Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 32 (2015)

Graph Isomorphism, Color Refinement, and Compactness*

V. Arvind[1], Johannes Köbler[2], Gaurav Rattan[1], Oleg Verbitsky[2,3]

[1] The Institute of Mathematical Sciences, Chennai 600 113, India
{arvind,grattan}@imsc.res.in
[2] Institut für Informatik, Humboldt Universität zu Berlin, Germany
{koebler,verbitsk}@informatik.hu-berlin.de
[3] On leave from the Institute for Applied Problems of Mechanics and Mathematics, Lviv, Ukraine.

Abstract. *Color refinement* is a classical technique used to show that two given graphs G and H are non-isomorphic; it is very efficient, although it does not succeed on all graphs. We call a graph G *amenable* to color refinement if the color-refinement procedure succeeds in distinguishing G from any non-isomorphic graph H.
Tinhofer (1991) explored a linear programming approach to Graph Isomorphism and defined the notion of *compact graphs*: A graph is *compact* if its fractional automorphisms polytope is integral. Tinhofer noted that isomorphism testing for compact graphs can be done quite efficiently by linear programming. However, the problem of characterizing and recognizing compact graphs in polynomial time remains an open question. Our results are summarized below:
– We determine the exact range of applicability of color refinement by showing that amenable graphs are recognizable in time $O((n+m)\log n)$, where n and m denote the number of vertices and the number of edges in the input graph.
– We show that all amenable graphs are compact. Thus, the applicability range for Tinhofer's linear programming approach to isomorphism testing is at least as large as for the combinatorial approach based on color refinement.
– Exploring the relationship between color refinement and compactness further, we study related combinatorial and algebraic graph properties introduced by Tinhofer and Godsil. We show that the corresponding classes of graphs form a hierarchy and we prove that recognizing each of these graph classes is P-hard. In particular, this gives a first complexity lower bound for recognizing compact graphs.

1 Introduction

The well-known *color refinement* (also known as *naive vertex classification*) procedure for Graph Isomorphism works as follows: it begins with a uniform coloring of the vertices of two graphs G and H and refines the vertex coloring step by step. In a refinement step, if two vertices have identical colors but differently colored neighborhoods (with the multiplicities of colors counted), then these vertices get new different colors. The procedure terminates when no further refinement of the vertex color classes is possible. Upon termination, if the multisets of vertex colors in G and H are different, we can correctly conclude that they are not isomorphic. However, color refinement sometimes

* This work was supported by the Alexander von Humboldt Foundation in its research group linkage program. The second author and the fourth author were supported by DFG grants KO 1053/7-2 and VE 652/1-2, respectively.

ISSN 1433-8092

fails to distinguish non-isomorphic graphs. The simplest example is given by any two non-isomorphic regular graphs of the same degree with the same number of vertices. Nevertheless, color refinement turns out to be a useful tool not only in isomorphism testing but also in a number of other areas; see [13, 16, 21] and references there.

For which pairs of graphs G and H does the color refinement procedure succeed in solving Graph Isomorphism? Mainly this question has motivated the study of color refinement from different perspectives.

Immerman and Lander [14], in their highly influential paper, established a close connection between color refinement and 2-variable first-order logic with counting quantifiers. They show that color refinement distinguishes G and H if and only if these graphs are distinguishable by a sentence in this logic.

A well-known approach to tackling intractable optimization problems is to consider an appropriate linear programming relaxation. A similar approach to isomorphism testing, based on the notion of a fractional isomorphism (introduced by Tinhofer [22] using the term doubly stochastic isomorphism), turns out to be equivalent to color refinement. Building on Tinhofer's work [22], it is shown by Ramana, Scheinerman and Ullman [20] (see also Godsil [11]) that two graphs are indistinguishable by color refinement if and only if they are fractionally isomorphic.

We say that color refinement *applies* to a graph G if it succeeds in distinguishing G from any non-isomorphic H. A graph to which color refinement applies is called *amenable*. There are interesting classes of amenable graphs:

1. An obvious class of graphs to which color refinement is applicable is the class of *unigraphs*. Unigraphs are graphs that are determined up to isomorphism by their degree sequences; see, e.g., [5, 26].
2. Trees are amenable (Edmonds [6, 27]).
3. It is easy to see that all graphs for which the color refinement procedure terminates with all singleton color classes (i.e. the color classes form the discrete partition) are amenable. Babai, Erdös, and Selkow [2] have shown that a random graph $G_{n,1/2}$ has this property with high probability. Moreover, the discrete partition of $G_{n,1/2}$ is reached within at most two refinement steps. This implies that graph isomorphism is solvable very efficiently in the average case (see [3]).

The concept of a fractional isomorphism was used by Tinhofer in [24] as a basis for yet another linear-programming approach to isomorphism testing. Tinhofer calls a graph G *compact* if the polytope of all its fractional automorphisms is integral; more precisely, if A is the adjacency matrix of G, then the polytope in $\mathbb{R}^{n^2}$ consisting of the doubly stochastic matrices X such that $AX = XA$ has only integral extreme points (i.e. all coordinates of these points are integer).

If a compact graph G is isomorphic to another graph H, then the polytope of fractional isomorphisms from G to H is also integral. If G is not isomorphic to H, then this polytope has *no* integral extreme point. Thus, isomorphism testing for compact graphs can be done in polynomial time by using linear programming to compute an extreme point of the polytope and testing if it is integral.

Before testing isomorphism of given G and H in this way, we need to know that G is compact. Unfortunately, no efficient characterization of these graphs is currently known.

Our results

What is the class of graphs to which color refinement applies? The logical and linear programming based characterizations of color refinement do not provide any efficient criterion answering this question.

We aim at determining the exact range of applicability of color refinement. We find an efficient characterization of the entire class of amenable graphs, which allows for a quasilinear-time test whether or not color refinement applies to a given graph. This result is shown in Section 5, after we unravel the structure of amenable graphs in Sections 3 and 4. We note that a weak *a priori* upper bound for the complexity of recognizing amenable graphs is $\mathsf{coNP}^{\mathsf{GI}[1]}$, where the superscript means the one-query access to an oracle solving the graph isomorphism problem. To the best of our knowledge, no better upper bound was known before.

Combined with the Immerman-Lander result [14] mentioned above, it follows that the class of graphs definable by first-order sentences with 2 variables and counting quantifiers is recognizable in polynomial time.

As our second main result, in Sections 6 and 7 we show that all amenable graphs are compact. Thus, Tinhofer's approach to Graph Isomorphism [24] has at least as large an applicability range as color refinement. In fact, the former approach is even more powerful because it is known that the class of compact graphs contains many regular graphs (for example, all cycles [22]), for which no nontrivial color refinement is possible.

In Section 8, we look at the relationship between the concepts of compactness and color refinement also from the other side. Let us call a graph G *refinable* if the color partition produced by color refinement coincides with the orbit partition of the automorphism group of G. It is interesting to note that the color-refinement procedure gives an efficient algorithm to check if a given refinable graph has a nontrivial automorphism. It follows from the results in [24] that all compact graphs are refinable. The inclusion $\mathsf{Amenable} \subset \mathsf{Compact}$, therefore, implies that all amenable graphs are refinable as well. The last result is independently obtained in [17] by a different argument. In the particular case of trees, this fact was observed long ago by several authors; see a survey in [25].

Taking a finer look at the inclusion $\mathsf{Compact} \subset \mathsf{Refinable}$, in Section 8 we discuss algorithmic and algebraic graph properties that were introduced by Tinhofer [24] and Godsil [11]. We note that, along with the other graph classes under consideration, the corresponding classes $\mathsf{Tinhofer}$ and Godsil form a hierarchy under inclusion:

$$\mathsf{Discrete} \subset \mathsf{Amenable} \subset \mathsf{Compact} \subset \mathsf{Godsil} \subset \mathsf{Tinhofer} \subset \mathsf{Refinable}. \qquad (1)$$

We show the following results on these graph classes:

- The hierarchy (1) is strict.

- Testing membership in any of these graph classes is P-hard.

We prove the last fact by giving a suitable uniform AC^0 many-one reduction from the P-complete monotone boolean circuit-value problem MCVP. More precisely, for a given MCVP instance (C, x) our reduction outputs a vertex-colored graph $G_{C,x}$ such that if $C(x) = 1$ then $G_{C,x}$ is discrete and if $C(x) = 0$ then $G_{C,x}$ is not refinable. In particular, the graph classes Discrete and Amenable are P-complete. We note that Grohe [12] established, for each $k \geq 2$, the P-completeness of the equivalence problem for first-order k-variable logic with counting quantifiers; according to [14], this implies the P-completeness of indistinguishability of two input graphs by color refinement. We adapt the gadget constructions in [12] to show our P-hardness results.

Note. At the same time as our result appeared in an e-print [1], Sandra Kiefer, Pascal Schweitzer, and Erkal Selman announced independently a similar characterization of amenable graphs that subsequently appeared as an e-print [17].

Notation. The vertex set of a graph G is denoted by $V(G)$. The vertices adjacent to a vertex $u \in V(G)$ form its neighborhood $N(u)$. A set of vertices $X \subseteq V(G)$ induces a subgraph of G, that is denoted by $G[X]$. For two disjoint sets X and Y, $G[X, Y]$ is the bipartite graph with vertex classes X and Y formed by all edges of G connecting a vertex in X with a vertex in Y. The vertex-disjoint union of graphs G and H will be denoted by $G + H$. Furthermore, we write mG for the disjoint union of m copies of G. The *bipartite complement* of a bipartite graph G with vertex classes X and Y is the bipartite graph G' with the same vertex classes such that $\{x, y\}$ with $x \in X$ and $y \in Y$ is an edge in G' if and only if it is not an edge in G. We use the standard notation K_n for the complete graph on n vertices, $K_{s,t}$ for the complete bipartite graph whose vertex classes have s and t vertices, and C_n for the cycle on n vertices.

2 Basic definitions and facts

For convenience, we will consider graphs to be vertex-colored in the paper. A *vertex-colored graph* is an undirected simple graph G endowed with a vertex coloring $c : V(G) \to \{1, \ldots, k\}$. Automorphisms of a vertex-colored graph and isomorphisms between vertex-colored graphs are required to preserve vertex colors. We get usual graphs when c is constant.

Given a graph G, the *color-refinement* algorithm (to be abbreviated as *CR*) iteratively computes a sequence of colorings C^i of $V(G)$. The initial coloring C^0 is the vertex coloring of G, i.e., $C^0(u) = c(u)$. Then,

$$C^{i+1}(u) = (C^i(u), \{\!\{ C^i(a) : a \in N(u) \}\!\}), \tag{2}$$

where $\{\!\{\ldots\}\!\}$ denotes a multiset.

The partition $\mathcal{P}^{i+1}$ of $V(G)$ into the color classes of C^{i+1} is a refinement of the partition $\mathcal{P}^i$ corresponding to C^i. It follows that, eventually, $\mathcal{P}^{s+1} = \mathcal{P}^s$ for some s; hence, $\mathcal{P}^i = \mathcal{P}^s$ for all $i \geq s$. The partition $\mathcal{P}^s$ is called the *stable partition* of G and denoted by $\mathcal{P}_G$.

Given a partition $\mathcal{P}$ of the vertex set of a graph G, we call its elements *cells*. We call $\mathcal{P}$ *equitable* if:

(i) Each cell $X \in \mathcal{P}$ is monochromatic, i.e., all vertices $u, v \in X$ have the same color $c(u) = c(v)$.

(ii) For any cell $X \in \mathcal{P}$ the graph $G[X]$ induced by X is *regular*, that is, all vertices in $G[X]$ have equal degrees.

(iii) For any two cells $X, Y \in \mathcal{P}$ the bipartite graph $G[X, Y]$ induced by X and Y is *biregular*, that is, all vertices in X have equally many neighbors in Y and vice versa.

It is easy to see that the stable partition of G is equitable; our analysis in the next section will make use of this fact.

A straightforward inductive argument shows that the colorings C^i are preserved under isomorphisms.

