

A Classification of Long-Refinement Graphs for Colour Refinement

Sandra Kiefer
University of Oxford
sandra.kiefer@cs.ox.ac.uk

T. Devini de Mel
University of Oxford
devini.demel@gmail.com

October 24, 2025

Abstract

The Colour Refinement algorithm is a classical procedure to detect symmetries in graphs, whose most prominent application is in graph-isomorphism tests. The algorithm and its generalisation, the Weisfeiler–Leman algorithm, evaluate local information to compute a colouring for the vertices in an iterative fashion. Different final colours of two vertices certify that no isomorphism can map one onto the other.

The number of iterations that the algorithm takes to terminate is its central complexity parameter. For a long time, it was open whether graphs that take the maximum theoretically possible number of Colour Refinement iterations actually exist. Starting from an exhaustive search on graphs of low degrees, Kiefer and McKay proved the existence of infinite families of such long-refinement graphs with degrees 2 and 3, thereby showing that the trivial upper bound on the iteration number of Colour Refinement is tight.

In this work, we provide a complete characterisation of the long-refinement graphs with low (or, equivalently, high) degrees. We show that, with one exception, the aforementioned families are the only long-refinement graphs with maximum degree at most 3, and we fully classify the long-refinement graphs with maximum degree 4. To this end, via a reverse-engineering approach, we show that all low-degree long-refinement graphs can be represented as compact strings, and we derive multiple structural insights from this surprising fact. Since long-refinement graphs are closed under taking edge complements, this also yields a classification of long-refinement graphs with high degrees.

Kiefer and McKay initiated a search for long-refinement graphs that are only distinguished in the last iteration of Colour Refinement before termination. We conclude it in this submission by showing that such graphs cannot exist.

1 Introduction

Colour Refinement, also known as Naïve Vertex Classification or the 1-dimensional Weisfeiler–Leman algorithm, is a classical combinatorial procedure to detect symmetries in graphs. The algorithm computes in iterations a colouring of the vertices that reflects how the vertices are connected to the rest of the graph. By assessing local structure, it refines in each iteration the colouring from the previous one, until it reaches a stable state, which is when it terminates and returns the current colouring. The colour of each vertex in an iteration consists of (an encoding of) the colours from the previous iteration of the vertex and the vertices in its neighbourhood. The colourings are isomorphism-invariant: if two vertices obtain different colours, then there is no isomorphism between the graphs mapping one vertex to the other. This way, the colourings can decrease the size of the search space for isomorphisms drastically, and, in fact, the most prominent application of the algorithm is in graph-isomorphism tests and graph-canonisation approaches.

Colour Refinement can be implemented to run in time $O((m+n)\log(n))$ [9, 32], where n and m are the number of vertices and edges, respectively, of the input graph, and efficient implementations are used in all competitive isomorphism solvers, see e.g. [1, 10, 22, 23, 32, 33].

Besides the practical implementations, the algorithm has also fascinated theoreticians over decades. The main reason for this is that it keeps revealing more and more links between various areas in computer science. For example, the distinguishing power of the algorithm corresponds to the expressivity of graph neural networks [34], and there are machine-learning kernel functions based on the algorithm [30, 36]. Two graphs have the same final colourings with respect to the algorithm if and only if there is a fractional isomorphism between them [16, 35, 37]. This is equivalent to no sentence in the 2-variable fragment of the counting logic C distinguishing the graphs [8], and also to the graphs agreeing in their homomorphism counts for all trees [11]. Remarkably, these last two results generalise naturally to higher-dimensional versions of the Weisfeiler–Leman algorithm ([12, 21], see also [39] for the original 2-WL algorithm), whose expressive power is a very active area of research. Most notably, the algorithm serves as a central subroutine in Babai’s celebrated quasipolynomial-time graph-isomorphism algorithm [5].

The connections that the algorithm exhibits to other areas such as logic, homomorphism counting, and machine learning extend the toolbox of proof methods to understand the algorithm better. While the expressive power of Colour Refinement has been completely characterised [2, 27], its computational complexity still sparks research. The central complexity measure that has tight links in the related fields is the number of iterations that the algorithm uses to reach its output. This number of colour recomputations is crucial for the parallel running time of Colour Refinement, since ℓ iterations can be simulated with $O(\ell)$ steps on a PRAM with $O(n)$ processors [20, 28]. Also, when run on two graphs in parallel, the iteration number corresponds to a bound on the descriptive complexity of the difference between the graphs. More precisely, the number of iterations until the graphs obtain different colourings equals the minimum quantifier depth of a distinguishing sentence in the counting logic C^2 [8, 21]. There are also interesting connections to proof complexity [3, 4] and database theory [14].

The algorithm proceeds by iteratively enriching the colouring from the previous iteration with more information, yielding with each iteration a finer partition of the vertices into colour classes. A graph that requires the maximum theoretically possible number of Colour Refinement iterations needs to provide as little “progress” as possible in each iteration, implying that the initial partition into colour classes needs to be the trivial one with all vertices of the same colour, whereas in the final partition, every vertex has to have its own distinct colour. Hence, the trivial upper bound on the iteration number on n -vertex graphs is $n - 1$, with the number of colour classes growing by 1 in each iteration. We call graphs that achieve this bound *long-refinement graphs*. On random graphs, the iteration number is asymptotically almost surely 2 [6]. On the other hand, Krebs and Verbitsky managed to construct graphs with an iteration number of $n - O(\sqrt{n})$ [29]. For a long time, no progress on the trivial upper bound $n - 1$ was made, and it was open whether long-refinement graphs actually exist. Inspired by work due to Gödicke [15] and starting from an exhaustive search on graphs of low degrees, Kiefer and McKay proved the existence of infinite families of long-refinement graphs with degrees 2 and 3, thereby showing that the trivial upper bound on the iteration number of Colour Refinement on n -vertex graphs is tight [25].

The main motivation for the works [15] and [25] was to find out whether long-refinement graphs exist. With the discovered families, various distinctive properties of long-refinement graphs with degrees 2 and 3 became apparent. Those allowed the authors to find compact representations for the graph families, which ultimately made it possible to show that the families are infinite. The hope was to use these insights to construct new families of long-refinement graphs of degrees 2 and 3 that would cover all possible graph sizes (starting from the smallest possible n , which is 10).

However, infinitely many gaps have persisted since then, and in particular, it has remained open for which of the remaining values of n there exist long-refinement n -vertex graphs at all.

Our Contribution In this work, we provide a complete characterisation of the long-refinement graphs with small (or, equivalently, high) degrees. More concretely, we show that, with one exception of a graph with minimum degree 1, the aforementioned families are the only long-refinement graphs with maximum degree at most 3, and we fully characterise the long-refinement graphs with maximum degree at most 4. Since complements of long-refinement graphs are long-refinement graphs as well, this also yields a characterisation of the long-refinement graphs with high degrees (i.e. with minimum degree at least $n - 5$). While it clearly strengthens and generalises the results from [15] and [25], our approach does not use experimental results and is entirely theoretical.

By the definition of Colour Refinement, it is easy to see that every long-refinement graph has exactly two different vertex degrees. From the facts that the algorithm iteratively refines a partition of the vertex set and that long-refinement graphs need to end up with the finest partition possible, we deduce that when executing Colour Refinement on a long-refinement graph, there will be a stage in which each colour class has size at most 2. We analyse the possible connections between these pairs and singletons, depending on the indices of the iterations in which the pairs split into singletons. This generalises the analyses from [25], since our results hold for any degree pairs. Our structural study gives us a clear understanding of the order and edges between splitting pairs in a long-refinement graph. We then head off to reverse-engineer the larger colour classes from earlier iterations step by step. Again, our results hold for long-refinement graphs with no degree restrictions, but the analysis becomes more and more complex with larger colour classes.

Restricting ourselves to small degrees, we obtain our classification of long-refinement graphs of maximum degree at most 4. With this, we manage to show that all even-order long-refinement graphs G with vertex degrees 2 and 3 have a stage in which each colour class has size 2, and all odd-order long-refinement graphs with these degree restrictions have a stage in which each colour class has size at most 2 and there is only one singleton colour class. This is a surprising insight because the corresponding statement was phrased as a requirement in [25], while we show that it is in fact a general property. It implies that *every* long-refinement graph with degrees 2 and 3 has a simple compact string representation. It also allows us to capture more restrictions on symmetries in the strings that represent long-refinement graphs, ultimately yielding that the families from [25] are exhaustive and that the gaps in sizes that the work could not cover occurred because there are no long-refinement graphs with those sizes and degrees simultaneously. From the characterisation via string representations, we can then easily deduce that there is exactly one long-refinement graphs with minimum degree 1 and maximum degree at most 3.

Since Colour Refinement is usually applied to two graphs with the purpose of distinguishing them through their colourings, a natural question to ask – also in the light of the lower bound from [29] – is whether there are actually *pairs of graphs* that take as many refinements as theoretically possible to be distinguished. This question was initiated in [25] and has been open since then. Our aforementioned results give us a complete understanding of the structure and splitting dynamics of long-refinement graphs with low (or high) degrees. No two of those long-refinement graphs are distinguished only in the last iteration of Colour Refinement on them, but there could in principle be higher-degree long-refinement graphs with this property. In this submission, we conclude the search by proving that there cannot be a “long-refinement pair” of graphs, i.e. a pair of long-refinement graphs (of any degrees) that is only distinguished in the last iteration of Colour Refinement before termination.

Further Related Work It is worth mentioning that just as for Colour Refinement, for general k , the iteration number $WL_k(n)$ of the k -dimensional Weisfeiler–Leman algorithm (k -WL) on n -vertex graphs is being studied. For 2-WL, the paper [26] used a simple combinatorial game to show that the trivial upper bound of $WL_2(n) \leq n^2 - 1$ on the iteration number on n -vertex graphs (as well as the improved bound one obtains merely accounting for the fact that the initial partition for k -WL with $k \geq 2$ always has multiple colour classes) is asymptotically not tight. In [31], algebraic machinery was employed to show that $WL_2(n) \in O(n \log(n))$. This was recently generalised to $O(kn^{k-1} \log(n))$ for $k \geq 3$ in [18]. Concerning lower bounds, Fürer’s $WL_k(n) \in \Omega(n)$ remained unbeaten over decades [13]. A few years ago, significant progress was made when proceeding to relational structures of higher arity [7] (see also [18]). Only very recently, Fürer’s bound was also improved for graphs, with $\Omega(n^{k/2})$ being the currently best bound [19].

Concerning the number of iterations until two graphs are distinguished, it has been shown that on graphs of bounded treewidth, on planar graphs, and on interval graphs, there is a constant dimension of the Weisfeiler–Leman algorithm that distinguishes a graph from such a class from every second one in $O(\log(n))$ iterations [17, 20, 38].

2 Overview of Results and Methods

This section serves as an overview of our main results and the methods that we use. All results formally stated in this section as well as the short proofs for Lemmas 2.10 and 2.11 are collected from the later sections.

Given a (vertex-coloured or monochromatic) input graph, Colour Refinement uses a simple local criterion to iteratively refine the partition of the vertices induced by the current colouring. As soon as a stable state with respect to the criterion is reached, the algorithm returns the graph together with the current vertex colouring.

Definition 2.1 (Colour Refinement). Let G be a graph with vertex set $V(G)$, edge set $E(G)$, and vertex-colouring $\lambda: V(G) \rightarrow \mathcal{C}$. On input (G, λ) , the Colour Refinement algorithm computes its output recursively with the following update rules. Letting $\chi^0 := \lambda$, for $i \in \mathbb{N}$, the *colouring after i iterations of Colour Refinement* is

$$\chi^i(v) := (\chi^{i-1}(v), \{\!\!\{ \chi^{i-1}(u) \mid \{u, v\} \in E(G) \}\!\!\}).$$

Thus, two vertices obtain different colours in iteration i precisely if they already had different colours in iteration $i - 1$ or if they differ in the χ^{i-1} -colour multiplicities in iteration $i - 1$ among their neighbours.

Letting $\pi^i := \pi_{G, \lambda}(\chi^i)$ be the vertex partition induced by the colouring after i iterations of Colour Refinement, it is immediate that π^{i+1} is at least as fine as π^i (we write $\pi^{i+1} \preceq \pi^i$), for every $i \in \mathbb{N}$, and that, hence, there is a minimal $j \leq |V(G)| - 1$ such that $\pi^{j+1} = \pi^j$. With this choice of j , the output of Colour Refinement on input (G, λ) is $\chi_{G, \lambda} := \chi^j$.

This paper studies long-refinement graphs, i.e. graphs G on which the Colour Refinement procedure takes $n - 1$ iterations to terminate, where $n = |V(G)|$. It is easy to see that this is the maximum theoretically possible number of iterations on an n -vertex graph. To analyse long-refinement graphs, we first recall some very basic insights about the refinement iterations on them. Here, since Colour Refinement refines the colour of each vertex based on the multiplicity of colour occurrences in its neighbourhood, the degrees of vertices (i.e. their numbers of neighbours) are crucial.

Colour Refinement terminates as soon as the vertex partition induced by the colours is not strictly finer than the one in the previous iteration. Hence, to obtain $n - 1$ iterations until termination, every iteration on a long-refinement graph must increase the number of colour classes by

exactly 1, starting with a single colour class containing all vertices and ending with a partition in which each vertex forms its own colour class. This implies in particular that there are precisely two vertex degrees in the graph and that in each iteration exactly one colour class is split into two colour classes, which then induces the next split. In Section 4, with Lemma 4.4, we formalise the notion of *inducing a split* by introducing *unbalanced* pairs of colour classes. The lemma is the central ingredient for the structural arguments in Sections 5 and 6.

In Section 5, we first observe that every long-refinement graph must have a stage in which every colour class has size at most 2, call that iteration p and the induced vertex partition π^p . From then on, there is a linear splitting order among the classes of size 2, the *pairs*. The first part of Section 5 analyses the connections between these classes, completely dropping the degree requirements from [25] and sharpening the insights about the partitions. The first split of a pair must be induced by the splitting of a class of size 4 into two pairs, which we denote P_a and P_b . See Figure 1 for the situation when P_a and P_b are successive in the linear order of pairs, and when P_b is the successor of the successor of P_a .

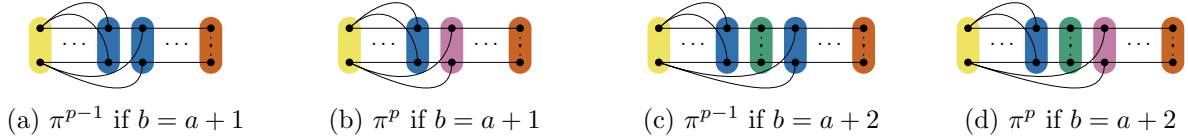


Figure 1: The partitions π^{p-1} and π^p , where the splitting order of pairs is from left to right. P_a is blue, P_b is first blue and then pink, respectively. For clarity, we omit all edges between non-consecutive pairs apart from the ones connecting P_a and P_b to the minimum of the linear order.

The iterations prior to this see the pairs symmetrically preceding P_a and succeeding P_b resulting from splits of classes of 4 themselves. We thus reverse-engineer the first partition to contain no classes larger than 4. This partition may or may not contain a class of size 3, and the analysis of partitions in the preceding iterations – wherein we reverse-engineer the classes of size 6 – considers both cases. The statements in the second part of Section 5 read more technical, due in part to the case distinction required for if there is a class of size 3, but a closer look at them lets us understand a similar argument and a similar symmetry between the iterations where classes of size 4 split into two pairs, and those where classes of size 6 split into a class of size 4 and a pair. Throughout Section 5, arguments about the classes allow us to make observations about the edges incident to the vertices within them, and we gain general insights about the structure of all long-refinement graphs, without any assumptions on degree.

We employ these insights in Section 6. Here, we first revisit the compact string notation for long-refinement graphs of degrees 2 and 3 that was introduced in [25].

Notation 2.2. Let $\Sigma := \{S, X, X_2, 0, 0_2, 1, 1_2\}$, and let n_P be the number of colour classes of size 2 in the minimal iteration p in which all colour classes have size at most 2. By π_S^p , denote the set of colour singletons in iteration p .

We define $\sigma: [n_P] \rightarrow \Sigma$ as the function that maps i to the element in Σ that matches the type of P_i . More precisely, $\sigma(i)$ is defined as follows:

- S , if $P_i = \min(<) = P_1$.
- X , if $P_i \in \{P_a, P_b\}$ and P_i is not adjacent to any singleton $S \in \pi_S^p$.
- X_2 , if $P_i \in \{P_a, P_b\}$ and P_i is adjacent to a singleton $S_i \in \pi_S^p$.
- 0 , if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 2$ and P_i is not adjacent to any singleton $S \in \pi_S^p$.

- 0_2 , if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 2$, and P_i is adjacent to a singleton $S_i \in \pi_S^p$.
- 1 , if $P_i \notin \{P_1, P_a, P_b\}$ with $\deg(P_i) = 3$, such that P_i is not adjacent to any singleton $S \in \pi_S^p$.
- 1_2 , if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 3$, and P_i is adjacent to a singleton $S_i \in \pi_S^p$.

Since $\prec$ is a linear order, we can define a string representation based on a long-refinement graph G , where the i -th letter in the string is $\sigma(i)$ and corresponds to P_i , the i -th element of $\prec$. We call a string defined in this way a long-refinement string.

Since the paper [25] was concerned with proving the existence of long-refinement graphs, it sufficed to use the string representations to find infinitely many long-refinement graphs. Here, we are concerned with a *complete characterisation*, and to this end, we use the results from Section 5 to detect precisely which strings over the alphabet yield long-refinement graphs. We find, for example, that every long-refinement graph with degrees 2 and 3 has a stage in which all classes have size 2 except for one singleton in the case where the order of the graph is odd. This yields the helpful insight that we can use the above alphabet to represent all long-refinement graphs with degrees 2 and 3 unambiguously, since the strings describe the exact neighbourhood of each pair and each singleton, as long as there is at most one singleton. Thus, up to isomorphism, every string Ξ represents exactly one graph, which we denote by $G(\Xi)$.

Using the technical insights from Section 5, Section 6 generalises the results beyond degrees 2 and 3, namely to maximum degree at most 4. This yields the first comprehensive classification of such long-refinement graphs. Through the symmetries defined in Section 5, we are able to divide this argument into cases depending on the number of pairs that remain in the partition, and not merged into larger classes.

In Section 7, with Theorems 7.1 – 7.6, we collect all results about the structure of long-refinement graphs of low degrees to obtain a compact classification, ordered by degree pairs. Moreover, we provide an alternative and simple proof that there is exactly one long-refinement graph with minimum degree 1 and maximum degree at most 4, which is the graph in Figure 2. We append these results in the following subsection (Subsection 2.1), with references to the figures and tables that represent the graphs. Note that, since a graph is a long-refinement graph if and only if the graph obtained by complementing the edge relation is a long-refinement graph, the results directly yield a classification of long-refinement graphs with high degrees.

Finally, in Section 8, we turn towards the distinguishability of long-refinement graphs. In its applications, Colour Refinement is used to distinguish two graphs. Therefore, given our insights about individual long-refinement graphs, it is natural to ask for graphs that take $n - 1$ iterations to be distinguished by Colour Refinement, i.e. to obtain distinct colourings (which implies that they are not isomorphic). The strongest known lower bound on the number of iterations to distinguish two graphs with Colour Refinement is $n - O(\sqrt{n})$ [29] and it is from a decade ago [29]. With the results from [25], it is plausible to search for long-refinement pairs of graphs, and hence this was posed as an open question in the same paper. Again using the fact that the last split in both graphs must be a pair splitting into singletons, we show that such pairs of graphs do not exist.

Theorem 2.3. *For $n \geq 3$, there is no pair of graphs G, H with $|V(H)| \leq |V(G)| = n$ that Colour Refinement distinguishes only after $n - 1$ iterations.*

2.1 Our Classification of Long-Refinement Graphs with Small Degrees

This subsection lists our main classification results, see also Theorems 7.1 – 7.6. Since our entire analysis is up to isomorphism, we will not always distinguish explicitly between graphs from the same isomorphism class.

We write $N_G(v)$ for the neighbourhood of vertex v in the graph G , and $\deg(G)$ for the set of vertex degrees in G .

Recall Notation 2.2. The work [25] identified families of long-refinement graphs G with $\deg(G) = \{2, 3\}$, but left as an open problem whether they are comprehensive. We answer that question affirmatively.

Theorem 2.4. *The even-order long-refinement graphs G with $\deg(G) = \{2, 3\}$ are exactly the ones represented by strings contained in the following sets:*

- $\{S011XX\}$
- $\{S1^k 001^k X1X1^k 0 \mid k \in \mathbb{N}_0\}$
- $\{S1^k 11001^k XX1^k 0 \mid k \in \mathbb{N}_0\}$
- $\{S1^k 0011^k XX1^k 10 \mid k \in \mathbb{N}_0\}$
- $\{S011(011)^k 00(110)^k XX(011)^k 0 \mid k \in \mathbb{N}_0\}$
- $\{S(011)^k 00(110)^k 1X0X1(011)^k 0 \mid k \in \mathbb{N}_0\}$

The odd-order long-refinement graphs G with $\deg(G) = \{2, 3\}$ are exactly the ones represented by strings contained in the following sets:

- $\{S1_211XX\}$
- $\{S0X1X_2\} \cup \{S1^k 1011^k X1X1^k 1_2 \mid k \in \mathbb{N}_0\}$
- $\{S110XX_2\} \cup \{S111^k 1011^k XX1^k 1_2 \mid k \in \mathbb{N}_0\}$
- $\{S1^k 01^k 1XX1^k 1_2 \mid k \in \mathbb{N}_0\}$
- $\{S(011)^k 00(110)^k X1_2X(011)^k 0 \mid k \in \mathbb{N}_0\}$

In particular, the theorem implies that the gaps that the graphs found in [24] leave open cannot be covered with degrees 2 and 3.

Corollary 2.5. *There are no long-refinement graphs G with $\deg(G) = \{2, 3\}$ and order $n \neq 12$ such that $n \bmod 18 \in \{6, 12\}$.*

Theorem 2.4 and Corollary 2.5 completely classify the long-refinement graphs G with $\deg(G) = \{2, 3\}$. Expanding to $\max \deg(G) \leq 4$, we obtain the following further parts of our classification of long-refinement graphs with small degrees (using the results in Section 6).

Theorem 2.6. *There is no long-refinement graph G with $\deg(G) = \{1, 2\}$ or $\deg(G) = \{1, 4\}$.*

Theorem 2.7. *The long-refinement graphs G with $\deg(G) = \{2, 4\}$ are precisely the ones in Figures 3b, 3c, 3e, 3h, 3j, 3k, 3l, 3n, 3q, 3r, 3s, and 3t.*

Theorem 2.8. *The long-refinement graphs G with $\deg(G) = \{3, 4\}$ are precisely the ones in Figures 3a, 3d, 3f, 3g, 3i, 3m, 3o, 3p, as well as Tables 1, 2, and 3.*

Theorem 2.9. *The graph in Figure 2 is the only long-refinement graph G with $\deg(G) = \{1, 3\}$.*

While Theorems 2.6 – 2.9 are proved through an analysis of the refinement iterations in Section 6, we now present a simple alternative proof for Theorem 2.9, which relies merely on Theorem 2.4. By [25, Lemma 13], a long-refinement graph G with $\deg(G) = \{1, 3\}$ has at most two vertices of degree 1. We start by treating the case where G has exactly one vertex of degree 1.

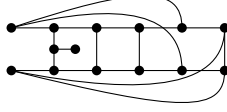


Figure 2: The unique long-refinement graph with degrees $\{1, 3\}$.

Lemma 2.10. *The graph in Figure 2 is the only long-refinement graph G with $\deg(G) = \{1, 3\}$ in which $|\{v \in V(G) : \deg(v) = 1\}| = 1$.*

Proof. Let G be any long-refinement graph with $\deg(G) = \{1, 3\}$ and the set of vertices of degree 1 being a singleton $\{v_1\}$. Then the partition after the first Colour Refinement iteration on G is $\pi_G^1 = \{\{v_1\}, V(G) \setminus \{v_1\}\}$. The graph $G' := G[V(G) \setminus \{v_1\}]$, i.e. the subgraph of G induced by $V(G) \setminus \{v_1\}$, satisfies $\deg(G') = \{2, 3\}$ and $\{v \in V(G') : \deg(v) = 2\} = \{v_2\}$, where $\{v_2\} = N_G(v_1)$. Also, Colour Refinement takes $n - 2$ iterations to terminate on G' , otherwise G would not be a long-refinement graph. Hence, G' is a long-refinement graph as well, and since $|V(G)|$ is even, $|V(G')|$ is odd. By our classification in Theorem 2.4, G' must be $G(\text{S1}_2\text{11XX})$, since that is the only long-refinement graph with degrees 2 and 3 in which there is only one vertex of degree 2. So, by degree arguments, G must be the graph in Figure 2. $\square$

Now we consider the situation in which there are exactly two vertices of degree 1.

