and y'', respectively. $\square$

Our use of Lemmas 5.1 and 5.2 is to construct an induced decomposition of a weighted twin graph T in a recursive fashion: We take an independent set S whose set N of neighbors is independent, too. Note, in particular, that N is a separator, or $S \cup N$ is the entire vertex set. Due to Lemma 5.2 we can take away an arbitrary portion of the total edge weight of $T[S \cup N]$, by cutting off some subset S' of S and all incident edges, where at most one vertex of S is split before this removal. Then we put aside $T[S' \cup N]$, and we recursively construct an induced decomposition of the remaining weighted twin graph (which still contains N as well). Lemma 5.1 guarantees the correctness of the construction, that is, we eventually obtain an induced decomposition of T.

A first obvious consequence is:

Theorem 5.3. *For every k, every bipartite twin graph T, with t vertices of arbitrary positive weights, has an induced decomposition into k induced subgraphs whose total edge weights are any k prescribed numbers that sum up to the total edge weight of T. Furthermore, the twin graph of the decomposition has at most $t + k - 1$ vertices.*

Problem C. For the simplest bipartite twin graph $T = K_2$ we had seen that $2\sqrt{k}(1+o(1))$ vertices are enough. Does every fixed bipartite graph T allow a decomposition into k induced subgraphs with equal edge weights and with only $o(k)$ vertices?

6 Odd Cycles as Twin Graphs

Recall that C_n is the cycle of n vertices. We also use the standard notation P_n for a path of n vertices. It is well known that bipartite graphs are exactly those graphs containing no (induced) subgraph C_n for odd n. We have just seen that bipartite graphs can be divided into an arbitrary number of induced subgraphs with (e.g.) equal total edge weights, and we have already studied the twin graph C_3 separately. Now it is natural to look at the larger odd cycles. In this section we will see that we can decompose them equitably as well, using independent separators.

Theorem 6.1. *For every $k \geq 3$ there exists an induced decomposition of the weighted twin graph C_5 (given with arbitrary vertex weights) into k induced subgraphs with equal total edge weights and with at most $k + 4$ vertices.*

Proof. We denote the edges in cyclic order by $e_1, \ldots, e_5$, and we let v_i denote the vertex incident to e_i and e_{i+1}. The starting point is arbitrary, that is, we can apply an arbitrary cyclic shift to the indices. Also recall that we can normalize the total edge weight to 1.

We prove the assertion by induction on k, where the intricate part is the induction base $k = 3$.

Suppose that two edges exist with weights larger than 1/3. No matter whether these two edges are incident or not, we can cyclically shift the indices such that: e_1 and e_2 together have weight larger than 1/3, and e_3 and e_4 together have weight larger than 1/3. Now we apply Lemma 5.2, once with $S = \{v_1\}$ and $N = \{v_5, v_2\}$, and once with $S = \{v_3\}$ and $N = \{v_2, v_4\}$. That is, we divide v_1 and v_3, respectively, and

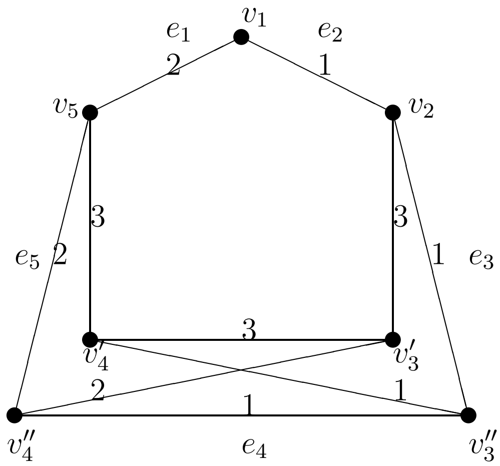

Figure 1: This shows the division of the twin graph C_5 into 3 induced subgraphs. A label i at an edge indicates that this edge belongs to H_i.

remove twice an induced subgraph with total edge weight 1/3. Hence the remaining graph has total edge weight 1/3 as well. Due to Lemma 5.1 (applied twice) these 3 subgraphs form an induced decomposition.

The other case is that at most one edge has a weight larger than 1/3. Then let e_1, e_2, e_3 be consecutive edges, each with weights at most 1/3. The edges e_3, e_4, e_5, e_1 have together a weight at least 2/3. We can assume that e_3 and e_4 have together a weight at least 1/3. (The case with e_5 and e_1 is symmetric.)