Lemma 1. *If ϕ is an isomorphism from G to H, then $C^i(u) = C^i(\phi(u))$ for any vertex u of G.*

Lemma 1 readily implies that, if graphs G and H are isomorphic, then

$$\{\!\{ C^i(u) : u \in V(G) \}\!\} = \{\!\{ C^i(v) : v \in V(H) \}\!\} \tag{3}$$

for all $i \geq 0$. When used for isomorphism testing, the CR algorithm accepts two graphs G and H as isomorphic exactly when the above condition is met on input $G + H$. Note that this condition is actually finitary: If Equality (3) is false for some i, it must be false for some $i < 2n$, where n denotes the number of vertices in each of the graphs. This follows from the observation that the partition $\mathcal{P}^{2n-1}$ induced by the coloring C^{2n-1} must be the stable partition of the disjoint union of G and H. In fact, Equality (3) holds true for all i if it is true for $i = n$; see, e.g., [19]. Thus, it is enough that CR verifies (3) for $i = n$.

Note that computing the vertex colors literally according to (2) would lead to an exponential growth of the lengths of color names. This can be avoided by renaming the colors after each refinement step. Then CR never needs more than n color names (appearance of more than n colors is an indication that the graphs are non-isomorphic).

Definition 2. *We call a graph G* amenable *if CR works correctly on the input G, H for every H, that is, Equality (3) is false for $i = n$ whenever $H \not\cong G$.*

3 Local structure of amenable graphs

Consider the stable partition $\mathcal{P}_G$ of an amenable graph G. The following lemma gives a list of all possible regular and biregular graphs that can occur, respectively, as $G[X]$ and $G[X, Y]$ for cells X, Y of $\mathcal{P}_G$.

Lemma 3. *The stable partition $\mathcal{P}_G$ of an amenable graph G fulfills the following properties:*

(A) *For any cell $X \in \mathcal{P}_G$, $G[X]$ is an empty graph, a complete graph, a matching graph mK_2, the complement of a matching graph, or the 5-cycle;*

(B) *For any two cells $X, Y \in \mathcal{P}_G$, $G[X, Y]$ is an empty graph, a complete bipartite graph, a disjoint union of stars $sK_{1,t}$ where X and Y are the set of s central vertices and the set of st leaves, or the bipartite complement of the last graph.*

The proof of Lemma 3 is based on the following facts.

Lemma 4 (Johnson [15]). *A regular graph of degree d with n vertices is a unigraph if and only if $d \in \{0, 1, n-2, n-1\}$ or $d = 2$ and $n = 5$.*[4]

Lemma 5 (Koren [18]). *A bipartite graph is determined up to isomorphism by the conditions that every of the m vertices in one part has degree c and every of the n vertices in the other part has degree d if and only if $c \in \{0, 1, n-1, n\}$ or $d \in \{0, 1, m-1, m\}$.*

If G contains a subgraph $G[X]$ or $G[X, Y]$ that is induced by some $X, Y \in \mathcal{P}_G$ but not listed in Lemma 3, then Lemmas 4 and 5 imply that this subgraph can be replaced by a non-isomorphic regular or biregular graph with the same parameters. Hence, in order to prove Lemma 3 it suffices to show that the resulting graph H is indistinguishable from G by color refinement. The graphs G and H in the following lemma have the same vertex set. Given a vertex u, we distinguish its neighborhoods $N_G(u)$ and $N_H(u)$ and its colors $C^i_G(u)$ and $C^i_H(u)$ in the two graphs.

Lemma 6. *Let X and Y be cells of the stable partition of a graph G.*

(i) If H is obtained from G by replacing the edges of the subgraph $G[X]$ with the edges of an arbitrary regular graph (X, E) having the same degree, then $C^i_G(u) = C^i_H(u)$ for any $u \in V(G)$ and any i.

(ii) If H is obtained from G by replacing the edges of the subgraph $G[X, Y]$ with the edges of an arbitrary biregular graph with the same vertex partition such that the vertex degrees remain unchanged, then $C^i_G(u) = C^i_H(u)$ for any $u \in V(G)$ and any i.

Proof of Lemma 6. We proceed by induction on i. In the base case of $i = 0$ the claim is trivially true. Assume that $C^i_G(a) = C^i_H(a)$ for all $a \in V(G)$. We consider an arbitrary vertex u and prove that

$$C^{i+1}_G(u) = C^{i+1}_H(u). \tag{4}$$

From now on we treat Parts (i) and (ii) separately.

(i) Suppose first that $u \notin X$. Since the transformation of G into H does not affect the edges emanating from u, we have $N_G(u) = N_H(u)$. Looking at the definition (2), we immediately derive (4) from the induction assumption.

If $u \in X$, we only have the equality $N_G(u) \setminus X = N_H(u) \setminus X$, implying that

$$\{\!\!\{\, C^i_G(a) : a \in N_G(u) \setminus X \,\}\!\!\} = \{\!\!\{\, C^i_H(a) : a \in N_H(u) \setminus X \,\}\!\!\}. \tag{5}$$

[4] The last case, in which the graph is the 5-cycle, is missing from the statement of this result in [15, Theorem 2.12]. The proof in [15] tacitly considers only graphs with at least 6 vertices.

The equality $N_G(u) \cap X = N_H(u) \cap X$ is not necessarily true. However, u has equally many neighbors from X in G and in H. Furthermore, for any two vertices a and a' in X we have $C^i_G(a) = C^i_G(a')$ because X is a cell of G, and $C^i_H(a) = C^i_G(a) = C^i_G(a') = C^i_H(a')$ by the induction assumption. That is, all vertices in X have the same C^i-color both in G and in H. It follows that

$$\{\!\!\{\, C^i_G(a) : a \in N_G(u) \cap X \,\}\!\!\} = \{\!\!\{\, C^i_H(a) : a \in N_H(u) \cap X \,\}\!\!\} . \tag{6}$$

Combining (5) and (6), we conclude that (4) holds in any case.

(ii) If $u \notin X \cup Y$, we have $N_G(u) = N_H(u)$ and Equality (4) readily follows from the induction assumption.

Suppose that $u \in Y$. In this case we still have (5) and, exactly as in Part (i), we also derive (6). Equality (4) follows. The case of $u \in X$ is symmetric. □

Proof of Lemma 3. **(A)** If $G[X]$ is a graph not from the list, by Lemma 4, it is not a unigraph. Hence, we can modify G locally on X by replacing $G[X]$ with a non-isomorphic regular graph with the same parameters. Part (i) of Lemma 6 implies that the resulting graph H satisfies Equality (3) for any i, implying that CR does not distinguish between G and H. The graphs G and H are non-isomorphic because, by Part (i) of Lemma 6 and by Lemma 1, an isomorphism from G to H would induce an isomorphism from $G[X]$ to $H[X]$. This shows that G is not amenable.

(B) This condition follows, similarly to Condition **A**, from Lemma 5 and Part (ii) of Lemma 6. □

4 Global structure of amenable graphs

Recall that $\mathcal{P}_G$ is the stable partition of the vertex set of a graph G, and that elements of $\mathcal{P}_G$ are called cells. We define the auxiliary *cell graph* $C(G)$ of G to be the complete graph on the vertex set $\mathcal{P}_G$ with the following labeling of vertices and edges. A vertex X of $C(G)$ is called *homogeneous* if the graph $G[X]$ is either complete or empty and *heterogeneous* otherwise. An edge $\{X, Y\}$ of $C(G)$ is called *isotropic* if the bipartite graph $G[X, Y]$ is either complete or empty and *anisotropic* otherwise. A path $X_1 X_2 \ldots X_l$ in $C(G)$ where every edge $\{X_i, X_{i+1}\}$ is anisotropic will be referred to as an *anisotropic path.* If also $\{X_l, X_1\}$ is an anisotropic edge, we speak of an *anisotropic cycle.* In the case that $|X_1| = |X_2| = \ldots = |X_l|$, such a path (or cycle) is called *uniform.*

For graphs fulfilling Conditions **A** and **B** of Lemma 3 we refine the labeling of the vertices and edges of $C(G)$ as follows. A heterogeneous cell $X \in \mathcal{P}_G$ is called *matching*, *co-matching*, or *pentagonal* depending on the type of $G[X]$. Note that a matching or co-matching cell X always consists of at least 4 vertices. Further, an anisotropic edge $\{X, Y\}$ is called *constellation* if $G[X, Y]$ is a disjoint union of stars, and *co-constellation* otherwise (i.e., the bipartite complement of $G[X, Y]$ is a disjoint union of stars). Likewise, homogeneous cells X (and isotropic edges $\{X, Y\}$) are called *empty* if the graph $G[X]$ (resp. $G[X, Y]$) is empty, and *complete* otherwise.

Note that if an edge $\{X, Y\}$ of a uniform path/cycle is *constellation* (resp. *co-constellation*), then $G[X, Y]$ is a matching (resp. co-matching) graph.

Lemma 7. *The cell graph $C(G)$ of an amenable graph G has the following properties:*

(C) *$C(G)$ contains no uniform anisotropic path connecting two heterogeneous cells;*

(D) *$C(G)$ contains no uniform anisotropic cycle;*

(E) *$C(G)$ contains neither an anisotropic path $XY_1 \dots Y_l Z$ such that $|X| < |Y_1| = \dots = |Y_l| > |Z|$ nor an anistropic cycle $XY_1 \dots Y_l X$ such that $|X| < |Y_1| = \dots = |Y_l|$;*

(F) *$C(G)$ contains no anisotropic path $XY_1 \dots Y_l$ such that $|X| < |Y_1| = \dots = |Y_l|$ and the cell Y_l is heterogeneous.*

Proof. **(C)** Suppose that P is a uniform anisotropic path in $C(G)$ connecting two heterogeneous cells X and Y. Let $k = |X| = |Y|$. Complementing $G[A, B]$ for each co-constellation edge $\{A, B\}$ of P, in G we obtain k vertex-disjoint paths connecting X and Y. These paths determine a one-to-one correspondence between X and Y. Given $v \in X$, denote its mate in Y by v^*. Call P *conducting* if this correspondence is an isomorphism between $G[X]$ and $G[Y]$, that is, two vertices u and v in X are adjacent exactly when their mates u^* and v^* are adjacent. In the case that one of X and Y is matching and the other is co-matching, we call P *conducting* also if the correspondence is an isomorphism between $G[X]$ and the complement of $G[Y]$.

We construct a non-isomorphic graph H such that CR does not distinguish between G and H. Since Y is heterogeneous, we can replace the edges of the subgraph $G[Y]$ with the edges of an isomorphic but different subgraph (Y, E). Since also X is heterogeneous it follows that P is a conducting path in the resulting graph H if and only if P is a non-conducting path in G. Now, Part (i) of Lemma 6 implies that CR computes the same stable partition for G and H and does not distinguish between them. On the other hand, Lemma 1 implies that any isomorphism ϕ between G and H must map each cell to itself. As ϕ must also preserve the conducting property along the path P, it follows that G and H are not isomorphic. Hence, G is not amenable.

(D) Suppose that $C(G)$ contains a uniform anisotropic cycle Q of length m. All cells in Q have the same cardinality; denote it by k. Complementing $G[A, B]$ for each co-constellation edge $\{A, B\}$ of Q, in G we obtain the vertex-disjoint union of cycles whose lengths are multiples of m. As two extreme cases, we can have k cycles of length m each or we can have a single cycle of length km. Denote the isomorphism type of this union of cycles by $\tau(Q)$. Note that this type is isomorphism invariant: For an isomorphism ϕ from G to another graph H, $\tau(\phi'(Q)) = \tau(Q)$ for the induced isomorphism ϕ' from $C(G)$ to $C(H)$.

Let X and Y be two consecutive cells in Q. We can replace the subgraph $G[X, Y]$ with an isomorphic but different bipartite graph so that in the resulting graph H, $\tau(Q)$ becomes either kC_m or C_{km}, whatever we wish. In particular, we can replace the subgraph $G[X, Y]$ in such a way that $\tau(Q)$ is changed.

Similarly as for Condition **C**, we use Part (ii) of Lemma 6 to argue that CR does not distinguish between G and H. Furthermore, $G \ncong H$ because the types $\tau(Q)$ in G and H are different. Therefore, G is not amenable.

(E) Suppose that $C(G)$ contains an anisotropic path $P = XY_1 \dots Y_l Z$ such that $|X| < |Y_1| = \dots = |Y_l| > |Z|$ (for the case of a cycle, where $Z = X$, the argument is virtually the same). Let $G[X, Y_1] = sK_{1,t}$ and $G[Z, Y_l] = aK_{1,b}$, where $s, a, t, b \geq 2$ (if any of these subgraphs is a co-constellation, we consider its complement). Thus, $|X| = s$, $|Z| = a$, and $|Y_1| = |Y_l| = st = ab$.