Lemma 2.11. *There is no long-refinement graph G with $\deg(G) = \{1, 3\}$ in which $|\{v \in V(G) : \deg(v) = 1\}| = 2$.*

Proof. Suppose G is a long-refinement graph with $\deg(G) = \{1, 3\}$ for which $\{v \in V(G) : \deg(v) = 1\} = \{v_1, v'_1\}$ with $v_1 \neq v'_1$. We know that $\{v_1, v'_1\} \notin E(G)$, since G is connected. However, now the graph $G' := (V(G), E(G) \cup \{\{v_1, v'_1\}\})$ is a long-refinement graph with $\deg(G') = \{2, 3\}$ and $\{v \in V(G') : \deg(v) = 2\} = \{v_1, v'_1\}$. The degree property is clear. To see that G' is a long-refinement graph, note that inserting the new edge does not alter the partition into degrees, i.e. π^1 . Also, since in both graphs, it holds that $|N(\{v_1, v'_1\})| \leq 2$, the class $\{v_1, v'_1\}$ must be the last class that splits (in both graphs). Furthermore, $N_G(v_1) \setminus \{v'_1\} = N_{G'}(v_1) \setminus \{v'_1\}$ and $N_G(v'_1) \setminus \{v_1\} = N_{G'}(v'_1) \setminus \{v_1\}$. This implies that $\pi_G^i = \pi_{G'}^i$ for every $i \in \mathbb{N}$, so G' is a long-refinement graph if and only if G is one. By double counting edges, G (and hence also G') has an even number of vertices and hence G' is one of the graphs from the first list in Theorem 2.4. The only graph from that list with exactly two vertices of degree 2 is $G(\text{S011XX})$, but these vertices are not adjacent. Hence, G' does not exist, and therefore G does not exist. $\square$

This concludes the proof of Theorem 2.9, which together with Theorem 2.6 yields that the graph in Figure 2 is the only long-refinement graph G with $\min \deg(G) = 1$ and $\max \deg(G) \leq 4$.

Figure 3 lists all long-refinement graphs from our classification that are not part of infinite families, in the order in which they appear in our analysis in the remainder of this paper. Tables 1 – 3 list the infinite families.

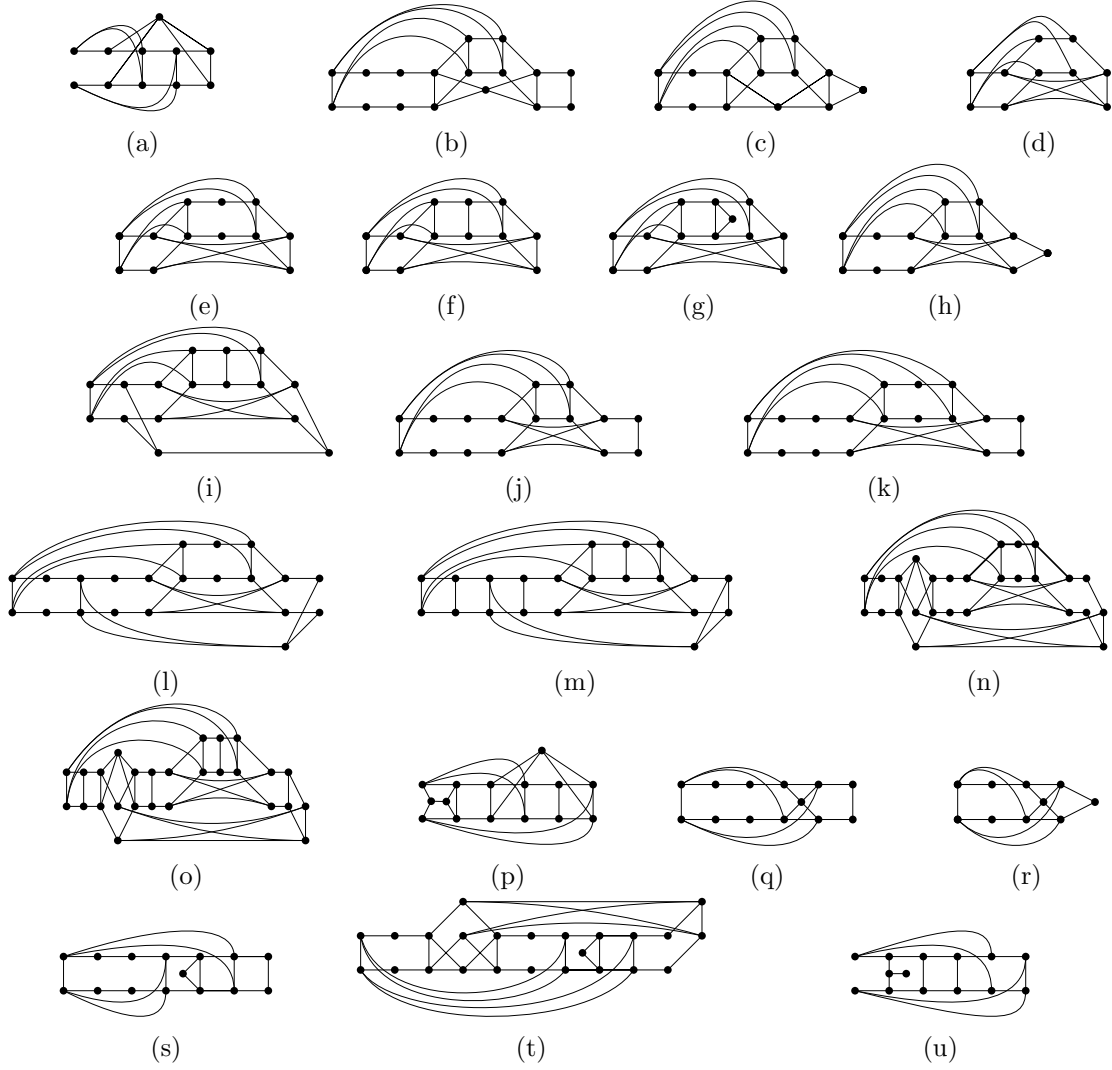


Figure 3: The long-refinement graphs G with $\deg(G) \neq \{2, 3\}$, omitting the infinite families from Tables 4 – 6.

vertex	adjacency 1 with $n \in \mathbb{N}$	adjacency 2 with $n \in \mathbb{N}_0$
0	$2n + 1, 2n + 2, 2n + 5, 2n + 6$	$2n + 1, 2n + 2, 2n + 5, 2n + 6$
1	$3, 4n + 7, 4n + 8$	$3, 4n + 7, 4n + 8$
2	$4, 4n + 9, 4n + 10$	$4, 4n + 9, 4n + 10$
odd $i \in I$	$i - 2, i + 1, i + 2$	$i - 2, i + 1, i + 2$
even $i \in I$	$i - 2, i - 1, i + 2$	$i - 2, i - 1, i + 2$
$2n + 1^*$	$0, 2n - 1, 2n + 3$	$2n + 1, 2n + 2, 2n + 5$
$2n + 2^*$	$0, 2n, 2n + 4$	$2n, 2n + 12n + 4$
$2n + 3$	$2n + 1, 2n + 5, 6n + 11, 6n + 12$	$0, 2n + 1, 2n + 4, 2n + 5$
$2n + 4$	$2n + 4, 2n + 8, 6n + 17, 6n + 18$	$0, 2n + 2, 2n + 3, 2n + 6$
$2n + 5$	$0, 2n + 3, 2n + 7$	$2n + 3, 2n + 6, 2n + 7$
$2n + 6$	$0, 2n + 4, 2n + 8$	$2n + 4, 2n + 5, 2n + 8$
$4n + 7$	$1, 4n + 5, 4n + 9$	$1, 4n + 5, 4n + 9$
$4n + 8$	$1, 4n + 6, 4n + 10$	$1, 4n + 6, 4n + 10$
$4n + 9$	$2, 4n + 7, 4n + 11$	$2, 4n + 7, 4n + 11$
$4n + 10$	$2, 4n + 8, 4n + 12$	$2, 4n + 8, 4n + 12$
$6n + 11$	$2n + 3, 2n + 4, 6n + 9$	$0, 6n + 12, 6n + 10$
$6n + 12$	$2n + 3, 2n + 4, 6n + 10$	$0, 6n + 11, 6n + 10$

Table 1: The two infinite families of long-refinement graphs G with order $6n+13$ for $\deg(G) = \{3, 4\}$ where $I = [3, 2n] \cup [2n + 7, 4n + 6] \cup [4n + 11, 6n + 10]$. Note also that $2n + 1, 2n + 2$ are omitted when $n = 0$.

vertex	adjacency 1	adjacency 2
0	$2k + 3, 2k + 4, 2k + 7, 2k + 8$	$2k + 5, 2k + 6, 6k + 13, 6k + 14$
1	$3, 4k + 9, 4k + 10$	$3, 4k + 9, 4k + 10$
2	$4, 4k + 11, 4k + 12$	$4, 4k + 11, 4k + 12$
odd $i \in I$	$i - 2, i + 1, i + 2$	$i - 2, i + 1, i + 2$
even $i \in I$	$i - 2, i - 1, i + 2$	$i - 2, i - 1, i + 2$
$2k + 3$	$0, 2k + 1, 2k + 5$	$2k + 1, 2k + 4, 2k + 5$
$2k + 4$	$0, 2k + 2, 2k + 6$	$2k + 2, 2k + 3, 2k + 6$
$2k + 5$	$2k + 3, 2k + 7, 6k + 13, 6k + 14$	$0, 2k + 3, 2k + 6, 2k + 7$
$2k + 6$	$2k + 4, 2k + 8, 6k + 13, 6k + 14$	$0, 2k + 4, 2k + 5, 2k + 8$
$2k + 7$	$0, 2k + 5, 2k + 9$	$2k + 5, 2k + 8, 2k + 9$
$2k + 8$	$0, 2k + 6, 2k + 10$	$2k + 6, 2k + 7, 2k + 10$
$4k + 9$	$1, 4k + 7, 4k + 11$	$1, 4k + 7, 4k + 11$
$4k + 10$	$1, 4k + 8, 4k + 12$	$1, 4k + 8, 4k + 12$
$4k + 11$	$2, 4k + 9, 4k + 13$	$2, 4k + 9, 4k + 13$
$4k + 12$	$2, 4k + 10, 4k + 14$	$2, 4k + 10, 4k + 14$
$6k + 13$	$2k + 5, 2k + 6, 6k + 11$	$0, 6k + 14, 6k + 11$
$6k + 14$	$2k + 5, 2k + 6, 6k + 12$	$0, 6k + 13, 6k + 12$

Table 2: All infinite families of long-refinement graphs G with order $6k + 15$ and $\deg(G) = \{3, 4\}$, where $I = [3, 2k + 2] \cup [2k + 9, 4k + 8] \cup [4k + 13, 6k + 12]$, defined for $k \in \mathbb{N}_0$.

vertex	adjacency 1	adjacency 2
0	$2k + 7, 2k + 8, 2k + 11, 2k + 12$	$2k + 5, 2k + 6, 6k + 17, 6k + 18$
1	$3, 4k + 13, 4k + 14$	$3, 4k + 13, 4k + 14$
2	$4, 4k + 15, 4k + 16$	$4, 4k + 15, 4k + 16$
odd $i \in I$	$i - 2, i + 1, i + 2$	$i - 2, i + 1, i + 2$
even $i \in I$	$i - 2, i - 1, i + 2$	$i - 2, i - 1, i + 2$
$2k + 7$	$0, 2k + 5, 2k + 9$	$2k + 5, 2k + 8, 2k + 9$
$2k + 8$	$0, 2k + 6, 2k + 10$	$2k + 6, 2k + 7, 2k + 10$
$2k + 9$	$2k + 7, 2k + 11, 6k + 17, 6k + 18$	$0, 2k + 7, 2k + 10, 2k + 11$
$2k + 10$	$2k + 8, 2k + 12, 6k + 17, 6k + 18$	$0, 2k + 8, 2k + 9, 2k + 12$
$2k + 11$	$0, 2k + 9, 2k + 13$	$2k + 9, 2k + 12, 2k + 13$
$2k + 12$	$0, 2k + 10, 2k + 14$	$2k + 10, 2k + 11, 2k + 14$
$4k + 13$	$1, 4k + 11, 4k + 15$	$1, 4k + 11, 4k + 15$
$4k + 14$	$1, 4k + 12, 4k + 16$	$1, 4k + 12, 4k + 16$
$4k + 15$	$2, 4k + 13, 4k + 17$	$2, 4k + 13, 4k + 17$
$4k + 16$	$2, 4k + 14, 4k + 18$	$2, 4k + 14, 4k + 18$
$6k + 17$	$2k + 9, 2k + 10, 6k + 15$	$0, 6k + 18, 6k + 15$
$6k + 18$	$2k + 9, 2k + 10, 6k + 16$	$0, 6k + 17, 6k + 16$

Table 3: All infinite families of long-refinement graphs G with order $6k + 19$ and $\deg(G) = \{3, 4\}$, where $I = [3, 2k + 6] \cup [2k + 13, 4k + 12] \cup [4k + 17, 6k + 16]$.

3 Preliminaries

With $\mathbb{N}$ being the set of natural numbers, we set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and for $a, b \in \mathbb{N}_0$, let $[a, b] := \{n \in \mathbb{N}_0 \mid a \leq n \leq b\}$ and $[b] := [1, b]$. Multisets are a generalisation of sets in which repeated elements are allowed, i.e. elements may occur more than once. To distinguish them from sets, we denote multisets using double braces $\{\!\{, \}\!\}$. Given a set S , a *partition* of S is a set of pairwise disjoint non-empty sets Π such that $\bigcup_{M \in \Pi} M = S$. Given two partitions Π, Π' of S , we say that Π *refines* Π' and write $\Pi \preceq \Pi'$ if every element of Π is a (not necessarily strict) subset of some element of Π' . We call a set with a single element a *singleton* and a set with two elements a *pair*. A partition that consists of singletons only is called *discrete*.

In our setting, all graphs are undirected, finite, and simple. In particular, they have no self-loops and between every pair of vertices only at most one edge. For a graph G , we use $V(G)$ and $E(G)$ to denote the set of vertices and the set of edges of G , respectively. We drop parameters if they are clear from or irrelevant in the context and just use V and E or write $G = (V, E)$. The *order* of G is $|G| := |V(G)|$. The *neighbourhood* of a vertex $v \in V(G)$ is $N(v) := \{u \in V(G) \mid \{u, v\} \in E\}$ and we extend the notion to sets of vertices S by setting $N(S) := \bigcup_{v \in S} (N(v) \setminus S)$. We consider v *adjacent* to S if $N(v) \cap S \neq \emptyset$ and say that C' is adjacent to C if $N(C) \cap C' \neq \emptyset$. The *degree* of v is $\deg(v) := |N(v)|$ and we set $\deg(G) := \{\deg(v) \mid v \in V(G)\}$. For $d \in \mathbb{N}_0$, the graph G is *d-regular* if $\deg(G) = \{d\}$, and *regular* if it is d -regular for some d . A *matching* is a 1-regular graph.

For non-empty $S, S' \subseteq V(G)$, the *induced subgraph* $G[S]$ is the graph (S, E) where $E = \{\{u, v\} \mid \{u, v\} \in E(G), u, v \in S\}$, and the subgraph $G[S, S']$ is the graph $(S \cup S', E')$ where $E' = \{\{u, v\} \mid \{u, v\} \in E(G), u \in S, v \in S'\}$. If $V(G) = S \cup S'$ and $S \cap S' = \emptyset$ and $E(G) \cap \{\{v, w\} \mid v, w \in S\} = \emptyset = E(G) \cap \{\{v, w\} \mid v, w \in S'\}$, then G is *bipartite (on bipartition (S, S'))*. If, additionally, $\{\{v, w\} \mid v \in S, w \in S'\} = E(G[S, S'])$, the graph G is *complete bipartite*.

A *path* (of length ℓ) from $u \in V(G)$ to $v \in V(G)$ is a sequence $u = v_0, v_1, \dots, v_\ell = v$ such that

$\{v_{i-1}, v_i\} \in E(G)$ and $v_i \neq v_j$ for all $i \in [\ell]$ and all $j \in [0, \ell]$ with $j \notin \{i, \ell\}$. The distance $d(u, v)$ between u and v is the length of a shortest path between u and v . A *vertex colouring* of G is a mapping λ from $V(G)$ into some set $\mathcal{C}$, whose elements we call *colours*. We call the pair (G, λ) a *vertex-coloured graph* and denote by $\pi_G(\lambda)$ the partition of $V(G)$ into the *colour classes*, i.e. into the maximal sets of vertices that have the same colour with respect to λ .

An *isomorphism* between graphs G and H is a bijective mapping $\varphi: V(G) \rightarrow V(H)$ such that for all $u, v \in V(G)$, it holds that $\{u, v\} \in E(G)$ if and only if $\{\varphi(u), \varphi(v)\} \in E(H)$. Isomorphisms between vertex-coloured graphs (G, χ_G) and (H, χ_H) must additionally satisfy $\chi_G(u) = \chi_H(\varphi(u))$.

4 Colour Refinement

In this section, we give a brief introduction to the Colour Refinement algorithm, which is also called the *1-dimensional Weisfeiler–Leman algorithm*, and state the main facts about it that we will use in our further analysis. Given a (vertex-coloured or monochromatic) input graph, Colour Refinement uses a simple local criterion to iteratively refine the partition of the vertices induced by the current colouring. As soon as a stable state with respect to the criterion is reached, the algorithm returns the graph together with the current vertex colouring.

Definition 4.1 (Colour Refinement). Let G be a graph with vertex colouring $\lambda: V(G) \rightarrow \mathcal{C}$. On input (G, λ) , the Colour Refinement algorithm computes its output recursively with the following update rules. Let $\chi^0 := \lambda$ and for $i \in \mathbb{N}$, the *colouring after i iterations of Colour Refinement* is

$$\chi^i(v) := (\chi^{i-1}(v), \{\!\!\{ \chi^{i-1}(u) \mid \{u, v\} \in E(G) \}\!\!\}).$$

Thus, two vertices obtain different colours in iteration i precisely if they already had different colours in iteration $i - 1$ or if they differ in the χ^{i-1} -colour multiplicities among their neighbours in iteration $i - 1$. Letting $\pi^i := \pi_{G, \lambda}(\chi^i)$ be the vertex partition induced by the colouring after i iterations of Colour Refinement, it is immediate that $\pi^{i+1} \preceq \pi^i$ holds for every $i \in \mathbb{N}$ and that, hence, there is a minimal $j \leq |G| - 1$ such that $\pi^{j+1} = \pi^j$. With this choice of j , the *output* of Colour Refinement on input (G, λ) is $\chi_{G, \lambda} := \chi^j$. We call j the *iteration number* of Colour Refinement on input (G, λ) and denote it by $\text{WL}_1(G, \lambda)$. If λ induces the trivial partition, i.e. (G, λ) is monochromatic, we set $\text{WL}_1(G) := \text{WL}_1(G, \lambda)$. We call (G, λ) and the induced partition $\pi(\lambda)$ *stable* if $\text{WL}_1(G, \lambda) = 0$, i.e. if $\pi(\lambda) = \pi(\chi_{G, \lambda})$. It follows from the definition of the Colour Refinement algorithm that for all $P, Q \in \pi(\chi_{G, \lambda})$ with $P \neq Q$, the graph $G[P]$ is regular and $H := G[P, Q]$ is *biregular*, i.e. there are $k, \ell \in \mathbb{N}$ such that for every $v \in P$, it holds that $|N_H(v)| = k$, and for every $w \in Q$, it holds that $|N_H(w)| = \ell$.

A *long-refinement graph* is a graph G for which $\text{WL}_1(G) = |G| - 1$, i.e. on which Colour Refinement takes the maximum possible number of iterations to compute its output. Note that, in our quest for long-refinement graphs, we do not have to consider vertex-coloured graphs, because, as observed in [25], all long-refinement graphs G of order at least 2 are monochromatic. Moreover, they are connected, have $|\deg(G)| = 2$, and Colour Refinement produces the discrete partition on them.

The only connected graphs with degrees 1 and 2 are paths, and a monochromatic path on n vertices takes $\lfloor \frac{n-1}{2} \rfloor$ iterations of Colour Refinement to stabilise (see, e.g., [25, Fact 5]). Solving $\lfloor \frac{n-1}{2} \rfloor = n - 1$ yields $n \leq 1$, but the graph consisting of a singleton vertex does not satisfy the degree restrictions. We hence obtain the following observation.

Observation 4.2 (cf. [25]). *There is no long-refinement graph G with $\deg(G) = \{1, 2\}$.*

Therefore, the smallest degree pairs with which there may be long-refinement graphs are $\{1, 3\}$ and $\{2, 3\}$. For both, examples are given in [25]. More precisely, they present one long-refinement graph G with $\deg(G) = \{1, 3\}$, one with $\deg(G) = \{1, 5\}$, and infinitely many with $\deg(G) = \{2, 3\}$. For specific ones of the infinite families with $\deg(G) = \{2, 3\}$, the proof of their Lemma 27 yields a simple construction to show the following.

Observation 4.3. *There are infinitely many sets S with $3 \in S$ and such that $\deg(G) = S$ for some long-refinement graph G .*

Proof. To see this, one just needs to verify that there are infinitely many numbers $3 \neq s \in \mathbb{N}$ such that there are long-refinement graphs G with $\deg(G) = \{2, 3\}$ and $s = |\{v \in V(G) \mid \deg(v) = 2\}|$, for which one can consider, for example, the last infinite family in [25, Theorem 24]. Now inserting a fresh vertex w with edges to all $v \in V(G)$ with $\deg(v) = 2$ yields a long-refinement graph G' with $\deg(G') = \{3, s\}$. $\square$

In the remainder of this section, we develop Lemma 4.4, which is the main structural insight about the partition classes in successive Colour Refinement iterations and which will allow us to pursue the following reverse-engineering approach. To study long-refinement graphs G with $|G| > 1$, we intend to construct the partition π^{i-1} from the partition π^i . Each partition π^i must have exactly one more colour class than π^{i-1} , see [25, Proposition 8]. So there must be disjoint A_i, B_i such that $\{A_i, B_i\} = \pi^i \setminus \pi^{i-1}$ and whose union is the unique C_{i-1} with $\{C_{i-1}\} = \pi^{i-1} \setminus \pi^i$. We say that C_{i-1} *splits into A_i and B_i* in iteration i . The class C_i splits immediately after and as a direct consequence of C_{i-1} splitting, so all vertices in C_i must have the same number of neighbours in C_{i-1} , but differ in their numbers of neighbours in A_i or in B_i . We formalise this by defining the *degree of vertex v with respect to set $C \subseteq V(G)$* as $\deg_C(v) := |N(v) \cap C|$ and, for vertex sets C, C' and $v \in C$, we define the *degree of C into C'* as follows:

$$\deg_{C'}(C) := \begin{cases} \deg_{C'}(v), & \text{if } \deg_{C'}(v) = \deg_{C'}(u) \text{ holds for every } u \in C \\ \perp, & \text{otherwise} \end{cases}$$

If $C' = V(G)$, we let $\deg(C) := \deg_{C'}(C)$. If $\deg_{C'}(C) = \perp$, then C is *unbalanced wrt C'* . Otherwise, C is *balanced wrt C'* . If we do not want to specify the class C' but just claim its existence, we call C simply *unbalanced*. With these notions, notice that every class in π^i with $i > 0$ is balanced wrt every class in π^{i-1} . In particular, C_i is balanced wrt C_{i-1} , and also with respect to each class in $\pi^i \cap \pi^{i-1}$. Since C_i is split in iteration $i + 1$, C_i cannot also be balanced wrt all classes in π^i , hence there must be a class in $\pi^i \setminus \pi^{i-1} = \{A_i, B_i\}$ wrt which C_i is unbalanced (cf. [25, Lemma 11]). In fact, since C_i is balanced wrt C_{i-1} , i.e. $A_i \cup B_i$, all vertices in C_i have the same degree into $A_i \cup B_i$, and therefore $\deg_{A_i}(C_i) = \perp$ holds if and only if $\deg_{B_i}(C_i) = \perp$. We summarise all these observations in the following lemma.

Lemma 4.4. *Let G be a long-refinement graph with $|G| > 1$. Then in any non-stable partition π^i on $V(G)$, there is exactly one unbalanced class $C_i \in \pi^i$.*

For the following, assume that, $i > 0$, that is, $\pi^i \neq \pi^0$. Then there are exactly two distinct options A_i, B_i for $C' \in \pi^i$ with $\deg_{C'} C_i = \perp$. Hence,

$$\deg_{A_i}(C_i) = \perp = \deg_{B_i}(C_i).$$

Furthermore, $C_i = A_{i+1} \cup B_{i+1}$ and

$$\deg_{A_i}(A_{i+1}) + \deg_{B_i}(A_{i+1}) = \deg_{C_{i-1}}(C_i) = \deg_{A_i}(B_{i+1}) + \deg_{B_i}(B_{i+1}) \in [1, |C_{i-1}| - 1]. \quad (1)$$

In particular, it holds that $\deg_{C_{i-1}}(C_i) \neq \perp$.

Finally, if $i > 1$, then $C_i \neq V(G)$ and, since the first iteration of Colour Refinement separates vertices by degree, $\deg(C_i) \neq \perp$ and thus $\deg(A_i) = \deg(B_i)$.

As a first application of our formal definition of classes being unbalanced wrt each other, we note that the class of long-refinement graphs is closed under edge complementation.