With these precautions we construct the following induced subgraphs. We divide v_3 into v_3' and v_3'', and we divide v_4 into v_4' and v_4''. The splitting of their vertex weights will be fixed later. Note that e_3 and e_5, respectively, is split in two edges, and e_5 is split in four edges. Let T^* be the resulting graph of 7 vertices. We define $H_1 := T^*[v_1, v_2, v_3'', v_4', v_4'']$, $H_2 := T^*[v_1, v_3', v_4'', v_5]$, and $H_3 := T^*[v_2, v_3', v_4']$. They form an induced decomposition; just check their pairwise intersections. Edge e_2 has a weight at most 1/3, and the sum of weights of e_3 and e_4 was at least 1/3. Hence we can choose the vertex weight of v_3'' such that H_1 gets the total edge weight 1/3. Similarly, e_2 has a weight at most 1/3, and the sum of weights of e_3, e_4, e_5, e_1 was at least 2/3. At most 1/3 (from e_3 and e_4) went into H_1. Hence we can choose the vertex weight of v_4'' such that H_2 gets the total edge weight 1/3, too. There remains exactly 1/3 for H_3. This finishes the proof for $k = 3$. Since this construction is hard to overlook in text, we also refer to Figure 1.

For the induction step, consider any fixed $k \geq 4$ and assume that the assertion is true for $k - 1$ subgraphs. Consider all 5 pairs of incident edges. Since every edge belongs to exactly 2 of them, the sum of their edge weights is 2. Hence there exist 2 incident edges of total weight at least $2/5 > 1/4 \geq 1/k$. Let e_2 and e_3 be these edges. We apply Lemma 5.2 with $S = \{v_2\}$ and $N = \{v_1, v_3\}$ and cut off an induced subgraph of total edge weight $1/k$. Due to Lemma 5.1 and the induction hypothesis,

C_5 has an induced decomposition into k induced subgraphs with equal total edge weights. Every step creates one new vertex. □

We also find that C_5 cannot always be divided into $k = 2$ induced subgraphs of equal weights, This negative result is based on the following observation: Since an induced subgraph H of a twin graph G consists of portions of the vertices of G, the graph H ignoring the weights is also an induced subgraph of G in the usual sense.

Now, the only types of induced subgraphs of C_5 (apart from isolated vertices) are P_2, P_3, P_4. Hence the only possible induced decompositions into $k = 2$ induced subgraphs consist of a P_3 and a P_4 that share their end vertices. Obviously, their total edge weights are in general different.

Dividing longer cycles is much easier because, loosely speaking, they have more independent sets.

Theorem 6.2. *For every $n \geq 6$ and $k \geq 2$ there exists an induced decomposition of the weighted twin graph C_n (given with arbitrary vertex weights) into k induced subgraphs with equal total edge weights and with at most $n + k - 1$ vertices.*

Proof. Twin graphs C_n with even n are already settled by Theorem 5.3, hence it suffices to consider odd cycles. Take any $j \geq 2$. In C_{4j-1}, consider all $4j - 1$ paths of $2j$ edges. Since every edge belongs to exactly $2j$ of them, the sum of their edge weights is $2j$. Hence there exists a path P of $2j$ edges with total weight at least $2j/(4j - 1) > 1/2$. The same double-counting argument shows that, in C_{4j+1}, some path P of $2j + 2$ edges has total weight at least $(2j + 2)/(4j + 1) > 1/2$.

Note that P has even length, and the end vertices of P are not adjacent. Thus we can apply Lemma 5.2, where the vertices of P are put alternatingly in N and S, starting and ending with a vertex in N. Since $1/k \leq 1/2$, we can cut off an induced subgraph U of total edge weight $1/k$. More specifically, we can successively process the vertices of S in their order on P, and put N and the visited vertices of S completely in U, until the total edge weight of U reaches $1/k$ (and at this moment at most one vertex of S is split). In particular, either some vertices of S end up completely in U, or we reach $1/k$ already by splitting the first vertex of S. In the former case we have "broken the cycle", that is, we get an induced decomposition of C_n into two paths of total edge weights $1/k$ and $1 - 1/k$, respectively. In the latter case we get an induced decomposition of C_n into P_3 and (still) C_n, with total edge weights $1/k$ and $1 - 1/k$, respectively. In both cases, the induced decomposition has at most $n + 1$ vertices.