Like in the proof of Condition **C**, the uniform anisotropic path $Y_1 \dots Y_l$ determines a one-to-one correspondence between the cells Y_1 and Y_l that can be used to make the identification $Y_1 = Y_l = \{1, 2, \dots, st\} = Y$. For each $x \in X$, let Y_x denote the set of vertices in Y adjacent to x. The set Y_z is defined similarly for each $z \in Z$. Note that for any $x \neq x'$ in X and $z \neq z'$ in Z,

$$|Y_x| = t, \quad |Y_z| = b, \quad Y_x \cap Y_{x'} = \emptyset, \quad \text{and} \quad Y_z \cap Y_{z'} = \emptyset.$$

We regard $\mathcal{Y}_G = \{\, Y_x \,\}_{x \in X} \cup \{\, Y_z \,\}_{z \in Z}$ as a hypergraph on the vertex set Y. Note that $\mathcal{Y}_G$ might be a multi-hypergraph as the two hyperedges Y_x and Y_z might coincide for some pairs $(x, z) \in X \times Z$. Without loss of generality, we can assume that the hyperedges Y_x, $x \in X$, form consecutive intervals in Y. We call the anisotropic path P *flat*, if there exists no pair $(x, z) \in X \times Z$ such that one of the two hyperedge Y_x and Y_z is contained in the other.

We construct a non-isomorphic graph H such that CR does not distinguish between G and H. If P is flat in G, we replace the edges of the subgraph $G[Z, Y_l]$ by the edges of an isomorphic but different biregular graph such that P becomes non-flat in the resulting graph H. More precisely, we replace the edges in such a way that all hyperedges of $\mathcal{Y}_H$ form consecutive intervals in Y by letting $\mathcal{Y}_H = \{\, Y_x \,\}_{x \in X} \cup \{\, Y_i \,\}_{i \in [a]}$, where $Y_i = \{(i-1)b+1, \dots, ib\}$. Likewise, if P is non-flat in G, we replace the edges of $G[Z, Y_l]$ such that P becomes flat in H by letting $\mathcal{Y}_H = \{\, Y_x \,\}_{x \in X} \cup \{\, Y_i \,\}_{i \in [a]}$, where $Y_i = \{i, i + a, i + (b - 1)a\}$.

Now, Part (i) of Lemma 6 implies that CR computes the same stable partition for G and H and does not distinguish between them. On the other hand, Lemma 1 implies that any isomorphism ϕ between G and H must map each cell to itself. As ϕ must also preserve the flatness property of the path P, it follows that G and H are not isomorphic. Hence, G is not amenable.

(F) Suppose that $C(G)$ contains an anisotropic path $XY_1 \dots Y_l$ where $|X| < |Y_1| = \dots = |Y_l|$ and Y_l is heterogeneous. Let $G[X, Y_1] = sK_{1,t}$ (in the case of a co-constellation, we consider the complement). Since $s, t \geq 2$ and $|Y_1| = st$, the cell Y_l cannot be pentagonal. Considering the complement if needed, we can assume without loss of generality that Y_l is matching. Like in the proof of Condition **E**, the uniform anisotropic path $Y_1 \dots Y_l$ determines a one-to-one correspondence between the cells Y_1 and Y_l that can be used to make the identification $Y_1 = Y_l = \{1, 2, \dots, st\} = Y$. Consider the hypergraph $\mathcal{Y}_G = \{\, Y_x \,\}_{x \in X} \cup E(G[Y_l])$, where $Y_x = N_G(x) \cap Y_1$ and $E(G[Y_l])$ denotes the edge set of $G[Y_l]$. Now, exactly as in the proof of Condition **E**, we can change the isomorphism type of $\mathcal{Y}_G$ by replacing the edges of the subgraph $G[X, Y_1]$ by the edges of an isomorphic biregular graph. This yields a non-isomorphic graph H that is indistinguishable from G by CR. □

It turns out that Conditions **A**–**F** are not only necessary for amenability (as shown in Lemmas 3 and 7) but also sufficient. As a preparation we first prove the following Lemma 8 that reveals a tree-like structure of amenable graphs. By an *anisotropic component* of the cell graph $C(G)$ we mean a maximal connected subgraph of $C(G)$ whose edges are all anisotropic. Note that if a vertex of $C(G)$ has no incident anisotropic edges, it forms a single-vertex anisotropic component.

Lemma 8. *Suppose that a graph G satisfies Conditions* **A**–**F**. *Then for any anisotropic component A of $C(G)$, the following is true.*

(G) *A is a tree with the following monotonicity property. Let R be a cell in A of minimum cardinality and let A_R be the rooted directed tree obtained from A by rooting A at R. Then $|X| \le |Y|$ for any directed edge (X, Y) of A_R.*

(H) *A contains at most one heterogeneous vertex. If R is such a vertex, it has minimum cardinality among the cells of A.*

Proof. **(G)** A cannot contain any uniform cycle by Condition **D** and any other cycle by Condition **E**. The monotonicity property follows from Condition **E**.

(H) Assume that A contains more than one heterogeneous cell. Consider two such cells S and T. Let $S = Z_1, Z_2, \ldots, Z_l = T$ be the path from S to T in A. The monotonicity property stated in Condition **G** implies that there is j (possibly $j = 1, l$) such that $|Z_1| \ge \ldots \ge |Z_j| \le \ldots \le |Z_l|$. Since the path cannot be uniform by Condition **C**, at least one of the inequalities is strict. However, this contradicts Condition **F**.

Suppose that R is a heterogeneous cell in A. Consider now a path $R = Z_1, Z_2, \ldots, Z_l = S$ in A where S is a cell with the smallest cardinality. By the monotonicity property and Condition **F**, this path must be uniform, proving that $|R| = |S|$. □

In combination with Conditions **A** and **B**, Conditions **G** and **H** on anisotropic components give a very stringent characterization of amenability.

Theorem 9. *For a graph G the following conditions are equivalent:*

(i) G is amenable.

(ii) G satisfies Conditions **A**–**F**.

(iii) G satisfies Conditions **A, B, G** *and* ***H***.

Proof. It only remains to show that any graph G fulfilling the Conditions **A, B, G** and **H** is amenable. Let H be a graph indistinguishable from G by CR. Then we have to show that G and H are isomorphic.

Consider the coloring C^s corresponding to the stable partition $\mathcal{P}^s$ of the disjoint union $G + H$. Since G and H satisfy Equality (3) for $i = s$, there is a bijection $f : \mathcal{P}_G \to \mathcal{P}_H$ matching each cell X of the stable partition of G to the cell $f(X) \in \mathcal{P}_H$ such that the vertices in X and $f(X)$ have the same C^s-color. Moreover, Equality (3) implies that $|X| = |f(X)|$. We claim that for any cells X and Y of G,

(a) $G[X] \cong H[f(X)]$ and

(b) $G[X, Y] \cong H[f(X), f(Y)]$,

implying that f is an isomorphism from $C(G)$ to $C(H)$.

Indeed, since X and $f(X)$ are cells of the stable partitions $\mathcal{P}_G$ and $\mathcal{P}_H$, both $G[X]$ and $H[f(X)]$ are regular. Since $X \cup f(X)$ is a cell of the stable partition $\mathcal{P}^s$ of $G+H$, the graphs $G[X]$ and $H[f(X)]$ have the same degree. By Condition **A**, $G[X]$ is a unigraph, implying Property (a). Property (b) follows from Condition **B** by a similar argument.

We now construct an isomorphism ϕ from G to H. By Lemma 1, we should have $\phi(X) = f(X)$ for each cell X. Therefore, we have to define the map $\phi : X \to f(X)$ on each X.

By Condition **H**, an anisotropic component A of the cell graph $C(G)$ contains at most one heterogeneous cell. Denote it by R_A if it exists. Otherwise fix R_A to be an arbitrary cell of the minimum cardinality in A.

For each A, define ϕ on $R = R_A$ to be an arbitrary isomorphism from $G[R]$ to $H[f(R)]$, which exists according to (a). After this, propagate ϕ to any other cell in A as follows. By Condition **G**, A is a tree. Let A_R be the directed rooted tree obtained from A by rooting it at R. Suppose that ϕ is already defined on X and (X, Y) is an edge in A. By the monotonicity property in Condition **G** and our choice of R, we can assume that $|X| \le |Y|$. Then ϕ can be extended to Y so that this is an isomorphism from $G[X, Y]$ to $H[f(X), f(Y)]$. This is possible by (b) due to the fact that all vertices in Y have degree 1 in $G[X, Y]$ or its bipartite complement (and the same holds for all vertices in $f(Y)$ in the graph $H[f(X), f(Y)]$).

It remains to argue that the map ϕ obtained in this way is indeed an isomorphism from G to H. It suffices to show that ϕ is an isomorphism between $G[X]$ and $H[f(X)]$ for each cell X of G and between $G[X, Y]$ and $H[f(X), f(Y)]$ for each pair of cells X and Y.

If X is homogeneous, $f(X)$ is homogeneous of the same type, complete or empty, according to (a). In this case, any ϕ is an isomorphism from $G[X]$ to $H[f(X)]$. If X is heterogeneous, the assumption of the lemma says that it belongs to a unique anisotropic component A (and $X = R_A$). Then ϕ is an isomorphism from $G[X]$ to $H[f(X)]$ by construction.

If $\{X, Y\}$ is an isotropic edge of $C(G)$, then (b) implies that $\{f(X), f(Y)\}$ is an isotropic edge of $C(H)$ of the same type, complete or empty. In this case, ϕ is an isomorphism from $G[X, Y]$ to $H[f(X), f(Y)]$, no matter how it is defined. If $\{X, Y\}$ is anisotropic, it belongs to some anisotropic component A, and ϕ is an isomorphism from $G[X, Y]$ to $H[f(X), f(Y)]$ by construction. $\square$

5 Examples and applications

Theorem 9 is a convenient tool for verifying amenability. For example, amenability of discrete graphs is a well-known fact. Recall that those are graphs whose stable partitions consist of singletons. As each cell is a singleton, any anisotropic component of a discrete graph consists of a single cell. Hence, Conditions **A** and **B** as well as Conditions **G** and **H** on anisotropic components are fulfilled by trivial reasons.

Checking these four conditions, we can also reprove the amenability of trees. Moreover, we can extend this result to the class of forests. This extension does not seem to be straightforward because the class of amenable graphs is not closed under disjoint unions. For example, $C_3 + C_4$ is indistinguishable by CR from C_7 and, hence, is not amenable.

Corollary 10. *All forests are amenable.*

Proof. A regular acyclic graph is either an empty or a matching graph. This implies Condition **A**. Condition **B** follows from the observation that biregular acyclic graphs are either empty or disjoint unions of stars.

Let $C^*(G)$ be the version of the cell graph $C(G)$ where all empty edges are removed. If $C^*(G)$ contains a cycle, G must contain a cycle as well. Therefore, if G is acyclic, then $C^*(G)$ is acyclic too, and any anisotropic component of $C(G)$ must be a tree. To prove the monotonicity property in Condition **G**, it suffices to show that $C(G)$ cannot contain an anisotropic path $XY_1 \dots Y_l Z$ with $|X| < |Y_1| = \dots = |Y_l| > |Z|$. But this easily follows since in this case each vertex of the induced subgraph $G[X \cup Y_1 \cup \dots \cup Y_l \cup Z]$ has degree at least 2 in G, contradicting the acyclicity of G.

To prove Condition **H**, suppose that $C(G)$ contains an anisotropic path $X_0, X_1, \dots, X_l$ connecting two heterogeneous cells X_0 and X_l. Then each vertex of the induced subgraph $G[X_0 \cup X_1 \cup \dots \cup X_{l-1} \cup X_l]$ has degree at least 2 in G, a contradiction. The same contradiction arises if such a path connects a heterogeneous cell X_0 with an arbitrary cell X_l, where $|X_l| < |X_{l-1}|$. Hence, X_0 must have minimum cardinality among all cells belonging to the same anisotropic component. ◻

Our characterization of amenable graphs via Conditions **A**, **B**, **G** and **H** leads to an efficient test for amenability of a given graph, that has the same time complexity as CR. It is known (Cardon and Crochemore [8]; see also [4]) that the stable partition of a given graph G can be computed in time $O((n + m) \log n)$. It is supposed that G is presented by its adjacency list.