Lemma 4.5. *A graph $G = (V, E)$ is a long-refinement graph if and only if the graph $G' = (V, E')$ is a long-refinement graph, where $E' = \{\{u, v\} \mid u, v \in V, \{u, v\} \notin E, u \neq v\}$.*

Proof. It suffices to show that $\pi_G^i = \pi_{G'}^i$ holds for every $i \in \mathbb{N}_0$. First note that for vertex sets $C, C' \subseteq V$ that are either equal or disjoint, we have that $\deg_{C'}(C) = \perp$ holds in G if and only if it holds in G' . Indeed, if $C = C'$, then $G[C]$ is a subgraph of the complete graph on C . The latter is $(|C| - 1)$ -regular, hence $G[C]$ is regular if and only if the “complement subgraph” is regular. Similarly, if $C \cap C' = \emptyset$, the graph $G[C, C']$ is a subgraph of the complete bipartite graph on bipartition (C, C') . The latter is biregular, meaning it satisfies $\deg_{C'}(C) \in \mathbb{N} \ni \deg_C(C')$. Hence, $\deg_{C'}(C) \neq \perp$ holds in $G[C, C']$ if and only if it holds in the “complement subgraph”, and the same is true for the condition $\deg_C(C') \neq \perp$.

We can now proceed by induction on i . It is clear that $\pi_G^0 = \pi_{G'}^0$. For the inductive step, suppose that $\pi_G^i = \pi_{G'}^i$ holds for some $i \geq 1$. It suffices to see that, up to possibly swapping the roles of A_i and B_i , the classes A_i, B_i , and C_i from Lemma 4.4 are the same in both graphs. Indeed, the graph G satisfies Lemma 4.4, so there is a unique unbalanced class $C_i \in \pi^i$ that splits into A_{i+1} and B_{i+1} in iteration $i + 1$. By the observation above, C_i is also unbalanced in G' , precisely to A_i and to B_i , and there is no other unbalanced class in G' in π^i (as otherwise the same would hold in the long-refinement graph G , contradicting Lemma 4.4). $\square$

5 Reverse-Engineering Long-Refinement Graphs

We now collect structural insights about long-refinement graphs with *arbitrary* degrees. We first treat the last iterations of Colour Refinement to deduce the connections between the partition classes in those iterations where we are guaranteed that all classes have size at most 2. We then reverse-engineer the previous, larger classes and their connections by analysing unions of the pairs and singletons. The structural insights allow us to deduce helpful facts about the arrangement of pairs and singletons, yielding necessary conditions for a graph to qualify as a long-refinement graph. We will use those in the Section 6 to show that long-refinement graphs with small degrees can be represented as strings and that the only strings whose graphs satisfy the necessary conditions are the ones found and presented in [25]. We also extend the classification to maximum degree 4 there.

Let G be a long-refinement graph of order $|G| > 1$, and let p be the minimal index of an iteration of Colour Refinement in which every element in the partition π^p has size at most 2, i.e. π^p contains only pairs and singletons. Note that $p < \infty$ since Colour Refinement must produce the discrete partition on G . Let $\pi_{\mathcal{S}}^p := \{M \in \pi^p \mid |M| = 1\}$ be the singletons in π^p , and $\pi_{\mathcal{P}}^p := \{M \in \pi^p \mid |M| = 2\}$ be the pairs in π^p . Note that for all $i \geq p$, the class C_i that splits in iteration $i + 1$ must be a pair $P \in \pi_{\mathcal{P}}^p$. This gives us a linear order $\prec$ on $\pi_{\mathcal{P}}^p$ by the iterations in which pairs split, and we denote the i -th element of $\prec$, the pair C_{p+i-1} , by P_i . Let $\alpha(P_i) := P_{i-1}$ and $\beta(P_i) := P_{i+1}$, respectively (for $i = \min(\prec)$, we let $\alpha(P_i) := \perp$, and for $i = \max(\prec)$, we let $\beta(P_i) := \perp$). We also call P, P' *consecutive pairs* if $P' = \beta(P)$. We let $n_{\mathcal{P}} := |\pi_{\mathcal{P}}^p|$ be the number of pairs.

We now analyse the structure of long-refinement graphs, in particular how the elements in $\pi_{\mathcal{P}}^p$ are connected. In the remainder of this paper, by a *pair*, we mean an element of $\pi_{\mathcal{P}}^p$. As the following

lemma states, consecutive pairs are always connected via a matching. The proof is identical to the proof of the same statement for graphs G with $\deg(G) = \{2, 3\}$, which is [25, Corollary 15].

Lemma 5.1. *For all $i \in [n_{\mathcal{P}} - 1]$, $\deg_{P_i}(P_{i+1}) = \deg_{P_{i+1}}(P_i) = 1$.*

Recall that C_p is the class that splits in iteration $p + 1$ and therefore is equal to P_1 , and that by definition, we have $A_p, B_p \in \pi^p$. By the choice of p , both A_p and B_p have cardinality at most 2.

By Lemma 4.4, we know $\deg_{C_{p-1}}(P_1) \neq \perp$ and $\deg_{A_p}(P_1) = \perp = \deg_{B_p}(P_1)$. Furthermore, a pair P is always balanced wrt itself, since $\deg_P(P) \in \{0, 1\}$. We obtain the following proposition.

Proposition 5.2. $P_1 \notin \{A_p, B_p\}$.

We now show that A_p and B_p are in fact pairs.

Lemma 5.3. $A_p, B_p \in \pi_{\mathcal{P}}^p$.

Proof. If $|A_p| = |B_p| = 1$, then $|C_{p-1}| = |A_p \cup B_p| = 2$, contradicting the minimality of p .

Let us next assume that $|A_p \cup B_p| = 3$, i.e. they are a singleton and a pair. Without loss of generality, let $A_p = \{s\}$ be the singleton. By Lemma 4.4, we know $\deg_{C_{p-1}}(P_1) \neq \perp$ and $\deg_{A_p}(P_1) = \perp = \deg_{B_p}(P_1)$. The latter yields that $G[A_p, P_1]$ contains a single edge. With Proposition 5.2 and since P_1 is the only class that splits in iteration $k + 1$, Lemma 4.4 gives us that B_p is balanced wrt P_1 , i.e. $\deg_{P_1}(B_p) \neq \perp$. Together with $\deg_{B_p}(P_1) = \perp$, this yields that the two vertices in B_p have the same (unique) neighbour in P_1 . But then $\deg_{C_{p-1}}(P_1) = \perp$, a contradiction.

Therefore, the only remaining possibility is that $|A_p| = |B_p| = 2$, which concludes the proof. $\square$

Let a and b be the positions of A_p and B_p , respectively, in $\prec$. That is, $A_p = P_a$ and $B_p = P_b$. By Proposition 5.2, we have $a > 1$ and $b > 1$. We can assume $b > a$ by symmetry. By Lemma 5.1, $\deg_{P_2}(P_1) = 1$ and $\deg_{A_p}(P_1) = \perp = \deg_{B_p}(P_1)$, and therefore $P_2 \notin \{P_a, P_b\}$. Thus, with Proposition 5.2, we obtain that $n_{\mathcal{P}} \geq 4$.

The following lemma describes how the first pair in $\prec$ is connected to the last classes that cause its splitting in iteration p . It generalises the same statement from degrees 2 and 3 to arbitrary degrees.

Lemma 5.4 (cf. [25], Lemma 16). *There exists $u \in P_1$ such that $G[\{u\}, P_a]$ and $G[P_1 \setminus \{u\}, P_b]$ are complete bipartite, and $E(G[\{u\}, P_b]) = E(G[P_1 \setminus \{u\}, P_a]) = \emptyset$.*

Proof. Recall again, by Lemma 4.4, that $\deg_{A_p}(P_1) = \perp = \deg_{B_p}(P_1)$. In particular, there is a vertex $u \in P_1$ such that $N(u) \cap P_a \neq \emptyset$. Furthermore, by Proposition 5.2 and again with Lemma 4.4, we must have $\deg_{P_1}(P_a) \neq \perp \neq \deg_{P_1}(P_b)$. The first inequality lets us conclude that $G[\{u\}, P_a]$ is complete bipartite and that $E(G[P_1 \setminus \{u\}, P_a]) = \emptyset$. With $\deg_{C_{p-1}}(P_1) \neq \perp$, we obtain the claimed statements about $G[P_1, P_b]$. $\square$

We already know that consecutive pairs are connected via a matching. The following lemma says that non-consecutive pairs (except for the case of P_a and P_b) as well as a pair and a singleton form either an empty or a complete bipartite graph, i.e. their connections are trivial.

Lemma 5.5. *Let P_i and P_j be pairs with $i + 1 < j$, and, if $i = 1$, then $j \notin \{a, b\}$. Let $S \in \pi_{\mathcal{S}}^p$. Then $\deg_S(P_i) \in \{0, 1\}$ and $\deg_{P_i}(P_j) = \deg_{P_j}(P_i) = \deg_{P_i}(S) \in \{0, 2\}$.*

Proof. Note that Lemma 5.3 says that C_{p-1} splits into two pairs and hence does not introduce new singletons in iteration p , i.e. $\pi_{\mathcal{S}}^p \subseteq \pi^{p-1}$. Also $\pi_{\mathcal{P}}^p \subseteq \pi^p$, so $S \in \pi^{p-1}$ and $P_i \in \pi^p$. Thus, $\perp \neq \deg_S(P_i) \in \{0, 1\}$ or, equivalently, $\deg_{P_i}(S) = \{0, 2\}$.

For the second part, choose u and v such that $\{u, v\} = P_i$ and $A_{p+i} = \{u\}$, and $B_{p+i} = \{v\}$. Naturally, $\{A_{p+i}, B_{p+i}\} \subseteq \pi^{p+i}$ and, since $j > i + 1$, $P_j \in \pi^{p+i+1}$. Therefore, $\deg_{A_{p+i}}(P_j) \neq \perp \neq \deg_{B_{p+i}}(P_j)$. By the fact that $|A_{p+i}| = |B_{p+i}| = 1$, we have that each of $\deg_{A_{p+i}}(P_j)$ and $\deg_{B_{p+i}}(P_j)$ is either 0 or 1. Now, we show that $\deg_{A_{p+i}}(P_j) = \deg_{B_{p+i}}(P_j)$. If $i = 1$, then by assumption $P_j \notin \{P_a, P_b\}$ and thus $P_j \in \pi^{p-1}$, while $P_i \in \pi^p$. If $i \neq 1$, then $P_j \in \pi^p$ and $P_i \in \pi^{p+1}$. In either case, since P_j splits later than P_i , this gives us $\deg_{P_i}(P_j) \neq \perp$ and therefore $\deg_{A_{p+i}}(P_j) = \deg_{B_{p+i}}(P_j) \in \{0, 1\}$. This tells us that either $G[P_i, P_j]$ is complete bipartite or $E(G[P_i, P_j]) = \emptyset$. $\square$

The following lemma states that the distance in $\prec$ between P_a and P_b , i.e. between the classes that induce the splitting of P_1 , is at most 2.

Lemma 5.6. $b \in \{a + 1, a + 2\}$.

Proof. Assume that $b > a + 2$. Then consider the set of distinct pairs $\Pi := \{P_{a-1}, P_{a+1}, P_{b-1}\}$. For each $P \in \Pi$, it holds that exactly one set of $\{P_a, P\}$ and $\{P_b, P\}$ consists of consecutive pairs. Therefore, by Lemmas 5.1 and 5.5, $C_{p-1} = P_a \cup P_b$ is unbalanced wrt P . Thus, the first iteration i for which any pair $P \in \Pi$ is in π^i must directly precede the iteration p in which C_{p-1} splits. We have that $|\pi^i \setminus \pi^{i-1}| = 2$ and $(\pi^p \setminus \pi^{p-1}) \cap \Pi = \emptyset$, so there must be a $P \in \Pi$ that is not in π^p , a contradiction. Thus, $a < b \leq a + 2$. $\square$

In the previous lemmas, we have analysed the splits that happen from the first moment on where the largest class has size 2. Our structural insights have revealed the relations between the pairs in π^p . Now we will work back in time from π^p , i.e. starting from π^p , we reverse-construct the partitions π^{p-i} for increasing i . Recall that a and b are the positions of A_p and B_p in the splitting order $\prec$ of the pairs, whose collection is π_p^p with size $n_p = |\pi_p^p|$, and that π_S^p is the collection of singletons in π^p . Define $\ell_{4 \rightarrow 2} := \min\{n_p - b, a - 1\}$. Then the technical statement in the lemma below expresses that each of the $\ell_{4 \rightarrow 2}$ iterations preceding π^p splits one class of four vertices into two pairs, which will be elements of π^p .

Lemma 5.7. For every $h \in [0, \ell_{4 \rightarrow 2}]$, the following holds:

$$\pi^{p-h-1} = \pi_S^p \cup \{P_{a-i} \cup P_{b+i} \mid i \in [0, h]\} \cup \{P_i \mid i \in [1, a - h - 1] \cup [a + 1, b - 1] \cup [b + h + 1, n_p]\}.$$

Proof. We prove the statement by induction on h . The case for $h = 0$ follows from Lemma 5.3 and the fact that $\pi^{p-1} = \pi^p \setminus \{A_p, B_p\} \cup \{A_p \cup B_p\}$. This is illustrated in Figure 4.

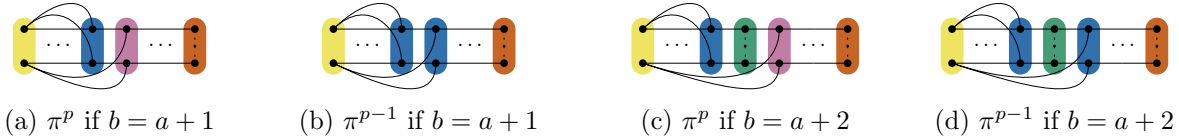


Figure 4: π^{p-1} when $b < n_p$. P_a is blue, P_b is pink and blue, respectively. For clarity, we omit all edges between non-consecutive pairs.

For the following arguments, see Figure 5 for a visualisation of the situation. Assume $h \in [\ell_{4 \rightarrow 2}]$ and suppose the claim holds for all values in $[0, h - 1]$. We show that it also holds for h then.

The definition of $\ell_{4 \rightarrow 2}$ ensures that P_{a-h} and P_{b+h} are well-defined, and the induction hypothesis for $h - 1$ ensures that they, as well as $P_{a-h+1} \cup P_{b+h-1}$, are elements of π^{p-h} .

By Lemma 5.1, $\deg_{P_{a-h}}(P_{a-h+1}) = 1 = \deg_{P_{b+h}}(P_{b+h-1})$. By Lemma 5.5, $\deg_{P_{b+h}}(P_{a-h+1})$ and $\deg_{P_{a-h}}(P_{b+h-1})$ are each either 0 or 2. But this implies that $P_{a-h+1} \cup P_{b+h-1}$ is unbalanced wrt

each of P_{a-h} and P_{b+h} . So, by Lemma 4.4, it follows that P_{a-h} and P_{b+h} must be in $\pi^{p-h} \setminus \pi^{p-h-1}$. That is, $\{P_{a-h}, P_{b+h}\} = \{A_{p-h}, B_{p-h}\}$, and hence $A_{p-h} \cup B_{p-h} = C_{p-h-1} \in \pi^{p-h-1}$. Moreover, Equation 1 yields that $\deg_{P_{b+h}}(P_{a-h+1}) = \deg_{P_{a-h}}(P_{b+h-1})$. Hence, $\pi^{p-h-1} = \{P_{a-h} \cup P_{b+h}\} \cup (\pi^{p-h} \setminus \{P_{a-h}, P_{b+h}\})$, and thus the desired result holds. $\square$



Figure 5: Induction step for the proof of Lemma 5.7 in the case that $b = a + 1$. The arrangement of the pairs follows the splitting order $\prec$, i.e. a pair further left has a lower index in $\prec$. The unique new pairs in π^{p-h+1} (in pink and brown) are in symmetric position to P_a and P_b (in blue) and united to a class of size 4 (in brown) in π^{p-h} .

For the purpose of better readability, let $\ell := \ell_{4 \rightarrow 2}$ in the remainder of this section. The above allows us to describe the partition $\pi^{p-\ell-1}$. With the following lemma, we describe how the classes in $\pi^{p-\ell-1}$ are connected to each other. By pattern matching with the classes in Lemma 5.7, the reader will be able to verify that it states the following. The classes of size 4 which split into pairs that are in π^p are connected non-trivially if and only if they split in consecutive iterations.

Lemma 5.8. *For every $h \in [0, \ell - 1]$, the following holds:*

$$\deg_{P_{a-h} \cup P_{b+h}}(P_{a-h-1} \cup P_{b+h+1}) = \deg_{P_{a-h-1} \cup P_{b+h+1}}(P_{a-h} \cup P_{b+h}) \in \{1, 3\}. \quad (2)$$

Moreover, for every $h \in [0, \ell - 1]$ and every $C \in \pi^{p-h-2} \setminus \{C_{p-h}, C_{p-h-1}, C_{p-h-2}\}$, the following hold:

$$\deg_C(P_{a-h} \cup P_{b+h}) \in \{0, |C|\} \quad \text{and} \quad \deg_{P_{a-h} \cup P_{b+h}}(C) \in \{0, 4\}. \quad (3)$$

Proof. Equation 2 follows directly from the proof of Lemma 5.7. Statement 3 follows from the fact that $C_{p-h-1} = P_{a-h} \cup P_{b+h}$ is in π^{p-h-1} and thus balanced wrt all classes in π^{p-h-2} . If $|C| = 1$, this is sufficient to prove $\deg_C(P_{a-h} \cup P_{b+h}) \in \{0, 1\}$ and the second, equivalent statement. If C is a pair, then because $P_{a-h-1}, P_{a-h+1}, P_{b+h-1}, P_{b+h+1}$ are not in π^{p-h-2} , Lemma 5.5 implies that $\deg_C(P_{a-h}) = \deg_C(P_{b+h}) \in \{0, 2\}$, giving the required result. Finally, if $|C| = 4$, then $C = P_{a-h'} \cup P_{b+h'}$ for some $h' \in [0, h - 2]$, and the above argument implies $\deg_C(P_{a-h}) = \deg_C(P_{b+h}) \in \{0, 4\}$, giving the desired result. $\square$

With the insights so far, we have collected information about the classes and their relations from the point on when all remaining classes of size 4 split consecutively into pairs until only pairs and singletons are left, with the pairs then consecutively splitting into singletons until a discrete partition is reached. We intend to carry this one step further and to analyse the split(s) before this moment. We are going to prove that $C_{p-\ell-2}$ is either a union of three pairs or a union of a certain pair and a singleton $S \in \pi_S^p$. Here, we need to perform a case distinction, and we introduce an auxiliary variable t to deal with the cases in parallel. We let $t := 0$ in the first case, and $t := 1$ in the second. Recall that p is the index of the first iteration in which all classes have size at most 2, and $p - \ell$ is the index of the first iteration in a cascade of splits of classes of size 4 into pairs (until iteration p). Depending on the case whether $t = 0$ or $t = 1$, we let $q(t) := p - \ell - 1 - t$, which will be the first iteration in which there are classes of size at most 4 (**quadruples**). Define $\pi_S^{q(t)} := \{M \in \pi^{q(t)} \mid |M| = 1\}$. We also define $c := a - \ell$, and $d := c - t - 1$ to simplify notation, which will play a similar role in future proofs that a and b did in proofs up to this point.

Lemma 5.9. $C_{p-\ell-2}$ is either $P_{c-1} \cup P_c \cup P_{n_p}$ or $P_{c-1} \cup S$ for a singleton $S \in \pi_S^p$. In the first case, set $t := 0$, in the second $t := 1$. Then,

$$\begin{aligned} \pi^{q(t)-1} = \pi_S^{q(t)} \cup \{P_{a-i} \cup P_{b+i} \mid i \in [0, \ell-1]\} \cup \{P_d \cup P_c \cup P_{n_p}\} \\ \cup \{P_i \mid i \in [c-t-2] \cup [a+1, b-1]\} \cup M_t, \end{aligned}$$

and also $M_0 = \emptyset$ and $M_1 = \{P_{c-1} \cup S\}$.

Proof. We first prove by contradiction that P_{c-1} exists. Assume that $\ell = a - 1$, i.e. $P_c = P_1$. Since $P_2 \prec P_a$, ℓ must be greater than 1, allowing us to apply Lemma 5.8, along with Lemma 5.4, to get $\deg_{P_a \cup P_b}(P_c) - \deg_{P_a \cup P_b}(P_{b+\ell}) \in \{-2, 2\}$. If $P_{b+\ell} \neq P_{n_p}$, then we have $\deg_{P_{b+\ell+1}}(P_c) - \deg_{P_{b+\ell+1}}(P_{b+\ell}) \in \{-1, 1\}$, by Lemmas 5.1 and 5.5. This contradicts Lemma 4.4. Assume then that $P_{b+\ell} = P_{n_p}$, implying there are no pairs in $\pi^{p-\ell-1}$. By Lemma 4.4, there must be a class C which satisfies $\deg_C(P_{a-\ell}) - \deg_C(P_{b+\ell}) \in \{-2, 2\}$. By Lemma 5.8, there are no classes of size 4 that can satisfy this. Thus, we must have $P_1 \prec P_{a-\ell}$ and $\ell = n_p - b$.

Then, $C_{p-\ell-1}$ is unbalanced by P_{c-1} , by Lemmas 5.5 and 5.1. Note that we have $\deg_{P_{c-1}}(P_c) - \deg_{P_{c-1}}(P_{n_p}) \in \{-1, 1\}$. Next, we look for a class $B_{p-\ell-1}$ with $\deg_{B_{p-\ell-1}}(P_c) - \deg_{B_{p-\ell-1}}(P_{b+\ell}) \in \{-1, 1\}$. By Lemma 5.8, no class $P_{a-h} \cup P_{b+h} \in \pi^{p-\ell-1}$ with $h < \ell$ can satisfy this requirement. By Lemma 5.5, no pair other than P_{c-1} can satisfy this equation. This leaves only $\pi_S^p \cup \{P_c \cup P_{n_p}\}$, proving the first statement of Lemma 5.9. This immediately gives us the desired result for $\pi^{q(t)-1}$ for $t = 0$.

For $t = 1$, the above yields $C_{q(t)} = P_{a-\ell-1} \cup S$, from which we can find $\pi^{q(t)}$, but we must still prove that $C_{q(t)-1}$ is $P_d \cup P_c \cup P_{n_p}$, in order to get the desired expression for $\pi^{q(t)-1}$. First, notice that $\deg_{P_c \cup P_{n_p}}(S) = 2$ and $\deg_{P_c \cup P_{n_p}}(P_{c-1}) \in \{1, 3\}$. Thus, without loss of generality, $A_{q(t)} = P_c \cup P_{n_p}$ and $B_{q(t)}$ satisfies $\deg_{B_{q(t)}}(S) - \deg_{B_{q(t)}}(P_{c-1}) \in \{-1, 1\}$. If $P_{c-1} = P_1$, then $P_{c-1} \cup S$ is unbalanced wrt $P_a \cup P_b$, but $P_a \cup P_b$ does not satisfy the above equation, so $P_1 \prec P_{c-1}$. Thus, P_{c-2} is well-defined and, by Lemmas 5.1 and 5.8, $\deg_{P_{c-2}}(P_{c-1}) = 1$ and $\deg_{P_{c-2}}(S) \in \{0, 2\}$. This satisfies the requirement for $B_{q(t)-1}$, so $\pi^{q(t)} = \{P_d \cup P_c \cup P_{n_p}\} \cup (\pi^{p-\ell-1} \setminus \{P_d, P_c \cup P_{n_p}\})$. $\square$

In the case that $t = 1$, we will let $S_{q(t)}$ denote the singleton S described in Lemma 5.9. Notice that $\pi_S^{q(t)} = \pi_S^p$ if $t = 0$, and $\pi_S^{q(t)} = \pi_S^p \setminus \{S_{q(t)}\}$ if $t = 1$. Also, for either value of t , let M_t be defined as in the lemma.

Having defined $C_{q(t)}$, where $t = 1$, we now make some observations about its neighbourhood.

Lemma 5.10. Suppose that $t = 1$, i.e. that $C_{p-\ell-2} = S_{q(t)} \cup P_{c-1}$. Then

$$\deg_{S_{q(t)} \cup P_{c-1}}(P_d \cup P_c \cup P_{n_p}) \in \{1, 2\}$$

and

$$\deg_{P_d \cup P_c \cup P_{n_p}}(S_{q(t)} \cup P_{c-1}) = 2 \cdot \deg_{S_{q(t)} \cup P_{c-1}}(P_d \cup P_c \cup P_{n_p}) \in \{2, 4\}.$$

Furthermore,

$$\deg_C(S_{q(t)} \cup P_{c-1}) \in \{0, |C|\} \quad \text{and, equivalently,} \quad \deg_{S_{q(t)} \cup P_{c-1}}(C) \in \{0, 3\}$$

for every $C \in \pi^{q(t)-\ell-2} \setminus \{S_{q(t)} \cup P_{c-1}, P_d \cup P_c \cup P_{n_p}\}$.

Proof. We get the values for $\deg_{P_d \cup P_c \cup P_{n_p}}(S_{q(t)} \cup P_{c-1})$ directly from the ones for $\deg_{P_c \cup P_{n_p}}(S_{q(t)})$, $\deg_{P_c \cup P_{n_p}}(P_{c-1})$, $\deg_{P_d}(S_{q(t)})$, and $\deg_{P_d}(P_{c-1})$. Since $P_d \cup P_c \cup P_{n_p}$ has double the vertices as

$S_{q(t)} \cup P_{c-1}$, is incident to the same number of edges, and both are balanced with respect to the other, we obtain $\deg_{S_{q(t)} \cup P_{a-\ell-1}}(P_d \cup P_c \cup P_{n_p}) \in \{1, 2\}$.