Now we are ready to prove the assertion by induction on k. For $k = 2$, both total edge weights are $1/2$, and the number of vertices is at most $n + 2 - 1$. Suppose that the assertion holds for $k - 1$. We split off U with total edge weight $1/k$ as described above. If the remaining weighted twin graph is a path P_t, then we can use Theorem 5.3 (since paths are bipartite) to obtain an induced decomposition of P_t into $k - 1$ induced subgraphs with equal total edge weights and at most $t + (k - 1) - 1$ vertices, in other words, with at most $k - 2$ additional vertices. If the remaining weighted twin graph is still C_n, then we temporarily scale the vertex weights such that the total edge weight becomes 1 again, and we apply the induction hypothesis. At most

$k-2$ new vertices are produced here, too. Clearly, after re-scaling to the original weights, all k induced subgraphs have the same total edge weights, and Lemma 5.1 ensures that they form an induced decomposition. □

Problem D. Does the necessary number of vertices of decompositions into k induced subgraphs with equal total edge weights grow linearly in k, for non-bipartite twin graphs T?

7 Using Homomorphisms

In this brief section we sketch an extension of some of the previous results.

A *graph homomorphism* $f : G \to H$ maps the vertices of a graph G to the vertices of a graph H such that $f(u), f(v)$ are adjacent in H whenever u, v are adjacent in G. In short: edges are mapped to edges. If a homomorphism $f : G \to H$ exists, then G is also called *H-colorable.* (Hence K_χ-colorability is the usual χ-colorability.) A homomorphism substitutes, in an obvious sense, the vertices of H by independent vertex sets in G, and the edges of H by induced bipartite subgraphs of G. Note that the mapping of a graph G to its twin graph T is a homomorphism, with the additional property that non-edges are mapped to non-edges, i.e., the said bipartite subgraphs are complete there.

The constructions in the proofs of Theorem 6.1 and 6.2 can also be applied to C_n-colorable twin graphs, to divide them into induced subgraphs with equal total edge weights. Only the number of vertices is more tricky and depends on the given twin graph. This implies the same divisibility results for C_n-colorable graphs, for every $n \geq 5$ and $k \geq 2$ (except $n=5$ and $k=2$).

Many graphs are known to be C_n-colorable. To give some examples: Hexagonal graphs are induced subgraphs of the triangular grid. All C_3-free hexagonal graphs are C_5-colorable [11], and elementary case distinctions show that all C_3-free graphs with up to 7 vertices are C_5-colorable, too.

8 Twin Graphs Without Short Odd Cycles

As we have seen, short odd cycles make it more difficult to divide the edge set of a twin graph equitably into induced subgraphs. Therefore it is interesting to notice the following result for a wider class of graphs where we just forbid the two shortest odd cycles. However we can prove it only from some large enough k on, depending on the smallest edge weight.

Theorem 8.1. *For every $\{C_3, C_5\}$-free weighted twin graph T with t vertices (given with arbitrary but fixed vertex weights such that the total edge weight equals 1), and for every k being at least the inverse of the smallest edge weight, there exists an induced decomposition into k induced subgraphs with equal total edge weights and with at most $t+k-1$ vertices.*

Proof. Due to Theorem 5.3 it suffices to consider non-bipartite graphs T. Let d denote the distance function in T, indicating the number of edges on a shortest path between any two vertices. We call a vertex *odd* if it is contained in some odd cycle. Since T is not bipartite, T contains some odd cycle, and hence some odd vertex v. Furthermore, there exists an edge xy with $d(v,x) = d(v,y)$. (To see this, do breadth-first-search in T starting at v, and consider the distance layers. Some edge xy has its two vertices in the same layer.) Finally, since T has no C_3 or C_5, this edge also satisfies $d(v,x) = d(v,y) \geq 3$.

Now we describe an algorithm that constructs the claimed induced decomposition. We use H as a variable for the induced subgraph of T that we construct as the next member of our induced decomposition. Also note that every edge of T has an edge weight of at least $1/k$.