Corollary 11. *The class of amenable graphs is recognizable in time $O((n + m) \log n)$, where n and m denote the number of vertices and edges of the input graph.*

Proof. Using known algorithms, we first compute the stable partition $\mathcal{P}_G = \{X_1, \dots, X_k\}$ of the input graph G. Let $C^*(G)$ be the version of the cell graph $C(G)$ where all empty edges are removed. We can compute the adjacency list of each vertex X_i of $C^*(G)$ by traversing the adjacency list of an arbitrary vertex $u \in X_i$ and listing all cells X_j that contain a vertex v adjacent to u. Simultaneously, we compute for each pair (i, j) such that $i = j$ or $\{X_i, X_j\}$ is an edge of $C^*(G)$ the number d_{ij} of neighbors in X_j of any vertex in X_i. Knowing the numbers $|X_i|$, $|X_j|$ and d_{ij} allows us to determine whether all the subgraphs $G[X_i]$ and $G[X_i, X_j]$ fulfill Conditions **A** and **B** of Lemma 3.

To check Conditions **G** and **H** we use breadth-first search in the graph $C^*(G)$ to find all anisotropic components A of $C(G)$ and, simultaneously, to check that each component A is a tree containing at most one heterogeneous cell. If we restart the search from an arbitrary cell in A having minimum cardinality, we can also check for each forward edge of the resulting search tree whether the monotonicity property of Condition **G** is fulfilled. $\square$

We conclude this section y considering logical aspects of our result. A *counting quantifier* $\exists^m$ opens a sentence saying that there are at least m elements satisfying some property. Immerman and Lander [14] discovered an intimate connection between color refinement and 2-variable first-order logic with counting quantifiers. This connection implies that amenability of a graph is equivalent to its definability in this logic. Thus, Corollary 11 asserts that the class of graphs definable by a first-order sentence with counting quantifiers and occurrences of just 2 variables is recognizable in polynomial time.

6 Amenable graphs are compact

An $n \times n$ real matrix X is *doubly stochastic* if its elements are nonnegative and all its rows and columns sum up to 1. Doubly stochastic matrices are closed under products and convex combinations. The set of all $n \times n$ doubly stochastic matrices forms the *Birkhoff polytope* $B_n \subset \mathbb{R}^{n^2}$. *Permutation matrices* are exactly 0-1 doubly stochastic matrices. By Birkhoff's Theorem, the $n!$ permutation matrices form precisely the set of all extreme points of B_n. Equivalently, every doubly stochastic matrix is a convex combination of permutation matrices.

Let G and H be graphs with vertex set $\{1, \ldots, n\}$. An isomorphism π from G to H can be represented by the permutation matrix $P_\pi = (p_{ij})$ such that $p_{ij} = 1$ if and only if $\pi(i) = j$. Denote the set of matrices P_π for all isomorphisms π by $\mathrm{Iso}(G, H)$, and let $\mathrm{Aut}(G) = \mathrm{Iso}(G, G)$.

Let A and B be the adjacency matrices of graphs G and H respectively. If the graphs are uncolored, a permutation matrix X is in $\mathrm{Iso}(G, H)$ if and only if $AX = XB$. For vertex-colored graphs, X must additionally satisfy the condition $X[u, v] = 0$ for all pairs of differently colored u and v, i.e., this matrix must be block-diagonal with respect to the color classes. We say that (vertex-colored) graphs G and H are *fractionally isomorphic* if $AX = XB$ for a doubly stochastic matrix X, where $X[u, v] = 0$ if u and v are of different colors. The matrix X is called a *fractional isomorphism.*

Denote the set of all fractional isomorphisms from G to H by $S(G, H)$ and note that it forms a polytope in $\mathbb{R}^{n^2}$. The set of isomorphisms $\mathrm{Iso}(G, H)$ is contained in $\mathrm{Ext}(S(G, H))$, where $\mathrm{Ext}(Z)$ denotes the set of all extreme points of a set Z. Indeed, $\mathrm{Iso}(G, H)$ is the set of integral extreme points of $S(G, H)$.

The set $S(G) = S(G, G)$ is the polytope of *fractional automorphisms* of G. A graph G is called *compact* [22] if $S(G)$ has no other extreme points than $\mathrm{Aut}(G)$, i.e., $\mathrm{Ext}(S(G)) = \mathrm{Aut}(G)$. Compactness of G can equivalently be defined by any of the following two conditions:

- The polytope $S(G)$ is integral;
- Every fractional automorphism of G is a convex combination of automorphisms of G, i.e., $S(G) = \langle \mathrm{Aut}(G) \rangle$, where $\langle Z \rangle$ denotes the convex hull of a set Z.

Example 12. *Complete graphs are compact as a consequence of Birkhoff's theorem. The compactness of trees and cycles is established in [22]. Matching graphs mK_2 are also compact. This is a particular instance of a much more general result by Tinhofer [24]: If G is compact, then mG is compact for any m. Tinhofer [24] also observes that compact graphs are closed under complement.*

For a negative example, note that the graph $C_3 + C_4$ is not compact. This follows from a general result in [24]: All regular compact graphs must be vertex-transitive (and $C_3 + C_4$ is not).

Tinhofer [24] noted that, if G is compact, then for any graph H, either all or none of the extreme points of the polytope $S(G, H)$ are integral. As mentioned in the introduction, this yields a linear-programming based polynomial-time algorithm to test if a compact graph G is isomorphic to any other given graph H. The following result shows that Tinhofer's approach works for all amenable graphs.

Theorem 13. *All amenable graphs are compact.*

We defer the proof to the next section. Theorem 13 unifies and extends several earlier results providing examples of compact graphs. In particular, it gives another proof of the fact that almost all graphs are compact, which also follows from a result of Godsil [11, Corollary 1.6]. Indeed, while Babai, Erdös, and Selkow [2] proved that almost all graphs are discrete, we already mentioned in Section 1 that all discrete graphs are amenable.

Furthermore, Theorem 13 reproves Tinhofer's result that trees are compact.[5] Since also forests are amenable [1], we can extend this result to forests. This extension is not straightforward as compact graphs are not closed under disjoint union; see Example 12. In [23], Tinhofer proves compactness for the class of *strong tree-cographs*, which includes forests only with pairwise non-isomorphic connected components.

Compactness of unigraphs, which also follows from Theorem 13, seems to be earlier never observed. Summarizing, we note the following result.

Corollary 14. *Discrete graphs, forests, and unigraphs are compact.*

7 Proof of Theorem 13

We will use a known fact on the structure of fractional automorphisms. For a partition $V_1, \ldots, V_m$ of $\{1, \ldots, n\}$ let $X_1, \ldots, X_m$ be matrices, where the rows and columns of X_i are indexed by elements of V_i. Then we denote the block-diagonal matrix with blocks $X_1, \ldots, X_m$ by $X_1 \oplus \cdots \oplus X_m$.

Lemma 15 (Ramana et al. [20]). *Let G be a (vertex-colored) graph on vertex set $\{1, \ldots, n\}$ and assume that the elements $V_1, \ldots, V_m$ of the stable partition $\mathcal{P}_G$ of G are intervals of consecutive integers. Then any fractional automorphism X of G has the form $X = X_1 \oplus \cdots \oplus X_m$.*

[5] The proof of Theorem 13 uses only compactness of complete graphs, matching graphs, and the 5-cycle.

Note that the assumption of the lemma can be ensured for any graph by appropriately renaming its vertices. An immediate consequence of Lemma 15 is that a graph G is compact if and only if it is compact with respect to its stable coloring.

Given an amenable graph G and a fractional automorphism X of G, we have to express X as a convex combination of permutation matrices in $\mathrm{Aut}(G)$. Our proof strategy consists in exploiting the structure of amenable graphs as described by Theorem 9. Given an anisotropic component A of the cell graph $C(G)$, we define the *anisotropic component* G_A *of* G as the subgraph of G induced by the union of all cells belonging to A. Our overall idea is to prove the claim separately for each anisotropic component G_A, applying an inductive argument on the number of cells in A. A key role will be played by the fact that, according to Theorem 9, A is a tree with at most one heterogeneous vertex.

By Lemma 15, we can assume that G is colored by the stable coloring. We first consider the case when G consists of a single anisotropic component. By Theorem 9, the corresponding cell graph $C(G)$ has at most one heterogeneous vertex, and the anisotropic edges form a spanning tree of $C(G)$. Without loss of generality, we can number the cells $V_1, \ldots, V_m$ of G so that V_1 is the unique heterogeneous cell if it exists; otherwise V_1 is chosen among the cells of minimum cardinality. Moreover, we can suppose that, for each $i \le m$, the cells $V_1, \ldots, V_i$ induce a connected subgraph in the tree of anisotropic edges of $C(G)$.

We will prove this case by induction on the number m of cells. In the base case of $m = 1$, our graph $G = G[V_1]$ is one of the graphs listed in Condition **A** of Theorem 9. All of them are known to be compact; see Example 12. As induction hypothesis, assume that the graph $H = G[V_1 \cup \cdots \cup V_{m-1}]$ is compact. For the induction step, we have to show that also $G = G[V_1 \cup \cdots \cup V_m]$ is compact.

Denote $D = V_m$. Since G has no more than one heterogeneous cell, $G[D]$ is complete or empty. It will be instructive to think of D as a "leaf" cell having a unique anisotropic link to the remaining part H of G. Let $C \in \{V_1, \ldots, V_{m-1}\}$ be the unique cell such that $\{C, D\}$ is an anisotropic edge of $C(G)$. To be specific, suppose that $G[C, D] \cong sK_{1,t}$. If $G[C, D]$ is the bipartite complement of $sK_{1,t}$, we can consider the complement of G and use the facts that the class of amenable graphs is closed under complementation and that complementation does not change fractional isomorphisms of the graph. By the monotonicity property stated in Condition **C** of Theorem 9, $|C| = s$ and $|D| = st$. Let $C = \{c_1, c_2, \ldots, c_s\}$ and, for each i, $N(c_i) \subset D$ be the neighborhood of c_i in $G[C, D]$. Thus, $D = \bigcup_{i=1}^{s} N(c_i)$.

Let X be a fractional automorphism of G. It is convenient to break it up into three blocks $X = X' \oplus Y \oplus Z$, where Y and Z correspond to C and D respectively, and X' is the rest. By induction hypothesis we have the convex combination

$$X' \oplus Y = \sum_{P' \oplus P \in \mathrm{Aut}(H)} \alpha_{P',P}\, P' \oplus P, \tag{7}$$

where $P' \oplus P$ are permutation matrices corresponding to automorphisms π of the graph H, such that the permutation matrix block P denotes the action of π on the color class C and P' the action on the remaining color classes of H.

We need to show that X is a convex combination of automorphisms of G. Let A denote the adjacency matrix of G, and $A_{S,T}$ denote the submatrix of A row-indexed by $S \subset V(G)$ and column-indexed by $T \subset V(G)$. Since X is a fractional automorphism of G, we have $XA = AX$. Recall that Y and Z are blocks of X corresponding to color classes C and D. Looking at the corner fragments of the matrices XA and AX, we get

$$\begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} \begin{pmatrix} A_{C,C} & A_{C,D} \\ A_{D,C} & A_{D,D} \end{pmatrix} = \begin{pmatrix} A_{C,C} & A_{C,D} \\ A_{D,C} & A_{D,D} \end{pmatrix} \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix},$$

which implies

$$Y A_{C,D} = A_{C,D}\, Z, \tag{8}$$
$$A_{D,C}\, Y = Z\, A_{D,C}. \tag{9}$$

Consider Z as an $st \times st$ matrix whose rows and columns are indexed by the elements of sets $N(c_1), N(c_2), \ldots, N(c_r)$ in that order. We can thus think of Z as an $s \times s$ block matrix of $t \times t$ matrix blocks $Z^{(k,\ell)}, 1 \le k, \ell \le s$. The next claim is a consequence of Equations (8) and (9).