The second statement holds by the fact that $S_{q(t)} \cup P_{c-1}$ is balanced wrt all classes $C \in \pi^{p-\ell-1} \setminus \{P_d, P_c \cup P_{n_p}\}$. Separately, $S_{q(t)}$ and P_{c-1} are trivially attached to the classes in $\pi^{p-\ell-1} \setminus \{P_d, P_c \cup P_{n_p}\}$. This holds by Lemma 5.5 for singletons and pairs, and by Lemma 5.8 for classes of four vertices. Since $S_{q(t)} \cup P_{c-1}$ is balanced wrt C , it holds that if $S_{q(t)}$ is adjacent to C , then P_{c-1} is as well, and vice versa. Thus $S_{q(t)} \cup P_{c-1}$ is trivially attached to C . $\square$

By a similar argument to that used in Lemma 5.7, we can work backwards from partition $\pi^{q(t)}$. Recall that $d = a - \ell - t - 1$. We define $\ell' := \min\{\ell, d - 1\}$. This definition guarantees that P_{d-h} and $P_{c+h} \cup P_{n_p-h}$ are elements of every $\pi^{q(t)-h}$ for $h \in [\ell']$.

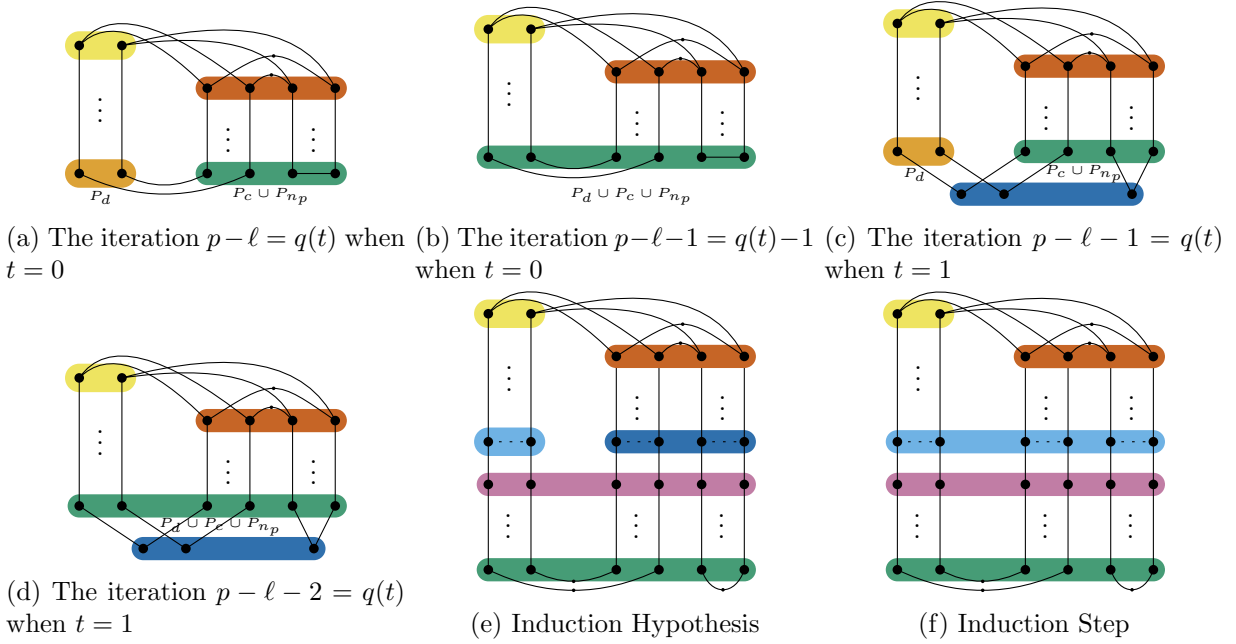


Figure 6: Diagrams depicting the base case and induction arguments for Lemma 5.11.

Lemma 5.11. For all $h \in [0, \ell']$,

$$\begin{aligned} \pi^{q(t)-h-1} = & \pi_S^{q(t)} \cup \{P_{c+i} \cup P_{n_p-i} \mid i \in [h+1, \ell']\} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_p-i} \mid i \in [0, h]\} \\ & \cup \{P_i \mid i \in [d-h-1] \cup [a+1, b-1]\} \cup M_t. \end{aligned}$$

Proof. By using the relations between the variables a, b, c, d , the base case of $h = 0$ is given by Lemma 5.9. Consider $h \in [\ell']$, with the above statement holding up to $h - 1$. Then, by Lemmas 5.1, 5.8, and 4.4, and by the same argument as used in Lemma 5.7, we obtain $C_{q(t)-h-1} = P_{d-h} \cup P_{c+h} \cup P_{n_p-h}$. $\square$

Corollary 5.12. For $h \in [0, \ell' - 1]$,

$$\deg_{C_{q(t)-h-2}}(C_{q(t)-h-1}) = \deg_{C_{q(t)-h-1}}(C_{q(t)-h-2}) \in \{1, 5\}. \quad (4)$$

Also,

$$\deg_{C_{q(t)-h-1}}(C_{q(t)-h-1}) \in \{0, 1, 4, 5\}. \quad (5)$$

For every $h \in [0, \ell' - 1]$ and $C \in \pi^{q(t)-h-2} \setminus \{C_{q(t)-h-2}, C_{q(t)-h-1}, C_{q(t)-h}\}$, it holds that

$$\deg_{C_{q(t)-h-1}}(C) \in \{0, 6\} \text{ and } \deg_C(C_{q(t)-h-1}) \in \{0, |C|\}. \quad (6)$$

Proof. Using Equation 1, we obtain

$$\perp \neq \deg_{C_{q(t)-h-1}}(C_{q(t)-h-2}) \in \{1, 5\}$$

and hence Equation 4.

To see Equation 5, recall that $C_{q(t)-h-1} \in \pi^{q(t)-h-1}$ is balanced wrt $C_{q(t)-h-1} \in \pi^{q(t)-h-2}$. Also P_{d-h} and $P_{c+h} \cup P_{n_p-h}$ are trivially attached, from Lemma 5.8. There is also $\deg_{P_{d-h}}(P_{d-h}) \in \{0, 1\}$. These ingredients together give $\deg_{C_{q(t)-h-1}}(P_{d-h}) \in \{0, 1, 4, 5\}$ and therefore, by balancedness $\deg_{C_{q(t)-h-1}}(C_{q(t)-h-1}) \in \{0, 1, 4, 5\}$.

Equation 6 follows from the fact that $P_{a-h} \cup P_{b+h} = C_{q(t)-h-1} \in \pi^{q(t)-h-1}$, and it is thus balanced wrt all classes in $\pi^{q(t)-h-2}$. If C is a singleton, this is sufficient to show $\deg_C(P_{a-h} \cup P_{b+h}) \in \{0, 1\}$ and the second, equivalent statement. If C is a pair, then by assumption $C \notin \{P_{a-h-1}, P_{a-h+1}, P_{b+h-1}, P_{b+h+1}\}$. Hence, Lemma 5.5 implies that $\deg_C(P_{a-h}) = \deg_C(P_{b+h}) \in \{0, 2\}$, giving the required result. Finally, if $|C| = 4$, then $C = P_{a-h'} \cup P_{b+h'}$ for some $h' \in [0, h-2]$ and the above argument implies $\deg_C(P_{a-h}) = \deg_C(P_{b+h}) \in \{0, 4\}$, giving the desired result. $\square$

Recall that by Lemma 5.6, we know that $b \in \{a+1, a+2\}$. Depending on which of the two situations occurs, the following lemma establishes the possible ranges for d . We will use this insight, and a case distinction on the values of d , in Section 6 to establish the existence and non-existence of long-refinement graphs and to classify them.

Lemma 5.13. *If $b = a+1$, then $d \in [\ell, \ell+4]$, and if $b = a+2$, then $d \in [\ell, \ell+2]$.*

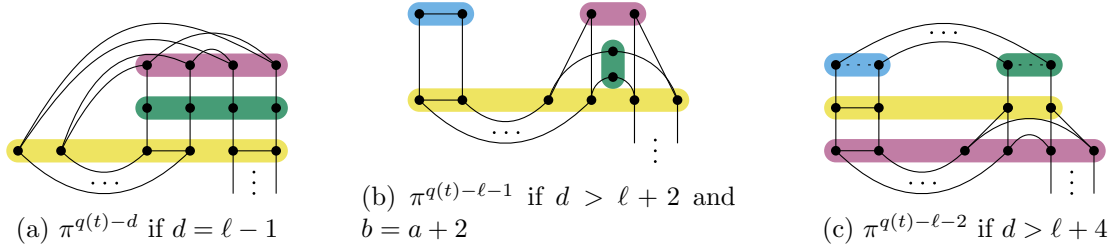


Figure 7: Examples of contradictions occurring in the proof of Lemma 5.13 for $d < \ell$, $d > \ell + 2$, and $d > \ell + 4$ when the degrees are 2 and 3. Notice that yellow classes are unbalanced wrt blue, green, and pink classes.

Proof. We prove this by showing that the assumption $d < \ell$ leads to a contradiction, as does $d > \ell + 2$ when $b = a + 2$ and $d > \ell + 4$ when $b = a + 1$. For the following, consider Figure 7.

Assume that $1 \leq d \leq \ell - 1$. Then $P_1 = P_{d-h}$, for some $h \leq \ell - 2$. Recall that $C_{q(t)-h-1} = P_1 \cup P_{c+h} \cup P_{n_p-h}$. See the lowest class in Figure 7a, which is unbalanced wrt all three classes. We have $\deg_{P_a \cup P_b}(P_1) = 2$. Since P_{c+h} and P_{n_p-h} are not consecutive to P_a or P_b , we have $\deg_{P_a \cup P_b}(P_{c+h} \cup P_{n_p-h}) \in \{0, 4\}$. Here, we have a difference of 2. We also know that

$$\deg_{P_{c+h+1} \cup P_{n_p-h-1}}(P_{c+h} \cup P_{n_p-h}) \in \{1, 3\},$$

whereas

$$\deg_{P_{c+h+1} \cup P_{n_p-h-1}}(P_1) \in \{0, 4\}.$$

We cannot get a difference of 2 from these values, hence we have a contradiction to Lemma 4.4, proving $d \geq \ell$.

Next, assume that $d \geq \ell + 3$ and $b = a + 2$. Then, $C_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$, with $d - \ell \geq 3$. See Figure 7b, where the yellow class $C_{q(t)-\ell-1}$ is unbalanced wrt the green class P_{a+1} , the blue class $P_{d-\ell-1}$ and the pink class P_1 . Formally, since $b = a + 2$, $\deg_{P_{a+1}}(P_a \cup P_b) = 1$ and, by Lemma 5.5, $\deg_{P_{a+1}}(P_{d-\ell}) \in \{0, 2\}$. Next, we have $\deg_{P_1}(P_a \cup P_b) = 1$ and, by Lemma 5.5, $\deg_{P_1}(P_{d-\ell}) \in \{0, 2\}$. Finally, we have $\deg_{P_{d-\ell-1}}(P_a \cup P_b) \in \{0, 2\}$ and $\deg_{P_{d-\ell-1}}(P_{d-\ell}) = 1$. Thus, as we have $C_{q(t)-\ell-1}$ unbalanced wrt at least three classes, we obtain a contradiction to Lemma 4.4 and have proved $d \leq \ell + 2$ for $b = a + 2$.

Finally, assume that $b = a + 1$ and, towards a contradiction, that $d \geq \ell + 5$. See Figure 7c. As before, we have $\deg_{P_1}(P_a \cup P_b) = 1$ and with Lemma 5.5, $\deg_{P_1}(P_{d-\ell}) \in \{0, 2\}$. Also, $\deg_{P_{d-\ell-1}}(P_a \cup P_b) \in \{0, 2\}$ and $\deg_{P_{d-\ell-1}}(P_{d-\ell}) = 1$. We must therefore have $C_{q(t)-\ell-2} = P_1 \cup P_{d-\ell-1}$. Then, $C_{q(t)-\ell-2}$ is unbalanced wrt P_2 , $P_{d-\ell-1}$, and $C_{q(t)-\ell-1}$. This is a contradiction and therefore $d \leq \ell + 4$. $\square$

With the reverse-engineering methods in this section, we have managed to reconstruct partitions during the execution of Colour Refinement in which the classes have size up to 6. This could be carried further, but the analysis would become more and more involved. However, when restricting the degrees to a maximum of 4, these last iterations comprise enough of the entire execution of Colour Refinement to yield characterisations and compact representations for all long-refinement graphs with small degrees. We will see this in Section 6, where we apply the ingredients collected in this section.

6 Long-Refinement Graphs with Maximum Degree at most 4

In this section, we establish a comprehensive list of all the long-refinement graphs with maximum degree at most 4. We use the observation from Lemma 5.13 that $d \in [\ell, \ell + 4]$ and consider the cases separately. Unifying previous results [25], we include a string notation to represent the long-refinement graphs G with $\deg(G) = \{2, 3\}$, where each letter represents a pair.

Notation 6.1. We define $\sigma: [n_{\mathcal{P}}] \rightarrow \Sigma$ as the function that maps i to the element in Σ that matches the type of P_i . More precisely, $\sigma(i)$ is defined as follows:

- S, if $P_i = \min(\prec) = P_1$.
- X, if $P_i \in \{P_a, P_b\}$ and P_i is not adjacent to any singleton $S \in \pi_S^p$.
- X₂, if $P_i \in \{P_a, P_b\}$ and P_i is adjacent to a singleton $S_i \in \pi_S^p$.
- 0, if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 2$ and P_i is not adjacent to any singleton $S \in \pi_S^p$.
- 0₂, if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 2$, and P_i is adjacent to a singleton $S_i \in \pi_S^p$.
- 1, if $P_i \notin \{P_1, P_a, P_b\}$ with $\deg(P_i) = 3$, such that P_i is not adjacent to any $S \in \pi_S^p$.
- 1₂, if $P_i \notin \{P_1, P_a, P_b\}$, $\deg(P_i) = 3$, and P_i is adjacent to a singleton $S_i \in \pi_S^p$.

Since $\prec$ is a linear order, we can define a string representation based on a long-refinement graph G , where the i -th character in the string is $\sigma(i)$ and corresponds to the i -th element of $\prec$, P_i . We call a string defined in this way a long-refinement string.

The strings describe the exact neighbourhood of each pair and each singleton, as long as there is at most one singleton. We will prove that this is the case. Thus, every string Ξ unambiguously represents exactly one graph, up to isomorphism, which we denote by $G(\Xi)$.

The limitation on degree allows for some immediate insights into the structure of long-refinement graphs. For example, Corollary 5.12 yields the following when the maximum degree is 4.

Observation 6.2. For $h \in [1, \ell' - 1]$, $C_{q(t)-h-1}$ (i.e. the class $P_{d-h} \cup P_{c+h} \cup P_{n_p-h}$) has degree 2 or 3, with $N(C_{q(t)-h-1}) \subseteq C_{q(t)-h-2} \cup C_{q(t)-h-1} \cup C_{q(t)-h}$.

With $h = 0$, Corollary 5.12 also describes the neighbourhood of $C_{q(t)-1}$ (i.e. $P_d \cup P_c \cup P_{n_p}$), which depends on the value of t . Figure 8 illustrates the possible edge layouts between $C_{q(t)}$ and $C_{q(t)-1}$ when $t = 0$ or $t = 1$, which are a result of the proof of Lemma 5.9. This yields the following observation.

Observation 6.3. If $t = 0$, then $\deg(C_{q(t)-1}) = 2$ with

$$\deg_{C_{q(t)-2}}(C_{q(t)-1}) = 1 = \deg_{C_{q(t)-1}}(C_{q(t)-1}),$$

and $\deg_C(C_{q(t)-1}) = 0$ for every other class $C \subseteq V(G)$.

If $t = 1$, then $\deg(C_{q(t)-1}) \in \{2, 3, 4\}$ with

$$\deg_{C_{q(t)-2}}(C_{q(t)-1}) = 1, \quad \deg_{C_{q(t)-1}}(C_{q(t)-1}) \in \{0, 1\}, \quad \deg_{C_{q(t)}}(C_{q(t)-1}) \in \{1, 2\},$$

and $\deg_C(C_{q(t)-1}) = 0$ for every other class $C \subseteq V(G)$.

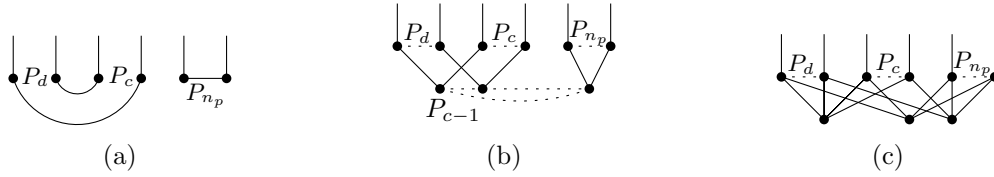


Figure 8: All possible layouts between $C_{q(t)-1}$ ($P_d \cup P_c \cup P_{n_p}$) and $C_{q(t)}$ ($P_c \cup P_{n_p}$ or $S_{q(t)} \cup P_{c-1}$) when the maximum degree is at most 4. Note

We begin by proving the following two technical lemmas, which will assist us in cases where ℓ can be arbitrarily large. Recall that $C_{q(t)-i-1} = P_{d-i} \cup P_{c+i} \cup P_{n_p-i}$ for $i \in [0, \ell']$ and $C_{q(t)-i-1} = P_{c-1} \cup S_{q(t)}$ for $i = -1$ when $t = 1$.

Lemma 6.4. Let j be an iteration $j > 0$ and let the following statements hold.

1. $\{C_{q(t)-i-1} \mid i \in [-t, \ell']\} \subseteq \pi^j$, with $\ell' + t > 0$.
2. B_j is $C_{q(t)-\ell'-1}$, and A_j is not contained in $\{C_{q(t)-i-1} \mid i \in [-t, \ell']\}$
3. $\deg_{C_{j-1}}(A_j) = \deg_{C_{j-1}}(B_j) + \deg_{C_{q(t)-\ell'}}(B_j)$, where $\deg_{C_{q(t)-\ell'}}(B_j) = 1$ (or 2 if $\ell' = 0$ and $t = 1$ as depicted in Figure 8c) and $\{\deg_{C_{j-1}}(A_j)\} = \deg(A_j)$.

Then, for $j' \in [0, \ell' + t]$, $\pi^{j-1-j'} \setminus \pi^j = \{C_{j-1-j'}\}$ and $C_{j-1-j'} = A_j \cup \left(\bigcup_{i \in [0, j']} C_{q(t)-\ell'+i-1} \right)$.

Proof. For $j' = 0$, we simply observe that $C_{j-1} = A_j \cup B_j$ is $A_j \cup C_{q(t)-\ell'-1}$, as required.

For $j' > 1$, we assume the statement holds up until $j' - 1$. We wish to show that

$$C_{j-j'} = C_{j-j'+1} \cup C_{q(t)-\ell'+j'-1}.$$

To do this, we prove that $C_{j-j'+1}$ is unbalanced wrt itself and $C_{q(t)-\ell'+j'-1}$. Since $A_j \subseteq C_{j-j'+1}$ and $N(A_j) \subseteq C_{j-1} \subseteq C_{j-j'+1}$, there are vertices $v \in A_j$ in $C_{j-j'+1}$ such that $\deg(v) = \deg_{C_{j-j'+1}}(v)$. The upper limit on j' gives $q(t) - \ell' + j' - 1 \leq q(t) - 1 + t$ so we can apply Corollary 5.12 and Lemma 5.10 to show that $\deg_{C_{q(t)-\ell'+j'-1}}(C_{q(t)-\ell'+j'-2}) = 1$ (or 2 if $\ell' = 0$, and $t = 1$ as depicted in Figure 8c). Since $C_{q(t)-\ell'+j'-2} \subseteq C_{j-j'+1}$, there are vertices u with $\deg_{C_{q(t)-\ell'+j'-1}}(u) = 1$ (or 2) and thus also $\deg_{C_{j-j'+1}}(u) < \deg(u)$. We thus get that $C_{j-j'+1}$ is unbalanced wrt itself and $C_{q(t)-\ell'+j'-1}$. By Lemma 4.4, this proves the statement for j' . $\square$

Lemma 6.5. *Let $j > 0$ be an iteration such that the following statements hold.*

1. $\{C_{q(t)-i-1} \mid i \in [-t, \ell']\} \subseteq \pi^j$, with $\ell' > 0$.
2. B_j is $C_{q(t)-\ell'-1}$, and A_j is not contained in $\{C_{q(t)-i-1} \mid i \in [-t, \ell']\}$.
3. There exists a class M with $\deg_{A_j}(M) = 2$, $\deg_M(A_j) = 1$, and $\deg_{B_j}(M) = \deg_M(B_j) = 0$.
4. $\{\deg_{C_{j-1}}(B_j) + \deg_{C_{q(t)-\ell'}}(B_j)\} = \deg(B_j)$ with $\deg_{C_{q(t)-\ell'}}(B_j) = 1$ (or 2 if $\ell' = 0$ and $t = 1$ as depicted in Figure 8c).

Then $\deg_{C_{q(t)-\ell'}}(B_j) = 1$ and, for $j' \in [0, \ell' - 1 + t]$, there are two classes $E_{j'}, D_{j'}$ in $\pi^{j-2-j'}$ which satisfy the following statements.

- $\pi^{j-2-j'} \setminus \pi^j = \{D_{j'}, E_{j'}\}$
- $D_{j'} = M \cup \bigcup_{i \in [0, \lfloor \frac{j'-1}{3} \rfloor]} C_{q(t)-\ell'+3i}$
- $E_{j'} = A_j \cup B_j \cup \left(\bigcup_{i \in [0, \lfloor \frac{j'-2}{3} \rfloor]} C_{q(t)-\ell'+1+3i} \right) \cup \left(\bigcup_{i \in [0, \lfloor \frac{j'-3}{3} \rfloor]} C_{q(t)-\ell'+2+3i} \right)$

Proof. We begin with the base case, $j' = 0$, by finding an expression for π^{j-2} . Given the assumed values of A_j and B_j , we can see that $\pi^{j-1} = \pi^j \setminus \{A_j, B_j\} \cup \{A_j \cup B_j\}$. Then, by Assumption 3, the class $C_{j-1} = A_j \cup B_j$ is clearly unbalanced wrt M . Since $\ell' + t > 0$, the class $C_{q(t)-\ell'}$ is in π^{j-1} , which implies that C_{j-1} is unbalanced wrt $C_{q(t)-\ell'}$. Thus, by Lemma 4.4, $C_{j-2} = M \cup C_{q(t)-\ell'}$. It is therefore apparent that $\pi^{j-2} \setminus \pi^j = \{C_{j-1}, C_{j-2}\}$ and that letting $D_0 := C_{j-2}$ and $E_0 := C_{j-1}$ satisfies all desired statements. Note that this also contradicts $\deg_{C_{q(t)-\ell'}}(B_j) = 2$, as it would not satisfy Lemma 4.4.

We next consider the induction step, where $j \in [1, \ell' - 1 + t]$. We assume that the statements hold up to $j' - 1$ (i.e. $\pi^{j-1-j'}$ and $D_{j'-1}, E_{j'-1}$ are as described) and prove the same for j' (i.e. $\pi^{j-2-j'}$ and $D_{j'}, E_{j'}$). We must consider three options separately: a) $C_{j-1-j'} = D_{j'-1}$, i.e. when $j' - 1 \equiv 1 \pmod{3}$; b) $C_{j-1-j'} = E_{j'-1}$ and $C_{j-j'} = D_{j'-1}$, i.e. $j' - 1 \equiv 2 \pmod{3}$; and c) $C_{j-1-j'} = E_{j'-1}$ with $C_{j-j'} \subseteq E_{j'-1}$ when $j' - 1 \equiv 0 \pmod{3}$.

In all three cases, since j' is no greater than $\ell' - 1 + t$, we may apply Corollary 5.12 and Lemma 5.10 to observe that $C_{q(t)-\ell'+j'-2}$ has degree 1 (or 2 in the case where $t = 1$ as depicted in Figure 8c), into $C_{q(t)-\ell'+j'-1}$. Therefore, $C_{j-1-j'}$ is unbalanced wrt $C_{q(t)-\ell'+j'-1}$, which we may denote $A_{j-1-j'}$. To construct $\pi^{j-2-j'}$, we need $B_{j-1-j'}$. The analysis differs by case, with arguments presented below for each.

In the first case, where $C_{j-1-j'} = D_{j'-1}$, we have by assumption $\deg_{E_{j'-1}}(M) = 2$ whereas $\deg_{E_{j'-1}}(C_{q(t)-\ell'+j'-2}) = 1$. This gives us that $B_{j-1-j'} = E_{j'-1}$ and thus

$$\begin{aligned} D_{j'} &= D_{j'-1}, \\ E_{j'} &= E_{j'-1} \cup C_{q(t)-\ell'+j'-1}, \end{aligned}$$

which is what we expect for $j' \equiv 2 \pmod{3}$, and also $\pi^{j-2-j'} \setminus \pi^j = \{D_{j'}, E_{j'}\}$.