We start with any odd vertex v of T that we mark *active*, and any edge vu. Our first subgraph H is induced by v and a portion of u. Obviously we can choose this portion such that H gets the edge weight $1/k$.

Our next subgraph H is also induced by v and another portion of u. If the rest of u is too small to achieve edge weight $1/k$, then we distinguish two cases:

(1) Another unused edge vu' exists in T. Then we put the rest of u and s suitable portion of u' in H. Since the edge vu' has also a weight of at least $1/k$, we can indeed finalize our current subgraph H. We also rename u' and let it be our next u.

(2) All edges incident to v are used up. Then let xy be some edge with $d(v,x) = d(v,y) \geq 3$. Note that $d(u',x) \geq 2$ and $d(u',y) \geq 2$. Hence the only possible edges in $T[v,u',x,y]$ are vu' and xy. We add to H the entire x and a suitable portion of y. Again, since $1/k$ is small enough, we can indeed finalize the current H. We also rename x to v and let it be our new active vertex, and we rename y to u. Furthermore, we remove the old active vertex v and all incident edges from T (as they are already assigned to earlier induced subgraphs in our induced decomposition).

Thanks to the renaming of vertices, this shows that we keep the following invariant after every construction of an induced subgraph H: There exists at most one active vertex v and am edge vu with the property that v together with some portion of u is already used, while all other edges of T are either completely used or completely unused.

Thus we can iterate these steps and split off induced subgraphs, each with total edge weight $1/k$. Whenever no vertex v as in the invariant exists, we can select a new active vertex v arbitrarily. As soon as the removal of edges makes the remaining twin graph bipartite, we abort the process and continue according to Theorem 5.3. In either case, we eventually divide all of T into induced subgraphs H, each of total edge weight $1/k$.

Finally we argue that they form an induced decomposition. Every H consists of an active vertex and some of its neighbors, or of two (consecutively) active vertices and some of their neighbors, and in both cases, all these neighbors form an independent set. Every vertex that was once active is deleted as soon as all its incident edges are used up. Hence any later subgraph H can intersect any earlier subgraph H only in some independent set.

The bound of $t+k-1$ vertices comes again from the fact that every new induced subgraph splits at most one further vertex of T. □

Acknowledgments

The author would like to thank the referees for the pointer to [1], various small corrections, and other valuable comments.

References

[1] M. Aigner, *Combinatorial Search*, Wiley (1988).

[2] Y. Aumann and Y. Dombb, The efficiency of fair division with connected pieces, *6th Int. Workshop on Internet and Network Economics WINE 2010*, LNCS 6484, 26–37.

[3] J. A. Bondy and J. L. Szwarcfiter, Induced decompositions of graphs, *J. Graph Theory* **72** (2013), 462–477.

[4] K. Cechlárová and E. Pillárová, On the computability of equitable divisions, *Discrete Optim.* **9** (2012), 249–257.

[5] F. Y. L. Chin, H. C. M. Leung and S. M. Yiu, Non-adaptive complex group testing with multiple positive sets, *Theoret. Comput. Sci.* **505** (2013), 11–18.

[6] J. Chodoriwsky and L. Moura, An adaptive algorithm for group testing for complexes, *Theoret. Comput. Sci.* **592** (2015), 1–8.

[7] P. Johann, A group testing problem for graphs with several defective edges, *Discrete Appl. Math.* **117** (2002), 99–108.

[8] J.S. Juan and G.J. Chang, Group testing in graphs, *J. Comb. Optim.* **14** (2007), 113–119.

[9] A. D. Procaccia, Cake-cutting algorithms, *Handbook of Computational Social Choice*, Cambridge Univ. Press (2015), 261–283.

[10] R. Reitzig and S. Wild, Efficient algorithms for envy-free stick division with fewest cuts, CoRR abs/1502.04048 (2015).

[11] P. Sparl and J. Zerovnik, Homomorphisms of hexagonal graphs to odd cycles, *Discrete Math.* **283** (2004), 273–277.

[12] E. Triesch, A group testing problem for hypergraphs of bounded rank, *Discrete Appl. Math.* **66** (1996), 185–188.

[13] R. M. Wilson, Decompositions of complete graphs into subgraphs isomorphic to a given graph, *Congr. Num.* **XV** (1976), 647–659.

(Received 14 May 2018; revised 13 Aug 2019)