Claim 16. *Each block $Z^{(k,\ell)}$ in Z is of the form*

$$Z^{(k,\ell)} = y_{k,\ell} W^{(k,\ell)}, \tag{10}$$

where $y_{k,\ell}$ is the $(k,\ell)^{th}$ entry of Y, and $W^{(k,\ell)}$ is a doubly stochastic matrix.

Proof. We first note from Equation (8) that the $(k,j)^{th}$ entry of the $s \times st$ matrix $YA_{C,D} = A_{C,D}Z$ can be computed in two different ways. In the left hand side matrix, it is $y_{k,\ell}$ for each $j \in N(c_\ell)$. On the other hand, the right hand side matrix implies that the same $(k,j)^{th}$ entry is also the sum of the j^{th} column of the $N(c_k) \times N(c_\ell)$ block $Z^{(k,\ell)}$ of the matrix Z.

We conclude, for $1 \le k, \ell \le s$, that each column in $Z^{(k,\ell)}$ adds up to $y_{k,\ell}$. By a similar argument, applied to Equation (9) this time, it follows, for each $1 \le k, \ell \le s$, that each *row* of any block $Z^{(k,\ell)}$ of Z adds up to $y_{k,\ell}$.

We conclude that, if $y_{k,\ell} \neq 0$, then the matrix $W^{(k,\ell)} = \frac{1}{y_{k,\ell}} Z^{(k,\ell)}$ is doubly stochastic. If $y_{k,\ell} = 0$, then (10) is true for any choice of $W^{(k,\ell)}$. ∎

For every $P = (p_{k\ell})$ appearing in an automorphism $P' \oplus P$ of H (see Equation (7)), we define the $st \times st$ doubly stochastic matrix W_P by its $t \times t$ blocks indexed by $1 \le k, \ell \le s$ as follows:

$$W_P^{(k,\ell)} = \begin{cases} W^{(k,\ell)} & \text{if } p_{k\ell} = 1, \\ 0 & \text{if } p_{k\ell} = 0. \end{cases} \tag{11}$$

Equations (7) and (10) imply that

$$X = X' \oplus Y \oplus Z = \sum_{P' \oplus P \in \mathrm{Aut}(H)} \alpha_{P',P}\, P' \oplus P \oplus W_P. \tag{12}$$

In order to see this, on the left hand side consider the $(k,\ell)^{th}$ block $Z^{(k,\ell)}$ of Z. On the right hand side, note that the corresponding block in each $P' \oplus P \oplus W_P$ is the matrix $W^{(k,\ell)}$. Clearly, the overall coefficient for this block equals the sum of $\alpha_{P',P}$ over all P' and P such that $p_{k,\ell} = 1$, which is precisely $y_{k,\ell}$ by Equation (7).

Since each $W^{(k,\ell)}$ is a doubly stochastic matrix, by Birkhoff's theorem we can write it as a convex combination of $t \times t$ permutation matrices $Q_{j,k,\ell}$, whose rows are indexed by elements of $N(c_k)$ and columns by elements of $N(c_\ell)$:

$$W^{(k,\ell)} = \sum_{j=1}^{t!} \beta_{j,k,\ell}\, Q_{j,k,\ell}.$$

Substituting the above expression in Equation (11), that defines the doubly stochastic matrix W_P, we express W_P as a convex combination of permutation matrices $W_P = \sum_Q \delta_{Q,P}\, Q$ where Q runs over all $st \times st$ permutation matrices indexed by the vertices in color class D. Notice here that $\delta_{Q,P}$ is nonzero only for those permutation matrices Q that have structure similar to that described in Equation (11): The block $Q^{(k,\ell)}$ is a null matrix if $p_{k\ell} = 0$ and it is some $t \times t$ permutation matrix if $p_{k\ell} = 1$. For each such Q, the $(s+st)\times(s+st)$ permutation matrix $P \oplus Q$ is an automorphism of the subgraph $G[C, D] = sK_{1,t}$ (because Q maps $N(c_i)$ to $N(c_j)$ whenever P maps c_i to c_j). Since $P \in \mathrm{Aut}(G[C])$ and D is a homogeneous set in G, we conclude that, moreover, $P \oplus Q$ is an automorphism of the subgraph $G[C \cup D]$.

Now, if we plug the above expression for each W_P in Equation (12), we will finally obtain the desired convex combination

$$X = \sum_{P',P,Q} \gamma_{P',P,Q}\, P' \oplus P \oplus Q.$$

It remains to argue that every $P' \oplus P \oplus Q$ occurring in this sum is an automorphism of G. Recall that a pair P', P can appear here only if $P' \oplus P \in \mathrm{Aut}(H)$. Moreover, if such a pair is extended to a matrix $P' \oplus P \oplus Q$, then $P \oplus Q \in \mathrm{Aut}(G[C \cup D])$. Since $G[B, D]$ is isotropic for every color class $B \neq D$ of G, we conclude that $P' \oplus P \oplus Q \in \mathrm{Aut}(G)$. This completes the induction step and finishes the case when G has one anisotropic component.

Next, we consider the case when $C(G)$ has several anisotropic components $T_1, \ldots, T_k$, $k \geq 2$. Let $G_1, \ldots, G_k$, where $G_i = G[\bigcup_{U \in V(T_i)} U]$, be the corresponding anisotropic components of G. By the proof of the previous case we already know that G_i is compact for each i.

Claim 17. *The automorphism group* $\mathrm{Aut}(G)$ *of* G *is the product of the automorphism groups* $\mathrm{Aut}(G_i)$*,* $1 \leq i \leq k$*.*

Proof. Recall that any automorphism of G must map each color class of G, which is a cell of the underlying amenable graph G', onto itself. Thus, any automorphism π of G is of the form $(\pi_1, \ldots, \pi_k)$, where π_i is an automorphism of the subgraph G_i. Now, for any two subgraphs G_i and G_j, we examine the edges between $V(G_i)$ and $V(G_j)$. For any color classes $U \subseteq V(G_i)$ and $U' \subseteq V(G_j)$, the edge $\{U, U'\}$ is isotropic because it is not contained in any anisotropic component of $C(G)$. Therefore, the bipartite graph $G[U, U']$ is either complete or empty. It follows that for any automorphisms π_i of G_i, $1 \le i \le k$, the permutation $\pi = (\pi_1, \ldots, \pi_k)$ is an automorphism of the graph G. ∎

As follows from Lemma 15, any fractional automorphism X of G is of the form $X = X_1 \oplus \cdots \oplus X_k$, where X_i is a fractional automorphism of G_i for each i. As each G_i is compact we can write each X_i as a convex combination $X_i = \sum_{\pi \in \mathrm{Aut}(G_i)} \alpha_{i,\pi} P_\pi$. This implies

$$I \oplus \cdots \oplus I \oplus X_i \oplus I \oplus \cdots \oplus I = \sum_{\pi \in \mathrm{Aut}(G_i)} \alpha_{i,\pi}\, I \oplus \cdots \oplus I \oplus P_\pi \oplus I \oplus \cdots \oplus I, \quad (13)$$

where block diagonal matrices in the above expression have X_i and P_π respectively in the i^{th} block (indexed by elements of $V(G_i)$) and identity matrices as the remaining blocks.

We now decompose the fractional automorphism X as a matrix product of fractional automorphisms of G

$$X = X_1 \oplus \cdots \oplus X_k = (X_1 \oplus I \oplus \cdots \oplus I) \cdot (I \oplus X_2 \oplus \cdots \oplus I) \cdot \cdots \cdot (I \oplus \cdots \oplus I \oplus X_k).$$

Substituting for $I \oplus \cdots \oplus I \oplus X_i \oplus I \oplus \cdots \oplus I$ from Equation (13) in the above expression and writing the product of sums as a sum of products, we see that X is a convex combination of permutation matrices of the form $P_{\pi_1} \oplus \cdots \oplus P_{\pi_k}$ where $\pi_i \in \mathrm{Aut}(G_i)$ for each i. By Claim 17, all the terms $P_{\pi_1} \oplus \cdots \oplus P_{\pi_k}$ correspond to automorphisms of G. Hence, G is compact, completing the proof of Theorem 13.

8 A color-refinement based hierarchy of graphs

Let $u \in V(G)$ and $v \in V(H)$ be vertices of two graphs G and H. By *individualization* of u and v we mean assigning the same *new color* to u and v, which makes them distinguished from the remaining vertices of G and H. Tinhofer [24] proved that, if G is compact, then the following polynomial-time algorithm correctly decides if G and H are isomorphic.

1. Run CR on G and H until the coloring of $V(G) \cup V(H)$ stabilizes.
2. If the multisets of colors in G and H are different, then output "non-isomorphic" and stop. Otherwise,
   (a) if all color classes are singletons in G and H, then if the map $u \mapsto v$ matching each vertex $u \in V(G)$ to the vertex $v \in V(H)$ of the same color is an isomorphism, output "isomorphic" and stop. Else output "non-isomorphic" and stop.
   (b) pick any color class with at least two vertices in both G and H, select an arbitrary $u \in V(G)$ and $v \in V(H)$ in this color class and individualize them. Goto Step 1.

If G and H are any two non-isomorphic graphs, then Tinhofer's algorithm will always output "non-isomorphic". However, it can fail for isomorphic input graphs, in general. We call G a *Tinhofer graph* if the algorithm works correctly on G and every H for all choices of vertices to be individualized. Thus, the result of [24] can be stated as the inclusion $\mathsf{Compact} \subseteq \mathsf{Tinhofer}$.

If G is a Tinhofer graph, then the above algorithm can be used to even find a canonical labeling of G. Using Theorem 13, we state the following fact.

Corollary 18. *Amenable and, more generally, compact graphs admit canonical labeling in polynomial time.*

Let A be a subgroup of the automorphism group $Aut(G)$ of a graph G. Then the partition of $V(G)$ into the A-orbits is called an *orbit partition* of G. Any orbit partition of G is equitable, but the converse is not true, in general. However, Godsil [11, Corollary 1.3] has shown that the converse holds for compact graphs. We define *Godsil graphs* as the graphs for which the two notions of an equitable and an orbit partition coincide. Thus, the result of [11] can be stated as the inclusion $\mathsf{Compact} \subseteq \mathsf{Godsil}$. Now, the inclusion $\mathsf{Compact} \subseteq \mathsf{Tinhofer}$ can easily be strengthened as follows.

Lemma 19. *Any Godsil graph is a Tinhofer graph.*

Proof. Assume that G is a Godsil graph. It suffices to show that Tinhofer's algorithm is correct whenever G and H are isomorphic. Let ϕ be an isomorphism from G to H. We will prove that, after the i-th refinement step made by the algorithm, there exists an isomorphism ϕ_i from G to H that preserves colors of the vertices. If this is true for each i, the algorithm terminates only if the discrete partition (i.e., the finest partition into singletons) is reached. Suppose that this happens in the k-th step. Then ϕ_k ensures that the algorithm decides isomorphism.

We prove the claim by induction on i. At the beginning, $\phi_1 = \phi$. Assume that an isomorphism ϕ_i exists and the partition is still not discrete. Suppose that now the algorithm individualizes $u \in V(G)$ and $v \in V(H)$. If $v = \phi_i(u)$, then $\phi_{i+1} = \phi_i$. Otherwise, consider the vertices u and $\phi_i^{-1}(v)$, which are in the same monochromatic class of G. Note that the partition of G produced in each refinement step is equitable. Since G is Godsil, there is an automorphism α preserving the partition such that $\alpha(u) = \phi_i^{-1}(v)$. We can, therefore, take $\phi_{i+1} = \phi_i \circ \alpha$. □

The orbit partition of G with respect to $Aut(G)$ is always a refinement of the stable partition $\mathcal{P}_G$ of G. We call G *refinable* if $\mathcal{P}_G$ is the orbit partition of $Aut(G)$. It is easy to show the following.

Lemma 20. *Any Tinhofer graph is refinable.*

Proof. Suppose that G is not refinable. Then G has vertices u and v that are in different orbits but not separated by the stable partition $\mathcal{P}_G$. This means that individualization of u and v in isomorphic copies G' and G'' of G gives non-isomorphic results. Therefore, if Tinhofer's algorithm is run on G' and G'' and individualizes u and v, it eventually decides that G' and G'' are non-isomorphic. □

Summarizing Theorem 13, Lemmas 19 and 20, and [11, Corollary 1.3], we state the following hierarchy result.