Now, consider the second case where $C_{j-1-j'} = E_{j'-1}$ and $C_{j-j'} = D_{j'-1}$. Consider an arbitrary $u \in B_j \subseteq E_{j'-1}$ which, by assumption, has degree 1 into $D_{j'-1}$ and degree $\deg(u) - 1$ into $E_{j'-1}$. Then $v \in C_{q(t)-\ell'+j'-2}$ has degree 1 into $D_{j'-1}$, degree 1 (or 2) into $A_{j-1-j'}$, and degree $\deg(v) - 2$ (or $\deg(v) - 3$) into $E_{j'-1}$. Hence, in this case, we obtain $B_{j-1-j'} = E_{j'-1}$ and

$$\begin{aligned} D_{j'} &= D_{j'-1}, \\ E_{j'} &= E_{j'-1} \cup C_{q(t)-\ell'+j'-1}, \end{aligned}$$

which is what we expect for $j' \equiv 0 \pmod{3}$.

Finally, we treat the case where $C_{j-1-j'} = E_{j'-1}$ with $C_{j-j'} \subseteq E_{j'-1}$. In this case, $C_{q(t)-\ell'+j'-2}$ has degree 0 into $D_{j'-1}$, whereas $B_j \subseteq E_{j'-1}$ has degree 1 into $D_{j'-1}$, making $B_{j-1-j'} = D_{j'-1}$. Then we have $B_{j-1-j'} = D_{j'-1}$ and

$$\begin{aligned} D_{j'} &= D_{j'-1} \cup C_{q(t)-\ell'+j'-1}, \\ E_{j'} &= E_{j'-1}, \end{aligned}$$

which is what we expect for $j' \equiv 1 \pmod{3}$. □

Having established these results, we now begin to look for long-refinement graphs with maximum degree at most 4. We make the following assumption throughout the rest of this section.

Assumption 6.6. G is a long-refinement graph with $\max \deg(G) \leq 4$.

Recall that, by Lemma 5.13, it holds that $d \geq \ell$. We first treat the case that $d = \ell$. In this case, ℓ' , which we recall is defined as $\min\{\ell, d - 1\}$, is thus $\ell - 1$. Thus, Lemma 5.11 with some substitutions yields the following equation.

$$\begin{aligned} \pi^{q(t)-\ell} &= \pi_S^{q(t)} \cup \{P_i \cup P_{a-i} \cup P_{b+i} \mid i \in [\ell - 1]\} \\ &\quad \cup \{P_i \mid i \in [a + 1, b - 1]\} \cup \{P_a \cup P_b\} \cup \{S \cup P_{c-1} \mid t = 1\} \end{aligned} \quad (7)$$

The next lemma will allow us to determine the possible expressions for $\pi^{q(t)-\ell-1}$.

Lemma 6.7. *Let $d = \ell$. Then (after possibly swapping the roles of $A_{q(t)-\ell}$ and $B_{q(t)-\ell}$) we have that $A_{q(t)-\ell} = P_a \cup P_b$, and either $B_{q(t)-\ell} \in \pi_S^{q(t)}$ or $B_{q(t)-\ell} = P_1 \cup P_{a-1} \cup P_{b+1}$, each with $\deg_{B_{q(t)-\ell}}(P_{a-1} \cup P_{b+1}) - \deg_{B_{q(t)-\ell}}(P_1) = 1$.*

Proof. Note the inclusion of $P_1 \cup P_{a-1} \cup P_{b+1}$ and $P_a \cup P_b$ in the expression for $\pi^{q(t)-\ell}$ (Equation 7). Lemma 5.4 yields $\deg_{P_a \cup P_b}(P_1) = 2$, and Lemma 5.8 with the degree restrictions on $P_a \cup P_b$ yields $\deg_{P_a \cup P_b}(P_{a-1} \cup P_{b+1}) = 1$. Then $P_1 \cup P_{a-1} \cup P_{b+1}$ is unbalanced wrt $P_a \cup P_b$ and therefore, by Lemma 4.4, either $A_{q(t)-\ell}$ or $B_{q(t)-\ell}$ is $P_a \cup P_b$. Let us fix $A_{q(t)-\ell} = P_a \cup P_b$, without loss of generality. By Lemma 4.4, the class $B_{q(t)-\ell} \in \pi^{q(t)-\ell} \setminus \{A_{q(t)-\ell}\}$ must satisfy $\deg_{B_{q(t)-\ell}}(P_{a-1} \cup P_{b+1}) - \deg_{B_{q(t)-\ell}}(P_1) = 1$. By Corollary 5.12, there is no class $C \in \{P_i \cup P_{a-i} \cup P_{b+i} \mid i \in [2, \ell - 1]\}$ that can satisfy this: by Observations 6.2 and 6.3, either $i = 2$ and $G[P_1 \cup P_{a-1} \cup P_{b+1}, C]$ is a matching, or $E(G[P_1 \cup P_{a-1} \cup P_{b+1}, C]) = \emptyset$. Lemma 5.10 discounts $B_{q(t)-\ell} = S \cup P_{c-1}$. So, observing the expression for $\pi^{q(t)-\ell}$, we see that $B_{q(t)-\ell} \in \pi_S^{q(t)}$ or $B_{q(t)-\ell+1} = P_1 \cup P_{a-1} \cup P_{b+1}$. □

Lemma 6.7 yields two possible expressions for $\pi^{q(t)-\ell-1}$. Taking $B_{q(t)-\ell} \in \pi_S^{q(t)}$, we fix one expression for $\pi^{q(t)-\ell-1}$. By analysing this partition, we can determine possible expressions of previous partitions and the graphs that realise them.

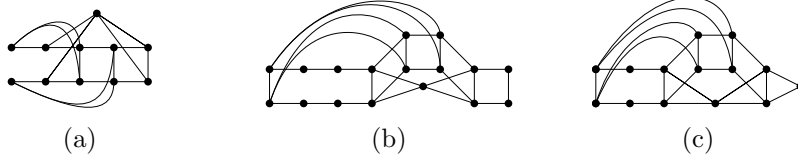


Figure 9: The long-refinement graphs with $d = \ell$ and $B_{q(t)-\ell} \in \pi_S^{q(t)}$.

Lemma 6.8. Assume $d = \ell$, $A_{q(t)-\ell} = P_a \cup P_b$, and $B_{q(t)-\ell} \in \pi_S^{q(t)}$. Then G is one of the graphs depicted in Figure 9.

Proof. The expression for $\pi^{q(t)-\ell}$ from Equation 7 and the values of $A_{q(t)-\ell}$ and $B_{q(t)-\ell}$ give the following expression for $\pi^{q(t)-\ell-1}$.

$$\begin{aligned} \pi^{q(t)-\ell-1} = \pi_S^{q(t)} \setminus \{B_{q(t)-\ell}\} \cup \{P_i \cup P_{a-i} \cup P_{b+i} \mid i \in [\ell-1]\} \\ \cup \{P_i \mid i \in [a+1, b-1]\} \cup \{P_a \cup P_b \cup B_{q(t)-\ell}\} \cup M_t \end{aligned}$$

Recall that $\deg_{B_{q(t)-\ell}}(P_{a-1} \cup P_{b+1}) - \deg_{B_{q(t)-\ell}}(P_1) = 1$. Since $|B_{q(t)-\ell}| = 1$, this implies that $\deg_{B_{q(t)-\ell}}(P_1) = 0$ and $\deg_{B_{q(t)-\ell}}(P_{a-1} \cup P_{b+1}) = 1$. Then, $\deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(B_{q(t)-\ell}) = 4$ and $\deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(P_a \cup P_b) = 2$ by Lemmas 5.4 and 5.5. Thus $C_{q(t)-\ell-1}$ is unbalanced wrt $P_1 \cup P_{a-1} \cup P_{b+1}$ so we can assume, without loss of generality, that $A_{q(t)-\ell-1} = P_1 \cup P_{a-1} \cup P_{b+1}$. By degree restrictions on $B_{q(t)-\ell}$, $\deg_{B_{q(t)-\ell}}(B_{q(t)-\ell-1}) = 0$, so to satisfy Lemma 4.4, $P_a \cup P_b$ will have degree 2 into $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$, and 0 into all other classes in $\pi^{q(t)-\ell-1}$. If $b = a + 2$, the degree of $P_a \cup P_b$ into class P_{a+1} is 1, so $b = a + 1$ and P_a, P_b must be consecutive pairs. As $P_a \cup P_b$ has non-zero degree into $B_{q(t)-\ell} \cup P_a \cup P_b$, the latter class must be $B_{q(t)-\ell-1}$. This implies that $P_a \cup P_b$ has degree 2 into $B_{q(t)-\ell} \cup P_a \cup P_b$ and thus that $P_a, P_b \in E(G)$. With this and the expression for $A_{q(t)-\ell-1}$, we get $C_{q(t)-\ell-2} = B_{q(t)-\ell} \cup P_1 \cup P_{a-1} \cup P_a \cup P_b \cup P_{b+1}$. Either $C_{q(t)-\ell-2}$ is $V(G)$ and $\deg(B_{q(t)-\ell} \cup P_a \cup P_b) \neq \deg(P_1 \cup P_{a-1} \cup P_{b+1})$, or all vertices in $C_{q(t)-\ell-2}$ have degree 4. In the first case, since vertices in P_1 , and therefore also $P_{a-1} \cup P_{b+1}$, have degree at least 3, G must be the graph illustrated in Figure 9a with $\deg(G) = \{3, 4\}$.

Now, consider the second case. Notice that $B_{q(t)-\ell} \cup P_a \cup P_b$ has degree 4 into $C_{q(t)-\ell-2}$, so must have degree 0 into all other classes. Thus, by Lemma 4.4, $P_1 \cup P_{a-1} \cup P_{b+1}$ must be adjacent to exactly one other class, which we may call $B_{q(t)-\ell-2}$ without loss of generality, and $\deg_{B_{q(t)-\ell-2}}(P_1 \cup P_{a-1} \cup P_{b+1})$ is at most 2. Degree restrictions on $B_{q(t)-\ell-2}$ imply that it is not a singleton. Therefore $B_{q(t)-\ell-2}$ is either $P_2 \cup P_{a-2} \cup P_{b+2}$ ($\ell > 1$) or $P_2 \cup S_{q(t)}$ ($\ell = 1$ and $t = 1$). By Observations 6.2 and 6.3, notice that $\deg(B_{q(t)-\ell-2}) = 4$ only in the case illustrated in Figure 8c, where $t = 1$ and $\deg_{P_d \cup P_c \cup P_{n\mathcal{P}}}(S \cup P_{c-1}) = 4$. Note, however, that this would result in a graph that is either 4-regular or disconnected. Therefore $C_{q(t)-\ell-3}$ must be $V(G)$, with either $B_{q(t)-\ell-2} = P_d \cup P_c \cup P_{n\mathcal{P}}$, as depicted in Figure 8a, or $S \cup P_{c-1}$, as depicted in Figure 8b. The resulting graphs are depicted in Figure 9b and 9c. $\square$

Now, consider the second case from Lemma 6.7, in which $B_{q(t)-\ell} = P_1 \cup P_{a-1} \cup P_{b+1}$, and thus consider the second possible expression for $\pi^{q(t)-\ell-1}$. There are infinitely many graphs which can realise this partition at iteration $q(t) - \ell - 1$. We describe them below.

Lemma 6.9. Assume $d = \ell$, $A_{q(t)-\ell} = P_a \cup P_b$ and $B_{q(t)-\ell} = P_1 \cup P_{a-1} \cup P_{b+1}$. If $\deg(G) = \{2, 3\}$, then G is represented by a string in one of the following sets:

- $\{S1^k 01^k 1XX1^k 1_2 \mid k \in \mathbb{N}_0\}$
- $\{S1^k 0011^k XX1^k 10 \mid k \in \mathbb{N}_0\}$
- $\{S(011)^k 00(110)^k 1X0X1(011)^k 0 \mid k \in \mathbb{N}_0\}$

If $\deg(G) \neq \{2, 3\}$, then G is one of the graphs in Figure 10 or in Table 4.

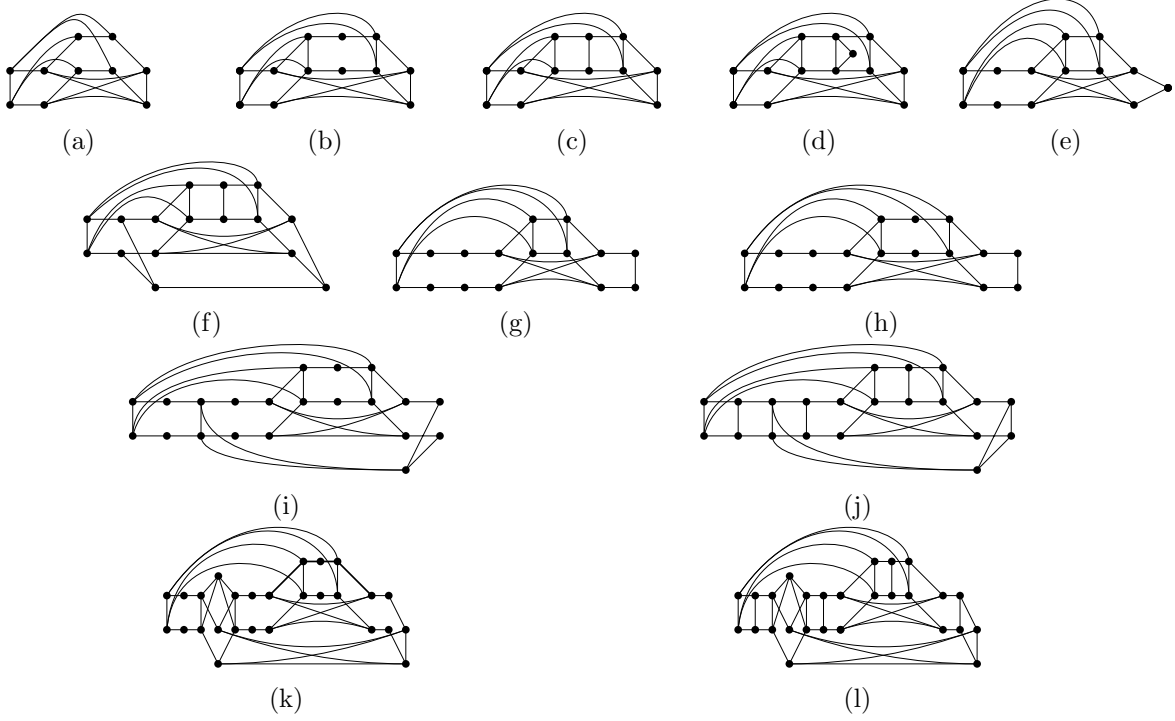


Figure 10: The long-refinement graphs G with $d = \ell$, $B_{q(t)-\ell} = P_1 \cup P_{a-1} \cup P_{b+1}$, and $\deg(G) \neq \{2, 3\}$.

Proof. The expression for $\pi^{q(t)-\ell}$ from Equation 7 and the assumed values of $A_{q(t)-\ell}$ and $B_{q(t)-\ell}$ yield the following expression for the previous partition.

$$\begin{aligned} \pi^{q(t)-\ell-1} = \pi_S^{q(t)} \cup \{P_i \cup P_{a-i} \cup P_{b+i} \mid i \in [2, \ell-1]\} \\ \cup \{P_i \mid i \in [a+1, b-1]\} \cup \{P_1 \cup P_{a-1} \cup P_a \cup P_b \cup P_{b+1}\} \cup M_t \end{aligned}$$

Recall that $\deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(P_{a-1} \cup P_{b+1}) - \deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(P_1) = 1$. The degrees of P_1 and $P_{a-1} \cup P_{b+1}$ into $P_1 \cup P_{a-1} \cup P_{b+1}$ depend on whether P_{a-1} is P_2 (which occurs when $\ell = 1$ and $t = 0$), whether $G[P_a, P_b]$ is empty or complete bipartite, and which among P_1, P_{a-1} , and P_{b+1} are in $E(G)$. Careful consideration of all cases yields four layouts for the edges incident to $P_1 \cup P_{a-1} \cup P_{b+1}$ which satisfy the degree difference of 1. These layouts are depicted in Figure 11 with $P_1 \cup P_{a-1} \cup P_{b+1}$ depicted in blue.

vertex	adjacency 1 with $n \in \mathbb{N}$	adjacency 2 with $n \in \mathbb{N}_0$
0	$2n+1, 2n+2, 2n+5, 2n+6$	$2n+1, 2n+2, 2n+5, 2n+6$
1	$3, 4n+7, 4n+8$	$3, 4n+7, 4n+8$
2	$4, 4n+9, 4n+10$	$4, 4n+9, 4n+10$
odd $i \in I$	$i-2, i+1, i+2$	$i-2, i+1, i+2$
even $i \in I$	$i-2, i-1, i+2$	$i-2, i-1, i+2$
$2n+1^*$	$0, 2n-1, 2n+3$	$2n+1, 2n+2, 2n+5$
$2n+2^*$	$0, 2n, 2n+4$	$2n, 2n+12n+4$
$2n+3$	$2n+1, 2n+5, 6n+11, 6n+12$	$0, 2n+1, 2n+4, 2n+5$
$2n+4$	$2n+4, 2n+8, 6n+17, 6n+18$	$0, 2n+2, 2n+3, 2n+6$
$2n+5$	$0, 2n+3, 2n+7$	$2n+3, 2n+6, 2n+7$
$2n+6$	$0, 2n+4, 2n+8$	$2n+4, 2n+5, 2n+8$
$4n+7$	$1, 4n+5, 4n+9$	$1, 4n+5, 4n+9$
$4n+8$	$1, 4n+6, 4n+10$	$1, 4n+6, 4n+10$
$4n+9$	$2, 4n+7, 4n+11$	$2, 4n+7, 4n+11$
$4n+10$	$2, 4n+8, 4n+12$	$2, 4n+8, 4n+12$
$6n+11$	$2n+3, 2n+4, 6n+9$	$0, 6n+12, 6n+10$
$6n+12$	$2n+3, 2n+4, 6n+10$	$0, 6n+11, 6n+10$

Table 4: The two infinite families of long-refinement graphs G with order $6n+13$ for $\deg(G) = \{3, 4\}$ where $I = [3, 2n] \cup [2n+7, 4n+6] \cup [4n+11, 6n+10]$. Note also that $2n+1, 2n+2$ are omitted when $n = 0$.

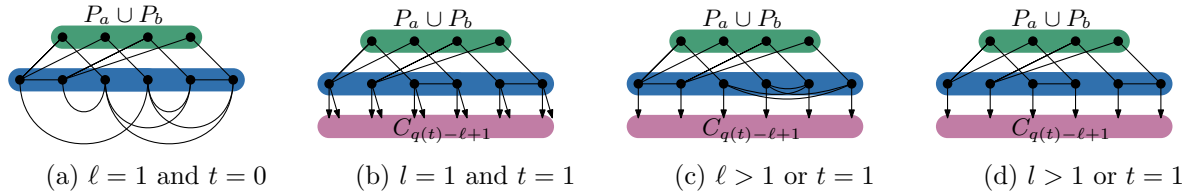


Figure 11: Edge layouts satisfying $\deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(P_{a-1} \cup P_{b+1}) - \deg_{P_1 \cup P_{a-1} \cup P_{b+1}}(P_1) = 1$. Note these depict colour classes as in $\pi^{q(t)-\ell}$ and only the edges incident to $P_1 \cup P_{a-1} \cup P_{b+1}$. The union of the green and blue classes is $C_{q(t)-\ell-1}$.

Consider Figure 11a, where $\ell = 1$ and $t = 0$. First, let $C_{q(t)-\ell-1} = V(G)$. Then $\deg(P_a \cup P_b) \neq \deg(P_1 \cup P_{a-1} \cup P_{b+1})$, so $\deg(P_a \cup P_b) \neq 4$. Since P_a and P_b have minimum degree 3, G must be as depicted in Figure 10a. Next, assume $C_{q(t)-\ell-1} \subsetneq V(G)$. Then $C_{q(t)-\ell-1}$ must be unbalanced wrt exactly two classes $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$. One of those classes must be $C_{q(t)-\ell-1}$, as it contains all the neighbours of $P_1 \cup P_{a-1} \cup P_{b+1}$ and we say without loss of generality that $A_{q(t)-\ell-1} = C_{q(t)-\ell-1}$. We can observe from Equation 8, taking $\ell = 1$ and $t = 0$, that the only other classes possible in $\pi^{q(t)-\ell-1}$ are P_{a+1} (if $b = a+2$) and singletons $\pi_S^{q(t)}$. We first assume that $B_{q(t)-\ell-1}$ is a singleton and prove a contradiction. Since $P_a \cup P_b$ is adjacent to but balanced wrt $B_{q(t)-\ell-1}$, $B_{q(t)-\ell-1}$ must be adjacent to all four vertices of $P_a \cup P_b$. Then all vertices in set $C_{q(t)-\ell-1} \cup B_{q(t)-\ell-1}$ have 4 neighbours in that same set, making G either disconnected or 4-regular. Thus $B_{q(t)-\ell-1}$ is P_{a+1} and $C_{q(t)-\ell-2} = P_1 \cup P_{a-1} \cup P_a \cup P_{a+1} \cup P_b \cup P_{b+1}$. Since $\ell = 1$ and $t = 0$, this contains all the pairs in π_P^p . If $\deg(P_{a+1}) \neq \{4\}$, then $V(G) = C_{q(t)-\ell-2}$ and G is as depicted in Figures 10b, where $\deg(P_{a+1}) = 2$, and 10c, where $\deg(P_{a+1}) = 3$. If $\deg(P_{a+1}) = 4$, then P_{a+1} can be adjacent to exactly one class in $\pi^{q(t)-\ell-2} \setminus \{C_{q(t)-\ell-2}\}$, i.e. a singleton $S \in \pi_S^{q(t)}$. We then get $C_{q(t)-\ell-3} = S \cup C_{q(t)-\ell-2}$. Since

$\deg_S(P_{a+1})$ can only be 1, we also need $P_{a+1} \in E(G)$ to satisfy $\deg(P_{a+1}) = \{4\}$. Singleton S then has degree 2 into P_{a+1} and shares no edges with the other pairs. Therefore, S can only have degree 4 if it is itself adjacent to two other singletons, at which point $C_{q(t)-\ell-3}$ would be unbalanced wrt three classes contradicting Lemma 4.4. Therefore $C_{q(t)-\ell-2} = V(G)$ and $\deg(P_{a+1}) = 2$ and G is as depicted in Figure 10d.

Next, consider Figure 11b. Here, $C_{q(t)-\ell-1} \neq V(G)$, so $\deg(P_a \cup P_b) = \deg(P_1 \cup P_{a-1} \cup P_{b+1}) = 4$. The vertices of $P_1 \cup P_{a-1} \cup P_{b+1}$ have two neighbours in $C_{q(t)-\ell+1}$, whereas $P_a \cup P_b$ have none by Observations 6.2 and 6.3. Thus, $C_{q(t)-\ell+1}$ is in $\pi^{q(t)-\ell-1} \setminus \pi^{q(t)-\ell-2}$ and without loss of generality we may denote it $A_{q(t)-\ell-1}$. Thus, to satisfy Lemma 4.4, we require the class $B_{q(t)-\ell-1}$ to satisfy $\deg_{B_{q(t)-\ell-1}}(P_a \cup P_b) - \deg_{B_{q(t)-\ell-1}}(P_1 \cup P_{a-1} \cup P_{b+1}) = 2$. Thus $B_{q(t)-\ell-1}$ can be neither a singleton (trivial), nor P_{a+1} (as $\deg_{P_{a+1}}(P_a \cup P_b) = 1$), nor any of the classes of size 6 (Observations 6.2, 6.3), and must therefore be $C_{q(t)-\ell-1}$. However, we then have the vertices of set $C_{q(t)-\ell-1} \cup C_{q(t)-\ell+1}$ each with 4 neighbours in $C_{q(t)-\ell-1} \cup C_{q(t)-\ell+1}$, implying that either G is 4-regular, or disconnected.