Theorem 21. *The classes of graphs under consideration form the inclusion chain*

$$\mathsf{Discrete} \subset \mathsf{Amenable} \subset \mathsf{Compact} \subset \mathsf{Godsil} \subset \mathsf{Tinhofer} \subset \mathsf{Refinable}. \tag{1}$$

Moreover, all of the inclusions are strict.

It is worth of noting that the hierarchy (1) collapses to Discrete if we restrict ourselves to only rigid graphs, i.e., graphs with trivial automorphism group.

The following separating examples prove that all inclusions are strict.

Separation of Discrete and Amenable: For $n \geq 2$, the complete graph K_n is amenable but not discrete.

Separation of Amenable and Compact: For $n \geq 6$, the cycles C_n are not amenable because they are indistinguishable from a pair $C_{n_1} + C_{n_2}$ of disjoint cycles on $n_1 + n_2 = n$ vertices. On the other hand, cycles are known to be compact graphs [22, Theorem 2].

Separation of Compact and Godsil: These classes are separated by the well-known Petersen graph. Evdokimov, Karpinski, and Ponomarenko [10, Corollary 5.4] prove that the Petersen graph is not compact. They explicitly give a fractional automorphism of the Petersen graph which cannot be written as a convex combination of its automorphisms. It remains to show that the Petersen graph belongs to the class Godsil.

This problem is solvable by modern computer algebra tools; see [29] where equitable and orbit partitions are counted for various strongly regular graphs, including the Petersen graph. We give a non-computer-assisted proof in Section A.1.

Separation of Godsil and Tinhofer: These classes are separated by the Johnson graphs $J(n, 2)$ for $n \geq 7$. The *Johnson graph* $J(n, k)$ has the k-element subsets of $[n] = \{1, \ldots, n\}$ as vertices; any two of them are adjacent if their intersection consists of $k-1$ elements. Note that $J(n, 1) = K_n$. Furthermore, the graph $J(n, 2)$ is the line graph of K_n: it has all 2-element subsets of $[n]$ as vertices and any two of them are adjacent if their intersection is non-empty. It is noticed in [9] that $J(n, 2)$ is not Godsil for $n \geq 7$. For establishing the separation, we show that $J(n, 2)$ is indeed Tinhofer. The proof is given in Section A.2.

Separation of Tinhofer and Refinable: Consider the gadget $\mathrm{CFI}(P_1, P_2, P_3)$ depicted in Figure 1, with two input pairs P_1 and P_2 and one output pair P_3. This gadget [7] has the property that any automorphism of it must flip an even number of the three pairs P_1, P_2, and P_3. We can combine $\mathrm{CFI}(P_1, P_2, P_3)$ with a second gadget $\mathrm{CFI}(P_1, P_2, P_4)$ with the same input pairs and a fresh output pair P_4. We assume that the four pairs P_1, P_2, P_3,

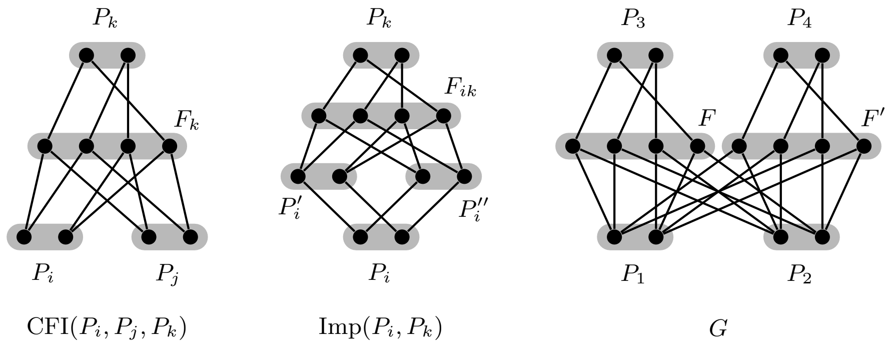

Fig. 1. The CFI(P_i, P_j, P_k)- and Imp(P_i, P_k)-gadgets and a graph G separating Refinable from Tinhofer

and P_4 and the intermediate sets of four connecting vertices, are all different color classes.
This defines a refinable graph G, also depicted in Figure 1, with four color classes P_1, P_2, P_3, and P_4 of size 2, and two color classes F and F' of size 4 corresponding to the orbit partition of G. The graph G has the property that any automorphism of it must flip either both pairs P_3 and P_3 or none of them. Now, if we run the Tinhofer procedure on two identical copies G' and G'' of G, it might individualize color class P_3 in the first round and color class P_4 in the second round in such a way that the resulting graphs are not isomorphic, since the partial isomorphism flips exactly one of the two color classes. Note that the vertex colors of G can be removed if we connect the four vertices in F by six edges and the two vertices in P_1 by one edge.

Finally, we show that testing membership in each of the graph classes in the hierarchy (1) is P-hard.

Theorem 22. *The recognition problem of each of the classes in the hierarchy (1) is* P*-hard under uniform AC^0 many-one reductions.*

Proof. We sketch the proof. Given a monotone boolean circuit C with and- and or-gates and constant input gates we construct a graph G as follows:
– For each gate g_k of C, G contains a pair $P_k = \{a_k, b_k\}$ of vertices.
– If g_k is a constant input gate with value 1, then a_k and b_k get different colors (i.e., they form singleton color classes); otherwise a_k and b_k both get the same color (i.e., they form a color class of size 2).
– For each and-gate g_k with input gates g_i and g_j, G additionally contains a color class F_k of size 4 that together with the two input pairs P_i and P_j as well as the output pair P_k forms a CFI(P_i, P_j, P_k)-gadget; see Figure 1.
– For each or-gate g_k with input gates g_i and g_j, G additionally contains two color classes F_{ik} and F_{jk} of size four, and four color classes P_i', P_i'', P_j', P_j'' of size 2. The color classes P_i', P_i'', P_k and F_{ik} form a CFI(P_i', P_i'', P_k)-gadget and each of the pairs P_i' and P_i'' is linked to P_i by two parallel edges. Henceforth, we denote this gadget by Imp(P_i, P_k); see Figure 1. Likewise, the color classes P_j', P_j'' and F_{jk} are used to form an Imp(P_j, P_k)-gadget.

A straightforward induction on the height of the and- and or-gates in C shows that CR on input G refines a color class P_k if and only if the corresponding gate g_k outputs value 1. This follows from the following observations.

- If g_k is an and-gate with input gates g_i and g_j, then the vertices in P_k get different C^{r+2} colors if and only if the vertices in P_i as well as the vertices in P_j have different C^r colors.
- If g_k is an or-gate with input gates g_i and g_j, then the vertices in P_k get different C^{r+3} colors if and only if either the vertices in P_i or the vertices in P_j have different C^r colors.

Now let G' be the graph that results from G by connecting the vertex pair P_l corresponding to the output gate g_l by two parallel edges with each pair P_k corresponding to a constant 0 input gate g_k. Then C evaluates to 1 if and only if G' is discrete (i.e., CR on input G' individualizes all vertices of G').

Moreover, if we connect the output pair P_l via an additional $\mathrm{Imp}(P_l, P_{l+1})$-gadget to a new vertex pair P_{l+1}, then the resulting graph G'' is not even refinable if C evaluates to 0. The reason is that no automorphism of G'' flips the pair P_{l+1}, but CR only refines the color class P_{l+1} if C evaluates to 1.

Hence, the mapping $C \mapsto G''$ simultaneously reduces MCVP to each of the graph classes in the hierarchy (1). □

We observe that the graph G'' used in the proof of the hardness results can be easily replaced by an uncolored graph. In fact, the vertex colors can be substituted by suitable graph gadgets in such a way that the automorphism group as well as the stable partition remain essentially unchanged (up to the addition of several singleton cells). Hence, the hardness results are also valid for the restricted versions of the classes in the hierarchy (1) where we consider only uncolored graphs.

References

1. V. Arvind, J. Köbler, G. Rattan, and O. Verbitsky. Graph isomorphism, color refinement, and compactness. *E-print*, `http:/arxiv.org/abs/1502.01255`, 4 February, 2015.
2. L. Babai, P. Erdös, and S. M. Selkow. Random graph isomorphism. *SIAM J. Comput.*, 9(3):628–635, 1980.
3. L. Babai and L. Kucera. Canonical labelling of graphs in linear average time. In *20th Annual Symposium on Foundations of Computer Science*, pages 39–46, Oct 1979.
4. C. Berkholz, P. Bonsma, and M. Grohe. Tight lower and upper bounds for the complexity of canonical colour refinement. In *Proceedings of 21st Annual European Symposium on Algorithms (ESA)*, volume 8125 of *Lecture Notes in Computer Science*, pages 145–156. Springer, 2013.
5. A. Borri, T. Calamoneri, and R. Petreschi. Recognition of unigraphs through superposition of graphs. *J. Graph Algorithms Appl.*, 15(3):323–343, 2011.
6. R. Busacker and T. Saaty. *Finite graphs and networks: an introduction with applications*. International Series in Pure and Applied Mathematics. McGraw-Hill Book Company, New York etc., 1965.
7. J.-Y. Cai, M. Fürer, and N. Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992.
8. A. Cardon and M. Crochemore. Partitioning a graph in $O(|A| \log_2 |V|)$. *Theor. Comput. Sci.*, 19:85–98, 1982.

9. A. Chan and C. D. Godsil. Symmetry and eigenvectors. In *Graph symmetry*, volume 497, pages 75–106. Kluwer Acad. Publ., Dordrecht, 1997.
10. S. Evdokimov, M. Karpinski, and I. N. Ponomarenko. Compact cellular algebras and permutation groups. *Discrete Mathematics*, 197-198:247–267, 1999.
11. C. Godsil. Compact graphs and equitable partitions. *Linear Algebra and its Applications*, 255(1–3):259 – 266, 1997.
12. M. Grohe. Equivalence in finite-variable logics is complete for polynomial time. *Combinatorica*, 19(4):507–532, 1999.
13. M. Grohe, K. Kersting, M. Mladenov, and E. Selman. Dimension reduction via colour refinement. In *Proceeding of 22th Annual European Symposium on Algorithms (ESA)*, volume 8737 of *Lecture Notes in Computer Science*, pages 505–516. Springer, 2014.
14. N. Immerman and E. Lander. Describing graphs: A first-order approach to graph canonization. In *Complexity Theory Retrospective*, pages 59–81. Springer, 1990.
15. R. Johnson. Simple separable graphs. *Pac. J. Math.*, 56:143–158, 1975.
16. K. Kersting, M. Mladenov, R. Garnett, and M. Grohe. Power iterated color refinement. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 1904–1910. AAAI Press, 2014.
17. S. Kiefer, P. Schweitzer, and E. Selman. Graphs identified by logics with counting. *E-print*, `http:/arxiv.org/abs/1503.08792`, 2015.
18. M. Koren. Pairs of sequences with a unique realization by bipartite graphs. *Journal of Combinatorial Theory, Series B*, 21(3):224 – 234, 1976.
19. A. Krebs and O. Verbitsky. Universal covers, color refinement, and two-variable logic with counting quantifiers: Lower bounds for the depth. *E-print*, `http:/arxiv.org/abs/1407.3175`, 2014.
20. M. V. Ramana, E. R. Scheinerman, and D. Ullman. Fractional isomorphism of graphs. *Discrete Mathematics*, 132(1-3):247–265, 1994.
21. N. Shervashidze, P. Schweitzer, E. J. van Leeuwen, K. Mehlhorn, and K. M. Borgwardt. Weisfeiler-Lehman graph kernels. *Journal of Machine Learning Research*, 12:2539–2561, 2011.
22. G. Tinhofer. Graph isomorphism and theorems of Birkhoff type. *Computing*, 36:285–300, 1986.
23. G. Tinhofer. Strong tree-cographs are Birkhoff graphs. *Discrete Applied Mathematics*, 22(3):275–288, 1989.
24. G. Tinhofer. A note on compact graphs. *Discrete Applied Mathematics*, 30(2-3):253–264, 1991.
25. G. Tinhofer and M. Klin. Algebraic combinatorics in mathematical chemistry. Methods and algorithms III. Graph invariants and stabilization methods, 1999.
26. R. Tyshkevich. Decomposition of graphical sequences and unigraphs. *Discrete Mathematics*, 220(1-3):201–238, 2000.
27. G. Valiente. *Algorithms on Trees and Graphs*. Springer, 2002.
28. H. Whitney. Congruent graphs and connectivity of graphs. *Amer. J. Math.*, 54:150–168, 1932.
29. M. Ziv-Av. Results of computer algebra calculations for triangle free strongly regular graphs. *E-print*, `http:/www.math.bgu.ac.il/~zivav/math/eqpart.pdf`, 2013.