Now, consider Figure 11c. Here, $C_{q(t)-\ell-1} \neq V(G)$, so $\deg(P_a \cup P_b) = \deg(P_1 \cup P_{a-1} \cup P_{b+1}) = 4$. First, consider the case where $b = a + 2$. We then have $\pi^{q(t)-\ell}$ satisfying the assumptions for π^j in Lemma 6.5, with $M = P_{a+1}$. Recall that $\ell' = \ell - 1$. Then, the results of the aforementioned lemma yield that

$$\pi^{q(t)-2\ell-t} \setminus \pi^{q(t)-\ell} = \{D_{\ell-2+t}, E_{\ell-2+t}\} \text{ with } P_1 \cup P_{a-1} \cup P_{b+1} = C_{q(t)-\ell} \subseteq E_{\ell-2+t}$$

We therefore require $\deg(E_{\ell-2+t}) = \deg(P_1 \cup P_{a-1} \cup P_{b+1}) = 4$, which implies that it cannot include any class $P_{d-h} \cup P_{c+h} \cup P_{n_{\mathcal{P}-h}} = C_{q(t)-h-1}$ where $h \in [1, \ell - 2]$, by Observations 6.2. This gives an upper limit to ℓ , yielding $\ell \in [1, 3]$. Where $\ell = 1$, t must be 1 as $C_{q(t)-\ell+1}$ exists. By Lemma 6.5, $\deg_{C_{q(t)-\ell+1}}(B_j) = 1$, so we must have edges depicted as in Figure 8b, and we have $C_{q(t)-\ell-2} = P_{a+1} \cup P_{c-1} \cup S_{q(t)}$. The graph with $\deg(C_{q(t)-\ell-2}) = 2$ is not a long-refinement graph as $C_{q(t)-\ell-2}$ is then balanced at iteration $q(t) - \ell - 2$. Degree $\deg(C_{q(t)-\ell-2}) = 4$ can only be achieved if $G[P_{c-1} \cup S_{q(t)}]$ is complete, P_{a+1} is in $E(G)$, and P_{a+1} is adjacent to a singleton. That singleton, however, would also need to have degree 4, which is not achievable, given the classes in $\pi^{q(t)-\ell-2}$ without violating Lemma 4.4. Finally, if $\deg(C_{q(t)-\ell-2}) = 3$, then $G[P_{c-1} \cup S_{q(t)}]$ is an empty graph but $P_{c-1} \cup S_{q(t)}$ is adjacent to a singleton which we denote $S_{q(t)-\ell-2}$. Degree restrictions on $S_{q(t)-\ell-2}$ imply it is not adjacent to P_{a+1} , so either $P_{a+1} \in E(G)$ or P_{a+1} is adjacent to another singleton $S'_{q(t)-\ell-2} \neq S_{q(t)-\ell-2}$. The former option yields the graph depicted in Figure 10f, while the latter leads to a contradiction as $C_{q(t)-\ell-3}$ would then be $S'_{q(t)-\ell-2} \cup S_{q(t)-\ell-2}$ implying $\deg(S'_{q(t)-\ell-2}) = 3$ and thus that it is adjacent to another singleton $S_{q(t)-\ell-3}$. As $\deg(G) = \{3, 4\}$, $S_{q(t)-\ell-3}$ cannot have a satisfying degree without violating Lemma 4.4, given the classes in $\pi^{q(t)-\ell-3}$. Where $\ell = 2$ and $t = 0$, by Lemma 6.3, we get the graph depicted in Figure 10h. Where $\ell = 2$ and $t = 1$, we have $C_{q(t)-2\ell-1} = E_{\ell-2+t} \cup P_{c-1} \cup S_{q(t)}$. Then $\deg(P_{c-1} \cup S_{q(t)}) = \deg(E_{\ell-2+t}) = 4$. This implies $P_{c-1} \cup S_{q(t)}$ has edge layout as depicted in Figure 8c, or as in Figure 8b with $G[P_{c-1} \cup S_{q(t)}]$ a complete graph. The former is not possible as both $C_{q(t)-2\ell+1}$ and $D_{\ell-2+t}$ are then unbalanced. The latter leads to the long-refinement graph depicted in Figure 10i when $\deg(D_{\ell-2+t}) = 2$ and Figure 10j when $\deg(D_{\ell-2+t}) = 3$. Finally, consider when $\ell = 3$ and therefore $\deg(C_{q(t)-1}) = 4$, where $C_{q(t)-1} = P_d \cup P_c \cup P_{n_{\mathcal{P}}}$. This can only be satisfied by $t = 1$, with edges as depicted in Figure 8c and $P_d, P_c, P_{n_{\mathcal{P}}}$. This case yields the graphs depicted in Figure 10k and 10l, for $\deg(D_{\ell-2+t})$ equal to 2 or 3, respectively.

Next, let us consider the case where $b = a + 1$. The class $P_1 \cup P_{a-1} \cup P_{b+1}$ has degree 1 into $C_{q(t)-\ell+1}$, whereas $P_a \cup P_b$ has degree 0 into $C_{q(t)-\ell+1}$. Thus, $C_{q(t)-\ell+1}$ is in $\pi^{q(t)-\ell-1} \setminus \pi^{q(t)-\ell-2}$ and without loss of generality we may denote it $A_{q(t)-\ell-1}$. Then, by Lemma 4.4, we require

the class $B_{q(t)-\ell-1}$ to satisfy $\deg_{B_{q(t)-\ell-1}}(P_a \cup P_b) - \deg_{B_{q(t)-\ell-1}}(P_1 \cup P_{a-1} \cup P_{b+1}) = 1$. This cannot hold if $B_{q(t)-\ell-1}$ is a class of size 6, by Corollary 5.12, or if it is the class $P_{c-1} \cup S_{q(t)}$, by Lemma 5.10. It also cannot be a singleton; we would then have $\deg(B_{q(t)-\ell-1}) = 4$, and since $A_{q(t)-\ell-1} \cup B_{q(t)-\ell-1} \subsetneq V(G)$, also $\deg(A_{q(t)-\ell-1}) = \{4\}$, which implies by Remark 6.2 and Observation 6.3 that $C_{q(t)-\ell+1}$ must be $P_d \cup P_c \cup P_{n\mathcal{P}}$, with edges as depicted in Figure 8c. However, we then get a graph that is either 4-regular or disconnected. Thus, $B_{q(t)-\ell-1}$ is $C_{q(t)-\ell-1}$. To satisfy $\deg(P_a \cup P_b) = 4$, we must have $P_a, P_b \in E(G)$ and $b = a + 1$. Now, either $\deg(C_{q(t)-\ell+1}) = 4$ or $V(G) = C_{q(t)-\ell-1} \cup C_{q(t)-\ell+1}$. The former can only be satisfied by $C_{q(t)-\ell+1}$ being $P_d \cup P_c \cup P_{n\mathcal{P}}$ with edges as depicted in Figure 8c. However, this results in G being 4-regular or disconnected. Thus, $V(G) = C_{q(t)-\ell-1} \cup C_{q(t)-\ell+1}$, with $C_{q(t)-\ell-1}$ either $P_d \cup P_c \cup P_{n\mathcal{P}}$ as depicted in Figure 8a, or $P_{c-1} \cup S_{q(t)}$, as in Figure 8b. These graphs are depicted in Figures 10e and 10g.

Finally, consider Figure 11d. Here, we have $\deg(P_1 \cup P_{a-1} \cup P_{b+1}) = 3$, and therefore also $\deg(P_a \cup P_b) = 3$. The class $P_a \cup P_b$ has degree 2 into $P_1 \cup P_{a-1} \cup P_{b+1}$, and degree 1 into either $P_a \cup P_b$ if $b = a + 1$, or into P_{a+1} if $b = a + 2$. In the former case, iteration $q(t) - \ell$ satisfies the assumptions for iteration j in Lemma 6.4. Thus, taking $j' = \ell' + t = \ell - 1 + t$, we get $C_{q(t)-2\ell+t} = \bigcup_{i \in [1, d] \cup [c, n\mathcal{P}]} P_i$. If $t = 0$, then $P_d \cup P_c \cup P_{n\mathcal{P}} \subsetneq C_{q(t)-2\ell-1}$ has degree 2 by Lemma 6.3, while $\deg(\bigcup_{i \in [1, d-1] \cup [c+1, n\mathcal{P}-1]} P_i) = 3$ and therefore $V(G) = \bigcup_{i=1}^{n\mathcal{P}} P_i$ and that G is a graph whose representative string is contained in the set $\{S1^k001^k1XX11^k0 \mid k \in \mathbb{N}_0\}$. If $t = 1$, then $C_{q(t)-2\ell-t} \subsetneq V(G)$ so $\deg(P_i) = 3$ for $i \in [1, d] \cup [c, n\mathcal{P}]$. The vertices $P_d \cup P_c \cup P_{n\mathcal{P}}$ have degree 1 into $P_{c-1} \cup S_{q(t)}$, where other vertices in $C_{q(t)-2\ell-1}$ have degree 0, and have all their neighbours within $C_{q(t)-2\ell-1}$. This implies that $\pi^{q(t)-2\ell-2} \setminus \pi^{q(t)-2\ell-1} = \{C_{q(t)-2\ell-1}, P_{c-1} \cup S\}$. If $C_{q(t)-2\ell-2} \neq V(G)$, then $\deg(P_{c-1} \cup S) = 3$, which can only be satisfied by having edges as depicted in Figure 8b and $P_{c-1} \cup S_{q(t)}$ also adjacent to a singleton S . Here, either the singleton has degree 3 and the graph is 3-regular, or the singleton has degree greater than 3 so there are vertices not in $C_{q(t)-2\ell-3} = C_{q(t)-2\ell-2} \cup S$, yet $\deg(C_{q(t)-2\ell-3}) = \perp$. We therefore let $C_{q(t)-2\ell-2} = V(G)$, in which case $\deg(P_{c-1} \cup S)$ is either 2 or 4. These can be achieved by edges as in 8b, with $\deg_{P_{c-1} \cup S}(P_{c-1} \cup S) \in \{0, 2\}$. We thus get three possible infinite families of graphs, either represented by the set $\{S1^k01^k1XX1^k1_2 \mid k \in \mathbb{N}_0\}$, or by Table 4.

Next, we consider the case where $b = a + 2$. Here, iteration $q(t) - \ell$ satisfies the assumptions for iteration j in Lemma 6.5, with $M = P_{a+1}$, thus yielding $\pi^{q(t)-2\ell-t} \setminus \pi^{q(t)-\ell} = \{D_{\ell-2+t}, E_{\ell-2+t}\}$. If $t = 0$, then $\deg(P_d \cup P_c \cup P_{n\mathcal{P}}) = 2$ by Observation 6.3, so cannot be a subset of $E_{\ell-2+t}$ with degree 3. It follows then that $\deg(D_{\ell-2+t}) = 2$, so $P \in E(G)$ for all pairs $P \in \pi_{\mathcal{P}}^p$ such that $P \subseteq E_{\ell-2+t} \setminus \{P_{a-1} \cup P_{b+1}\}$, and $P \notin E(G)$ if $P \subseteq D_{\ell-3} \setminus \{P_{n\mathcal{P}}\}$. We therefore get the graph represented by the set $\{S(011)^k00(110)^k1X0X1(011)^k0 \mid k \in \mathbb{N}_0\}$. Consider $t = 1$ with the layout of edges as depicted in Figure 8b. When $C_{q(t)-2\ell-t} = P_{c-1} \cup S_{q(t)} \cup D_{\ell-2+t}$, the argument that G cannot be a long-refinement graph is the same as that used for Figure 11c, when $\ell = 1$ and $t = 1$, except that since $E_{\ell-2+t}$ has degree 3, the case that yields the long-refinement graph there instead yields a 3-regular graph here. When $C_{q(t)-2\ell-t} = P_{c-1} \cup S_{q(t)} \cup E_{\ell-2+t}$, $\deg(C_{q(t)-2\ell-t}) = 3$ and hence $P_{c-1} \cup S_{q(t)}$ must be adjacent to a singleton, which we will denote $S_{q(t)-2\ell-t}$. If $C_{q(t)-2\ell+1-t} = D_{\ell-2+t}$, then $P_{c-1} \cup S_{q(t)}$ contradicts Lemma 4.4 having higher degree into $D_{\ell-2+t}$ and $S_{q(t)-2\ell-t}$. If $C_{q(t)-2\ell+1-t} = E_{\ell-2+t}$, then by Lemma 4.4, $C_{q(t)-2\ell-t-1} = S_{q(t)-2\ell-t} \cup D_{\ell-2+t}$, and $\deg(D_{\ell-2+t}) \geq 3$. If $\deg(D_{\ell-2+t}) = 3$, we have a 3-regular graph, whereas if $\deg(D_{\ell-2+t}) = 4$, then $\deg(P_{a+1}) = 4$ and must, by a previous argument, be adjacent to a singleton with degree 4 - which is not possible. Next, consider $t = 1$ with the layout of edges depicted in Figure 8c. Here, $C_{q(t)-2\ell-t} = P_{c-1} \cup S_{q(t)} \cup D_{\ell-2+t}$ since $\deg(E_{\ell-2+t}) = 3$. Thus $C_{q(t)-\ell+1} \subseteq D_{\ell-2+t}$ can only be $P_d \cup P_c \cup P_{n\mathcal{P}}$ or $P_{c-1} \cup S_{q(t)}$. The latter contradicts Lemma 6.5, the former contradicts Lemma 4.4. $\square$

Having identified the graphs that correspond to both cases from Lemma 6.7, this concludes the case that $d = \ell$. Now, since the remaining cases have $d > \ell$, we henceforth have $\ell' = \ell$ and thus Lemma 5.11 yields

$$\pi^{q(t)-\ell-1} = \pi_S^{q(t)} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_p-i} \mid i \in [0, \ell]\} \cup \{P_i \mid i \in [d-\ell-1] \cup [a+1, b-1]\} \cup M_t \quad (8)$$

We now consider the cases where $d = \ell + 1$ and $d = \ell + 2$ simultaneously. In Lemma 6.10, similarly to Lemma 6.7, we identify all the classes that may be $A_{q(t)-\ell-1}$ or $B_{q(t)-\ell-1}$, so as to determine all possible expression of $\pi^{q(t)-\ell-2}$ and treat them separately.

Lemma 6.10. *If $d \in \{\ell + 1, \ell + 2\}$, then $\{A_{q(t)-\ell-1}, B_{q(t)-\ell-1}\} \subseteq \{P_{d-\ell} \cup P_a \cup P_b, P_{a+1}\} \cup \pi_S^{q(t)}$.*

Proof. Observe the classes in Equation 8. By Remark 6.2, $C_{q(t)-\ell-1}$ is balanced wrt all classes in $\{P_{d-i} \cup P_{c+i} \cup P_{n_p-i} \mid i \in [0, \ell-1]\}$. Degree restrictions on $C_{q(t)-\ell-1}$ and $S \cup P_{c-1}$ give the same result for $S \cup P_{c-1}$ if $t = 1$. When $d = \ell + 2$, Lemmas 5.1 and 5.4 imply $C_{q(t)-\ell-1} = P_2 \cup P_a \cup P_b$ is balanced wrt P_1 . Thus, $C_{q(t)-\ell-1}$ may only be unbalanced wrt the classes $P_{d-\ell} \cup P_a \cup P_b$ and P_{a+1} , and the singletons in $\pi_S^{q(t)}$. $\square$

This result allows us to divide the search for long-refinement graphs where $d \in \{\ell + 1, \ell + 2\}$ into the following cases: $C_{q(t)-\ell-1}$ is either unbalanced wrt by a) two singletons, b) P_{a+1} and a singleton, c) $P_{d-\ell} \cup P_a \cup P_b$ and a singleton adjacent to $P_{d-\ell}$, d) $P_{d-\ell} \cup P_a \cup P_b$ and a singleton adjacent to $P_a \cup P_b$, and e) P_{a+1} and $P_{d-\ell} \cup P_a \cup P_b$. We treat these cases in the next five lemmas. We determine that the first two cases cannot be satisfied by any long-refinement graphs.

Lemma 6.11. *Let $d \in \{\ell + 1, \ell + 2\}$. There is no long-refinement graph with $A_{q(t)-\ell-1}, B_{q(t)-\ell-1} \in \pi_S^{q(t)}$.*

Proof. From Equation 8 and the values of $A_{q(t)-\ell-1}, B_{q(t)-\ell-1}$, we get the following.

$$\pi^{q(t)-\ell-2} = \pi_S^{q(t)} \setminus \{A_{q(t)-\ell-1}, B_{q(t)-\ell-1}\} \cup \{A_{q(t)-\ell-1} \cup B_{q(t)-\ell-1}\} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_p-i} \mid i \in [0, \ell]\} \cup \{P_i \mid i \in [d-\ell-1] \cup [a+1, b-1]\} \cup M_t$$

Since $|A_{q(t)-\ell-1}| = |B_{q(t)-\ell-1}| = 1$ and $C_{q(t)-\ell-1}$ is unbalanced wrt both classes, we must have $P_{d-\ell}$ adjacent to one singleton, and $P_a \cup P_b$ adjacent to the other. Without loss of generality, let $\deg_{A_{q(t)-\ell-1}}(P_{d-\ell}) = 1$ and $\deg_{B_{q(t)-\ell-1}}(P_a \cup P_b) = 1$. Then $\deg_{P_{d-\ell} \cup P_a \cup P_b}(A_{q(t)-\ell-1}) = 2$ and $\deg_{P_{d-\ell} \cup P_a \cup P_b}(B_{q(t)-\ell-1}) = 4$, and without loss of generality $A_{q(t)-\ell-2} = P_{d-\ell} \cup P_a \cup P_b$. By degree restrictions, $B_{q(t)-\ell-1}$ has degree 0 into $B_{q(t)-\ell-2}$, so $\deg_{B_{q(t)-\ell-2}}(A_{q(t)-\ell-1}) = 2$. All classes in $\pi^{q(t)-\ell}$ are balanced wrt $-$ and since $A_{q(t)-\ell-1}$ is a singleton, trivially connected to $-$ $A_{q(t)-\ell-1} \in \pi^{q(t)-\ell-1}$. The class $B_{q(t)-\ell-2}$ must therefore be a pair, so $d = \ell + 2$ and $B_{q(t)-\ell-2} = P_1$. From this and by Lemma 4.4, we get $C_{q(t)-\ell-3} = P_1 \cup P_2 \cup P_a \cup P_b$. With P_1 adjacent to singleton $A_{q(t)-\ell-1}$, it has degree 4, so $P_2 \cup P_a \cup P_b$ must also have degree 4. If $\ell = 0$ and $t = 0$, the graph would be 4-regular or disconnected, so $\ell > 0$ or $t = 1$ must hold. Thus $C_{q(t)-\ell}$ is either $P_3 \cup P_{a-1} \cup P_{b+1}$, or $P_3 \cup S_{q(t)}$. In either case, class $C_{q(t)-\ell-3}$ is unbalanced wrt itself and $C_{q(t)-\ell}$, so by Lemma 4.4, $C_{q(t)-\ell-4} = C_{q(t)-\ell-3} \cup C_{q(t)-\ell}$ and thus $C_{q(t)-\ell}$ must also have degree 4. This can only occur if $C_{q(t)-\ell}$ is $\{P_d, P_c, P_{n_p}\}$ or M_t , where the layout of edges is as depicted in Figure 8c. Again, in this case, the graph is 4-regular or disconnected. $\square$

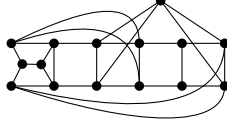


Figure 12: The long-refinement graph satisfying the prerequisites of Lemma 6.12.

Lemma 6.12. *Let $d \in \{\ell + 1, \ell + 2\}$ and $A_{q(t)-\ell-1} = P_{a+1}$ and $B_{q(t)-\ell-1} \in \pi_S^{q(t)}$. Then G is depicted in Figure 12.*

Proof. The expression for $\pi^{q(t)-\ell-1}$ from Equation 8 and the assumed values of $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$ yield the following expression for $\pi^{q(t)-\ell-2}$.

$$\pi^{q(t)-\ell-2} = \pi_S^{q(t)} \setminus \{B_{q(t)-\ell-1}\} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_{\mathcal{P}}-i} \mid i \in [0, \ell]\} \cup \{P_{a+1} \cup B_{q(t)-\ell-1}\} \cup \{P_i \mid i \in [d - \ell - 1]\} \cup M_t \quad (9)$$

Clearly $C_{q(t)-\ell-2} = P_{a+1} \cup B_{q(t)-\ell-1}$ is a strict subset of $V(G)$, so we require $C_{q(t)-\ell-2}$ to be unbalanced wrt two classes $A_{q(t)-\ell-2}$ and $B_{q(t)-\ell-2}$. We will show that there is no pair of classes in $\pi^{q(t)-\ell-2}$ that satisfy the requirements of $A_{q(t)-\ell-2}$ and $B_{q(t)-\ell-2}$.

First, consider the classes in $\{P_{d-i} \cup P_{c+i} \cup P_{n_{\mathcal{P}}-i} \mid i \in [0, \ell]\}$. Observe that $P_{d-\ell} \cup P_a \cup P_b$ ($i = \ell$) cannot be $A_{q(t)-\ell-2}$ or $B_{q(t)-\ell-2}$ since $\deg_{P_{d-\ell} \cup P_a \cup P_b}(C_{q(t)-\ell-2}) = 2$. Next, by Remark 6.2 and Observation 6.3, there are no edges between $C_{q(t)-\ell-2}$ and any class $P_{d-i} \cup P_{c+i} \cup P_{n_{\mathcal{P}}-i}$ where $i \in [0, \ell - 1]$, so such classes are not $A_{q(t)-\ell-2}$ or $B_{q(t)-\ell-2}$. By Lemma 5.10 and degree restrictions on $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$, $S_{q(t)} \cup P_{c-1} \notin \{A_{q(t)-\ell-2}, B_{q(t)-\ell-2}\}$.

Observing Equation 9, this leaves us with

$$\{A_{q(t)-\ell-2}, B_{q(t)-\ell-2}\} \subseteq \pi_S^{q(t)} \setminus \{B_{q(t)-\ell-1}\} \cup \{P_{a+1} \cup B_{q(t)-\ell-1}\} \cup \{P_i \mid i \in [d - \ell - 1]\}.$$

Let us assume that $P_1 = A_{q(t)-\ell-2}$ ($d = \ell + 2$) and prove a contradiction. In this case, $\deg_{P_1}(B_{q(t)-\ell-1}) - \deg_{P_1}(P_{a+1})$ must be 2, by Lemmas 4.4, 5.5, and degree restrictions on P_1 . However, no other class in $\pi_S^{q(t)} \setminus \{B_{q(t)-\ell-1}\} \cup \{P_{a+1} \cup B_{q(t)-\ell-1}\}$ can then satisfy Lemma 4.4: trivially, singleton classes are not large enough, while $\deg_{P_{a+1} \cup B_{q(t)-\ell-1}}(P_{a+1}) \in \{0, 1\}$ by degree restrictions on $B_{q(t)-\ell-1}$. By this and the symmetric argument for $B_{q(t)-\ell-2}$, $P_1 \notin \{A_{q(t)-\ell-2}, B_{q(t)-\ell-2}\}$. Thus, either $A_{q(t)-\ell-2} \cup B_{q(t)-\ell-2}$ is the union of two singletons $S_{q(t)-\ell-2}, S'_{q(t)-\ell-2}$, or of a singleton $S_{q(t)-\ell-2}$ and $P_{a+1} \cup B_{q(t)-\ell-1}$, when $P_{a+1} \in E(G)$.

Both cases, within a few simple applications of Lemma 4.4, lead to the class $C_{q(t)-\ell-4}$ in the first case, and $C_{q(t)-\ell-3}$ in the second case, being $P_1 \cup P_2 \cup P_a \cup P_b$ with $\deg(P_1 \cup P_2 \cup P_a \cup P_b) = 4$ and therefore $d = \ell + 2$. This requires the previous iteration to contain, by Lemma 4.4, either $P_1 \cup P_2 \cup P_a \cup P_b \cup M_t$ with M_t as depicted in Figure 8b with $G[M_t]$ a complete graph, or $P_1 \cup P_2 \cup P_3 \cup P_4 \cup P_a \cup P_b \cup P_{b+1}$, with all vertices of degree 4, therefore M_t as depicted in Figure 8c. The latter leads to contradictions, either leading to a 4-regular graph, or leading to a class with two degrees after iteration 1. This statement holds for the former case when both $A_{q(t)-\ell-2}$ and $B_{q(t)-\ell-2}$ are singletons. However, when $A_{q(t)-\ell-2}$ or $B_{q(t)-\ell-2}$ is P_{a+1} , we can have the graph depicted in Figure 12. $\square$

Now, we consider the cases where $A_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$ and $B_{q(t)-\ell-1} \in \pi_S^p$. Here we distinguish cases according to which of $A_{q(t)-\ell}$ or $B_{q(t)-\ell}$, i.e. which of $P_{d-\ell}$ and $P_a \cup P_b$, the singleton $B_{q(t)-\ell-1}$ is adjacent to. Although both cases have the same expression for $\pi^{q(t)-\ell-2}$, they require dissimilar proofs.

Lemma 6.13. *Let $d \in \{\ell+1, \ell+2\}$, $A_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$, and $B_{q(t)-\ell-1} \in \pi_S^{q(t)}$ with $B_{q(t)-\ell-1}$ adjacent to $P_a \cup P_b$. Then G is depicted in Figure 13.*



Figure 13: The two graphs satisfying the prerequisites from Lemma 6.13.

Proof. We know by the fact that $P_{a+1} \notin \{A_{q(t)-\ell-1}, B_{q(t)-\ell-1}\}$, that $b = a + 1$. Let us first assume that $d = \ell + 2$. Substituting 2 for $d - \ell$, we then get that $\deg_{P_2 \cup P_a \cup P_b}(P_2) > \deg_{P_2 \cup P_a \cup P_b}(P_a \cup P_b)$. However, no possible layout of edges satisfies this requirement without violating degree restrictions or Lemma 5.8. We must therefore have $d = \ell + 1$. There is a layout of edges that satisfies this, in which $P_1 \in E(G)$ and $P_a, P_b \notin E(G)$.