A Missing parts of the proof of Theorem 21

A.1 The Petersen Graph is Godsil

It is well-known that the Petersen graph, denoted by P, is isomorphic to the Kneser graph $K(5,2)$. The *Kneser graph* $K(n,k)$ has the k-element subsets of $[n] = \{1,\ldots,n\}$ as vertices and any two of them are adjacent if they are disjoint. An important fact about $K(5,2)$ is that its automorphism group is isomorphic to the symmetric group S_5 acting on the set $\{1,\ldots,5\}$. In fact, any automorphism of $K(5,2)$ can be realized by extending the action of a permutation $\pi \in S_5$ to the vertex set of $K(5,2)$ [28].

First, we state some useful facts about the Petersen graph.

Proposition 23. *The Petersen graph has the following properties:*
(i) There are no cycles of length 3, 4 and 7.
(ii) There are no independent sets of size greater than 4.
(iii) Any two adjacent vertices have no common neighbors and any two non-adjacent vertices have a unique common neighbor.

We will need some definitions regarding partitions of the vertex set of a graph $G = (V,E)$. Given a partition $\Sigma = \{S_1,\ldots,S_k\}$ of V, we refer to the sets $S_1,\ldots,S_k$ as the *cells* of Σ. If the size of a cell is k, we call it a *k-cell*. Two cells S and S' are said to be *compatible* if the induced bipartite graph $P[S,S']$ is biregular (it can be empty). Otherwise, we say they are *incompatible*. Recall that any cell S of an equitable partition induces a regular graph $G[S]$. Moreover, in that case, any two cells S,S' are compatible and the number of edges in the biregular graph $G[S,S']$ is a common multiple of $|S|$ and $|S'|$.

Now we are ready to prove the following theorem.

Theorem 24. *The Petersen graph P is* Godsil.

Proof. To prove the theorem, we will enumerate all equitable partitions of P. For each such partition Σ, we describe a subgroup of $Aut(P)$ such that its orbit partition coincides with Σ. We represent the vertices of P by the two-element subsets of the set $\Omega = \{a,b,c,d,e\}$, where two vertices are adjacent if they are disjoint. This representation allows us to describe any subgroup of $Aut(P)$ as a subgroup of the permutation group S_Ω on Ω.

The two trivial partitions of $V(P)$ into one set and into ten singleton sets are clearly orbit partitions, since the Petersen graph is vertex-transitive. For our case analysis, we classify the remaining non-trivial equitable partitions of P by the minimum size δ of the cells in the partition. Clearly, $\delta \le 5$. In the following claims we show for each $k \in \{1,2,3,4,5\}$ that any equitable partition of P with $\delta = k$ is an orbit partition of P.

Claim 25. *P does not have any equitable partition with $\delta = 3$.*

Proof. Suppose that there is an equitable partition Σ with $\delta = 3$ and let S be a 3-cell in it. Then Σ either has the form $\Sigma = \{S,T\}$, where $|T| = 7$, or the form $\Sigma = \{S,U,V\}$ where $|U| = 3$ and $|V| = 4$. The first case is ruled out since

$P[T]$ can never be regular (P has neither independent sets of size 7 nor cycles of size 7). Suppose the second case is possible. Then $P[S]$ and $P[U]$ must be empty (since P has no triangles). Furthermore, the bipartite graphs $P[S,V]$ and $P[U,V]$ must be both biregular. The graph $P[S,V]$ (likewise, $P[U,V]$) is empty or it has 12 edges. It is not possible that $P[S,V]$ has 12 edges because then $P[V]$ has only 3 edges and cannot be regular. If both $P[S,V]$ and $P[U,V]$ are empty then V is disconnected from the rest of the graph, which is a contradiction. ∎

Claim 26. *All equitable partitions of P with $\delta = 4$ are orbit partitions.*

Proof. We first show that any equitable partition Σ with $\delta = 4$ has one 4-cell S and one 6-cell T, where $P[S]$ is empty and $P[T]$ is a 3-matching (a matching with 3 edges). Clearly, Σ must be of the form $\{S,T\}$, where $|S| = 4$ and $|T| = 6$. Moreover, $P[S]$ must be empty (0-regular) or 2-matching (1-regular) since it cannot be a 4-cycle (2-regular). In fact, the case of 2-matching can also be ruled out by counting the number of edges as follows. For S and T to be compatible, there must be 12 edges in the graph $P[S,T]$. Then there is exactly one edge left in the induced graph $P[T]$ which is impossible. Therefore, $P[S]$ must be empty. This also implies that the graph $P[S,T]$ has $4 \times 3 = 12$ edges. Hence, $P[T]$ must be a 3-matching.

Now observe that any independent-set S of size 4 in P must be of the kind $S = \{ab, ac, ad, ae\}$ (up to automorphisms), implying that $T = \{bc, bd, be, cd, ce, de\}$. The partition $\{S,T\}$ can be easily verified to be equitable and that it is the orbit partition of the subgroup $S_{\{b,c,d,e\}}$. ∎

Claim 27. *All equitable partitions of P with $\delta = 5$ are orbit partitions.*

Proof. In this case Σ must have the form $\Sigma = \{S,T\}$ where $|S| = |T| = 5$. Moreover, since P does not have independent sets of size 5, $P[S]$ and $P[T]$ must be 5-cycles. Clearly, such partitions exist, and any such partition has a matching between sets S and T.

It remains to show that $\Sigma = \{S,T\}$ is indeed an orbit partition of some subgroup of $Aut(P)$. Denote the 5-cycle in S by 1-2-3-4-5. Let $1'$ be the matching partner of 1 in T and so on. Now, $1'$ and $2'$ cannot be adjacent, else there is a 4-cycle in P. The unique common neighbor of $1'$ and $2'$ must be $4'$, otherwise it is easy to verify that we will have a 4-cycle in P. The partners $3'$ and $5'$ can also be uniquely determined in T. The permutation $\pi = (12345)(1'2'3'4'5')$ can be verified to be an automorphism of P and the orbits of the subgroup generated by π are precisely $\{S,T\}$. ∎

Claim 28. *All equitable partitions of P with $\delta = 2$ are orbit partitions.*

Proof. Let Σ be an equitable partition of P with $\delta = 2$ and let $S = \{u,v\}$ be a 2-cell in it. We first show that uv must be an edge. This holds because any two non-adjacent vertices have a unique common neighbor x. The cell containing x can only be a singleton set, which contradicts $\delta = 2$.

Next we show that the neighborhood $N(S) = \bigcup_{x \in S} N(x) \setminus S$ of S is also a cell of Σ (see Figure 2). Since uv is an edge, there are no common neighbors of u and v. Therefore, $|N(S)| = 4$. Moreover, $N(S)$ is an independent set since

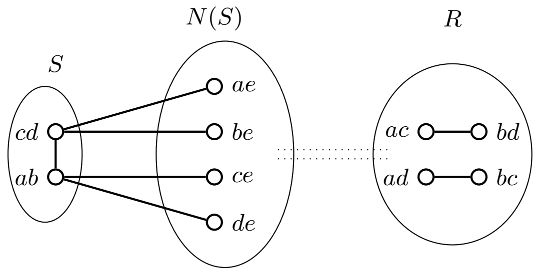

Fig. 2. The case $\delta = 2$.

any edge among vertices in $N(S)$ can be used to construct a cycle of length 3 or 4 passing through the edge uv. This is not possible by Proposition 23. Now let $R = V(P) \setminus (S \cup N(S))$ be the set of the four remaining vertices as shown in Figure 2. Observe that no cell can contain vertices from both $N(S)$ and R, since then it would be incompatible with S. Since $N(S)$ is an independent set, there cannot be a 2-cell inside $N(S)$. Clearly, there cannot be 1-cells and hence 3-cells inside $N(S)$. Therefore, $N(S)$ must indeed be a cell.

By accounting for edges of S and $N(S)$, it is easy to verify that R has exactly two edges, and hence $P[R]$ must be a 2-matching. Since $\delta = 2$, R does not contain any 1-cell and hence, any 3-cells. This leaves us with only two cases.

Case 1: R is a cell. We characterize all such partitions by naming a typical case. W.l.o.g, let $S = \{ab, cd\}$ since S is an edge. Then $N(S)$ must be $\{ae, be, ce, de\}$ and R must be $\{ac, ad, bc, bd\}$. The partition $\{ab, cd\}, \{ae, be, ce, de\}, \{ac, ad, bc, bd\}$ can be easily verified to be equitable. Moreover, it is easy to check that it is the orbit partition of the subgroup of all permutations in S_Ω which preserve the Ω-partition $\{ab\}, \{cd\}, \{e\}$. This is also the subgroup generated by the automorphisms $(ab), (cd), (ac)(bd)$.

Case 2: Σ partitions R in two sets A and B where $|A| = |B| = 2$. Since each 2-cell has to be an edge (see above), the sets A and B must be $\{ac, bd\}$ and $\{bc, ad\}$. The partition $\{ab, cd\}, \{ae, be, ce, de\}, \{ac, bd\}, \{ad, bc\}$ can be easily verified to be equitable. Moreover, it is easy to check that it is the orbit partition of the subgroup of all permutations in S_Ω which preserve the Ω-partition $\{ab\}, \{cd\}, \{e\}$ and additionally, stabilize the sets $\{ac, bd\}$ and $\{ad, bc\}$. This is also the subgroup generated by the automorphisms $(ac)(bd), (ad)(bc), (ab)(cd)$. ∎

Claim 29. *All equitable partitions of P with $\delta = 1$ are orbit partitions.*

Proof. Let S be a singleton set in such an equitable partition. Similar to a previous argument, a cell cannot have vertices from both $N(S)$ and $V \setminus N(S)$. Therefore, any equitable partition refines the partition $S, N(S), R$ (see Figure 3). Observe that $N(S)$ must be an independent set (otherwise there is a 3-cycle). Moreover, if we assume that $S = \{ab\}$, $N(S)$ must be $\{ce, de, cd\}$ and therefore, $R = \{ae, be, ac, bc, ad, bd\}$ forms a 6-cycle, as shown in the figure. We proceed by further classifying equitable partitions on the basis of the partition

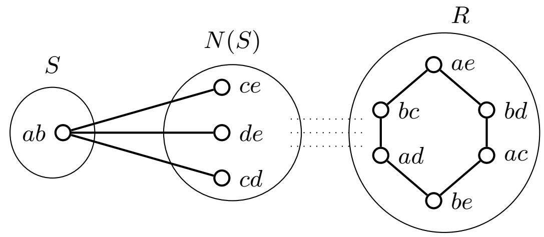

Fig. 3. The case $\delta = 1$.

induced by them inside $N(S)$. Since $|N(S)| = 3$, we have three possible cases. Either $N(S)$ is a cell, or it contains three 1-cells, or it contains one singleton and one 2-cell.

Case 1: $N(S)$ is a cell. We further classify the equitable partitions in this case on the basis of the partition induced on the set R. First, we examine the possible cells X in R which are compatible with $N(S)$. X cannot be of size 1 or 2, otherwise $P[N(S), X]$ has at most two edges. Also, X cannot be of size 4 or 5 since this would imply a cell of size 1 or 2 in R. Therefore, either R is a cell, or there are two 3-cells in R.

(a) R is a cell. The partition $\{ab\}$, $\{de, cd, ce\}$, $\{ac, ad, ae, bc, bd, be\}$ can be verified to be an equitable partition. Moreover, it is easy to check that it is the orbit partition of the subgroup $S_{\{c,d,e\}} \times S_{\{a,b\}}$.

(b) The partition induced on R is of the form $\{A, B\}$, where $|A| = |B| = 3$. Because of regularity, the only possible 3-cells in R are the independent sets $\{ad, ac, ae\}$ and $\{bc, bd, be\}$. The partition $\{ab\}$, $\{de, cd, ce\}$, $\{ad, ac, ae\}$, $\{bc, bd, be\}$ is clearly equitable. Moreover, it is easy to check that this partition is the orbit partition of the subgroup $S_{\{c,d,e\}}$.