The expression for $\pi^{q(t)-\ell-1}$ from Equation 8 and the values of $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$ yield

$$\begin{aligned} \pi^{q(t)-\ell-2} = \pi_S^{q(t)} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_P-i} \mid i \in [0, \ell]\} \\ \cup \{P_i \mid i \in [d - \ell - 1] \cup [a + 1, b - 1]\} \cup M_t \end{aligned}$$

By assumption, $\deg_{B_{q(t)-\ell-1}}(P_a \cup P_b) = 1$ and thus also $\deg_{P_a \cup P_b}(B_{q(t)-\ell-1}) = 4$. Therefore, since $C_{q(t)-\ell-2} \subsetneq V(G)$ (it must omit at least P_2) then $\deg(P_{d-\ell} \cup P_a \cup P_b) = 4$. Note that, now, the class that includes pair P_2 must either be the only other class, or must also have degree 4. From Lemma 6.3 and Remark 6.2, it is clear that, if the latter is true, G is either 4-regular or disconnected. Thus, the class containing P_2 is the only other class at iteration $q(t) - \ell - 2$ and can be either $P_2 \cup S_{q(t)}$ or $P_2 \cup P_3 \cup P_5$. This yields the two graphs in Figure 13. $\square$

Lemma 6.14. *Let $d \in \{\ell + 1, \ell + 2\}$. There is no long-refinement graph satisfying $A_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$ and $B_{q(t)-\ell-1} \in \pi_S^{q(t)}$ with $B_{q(t)-\ell-1}$ adjacent to $P_{d-\ell}$.*

Proof. From the expression for $\pi^{q(t)-\ell-1}$ in Equation 8 and the given values of $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$, we get the expression

$$\begin{aligned} \pi^{q(t)-\ell-2} = \pi_S^{q(t)} \setminus \{B_{q(t)-\ell-1}\} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_P-i} \mid i \in [0, \ell - 1]\} \\ \cup \{P_{d-\ell} \cup P_a \cup P_b \cup B_{q(t)-\ell-2}\} \cup \{P_i \mid i \in [d - \ell - 1] \cup [a + 1, b - 1]\} \cup M_t. \quad (10) \end{aligned}$$

First, assume that $d = \ell + 1$. Then, recalling that $a > 2$, we must have either $\ell > 0$ or $t = 1$, implying that $C_{q(t)-\ell} \in \{P_2 \cup P_{a-1} \cup P_{b+1}, P_2 \cup S\}$ is a class in $\pi^{q(t)-\ell-2}$. As the singleton $B_{q(t)-\ell-1}$ is only adjacent to $P_{d-\ell}$, it has degree 2 into $C_{q(t)-\ell-1}$. Meanwhile, $A_{q(t)-\ell-1}$ has degree 3 into $C_{q(t)-\ell-2}$ and 1 into $C_{q(t)-\ell}$. However, by Remark 6.2, $B_{q(t)-\ell-1}$ shares no edges with $C_{q(t)-\ell}$. This implies that $C_{q(t)-\ell-2}$ is unbalanced by both $C_{q(t)-\ell-2}$ and $C_{q(t)-\ell}$, with $A_{q(t)-\ell-1}$ having higher degree into both, contradicting Lemma 4.4.

Letting $d = \ell + 2$, $A_{q(t)-\ell-1}$ is $P_2 \cup P_a \cup P_b$. Consequently, $C_{q(t)-\ell-2}$ is unbalanced by P_1 as $\deg_{P_1}(P_2 \cup P_a \cup P_b) = 1$ by Lemmas 5.1 and 5.4, whereas $B_{q(t)-\ell-1}$ has degree 0 into P_1 by Lemma 5.5 and degree restrictions. If either $\ell > 0$ or $t = 1$, then $C_{q(t)-\ell}$ is another class in $\pi^{q(t)-\ell-2}$ which by Corollary 5.12 and Remark 6.2, $P_2 \cup P_a \cup P_b$ is adjacent to and $B_{q(t)-\ell-1}$ is not.

This contradicts Lemma 4.4. Thus, we can only have $\ell = 0$ and $t = 0$, in which case $\pi^{q(t)-\ell-2} = \pi_S^{q(t)} \setminus \{B_{q(t)-\ell-2}\} \cup \{P_1, P_2 \cup P_a \cup P_b \cup B_{q(t)-\ell-1}\}$. We can deduce that $\deg(P_2 \cup P_a \cup P_b \cup B_{q(t)-\ell-1}) = 3$, since $P_2 \cup P_a \cup P_b$ could not be adjacent to any singleton without contradicting the degree restrictions. Thus, we must have $B_{q(t)-\ell-1}$ adjacent to a singleton to satisfy $\deg(B_{q(t)-\ell-1}) = 3$ and thus $C_{q(t)-\ell-2}$ must be unbalanced wrt some singleton S' . Thus $C_{q(t)-\ell-2} = P_1 \cup S'$ with $C_{q(t)-\ell-2}$ unbalanced wrt $C_{q(t)-\ell-1}$ with P_1 having degree 3 into $C_{q(t)-\ell-1}$ and S' having degree 0. There are no other classes that can satisfy a degree difference of 2, so we have a final contradiction. $\square$

We now treat the final case for $d \in \{\ell + 1, \ell + 2\}$.

Lemma 6.15. *Let $d \in \{\ell + 1, \ell + 2\}$, $A_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$, and $B_{q(t)-\ell-1} = P_{a+1}$. If $\deg(G) = \{2, 3\}$, then G is represented by a long-refinement string in one of the following sets*

- $\{S1X1_2X\}$
- $\{S(011)^k 00(110)^k X1_2X(011)^k 0 \mid k \in \mathbb{N}_0\}$
- $\{S1^k 001^k X1X1^k 0 \mid k \in \mathbb{N}_0\}$
- $\{S0X1X_2\} \cup \{S1^k 1011^k X1X1^k 1_2 \mid k \in \mathbb{N}_0\}$

If $\deg(G) \neq \{2, 3\}$, then G is as represented in Figure 14 or is isomorphic to a graph in Table 5.

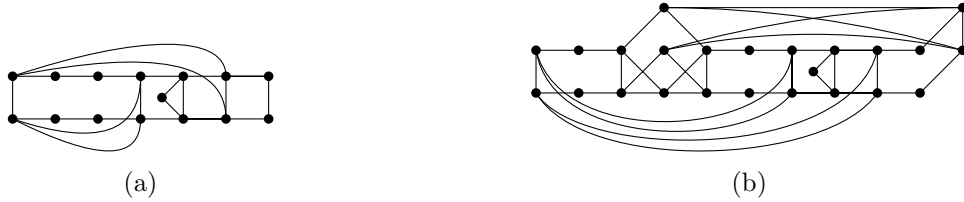


Figure 14: The long-refinement graphs G with $d \in \{\ell + 1, \ell + 2\}$, where $A_{q(t)-\ell-1} = P_{d-\ell} \cup P_a \cup P_b$, $B_{q(t)-\ell-1} = P_{a+1}$, and $\deg(G) \neq \{2, 3\}$.

Proof. Given the expression for $\pi^{q(t)-\ell-1}$ in Equation 8 and the assumed values of $A_{q(t)-\ell-1}$ and $B_{q(t)-\ell-1}$, we get the following expression for $\pi^{q(t)-\ell-2}$.

$$\begin{aligned} \pi^{q(t)-\ell-2} = \pi_S^{q(t)} \cup \{P_{d-\ell} \cup P_a \cup P_{a+1} \cup P_b\} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_P-i} \mid i \in [0, \ell-1]\} \\ \cup \{P_i \mid i \in [d-\ell-1]\} \cup M_t. \end{aligned} \quad (11)$$

Let us start by considering $d = \ell + 1$. Then $C_{q(t)-\ell-2}$ is $P_1 \cup P_a \cup P_{a+1} \cup P_b$ with $\deg(C_{q(t)-\ell-2}) \geq 3$. Since $P_2 \prec P_a$, we must have either $\ell > 0$ or $t = 1$. If $\ell = 0$ and $t = 1$ with edges as depicted in Figure 8c, then $\deg_{C_{q(t)-\ell}}(P_1 \cup P_a \cup P_b) - \deg_{C_{q(t)-\ell}}(P_{a+1}) = 2$. However, the other classes in $\pi^{q(t)-\ell-2}$ (the singletons in $\pi_S^{q(t)}$, $P_1 \cup P_a \cup P_{a+1} \cup P_b$) cannot satisfy Lemma 4.4 to complete $\{A_{q(t)-\ell-2}, B_{q(t)-\ell-2}\}$. In all cases where $\ell > 0$, $\deg_{C_{q(t)-\ell}}(P_1 \cup P_a \cup P_b) - \deg_{C_{q(t)-\ell}}(P_{a+1}) = 1$, by Lemmas 6.2 and 6.3. Without loss of generality, let $A_{q(t)-\ell-2} = C_{q(t)-\ell}$. To satisfy Lemma 4.4 $B_{q(t)-\ell-2}$ must satisfy $\deg_{B_{q(t)-\ell-2}}(P_1 \cup P_a \cup P_b) - \deg_{B_{q(t)-\ell-2}}(P_{a+1}) = -1$. Thus, by 6.2 and 6.3, $B_{q(t)-\ell-2}$ cannot be a class of size 6, nor M_t as there is no layout of edges that satisfies requirements. Thus $B_{q(t)-\ell-2}$ may only be $P_1 \cup P_a \cup P_{a+1} \cup P_b$, or in $\pi_S^{q(t)}$.

vertex	adjacency 1	adjacency 2
0	$2k + 3, 2k + 4, 2k + 7, 2k + 8$	$2k + 5, 2k + 6, 6k + 13, 6k + 14$
1	$3, 4k + 9, 4k + 10$	$3, 4k + 9, 4k + 10$
2	$4, 4k + 11, 4k + 12$	$4, 4k + 11, 4k + 12$
odd $i \in I$	$i - 2, i + 1, i + 2$	$i - 2, i + 1, i + 2$
even $i \in I$	$i - 2, i - 1, i + 2$	$i - 2, i - 1, i + 2$
$2k + 3$	$0, 2k + 1, 2k + 5$	$2k + 1, 2k + 4, 2k + 5$
$2k + 4$	$0, 2k + 2, 2k + 6$	$2k + 2, 2k + 3, 2k + 6$
$2k + 5$	$2k + 3, 2k + 7, 6k + 13, 6k + 14$	$0, 2k + 3, 2k + 6, 2k + 7$
$2k + 6$	$2k + 4, 2k + 8, 6k + 13, 6k + 14$	$0, 2k + 4, 2k + 5, 2k + 8$
$2k + 7$	$0, 2k + 5, 2k + 9$	$2k + 5, 2k + 8, 2k + 9$
$2k + 8$	$0, 2k + 6, 2k + 10$	$2k + 6, 2k + 7, 2k + 10$
$4k + 9$	$1, 4k + 7, 4k + 11$	$1, 4k + 7, 4k + 11$
$4k + 10$	$1, 4k + 8, 4k + 12$	$1, 4k + 8, 4k + 12$
$4k + 11$	$2, 4k + 9, 4k + 13$	$2, 4k + 9, 4k + 13$
$4k + 12$	$2, 4k + 10, 4k + 14$	$2, 4k + 10, 4k + 14$
$6k + 13$	$2k + 5, 2k + 6, 6k + 11$	$0, 6k + 14, 6k + 11$
$6k + 14$	$2k + 5, 2k + 6, 6k + 12$	$0, 6k + 13, 6k + 12$

Table 5: All infinite families of long-refinement graphs G with order $6k + 15$ and $\deg(G) = \{3, 4\}$, where $I = [3, 2k + 2] \cup [2k + 9, 4k + 8] \cup [4k + 13, 6k + 12]$, defined for $k \in \mathbb{N}_0$.

Let us first assume $B_{q(t)-\ell-2}$ is $P_1 \cup P_a \cup P_{a+1} \cup P_b$. Then $\deg(P_{a+1}) = 3$, as in this case P_{a+1} is not adjacent to a singleton. Furthermore, P_{a+1} has all its neighbours in its own class, which means that the partition $\pi^{q(t)-\ell-2}$ satisfies the requirements of partition π^j , which means one can apply Lemma 6.4 to yield $C_{j-1-j'} = C_{q(t)-2\ell-3-t} = A_j \cup \left(\bigcup_{i \in [0, \ell+t]} C_{q(t)-\ell-1+i} \right)$. If $t = 0$, then by the result from Remark 6.3 that $\deg(P_d \cup P_c \cup P_{np}) = 2$, $C_{q(t)-2\ell-3-t} = V(G)$ so we have exactly the graphs represented by the strings in the set $\{S1^k001^kX1X1^k0 \mid k \in \mathbb{N}_0\}$. If $t = 1$, then either $C_{q(t)-2\ell-3-t} = V(G)$ and $\deg(P_{c-1} \cup S_{q(t)}) \in \{2, 4\}$, or $C_{q(t)-2\ell-3-t} \subsetneq V(G)$ and $\deg(P_{c-1} \cup S_{q(t)}) = 3$. By an argument identical to that used in the proof Lemma 6.9, the former case yields exactly three infinite families of graphs – one represented by the strings in $\{S1^k1011^kX1X1^k1_2 \mid k \in \mathbb{N}_0\}$, and two represented in Table 5 – while the latter case yields none.

We return to the case where $B_{q(t)-\ell-2} \in \pi_S^{q(t)}$. Note that, in this case, $\pi^{q(t)-\ell-1}$ matches the requirements for partition j in Lemma 6.5 with M as the singleton $B_{q(t)-\ell-2}$. Note that $\deg B_{q(t)-\ell-2} = 2$, as it cannot be adjacent to the larger classes, by degree restrictions, and any singleton adjacent to $B_{q(t)-\ell-2}$ could not then have degree greater than 2, leading to a contradiction to Lemma 4.4 at the previous iteration. Now, $\deg P_{a+1}$ can be 3 or 4 depending on whether P_{a+1} is in $E(G)$. Where $\deg P_{a+1} = 4$, we can have at most $\ell = 2$ by Remark 6.2. We then get the graph in Figure 14a when $\ell = 1$ and the graph in Figure 14b when $\ell = 2$. When $\deg P_{a+1} = 3$, we obtain the graphs represented by the strings $\{S(011)^k00(110)^kX1_2X(011)^k0 \mid k \in \mathbb{N}_0\}$.

Now, consider the case with $d = \ell + 2$ and recall the partition $\pi^{q(t)-\ell-2}$ (Equation 11). The class P_1 is adjacent to $P_2 \cup P_a \cup P_b$, but not to P_{a+1} , by Lemma 5.5 and degree restrictions on P_1 . By Lemma 4.4, the same cannot be true of either $P_3 \cup P_{a-1} \cup P_{b+1}$ or $P_3 \cup S$, yet neither class can be adjacent to P_{a+1} by Lemmas 6.2 and 6.3, respectively, so neither class exists and thus $\ell = 0$ and $t = 0$. Taking, without loss of generality, $A_{q(t)-\ell-2} = P_1$, we then need $\deg_{B_{q(t)-\ell-2}}(P_{a+1}) - \deg_{B_{q(t)-\ell-2}}(P_2 \cup P_a \cup P_b) = 1$. This may be satisfied by any of the other classes in $\pi^{q(t)-\ell-2}$,

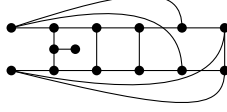


Figure 15: The unique long-refinement graph with $\deg(G) = \{1, 3\}$.

vertex	adjacency 1	adjacency 2
0	$2k + 7, 2k + 8, 2k + 11, 2k + 12$	$2k + 5, 2k + 6, 6k + 17, 6k + 18$
1	$3, 4k + 13, 4k + 14$	$3, 4k + 13, 4k + 14$
2	$4, 4k + 15, 4k + 16$	$4, 4k + 15, 4k + 16$
odd $i \in I$	$i - 2, i + 1, i + 2$	$i - 2, i + 1, i + 2$
even $i \in I$	$i - 2, i - 1, i + 2$	$i - 2, i - 1, i + 2$
$2k + 7$	$0, 2k + 5, 2k + 9$	$2k + 5, 2k + 8, 2k + 9$
$2k + 8$	$0, 2k + 6, 2k + 10$	$2k + 6, 2k + 7, 2k + 10$
$2k + 9$	$2k + 7, 2k + 11, 6k + 17, 6k + 18$	$0, 2k + 7, 2k + 10, 2k + 11$
$2k + 10$	$2k + 8, 2k + 12, 6k + 17, 6k + 18$	$0, 2k + 8, 2k + 9, 2k + 12$
$2k + 11$	$0, 2k + 9, 2k + 13$	$2k + 9, 2k + 12, 2k + 13$
$2k + 12$	$0, 2k + 10, 2k + 14$	$2k + 10, 2k + 11, 2k + 14$
$4k + 13$	$1, 4k + 11, 4k + 15$	$1, 4k + 11, 4k + 15$
$4k + 14$	$1, 4k + 12, 4k + 16$	$1, 4k + 12, 4k + 16$
$4k + 15$	$2, 4k + 13, 4k + 17$	$2, 4k + 13, 4k + 17$
$4k + 16$	$2, 4k + 14, 4k + 18$	$2, 4k + 14, 4k + 18$
$6k + 17$	$2k + 9, 2k + 10, 6k + 15$	$0, 6k + 18, 6k + 15$
$6k + 18$	$2k + 9, 2k + 10, 6k + 16$	$0, 6k + 17, 6k + 16$

Table 6: All infinite families of long-refinement graphs G with order $6k + 19$ and $\deg(G) = \{3, 4\}$, where $I = [3, 2k + 6] \cup [2k + 13, 4k + 12] \cup [4k + 17, 6k + 16]$.

$B_{q(t)-\ell-2} = P_2 \cup P_a \cup P_{a+1} \cup P_b$ or $B_{q(t)-\ell-2} \in \pi_S^{q(t)}$, but the former results in a 3-regular or disconnected graph. The latter yields the long-refinement graph $G(\text{S1X1}_2\text{X})$. $\square$

Having addressed every case from Lemma 6.10, this concludes the case for both $d = \ell + 1$ and $d = \ell + 2$. We now consider the remaining cases for the value of d . As per Lemma 5.13, those are $d \in \{\ell + 3, \ell + 4\}$, in which case $b = a + 1$.

Lemma 6.16. *Let $d \in \{\ell + 3, \ell + 4\}$. Then G is either the graph depicted in Figure 15 or isomorphic to a graph from Table 6 or represented by a string in one of the following sets.*

- $\{\text{S111}^k\text{001}^k\text{XX1}^k\text{0} \mid k \in \mathbb{N}_0\}$
- $\{\text{S110XX}_2\} \cup \{\text{S111}^k\text{1011}^k\text{XX1}^k\text{1}_2 \mid k \in \mathbb{N}_0\}$
- $\{\text{S011(011)}^k\text{00(110)}^k\text{XX(011)}^k\text{0} \mid k \in \mathbb{N}_0\}$
- $\{\text{S11XX}, \text{S011XX}, \text{S1}_2\text{11XX}\}$

Proof. We are given that $d \in \{\ell + 3, \ell + 4\}$ and therefore have, from Equation 8, an expression for $\pi^{q(t)-\ell-1}$. Since $d > \ell + 2$, the pairs $P_{d-\ell-1}$ and P_1 are distinct. By Lemmas 5.1 and 5.4,

$\deg_{P_{d-\ell-1}}(P_{d-\ell}) = 1 = \deg_{P_1}(P_a \cup P_b)$, whereas $\deg_{P_1}(P_{d-\ell}) \in \{0, 2\}$ and $\deg_{P_{d-\ell-1}}(P_{d-\ell}) \in \{0, 2\}$ by Lemma 5.5. Thus, by Lemma 4.4, we must have $C_{q(t)-\ell-2} = P_1 \cup P_{d-\ell-1}$ and

$$\pi^{q(t)-\ell-2} = \pi_S^{q(t)} \cup \{P_{d-i} \cup P_{c+i} \cup P_{n_P-i} \mid i \in [0, \ell]\} \cup \{P_1 \cup P_{d-\ell-1}\} \cup \{P_i \mid i \in [2, d-\ell-2]\} \cup M_t. \quad (12)$$

By Lemma 5.5 and degree restrictions, we have $\deg_{P_1}(P_{d-\ell}) = \deg_{P_{d-\ell-1}}(P_{d-\ell}) = 0$.

Now, we examine the neighbourhoods of $P_{d-\ell-1}$ and P_1 and their classes at iteration $q(t) - \ell - 2$. Notice that $\deg_{P_{d-\ell} \cup P_a \cup P_b}(P_1) = 2$ and $\deg_{P_{d-\ell} \cup P_a \cup P_b}(P_{d-\ell-2}) = 1$ and so $P_{d-\ell} \cup P_a \cup P_b \in \pi^{q(t)-\ell-2} \setminus \pi^{q(t)-\ell-3}$ and without loss of generality we can denote it $A_{q(t)-\ell-2}$. Then we need another class $B_{q(t)-\ell-2}$ such that $\deg_{B_{q(t)-\ell-2}}(P_{d-\ell-1}) - \deg_{B_{q(t)-\ell-2}}(P_1) = 1$. By Lemma 5.1, Remarks 6.2 and 6.3, and Lemma 5.10 respectively, we can discard P_2 (when $d = \ell + 4$), the classes of six, and where $t = 1$ the class $C_{q(t)}$. We are left with $B_{q(t)-\ell-2} \in \pi_S^{q(t)}$ or $B_{q(t)-\ell-2} = P_1 \cup P_{d-\ell-1}$. We show that the former leads to a contradiction, since it would lead to $C_{q(t)-\ell-3} = P_{d-\ell} \cup P_a \cup P_b \cup S$ for some singleton S . We get a contradiction when $\ell + t > 0$, as $P_{d-\ell} \cup P_a \cup P_b$ has more neighbours in both $P_{d-\ell} \cup P_a \cup P_b \cup S$ and $C_{q(t)-\ell}$ than S , contradicting Lemma 4.4. We get a contradiction when $\ell + t = 0$ because then $\deg_{P_1 \cup P_{d-\ell}}(S) - \deg_{P_1 \cup P_{d-\ell}}(P_{d-\ell} \cup P_a \cup P_b) = 1$ whereas $\deg_{P_1 \cup P_{d-\ell}}(S) - \deg_{P_1 \cup P_{d-\ell}}(P_{d-\ell} \cup P_a \cup P_b) = -2$, contradicting Lemma 4.4. As such, we must have $C_{q(t)-\ell-3} = C_{q(t)-\ell-2} \cup C_{q(t)-\ell-1}$, with possible edge layouts within $C_{q(t)-\ell-3}$ depicted in Figure 16. We work

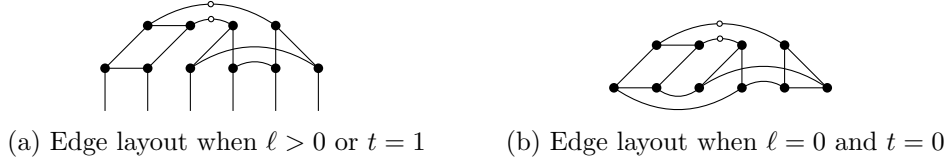


Figure 16: Diagram showing vertices $C_{q(t)-\ell-3}$ and all edges incident to it. Note the empty disks representing P_2 only when $d = \ell + 4$.

backward from each of the two possible expression for partition $\pi^{q(t)-\ell-3}$ to determine all long-refinement graphs in which $d \in \{\ell + 3, \ell + 4\}$. We know that $C_{q(t)-\ell-3} = P_1 \cup P_{d-\ell-1} \cup P_{d-\ell} \cup P_a \cup P_b$ with (up to swapping A and B) $A_{q(t)-\ell-2} = P_1 \cup P_{d-\ell-1}$ and $B_{q(t)-\ell-2} = P_{d-\ell} \cup P_a \cup P_b$. Let us consider the cases according to the value of $t + \ell$, which determines whether $P_{d-\ell}$ and P_a are subsequent pairs.

If $t + \ell = 0$ and $d = \ell + 3$, then $C_{q(t)-\ell-3}$ includes all the pairs in $\mathcal{P}$. $B_{q(t)-\ell-2}$ cannot be adjacent to any singletons, as we would then get a vertex with degree at least 6. Thus, either $C_{q(t)-\ell-3} = V(G)$ and the graph is 3-regular, or $C_{q(t)-\ell-3} \subsetneq V(G)$ and $\deg(C_{q(t)-\ell-3}) = 2$. Either is a contradiction.

If $t + \ell = 0$ and $d = \ell + 4$, then $C_{q(t)-\ell-3}$ includes all the pairs in $\mathcal{P}$, except P_2 . As $P_1 \cup P_{d-\ell}$ has degree 1 to P_2 whereas the remaining vertices in $C_{q(t)-\ell-3}$ have all their neighbours in $C_{q(t)-\ell-3}$, then $C_{q(t)-\ell-4} = C_{q(t)-\ell-3} \cup P_2 = \bigcup_{P \in \pi_P^p} P$, by Lemma 4.4. If $C_{q(t)-\ell-4} = V(G)$, then $\deg(P_2) \neq 3$ and thus we get $G(S011XX)$. If $C_{q(t)-\ell-4} \subsetneq V(G)$, then $\deg(P_2) = 3$, so P_2 must have degree 1 into some singleton S and by the same argument as above, we get $C_{q(t)-\ell-5} = \bigcup_{P \in \pi_P^p} P \cup S$. If $C_{q(t)-\ell-5} = V(G)$, then $\deg(S) \neq 3$, which is satisfied by $G(S1211XX)$. If $C_{q(t)-\ell-5} \subsetneq V(G)$, then $\deg(S) = 3$, so it must itself be adjacent to a singleton, which we denote $S_{q(t)-\ell-5}$. We cannot have $\deg(S_{q(t)-\ell-5}) = 3$ unless there are two singletons in $\pi^{q(t)-\ell-6}$ adjacent to $S_{q(t)-\ell-5}$ - neither of which could be adjacent to $\bigcup_{P \in \pi_P^p} P$, contradicting Lemma 4.4. Thus, $C_{q(t)-\ell-5} = V(G)$ and the last long-refinement graph in this case is that depicted in Figure 15.