Case 2: $N(S)$ contains three 1-cells. Again, we classify the equitable partitions on the basis of the partition induced on the set R. We can check that a cell of size more than two in R will have at least one edge to some singleton in $N(S)$, and will be incompatible with that singleton. Therefore, cells in R must have size at most 2. Moreover, any 2-cell must be of the form $\{ax, bx\}$ for some $x \in \{d, c, e\}$ since all other 2-cells can be seen to be incompatible with some singleton cell in $N(S)$. Finally, it can be seen that every possible 1-cell is incompatible with these three 2-cells. Hence, R must consist of three cells of size 2, namely $\{ad, bd\}$, $\{ac, bc\}$, $\{ae, be\}$. The partition $\{ab\}$, $\{cd\}$, $\{ce\}$, $\{de\}$, $\{ad, bd\}$, $\{ac, bc\}$, $\{ae, be\}$ can be easily seen to be equitable. Moreover, it is easy to check that it is the orbit partition of the subgroup $S_{\{a,b\}}$.

Case 3: $N(S)$ contains a 2-cell $U = \{ce, de\}$ and a 1-cell $V = \{cd\}$. Again, we need to classify the equitable partitions on the basis of the partition induced on the set R. First, we examine the possible cells X in R which are compatible with U and V. Clearly, X cannot be a 5-cell since $P[X]$ cannot be regular. It cannot be a 3-cell as well since the two candidate 3-cells are the independent sets $\{ad, ac, ae\}$ and $\{bc, bd, be\}$. Neither of them can be compatible with the singleton set V. Also, R cannot be a cell since it is incompatible with the singleton set V. Moreover, the only possible 4-cell is the neighborhood of the set U, i.e. $\{ac, bd, ad, bc\}$. Any other 4-cell is incompatible with U. Overall, we have no cells of size 3, 5, or 6 in R. Therefore, we have only the following four remaining subcases.

(a) R consists of one 4-cell and two 1-cells. This case is not possible since a 1-cell cannot be compatible with a 4-cell.

(b) R consists of one 4-cell and one 2-cell. The cells are $\{ac, bd, ad, bc\}$ and $\{ae, be\}$. The partition $\{ab\}$, $\{cd\}$, $\{ce, de\}$, $\{ae, be\}$, $\{ac, bd, ad, bc\}$ can be verified to be an equitable partition. Moreover, it is easy to check that it is the orbit partition of the subgroup $S_{\{a,b\}} \times S_{\{c,d\}}$

(c) R consists of three 2-cells. First, ae and be must be in the same 2-cell, otherwise the cell containing any of them would be incompatible with V. For the remaining vertices ac, ad, bc, bd, we can pair them up in three ways: (i) ac, ad and bc, bd, (ii) ac, bc and ad, bd, or (iii) ac, bd and ad, bc The first case is not possible since $\{ae, be\}$ and $\{ac, ad\}$ are not compatible. The second case is not possible because $\{ac, bc\}$ and $U = \{ce, de\}$ are not compatible. The third case gives an equitable partition $\{ab\}$, $\{cd\}$, $\{ce, de\}$, $\{ae, be\}$, $\{ac, bd\}$, $\{ad, bc\}$. Moreover, it is easy to check that it is the orbit partition of the subgroup generated by $(ab)(cd)$.

(d) R consists of a bunch of 1-cells and 2-cells. Clearly, the vertices ac, ad, bc, bd cannot form a singleton cell, since such a 1-cell will not be compatible with U. Therefore, $\{ae\}$ and $\{be\}$ are the only possible singleton cells. Neither of them can pair up with one of ac, ad, bc, bd since that cell would be incompatible with V. Therefore, they are forced to be singleton cells. It remains to partition ac, ad, bc, bd into two 2-cells. The vertex ac cannot be paired up with bd or bc since it will be incompatible with be. Therefore, the only possible case is to have 2-cells $\{ac, ad\}$ and $\{bc, bd\}$. The partition $\{ab\}$, $\{cd\}$, $\{ce, de\}$, $\{ae\}$, $\{be\}$, $\{ac, ad\}$, $\{bc, bd\}$ can be verified to be equitable. Moreover, it is easy to check that it is the orbit partition of the subgroup $S_{\{c,d\}}$. (This case is identical to Case 2(b)). ■ □

A.2 The Johnson Graphs $J(n, 2)$ are Tinhofer

In this section, we show that the Johnson graphs $J(n, 2)$ are Tinhofer. We begin with some necessary definitions. Let G be a graph and denote the automorphism group of G by A. For $v \in V(G)$, by A_v we denote the stabilizer subgroup of A that fixes the vertex v. Furthermore, for a subset $F \subset V(G)$, let $A_F = \bigcap_{v \in F} A_v$. Let $\mathcal{P}_F$ denote the stable partition of the colored version of G where each vertex in F is individualized. Then the orbit partition of A_F is a subpartition of $\mathcal{P}_F$. Note that G is Tinhofer if and only if, for every F, the orbit partition of A_F coincides with $\mathcal{P}_F$.

One way to prove that the two partitions coincide is to show that each orbit O of A_F is definable in terms of F in two-variable first-order logic. Here, "in terms of F" means that a defining formula $\Phi_O(x)$ can use constant symbols (names) for each vertex in F. Furthermore, $\Phi_O(x)$ contains occurrences of only two variables, x and y. At least one occurrence of x is free. $\Phi_O(x)$ uses two binary relation symbols $\sim$ and $=$ for adjacency and equality of vertices. This formula is true on G for $x = v$ exactly when $v \in O$.

Once $\Phi_O(x)$ is found for each O, the equality of the partitions follows by a similar argument as in [14, Theorem 1.8.1] or directly from the definitions of orbits, as those will imply that any two orbits are separated by color refinement starting from the individualization of F. The number of refinement steps sufficient to separate O from any other orbit can be only one greater than the quantifier depth of $\Phi_O(x)$.

In order to implement this scenario for $G = J(n, 2)$, it will be convenient to assume that $V(G) = \binom{[n]}{2}$ (note, however, that the formulas $\Phi_O(x)$ do not involve variables over $[n]$). Given $\alpha \in S_n$, by $\ell(\alpha)$ we denote the corresponding permutation of $\binom{[n]}{2}$. Obviously, every $\ell(\alpha)$ is an automorphism of G, and the automorphism group A contains nothing else by the Whitney theorem [28].

Before designing the definitions $\Phi_O(x)$, we will need to make two preliminary steps: Describe A_F and, then, describe the orbits of A_F (first irrespectively of any logical formalism; expressing these descriptions in two-variable first-order logic will be the next task).

We now proceed to the detailed proof.

Theorem 30. *$J(n, 2)$ is a Tinhofer graph for all n.*

Proof. Note that $J(2, 2) = K_1$, $J(3, 2) = K_3$, and $J(4, 2)$ is the octahedral graph, whose complement is $K(4, 2) = 3K_2$. Thus, these three graphs are amenable and, hence, Tinhofer. We can, therefore, assume that $n \geq 5$.

Call a fixed vertex $p \in F$ *isolated* if F contains no vertex adjacent to p. Let $F = F_1 \cup F_2$ be the partition of F into non-isolated and isolated vertices. Furthermore, we define the partition

$$[n] = W_1 \cup W_2 \cup W_3$$

as follows: W_1 is the union of all non-isolated pairs p (i.e., all p in F_1), and W_2 is the union of all isolated pairs p (i.e., all p in F_2). Thus, W_3 consists of the points of $[n]$ that are not included in any fixed pair.

Note now that $\ell(\alpha) \in A_F$ if and only if α either fixes or transposes the two points in each fixed pair. It follows that $\ell(\alpha) \in A_F$ exactly when

- $\alpha(w) = w$ for every $w \in W_1$ and
- $\alpha(p) = p$ for every $p \in F_2$.

Given a vertex $u = \{a, b\}$ of G, let $O(u)$ denote its orbit with respect to A_F. There are six kinds of orbits. Below we describe all of them along with providing suitable formal definitions $\Phi_{O(u)}(x)$.

Case 1: $\{a, b\} \subseteq W_1$. Then $O(u) = \{u\}$. *Formal definition:* $x = u$.

Case 2: $\{a, b\} \subseteq W_2$. Here we have two subcases. If $u \in F_2$, then $O(u) = \{u\}$ again. Otherwise, F_2 contains two pairs $p_1 = \{a, a'\}$ and $p_2 = \{b, b'\}$. In this subcase,

$$O(u) = \{\{a, b\}, \{a', b\}, \{a, b'\}, \{a', b'\}\},$$

which is exactly the common neighborhood of p_1 and p_2. *Formal definition:* $x \sim p_1 \wedge x \sim p_2$.

Case 3: $\{a, b\} \subseteq W_3$. Now $O(u) = \binom{W_3}{2}$, which are exactly the non-fixed vertices with no neighbor in F. *Formal definition:* $\bigwedge_{p \in F}(x \neq p \wedge x \not\sim p)$.

Case 4: $a \in W_1$, $b \in W_2$. Let $p = \{b, b'\}$ be the pair in F_2 containing b. Then,

$$O(u) = \{\{a, b\}, \{a, b'\}\}.$$

To give a formal definition of $O(u)$, we consider two subcases.

(i) a belongs to two adjacent vertices $q_1 = \{a, a_1\}$ and $q_2 = \{a, a_2\}$ in F_1. *Formal definition:* $x \sim p \wedge x \sim q_1 \wedge x \sim q_2$. Indeed, the condition $x \sim p$ forces x to contain either b or b'. This excludes the possibility that $x = \{a_1, a_2\}$ and, therefore, x is forced to contain a by the adjacency to q_1 and q_2.

(ii) a belongs to a single vertex $q_1 = \{a, a'\}$ in F_1. By definition, F_1 contains also a vertex $q_2 = \{a', a''\}$. *Formal definition:* $x \sim p \wedge x \sim q_1 \wedge x \not\sim q_2$.

Case 5: $a \in W_1$, $b \in W_3$. Then

$$O(u) = \{\,\{a, b'\} : b' \in W_3\}\,.$$

Similarly to the preceding case, we distinguish two subcases.

(i) a belongs to two adjacent vertices $q_1 = \{a, a_1\}$ and $q_2 = \{a, a_2\}$ in F_1. *Formal definition:* First of all, we say that $x \sim q_1 \wedge x \sim q_2$. It remains to exclude the possibility that $x \subseteq W_1 \cup W_2$ (in particular, this will exclude $x = \{a_1, a_2\}$ and force x to contain a). We do this by adding the following expression

$$\bigwedge_{p \in F} x \neq p \wedge \bigwedge_{p,q \in F, p \not\sim q} \neg(x \sim p \wedge x \sim q) \\ \wedge \bigwedge_{p,q \in F_1, p \sim q} (x \sim p \wedge x \sim q \to \exists y\, (y \sim x \wedge y \sim p \wedge y \sim q)). \quad (14)$$

The first conjunctive term prevents x to be one of the pairs in F. The second term excludes the case that x is covered by two disjoint pairs p and q in F. The third term excludes the case that x is covered by two intersecting pairs p and q in F or, equivalently, the case where x, p, and q form a triangle. It would be not enough just to forbid x, p, and q from forming a clique because this could also exclude a permissible case where x, p, and q form a star (which is captured by the subformula beginning with $\exists y$). Note, that we need the assumption $n \geq 5$ in this place.

(ii) a belongs to a single vertex $q_1 = \{a, a'\}$ in F_1, and $q_2 = \{a', a''\}$ is another vertex in F_1. *Formal definition:* $x \sim q_1 \wedge x \not\sim q_2 \wedge x \not\subseteq W_1 \cup W_2$, the last being expressed by the formula (14).

Case 6: $a \in W_2$, $b \in W_3$. In this case, F_2 contains a pair $p = \{a, a'\}$ and

$$O(u) = \{\,\{a, b'\} : b' \in W_3\} \cup \{\,\{a', b'\} : b' \in W_3\}\,.$$

Formal definition: $x \sim p \wedge x \not\subseteq W_1 \cup W_2$, the latter being expressed by (14).

The proof is complete. □

ECCC ISSN 1433-8092

http://eccc.hpi-web.de