If $d = \ell + 3$ and $t + \ell > 0$, then the partition $\pi^{q(t)-\ell-2}$ satisfies the requirements for π^j in Lemma 6.4, since $A_{q(t)-\ell-2}$ has all its vertices in $C_{q(t)-\ell-3}$. Applying that lemma gives us that $C_{q(t)-2\ell-3-t}$ is a superset of all the classes of $C_{q(t)-i}$ for $i \in [0, \ell + 1]$ with $A_{q(t)-2\ell-2-t} = C_{q(t)-1+t}$. If $t = 0$ then by Remark 6.3, $A_{q(t)-2\ell-2-t} = C_{q(t)-1+t}$ must have degree 2 yielding only the graphs represented by strings in the set $\{S111^k001^kXX1^k0 \mid k \in \mathbb{N}_0\}$. If $t = 1$, the edges between the vertices of $C_{q(t)}$ and $C_{q(t)-1}$ can be either as in Figure 8b or as in 8c. In the former case, we see that if $V(G) = C_{q(t)-2\ell-3-t}$, then $\deg(P_{c-1} \cup S_{q(t)})$ can be 2 – which results in the graphs represented by strings $\{S1111^k011^kXX1^k1_2 \mid k \in \mathbb{N}_0\}$ – or 3 – only if it is adjacent to a singleton S with $\deg(S) = 3$ which, without contradicting Lemma 4.4, yields only 3-regular or disconnected graphs – or 4 – which yields a graph in the right column of Table 6. In the latter case, the vertices of $P_{c-1} \cup S_{q(t)}$ have degree 4, so $C_{q(t)-2\ell-3-t} = V(G)$ and we obtain the graphs described in the left column of Table 6.

If $d = \ell + 4$ and $t + \ell > 0$, then we can apply Lemma 6.5, with $j = q(t) - \ell - 2$, to show that $\pi^{q(t)-2\ell-3-t} \setminus \pi^{q(t)-\ell-2} = \{D_{\ell-1+t}, E_{\ell-1+t}\}$.

If $D_{\ell-1+t} \cup E_{\ell-1+t} = V(G)$, then $\deg(D_{\ell-1+t}) \neq \deg(E_{\ell-1+t}) = 3$. However, we cannot achieve $\deg(D_{\ell-1+t}) = 4$ without P_2 adjacent to some singleton, which would contradict $D_{\ell-1+t} \cup E_{\ell-1+t} = V(G)$. Thus $\deg(D_{\ell-1+t}) = 2$. If $t = 0$, then $\deg(P_d \cup P_c \cup P_{n_p}) = 2$. Thus $C_{q(t)-1} \subsetneq D_{\ell-1+t}$, and therefore we narrow down to the graphs represented by strings in the set $\{S011(011)^k00(110)^kXX(011)^k0 \mid k \in \mathbb{N}_0\}$. If $t = 1$, we must again have $\deg(D_{\ell-1+t}) \in \{2, 4\}$. The latter, once again, cannot be achieved without a singleton not in $D_{\ell-1+t} \cup E_{\ell-1+t}$. The former leads to a stable graph.

If $D_{\ell-1+t} \cup E_{\ell-1+t} \subsetneq V(G)$, then t must be 1 as otherwise, as previously argued $C_{q(t)-2\ell-3-t} = D_{\ell-1+t}$ and $N(D_{\ell-1+t}) \subseteq D_{\ell-1+t} \cup E_{\ell-1+t}$, so G is disconnected. With $t = 1$, we could have a singleton S adjacent to $P_{c-1} \cup S_{q(t)}$, or where $P_2 \in \{A_{q(t)-2\ell-2-t}, B_{q(t)-2\ell-3-t}\}$, adjacent to P_2 . The former case can only occur when $t = 1$, as in Figure 8b, with $P_{c-1} \cup S_{q(t)}$ adjacent to a singleton. Without contradicting Lemma 4.4, this can only give us 3-regular or disconnected graphs. The latter case is also quickly dismissed: either $\deg(P_2)$ is 3 – so $C_{q(t)-1+t}$ must be $P_{c-1} \cup S_{q(t)}$ adjacent to a singleton S' yielding $C_{q(t)-2\ell-3-t} = S \cup S'$ which causes contradictions within two applications of Lemma 4.4 – or it is 4 – and $\deg S = 4$, which is impossible considering the classes in partition $\pi^{q(t)-2\ell-3-t}$, without contradicting Lemma 4.4. $\square$

This concludes our case analysis of long-refinement graphs with maximum degree at most 4.

7 The Classification

We now bring together the results that we have found in the Section 6 for our final classification.

Recall Notation 6.1. The work [25] identified families of long-refinement graphs G with $\deg(G) = \{2, 3\}$, but left as an open problem whether they are comprehensive. Here, we answer their question affirmatively.

Theorem 7.1. *The even-order long-refinement graphs G with $\deg(G) = \{2, 3\}$ are exactly the ones represented by strings contained in the following sets:*

- $\{S011XX\}$
- $\{S1^k001^kX1X1^k0 \mid k \in \mathbb{N}_0\}$
- $\{S1^k11001^kXX1^k0 \mid k \in \mathbb{N}_0\}$
- $\{S1^k0011^kXX1^k10 \mid k \in \mathbb{N}_0\}$

- $\{S011(011)^k00(110)^kXX(011)^k0 \mid k \in \mathbb{N}_0\}$
- $\{S(011)^k00(110)^k1X0X1(011)^k0 \mid k \in \mathbb{N}_0\}$

The odd-order long-refinement graphs G with $\deg(G) = \{2, 3\}$ are exactly the ones represented by strings contained in the following sets:

- $\{S1_211XX\}$
- $\{S0X1X_2\} \cup \{S1^k1011^kX1X1^k1_2 \mid k \in \mathbb{N}_0\}$
- $\{S110XX_2\} \cup \{S111^k1011^kXX1^k1_2 \mid k \in \mathbb{N}_0\}$
- $\{S1^k01^k1XX1^k1_2 \mid k \in \mathbb{N}_0\}$
- $\{S(011)^k00(110)^kX1_2X(011)^k0 \mid k \in \mathbb{N}_0\}$

In particular, the theorem implies that the gaps that the families found in [24] leave open cannot be covered with degrees 2 and 3.

Corollary 7.2. *There is no long-refinement graph G with $\deg(G) = \{2, 3\}$ and order $n \neq 12$ such that $n \bmod 18 \in \{6, 12\}$.*

Theorem 7.1 and Corollary 7.2 completely classify the long-refinement graphs G with $\deg(G) = \{2, 3\}$. Expanding to $\max \deg(G) \leq 4$, we obtain the following further parts of our classification of long-refinement graphs with small degrees from the results in section 6.

Theorem 7.3. *There is no long-refinement graph G with $\deg(G) = \{1, 2\}$ or $\deg(G) = \{1, 4\}$.*

Theorem 7.4. *The long-refinement graphs G with $\deg(G) = \{2, 4\}$ are precisely the ones in Figures 9b, 9c, 10b, 10e, 10g, 10h, 10i, 10k, 13a, 13b, 14a, and 14b.*

Theorem 7.5. *The long-refinement graphs G with $\deg(G) = \{3, 4\}$ are precisely the ones in Figures 9a, 10a, 10c, 10d, 10f, 10j, 10l, 12 as well as Tables 4, 5, and 6.*

Theorem 7.6. *The graph in Figure 15 is the only long-refinement graph G with $\deg(G) = \{1, 3\}$.*

While Theorems 7.3 – 7.6 are proved through an analysis of the refinement iterations in Section 6, we now present a simple alternative proof for Theorem 7.6, which relies merely on Theorem 7.1.

By [25, Lemma 13], a long-refinement graph G with $\deg(G) = \{1, 3\}$ has at most two vertices of degree 1. We start by treating the case where G has exactly one vertex with degree 1.

Lemma 7.7. *The graph in Figure 15 is the only long-refinement graph G with $\deg(G) = \{1, 3\}$ in which $|\{v \in V(G) : \deg(v) = 1\}| = 1$.*

Proof. Let G be any long-refinement graph with $\deg(G) = \{1, 3\}$ and the set of vertices of degree 1 being a singleton $\{v_1\}$. Then the partition after the first Colour Refinement iteration on G is $\pi_G^1 = \{\{v_1\}, V(G) \setminus \{v_1\}\}$. The graph $G' := G[V(G) \setminus \{v_1\}]$, i.e. the subgraph of G induced by $V(G) \setminus \{v_1\}$, satisfies $\deg(G') = \{2, 3\}$ and $\{v \in V(G') : \deg(v) = 2\} = \{v_2\}$, where $\{v_2\} = N_G(v_1)$. Also, Colour Refinement takes $n - 2$ iterations to terminate on G' , otherwise G would not be a long-refinement graph. Hence, G' is a long-refinement graph as well, and since $|V(G)|$ is even, $|V(G')|$ is odd. By our classification in Theorem 7.1, G' must be $G(S1_211XX)$, since that is the only long-refinement graph with degrees 2 and 3 in which there is only one vertex of degree 2. So, by degree arguments, G must be the graph in Figure 15. $\square$

Now we consider the situation in which there are exactly two vertices of degree 1.

Lemma 7.8. *There is no long-refinement graph G with $\deg(G) = \{1, 3\}$ in which $|\{v \in V(G) : \deg(v) = 1\}| = 2$.*

Proof. Suppose G is a long-refinement graph with $\deg(G) = \{1, 3\}$ for which $\{v \in V(G) : \deg(v) = 1\} = \{v_1, v'_1\}$ with $v_1 \neq v'_1$. We know that $\{v_1, v'_1\} \notin E(G)$, since G is connected. However, now the graph $G' := (V(G), E(G) \cup \{\{v_1, v'_1\}\})$ is a long-refinement graph with $\deg(G') = \{2, 3\}$ and $\{v \in V(G') : \deg(v) = 2\} = \{v_1, v'_1\}$. The degree property is clear. To see that G' is a long-refinement graph, note that inserting the new edge does not alter the partition into degrees, i.e. π^1 . Also, since in both graphs, it holds that $|N(\{v_1, v'_1\})| \leq 2$, the class $\{v_1, v'_1\}$ must be the last class that splits (in both graphs). Furthermore, $N_G(v_1) \setminus \{v'_1\} = N_{G'}(v_1) \setminus \{v'_1\}$ and $N_G(v'_1) \setminus \{v_1\} = N_{G'}(v'_1) \setminus \{v_1\}$. This implies that $\pi_G^i = \pi_{G'}^i$ for every $i \in \mathbb{N}$, so G' is a long-refinement graph if and only if G is one. By double counting edges, G (and hence also G') has an even number of vertices and hence G' is one of the graphs from the first list in Theorem 7.1. The only graph from that list with exactly two vertices of degree 2 is $G(\text{S011XX})$, but these vertices are not adjacent. Hence, G' does not exist, and therefore G does not exist. $\square$

This concludes the proof of Theorem 7.6, which together with Theorem 7.3 yields that the graph in Figure 15 is the only long-refinement graph G with $\min \deg(G) = 1$ and $\max \deg(G) \leq 4$.

8 Distinguishability of Long-Refinement Graphs

To conclude and to address another open question from [25], we consider the situation when Colour Refinement is applied to two graphs G and H in parallel, for example to check whether they are isomorphic. Formally, Colour Refinement is applied to the graph consisting of the disjoint union U of G and H (ensuring that their vertex sets are disjoint). Recall that χ_U^i denotes that colouring that Colour Refinement has computed on U after i iterations. We say that Colour Refinement *distinguishes* G and H in i iterations if $\{\{\chi_U^i(v) \mid v \in V(G)\}\} \neq \{\{\chi_U^i(v) \mid v \in V(H)\}\}$. If there is no such i , we say that G and H are *equivalent* with respect to Colour Refinement.

Theorem 8.1. *For $n \geq 3$, there is no pair of graphs G, H with $|H| \leq |G| = n$ that Colour Refinement distinguishes only after $n - 1$ iterations.*

Proof. Graphs of different orders are distinguished within one iteration of Colour Refinement, so any counterexample to the theorem would have to consist of graphs G, H with $|G| = |H|$ and both graphs would be long-refinement graphs. In particular, Colour Refinement computes the discrete partition on both graphs and, hence, the last colour class that splits in each graph is a pair P_G and P_H , respectively. Let U be the disjoint union of G and H . Since P_G and P_H must have the same colour in iteration $n - 2 \geq 1$, it holds that $P_G \in E(G)$ if and only if $P_H \in E(H)$. Let $\{v_1, v_2\} := P_G$. Since in π_U^{n-2} , all vertices in $V(G) \setminus P_G$ and in $V(H) \setminus P_H$ form singleton colour classes, there has to be a vertex $w_1 \in P_H$ that is adjacent to exactly the same colour classes from iteration $n - 2$ as v_1 , and similarly, the second vertex $w_2 \in P_H$ must be adjacent to exactly the same singleton classes as v_2 . However, then for each $C \in \pi_U^{n-2}$ and for $i \in \{1, 2\}$, it holds that $\deg_C(v_i) = \deg_C(w_i)$. But then $\chi_U^{n-1}(v_i) = \chi_U^{n-1}(w_i)$. Thus, G and H are equivalent with respect to Colour Refinement and in particular not distinguished after $n - 1$ iterations. $\square$

The theorem gives a negative answer to the question whether there are pairs of graphs that are only distinguished after the maximum possible number of iterations of Colour Refinement. The above argument might be pushed a little further to exclude other strong lower bounds on the number of iterations for distinguishability. Note, however, that it may still be possible that there

exists a constant c such that for all $n \in \mathbb{N}$ (from some start value on), there are pairs of graphs G and H with $|G| = |H| = n$ which are only distinguished in $n - c$ iterations of Colour Refinement.

9 Conclusion

Long-refinement graphs are graphs which take the maximum theoretically possible number of Colour Refinement iterations to stabilise. Even though the Colour Refinement procedure is simple and its power and mechanics are well-understood (see, for example, [24]), the fact that long-refinement graphs actually exist is surprising, given that this contrasts the situation for higher dimensions of the Weisfeiler–Leman algorithm. The original approaches to show their existence were empirical, i.e. to test candidate graphs [15, 25]. A theoretical analysis of the results then yielded infinite families of long-refinement graphs with small degrees, but also left infinitely many gaps with respect to graph sizes.

In this work, we have shown that for minimal degrees, those gaps cannot be closed. More precisely, we have proved that Theorems 18 and 24 from [25] are actually a classification, meaning those graphs are *the only long-refinement graphs with degrees $\{2, 3\}$ that exist*, for arbitrarily large n . We have extended this to the first classification of long-refinement graphs of maximum degree 4. While previous searches for long-refinement graphs started from computational data, our reverse-engineering method is a purely theoretically-motivated approach. Our analysis shows that *all* families of long-refinement graphs with minimal degrees (i.e. with maximum degree at most 3) can be captured with string representations. The characterisation that falls out of this insight yields a simple proof that there is only one long-refinement graph with minimum degree 1 and maximum degree at most 4.

As witnessed in Lemma 4.5, the graph obtained from a long-refinement graph by taking the complement of the edge set is again a long-refinement graph. Thus, our classification of long-refinement graphs with maximum degree at most 4 also immediately yields a classification of long-refinement graphs with minimum degree at least $n - 5$. Extending the analyses to maximum degree larger than 4 and minimum degree smaller than $n - 5$ remains an interesting project for future work. The analysis in Section 5 describes at least the final five Colour Refinement iterations on long-refinement graphs. For the long-refinement graphs with maximum degree at most 4, it accounts for at least the final half of all iterations. Previous study of long-refinement graphs with general degree has consisted in computing possible *split procedures* in a brute-force-like fashion [15]. Assuming a relatively small number of singletons (or equivalently, a relatively high number of pairs) in the first iteration in which every class has size at most 2, our structural results have the potential to significantly reduce the search space for split procedures, as they provide additional constraints that the previous approach did not exploit. This would potentially allow us to compute larger long-refinement graphs of arbitrary degrees.

Given the existence of infinite families of long-refinement graphs, in the light of the practical applications of Colour Refinement, it is an intriguing question to ask for pairs of long-refinement graphs that are only distinguished (i.e. receive different colourings) in the last Colour Refinement iteration on them. This search was initiated in [25] and has been ongoing since then. In Section 8, we close it by showing that such graphs do not exist. Our technique to capture long-refinement graphs via string representations opens the door to generalisations, e.g. to the construction of long-refinement graphs which take long to be distinguished by Colour Refinement. Without the requirement to be long-refinement graphs, the currently best known lower bound is $n - 8\sqrt{n}$ [29, Theorem 4.6]. The task to improve on this lower bound to $n - c$ for a constant c (which must be larger than 1, due to Theorem 8.1) is particularly captivating given that such pairs of graphs

do not exist for higher dimensions of the Weisfeiler–Leman algorithm [18]. It is plausible that the bound of $n - 1$ in the statement of Theorem 8.1 could be strengthened to $n - c$ for some small c . For such graphs, there can be at most $c - 1$ – hence very few – iterations in which the number of colours increases by more than 1, i.e. which differ from splittings in long-refinement graphs. Since the number of iterations is crucial for the parallelisability of the algorithm, improved upper bounds mean an improved complexity of a parallel implementations.

References

- [1] Markus Anders and Pascal Schweitzer. Engineering a fast probabilistic isomorphism test. In Martin Farach-Colton and Sabine Storandt, editors, *Proceedings of the Symposium on Algorithm Engineering and Experiments, ALENEX 2021, Virtual Conference, January 10-11, 2021*, pages 73–84. SIAM, 2021.
- [2] Vikraman Arvind, Johannes Köbler, Gaurav Rattan, and Oleg Verbitsky. Graph isomorphism, color refinement, and compactness. *Computational Complexity*, 26(3):627–685, 2017.
- [3] Albert Atserias and Elitza N. Maneva. Sherali-Adams relaxations and indistinguishability in counting logics. *SIAM J. Comput.*, 42(1):112–137, 2013.
- [4] Albert Atserias and Joanna Ochremiak. Definable ellipsoid method, sums-of-squares proofs, and the isomorphism problem. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*, pages 66–75. ACM, 2018.
- [5] László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 684–697. ACM, 2016.
- [6] László Babai, Paul Erdős, and Stanley M. Selkow. Random graph isomorphism. *SIAM Journal on Computing*, 9(3):628–635, 1980.
- [7] Christoph Berkholz and Jakob Nordström. Near-optimal lower bounds on quantifier depth and Weisfeiler-Leman refinement steps. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 267–276, 2016.
- [8] Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992.
- [9] Alain Cardon and Maxime Crochemore. Partitioning a graph in $O(|A| \log |A|)$. *Theoretical Computer Science*, 19:85–98, 1982.
- [10] Paul T. Darga, Mark H. Liffiton, Karem A. Sakallah, and Igor L. Markov. Exploiting structure in symmetry detection for CNF. In *Proceedings of the 41th Design Automation Conference, DAC 2004, San Diego, CA, USA, June 7-11, 2004*, pages 530–534. ACM, 2004.
- [11] Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets Weisfeiler and Leman. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July*

- 9-13, 2018, Prague, Czech Republic, volume 107 of *LIPICs*, pages 40:1–40:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.
- [12] Zdenek Dvorák. On recognizing graphs by numbers of homomorphisms. *J. Graph Theory*, 64(4):330–342, 2010.
 - [13] Martin Fürer. Weisfeiler-Lehman refinement requires at least a linear number of iterations. In *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 322–333. Springer, 2001.
 - [14] Andreas Göbel, Leslie Ann Goldberg, and Marc Roth. The Weisfeiler-Leman dimension of conjunctive queries. *Proc. ACM Manag. Data*, 2(2):86, 2024.
 - [15] Maximilian Gödicke. The iteration number of the Weisfeiler-Lehman-algorithm. Master’s thesis, RWTH Aachen University, 2015.
 - [16] Christopher D. Godsil. Compact graphs and equitable partitions. *Linear Algebra Appl.*, 255:259–266, 1997.
 - [17] Martin Grohe and Sandra Kiefer. Logarithmic Weisfeiler-Leman identifies all planar graphs. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPICs*, pages 134:1–134:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.
 - [18] Martin Grohe, Moritz Lichter, and Daniel Neuen. The iteration number of the Weisfeiler-Leman algorithm. *ACM Trans. Comput. Log.*, 26(1):6:1–6:31, 2025.
 - [19] Martin Grohe, Moritz Lichter, Daniel Neuen, and Pascal Schweitzer. Compressing CFI graphs and lower bounds for the Weisfeiler-Leman refinements. *J. ACM*, 72(3):21:1–21:27, 2025.
 - [20] Martin Grohe and Oleg Verbitsky. Testing graph isomorphism in parallel by playing a game. In *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10–14, 2006, Proceedings, Part I*, pages 3–14, 2006.
 - [21] Neil Immerman and Eric Lander. Describing graphs: A first-order approach to graph canonization. In Alan L. Selman, editor, *Complexity theory retrospective*, pages 59–81. Springer, 1990.
 - [22] Tommi A. Junttila and Petteri Kaski. Engineering an efficient canonical labeling tool for large and sparse graphs. In *Proceedings of the Nine Workshop on Algorithm Engineering and Experiments, ALENEX 2007, New Orleans, Louisiana, USA, January 6, 2007*. SIAM, 2007.
 - [23] Tommi A. Junttila and Petteri Kaski. Conflict propagation and component recursion for canonical labeling. In Alberto Marchetti-Spaccamela and Michael Segal, editors, *Theory and Practice of Algorithms in (Computer) Systems - First International ICST Conference, TAPAS 2011, Rome, Italy, April 18-20, 2011. Proceedings*, volume 6595 of *Lecture Notes in Computer Science*, pages 151–162. Springer, 2011.
 - [24] Sandra Kiefer. *Power and Limits of the Weisfeiler-Leman Algorithm*. PhD thesis, RWTH Aachen University, Aachen, 2020.

- [25] Sandra Kiefer and Brendan D. McKay. The Iteration Number of Colour Refinement. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 73:1–73:19, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- [26] Sandra Kiefer and Pascal Schweitzer. Upper bounds on the quantifier depth for graph differentiation in first-order logic. *Logical Methods in Computer Science*, 15(2), 2019.
- [27] Sandra Kiefer, Pascal Schweitzer, and Erkal Selman. Graphs identified by logics with counting. *ACM Trans. Comput. Log.*, 23(1):1:1–1:31, 2022.
- [28] Johannes Köbler and Oleg Verbitsky. From invariants to canonization in parallel. In *Computer Science - Theory and Applications, Third International Computer Science Symposium in Russia, CSR 2008, Moscow, Russia, June 7-12, 2008, Proceedings*, volume 5010 of *Lecture Notes in Computer Science*, pages 216–227. Springer, 2008.
- [29] Andreas Krebs and Oleg Verbitsky. Universal covers, color refinement, and two-variable counting logic: Lower bounds for the depth. In *Proceedings of the Thirtieth Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 689–700, 2015.
- [30] Wenchao Li, Hossein Saidi, Huascar Sanchez, Martin Schäfer, and Pascal Schweitzer. Detecting similar programs via the Weisfeiler-Leman graph kernel. In *Software Reuse: Bridging with Social-Awareness - 15th International Conference, ICSR 2016, Limassol, Cyprus, June 5-7, 2016, Proceedings*, pages 315–330, 2016.
- [31] Moritz Lichter, Ilia N. Ponomarenko, and Pascal Schweitzer. Walk refinement, walk logic, and the iteration number of the Weisfeiler-Leman algorithm. In *Proceedings of the Thirty-Fourth Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 1–13, June 2019.
- [32] Brendan D. McKay. Practical graph isomorphism. *Congressus Numerantium*, 30:45–87, 1981.
- [33] Brendan D. McKay and Adolfo Piperno. Practical graph isomorphism, II. *Journal of Symbolic Computation*, 60:94–112, 2014.
- [34] Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan E. Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and Leman go neural: Higher-order graph neural networks. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, January 2019.
- [35] Motakuri V. Ramana, Edward R. Scheinerman, and Daniel Ullman. Fractional isomorphism of graphs. *Discrete Mathematics*, 132(1–3):247–265, 1994.
- [36] Nino Shervashidze, Pascal Schweitzer, Erik Jan van Leeuwen, Kurt Mehlhorn, and Karsten M. Borgwardt. Weisfeiler-Lehman graph kernels. *Journal of Machine Learning Research*, 12:2539–2561, 2011.
- [37] Gottfried Tinhofer. A note on compact graphs. *Discrete Applied Mathematics*, 30(2–3):253–264, 1991.
- [38] Steffen van Bergerem, Martin Grohe, Sandra Kiefer, and Luca Oeljeklaus. Simulating logspace-recursion with logarithmic quantifier depth. In *38th Annual ACM/IEEE Symposium on Logic*

in Computer Science, LICS 2023, Boston, MA, USA, June 26-29, 2023, pages 1–13. IEEE, 2023.

- [39] Boris Weisfeiler and Andrei Leman. The reduction of a graph to canonical form and the algebra which appears therein. *NTI, Series 2*, 1968. English translation by G. Ryabov available at https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf.