Graph Isomorphism, Color Refinement, and Compactness

V. Arvind, Johannes Köbler, Gaurav Rattan, Oleg Verbitsky

Abstract

Color refinement is a classical technique to show that two given graphs G and H are non-isomorphic; it is very efficient, even if incomplete in general. We call a graph G *amenable* to color refinement if this algorithm succeeds in distinguishing G from any non-isomorphic graph H. Babai, Erdős, and Selkow (1982) proved that almost all graphs G are amenable. We here determine the exact range of applicability of color refinement by showing that the class of all amenable graphs is 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.

Furthermore, we prove that amenable graphs are compact in the sense of Tinhofer (1991), that is, their polytopes of fractional automorphisms are integral. The concept of compactness was introduced in order to identify the class of graphs G for which isomorphism $G \cong H$ can be decided by computing an extreme point of the polytope of fractional isomorphisms from G to H and checking if this point is integral. Our result implies that this linear programming approach to isomorphism testing has the applicability range at least as large as the combinatorial approach based on color refinement.

1 Introduction

The well-known *color refinement* (or *naive vertex classification*) algorithm begins with a uniform coloring of the vertices of two graphs G and H and refines it step by step so that, if two vertices have equal colors but differently colored neighborhoods (with the multiplicities of colors counted), then these vertices get new different colors in the next refinement step. The algorithm terminates as soon as no further refinement is possible and concludes that G and H are non-isomorphic if the multisets of colors occurring in these graphs are different. If this happens, the conclusion is correct. However, color refinement cannot sometimes 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. We say that color refinement *applies* to a graph G if it succeeds in distinguishing G from any non-isomorphic H. The most obvious class of graphs to which color refinement is applicable is formed by *unigraphs*. Those are the graphs which are determined up to isomorphism by their degree sequences; see, e.g., [4, 24]. Another class where color refinement works successfully consists of trees (Edmonds [5, 25]). Babai, Erdös, and Selkow [2] proved that the method is actually very powerful as it applies to almost all graphs. A key observation here is that color refinement secceeds on G and H whenever the vertices of G obtain pairwise different colors, and it is proved in [2] that almost all G have this property.

We aim at determining the exact range of applicability of color refinement. Call a graph to which color refinement applies *amenable*. 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. Note that an a priori upper bound for the complexity of this decision problem is $\mathsf{coNP}^{\mathsf{GI}}$, where the superscript means the one-query access to an oracle solving the graph isomorphism problem.

Immerman and Lander [12] established a close connection between color refinement and 2-variable first-order logic with counting quantifiers, which implies that a graph is amenable iff it is definable in this logic. Thus, the class of graphs definable by sentences with 2 variables and counting quantifiers proves to be recognizable in polynomial time. This result generalizes to structures over any binary relational vocabulary (see Corollary 2.12).

A standard approach to solving a computationally hard problem consists in considering an appropriate linear programming relaxation. One such approach to isomorphism testing, based on the concept of fractional isomorphism (see Section 3.1), is known to be equivalent to color refinement. In our terminology, a graph G is amenable iff the isomorphism relation $G \cong H$ is equivalent to the weaker, efficiently verifiable relation of fractional isomorphism $G \cong_f H$. The tight connection between color refinement and fractional isomorphism was discovered by Tinhofer [20] and studied by Ramana, Scheinerman and Ullman [17] (see also Godsil [9]).

Another linear programming approach to isomorphism testing was sugested by Tinhofer in [22]. He calls a graph *compact* if the polytope of all its fractional automorphisms is integral, that is, all extreme points of this polytope have integer coordinates. If G is compact, it has the following remarkable property: If $G \cong H$, then the polytope of fractional isomorphism from G to H is integral while if $G \not\cong H$, then this polytope has no integer extreme point (in particular, it can be empty). It follows that the relation $G \cong H$ can be decided by computing just one extreme point of the polytope and checking if it is integral.

We prove that all amenable graphs are compact. This means that Tinhofer's approach has at least as large applicability range as color refinement. It remains an intriguing open problem to find an efficient characterization for the whole class of compact graphs (or to show that its decision problem is hard).

Finally, we consider the class of graphs for which color refinement succeeds in solving the *automorphism* problem. Specifically, we call a graph G *refinable* if the color partition produced by color refinement coincides with the orbit partition of G. The fact that all trees are refinable was observed independently by several authors; see a survey in [23]. Our results imply that all amenable graphs are refinable. In Section 4 we discuss structural and algorithmic properties of graphs that were introduced by Tinhofer [22] and Godsil [9]. We note that these properties, along with compactness, define a hierarchy of graph classes between the amenable and the refinable graphs.

Related work. Color refinement turns out to be a useful tool not only in isomorphism testing but also in a number of other areas; see [10, 14, 19] and references there. The concept of compactness is generalized to *weak compactness* in [7, 8]. The linear programming approach of [20, 17] to isomorphism testing is extended in [1, 11], where it is shown that this extension corresponds to the k-dimensions Weisfeiler-Lehman algorithm (which is just color refinement if $k = 1$).

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 complement of a graph G is denoted by $\overline{G}$. 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' iff 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 Amenability

2.1 Basic definitions and facts

Given a graph G, the *color refinement* algorithm (to be abbreviated as *CR*) recursively computes a sequence of colorings C^i of $V(G)$. The initial coloring C^0 is uniform. Then,

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

where $\{\!\!\{ \ldots \}\!\!\}$ denotes a multiset. Note that $C^1(u) = C^1(v)$ iff the two vertices have the same degree.

A simple inductive argument shows that $C^{i+1}(u) = C^{i+1}(v)$ implies $C^i(u) = C^i(v)$. Therefore, C^{i+1} is a refinement of C^i, where C^i is regarded as a partition of $V(G)$ (consisting of the monochromatic classes of vertices). It follows that, eventually, $C^{s+1} = C^s$ for some s; hense, $C^i = C^s$ for all $i \geq s$. The partition C^s is called *stable*.

A partition $\mathcal{P}$ of $V(G)$ is called *equitable* if, for any elements $X \subseteq V(G)$ and $Y \subseteq V(G)$ of $\mathcal{P}$, any two vertices in X have the same number of neighbors in Y. A trivial example of an equitable partition is the partition of $V(G)$ into singletons, which we call *discrete*. There is a unique equitable partition $\mathcal{P}_G$ that is the *coarsest* in the sense that any other equitable partition $\mathcal{P}$ of G is a subpartition of $\mathcal{P}_G$. It is easy to see that the stable partition of G is equitable, and an inductive argument shows that it is actually the coarsest [6, Lemma 1].

A straightforward inductive argument shows that the coloring (1) is preserved under isomorphisms.

Lemma 2.1. *If ϕ is an isomorphism from G to H, then $C^i(u) = C^i(\phi(u))$ for any vertex u of G.*

Lemma 2.1 readily implies that, if graphs G and H are isomorphic, then

$$\{\!\{ C^i(u) : u \in V(G) \}\!\} = \{\!\{ C^i(v) : v \in V(H) \}\!\} \tag{2}$$

for all $i \geq 0$. When used for isomorphism testing, the CR algorithm accepts two input graphs G and H as isomorphic exactly when the last condition is met. Note that this condition is actually finitary: If Equality (2) 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 C^{2n-1} must be a stable partition of the disjoint union of G and H. In fact, (2) holds true for all i iff it is true for $i = n$; see [16, Lemma 3.5]. Thus, it is enough that CR verifies Equality (2) for $i = n$.

Note that computing the vertex colors literally according (1) would lead to 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).

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

2.2 Local structure of amenable graphs

Given an equitable partition $\mathcal{P}$ of a graph G, we call its elements *cells*. The definition of an equitable partition implies that any graph $G[X]$ induced by a cell X is *regular*, that is, all vertices in $G[X]$ have equal degrees. Furthermore, the bipartite graph $G[X, Y]$ induced by two distinct cells X and Y is *biregular*, that is, all vertices in each vertex class have equal degrees: The vertices in X have equally many neighbors in Y and vice versa. The following lemma gives a list of all possible regular and biregular graphs that can occur in the stable (i.e., the coarsest equitable) partition of an amenable graph.

Lemma 2.2. *Suppose that G is an amenable graph. Let $\mathcal{P}_G$ be the stable partition of G.*

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

- **(B)** *If $X, Y \in \mathcal{P}_G$, then $G[X, Y]$ is an empty graph, a complete bipartite graph, the 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.*

For the proof we need the following facts.

Lemma 2.3 (Johnson [13]). *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 $n = 5$.*[1]

Lemma 2.4 (Koren [15]). *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 $d \in \{0, 1, m-1, m\}$.*

The graphs G and H in the following lemma have the same vertex set. Given a vertex u, we distunguish its neighborhoods $N_G(u)$ and $N_H(u)$ and it colors $C^i_G(u)$ and $C^i_H(u)$ in the two graphs.

Lemma 2.5. *Suppose that X and Y are cells of the stable partition of a graph G.*

1. *Let H be obtained from G by changing the edges between the vertices in X so that $G[X]$ is replaced with a regular graph of the same degree. Then $C^i_G(u) = C^i_H(u)$ for any $u \in V(G)$ and any i.*
2. *Let H be obtained from G by changing the edges between X and Y so that $G[X, Y]$ is replaced with a biregular graph having the same vertex degrees. Then $C^i_G(u) = C^i_H(u)$ for any $u \in V(G)$ and any i.*

Proof. 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 $a \in V(G)$. Consider an arbitrary vertex u and prove that

$$C^{i+1}_G(u) = C^{i+1}_H(u). \tag{3}$$

From now on we treat Parts 1 and 2 separately.

1. 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 (1), we immediately derive (3) from the induction assumption.

The case of $u \in X$ is a bit more complicated. Now we have only equality $N_G(u) \setminus X = N_H(u) \setminus X$, which implies

$$\{\!\{\, C^i_G(a) : a \in N_G(u) \setminus X \,\}\!\} = \{\!\{\, C^i_H(a) : a \in N_H(u) \setminus X \,\}\!\}. \tag{4}$$

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

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

[1] The last case, in which the graph is the 5-cycle, is missing from the statement of this result in [13, Theorem 2.12]. The proof in [13] tacitly considers only graphs with at least 6 vertices.

Combining (4) and (5), we conclude that (3) holds in any case.

2. If $u \notin X \cup Y$, we have $N_G(u) = N_H(u)$ and Equality (3) readily follows from the induction assumption.

Suppose that $u \in Y$. In this case we still have (4) and, exactly as in Part 1, we also derive (5). Equality (3) follows.

The case of $u \in X$ is symmetric. ∎

Proof of Lemma 2.2. **A.** If $G[X]$ is a graph not from the list, by Lemma 2.3, it is not a unigraph. We, therefore, can modify G locally on X by replacing $G[X]$ with a non-isomorphic regular graph with the same parameters. Lemma 2.5.1 implies that the resulting graph H satisfies Equality (2) for any i, that is, CR does not distinguishes between G and H. The graphs G and H are non-isomorphic because, by Lemmas 2.5.1 and 2.1, 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 part follows, similarly to Part A, from Lemmas 2.4 and 2.5.2. ∎

2.3 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* $C(G)$ to be the complete graph on the vertex set $\mathcal{P}_G$ with the following labeling of vertices and edges. A vertex $X \in \mathcal{P}_G$ is either *complete*, *empty*, *matching*, *co-matching*, or *pentagonal* depending on the type of $G[X]$; see Condition A in Lemma 2.2. An edge $\{X, Y\}$ is either *complete*, *empty*, *constellation*, or *co-constellation*, depending on the type of $G[X, Y]$; see Condition B.

Each vertex and edge has an exactly one label. This requires removing an ambiguity in several cases. First, every single-element cell $X = \{u\}$ is labeled as *empty* (rather than *complete*). Note that the complete graph K_2 is as well a matching graph. This is resolved by labeling each two-element cell $X = \{u, v\}$ as *complete* or *empty* (rather than *matching* or *co-matching*) depending on whether or not u and v are adjacent in G. Thus, a matching or co-matching X always consists of at least 4 vertices. Furthermore, the edges $\{X, Y\}$ of the cell graph such that $G[X, Y] \cong K_{1,t}$ are labeled as *complete* (rather than *constellation*). If $G[X, Y]$ is the bipartite complement of $K_{1,t}$, then $\{X, Y\}$ is labeled as *empty* (rather than *co-constellation*). Another source of ambiguity is that $2K_{1,t}$ is isomorphic to its bipartite complement. The edges $\{X, Y\}$ corresponding to $G[X, Y] \cong 2K_{1,t}$ are labeled as *constellation* (rather than *co-constellation*).

A vertex of the cell graph is called *homogeneous* if it is labeled as *complete* or *empty* and *heterogeneous* in any of the other three cases. An edge of the cell graph is called *isotropic* if it is labeled as *complete* or *empty* and *anisotropic* if it is labeled as *constellation* or *co-constellation*.

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) will be called *uniform*. Note that if an edge $\{X_i, X_{i+1}\}$ of a uniform path/cycle is *constellation* (resp. *co-constellation*), then $G[X_i, X_{i+1}]$ is a matching (resp. co-matching) graph.

Lemma 2.6. *If G is amenable, then*

(C) *the cell graph $C(G)$ contains no uniform anisotropic path connecting two heterogeneous vertices;*

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

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

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

Proof. **C.** Suppose that P is a uniform anisotropic path in $C(G)$ connecting heterogeneous vertices 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 $G[Y]$.

Note that an isomorphism ϕ from G to another graph H preserves the mate relation, which is definable on pairs in $\phi(X) \times \phi(Y)$ in the same vein. In other terms, $\phi(v^*) = \phi(v)^*$ for any $v \in X$. It follows that, for any $u, v \in X$, $\phi(u)$ and $\phi(v)$ are adjacent exactly when $\phi(u)^*$ and $\phi(v)^*$ are adjacent, which means that ϕ preserves also the conducting property. More precisely, ϕ induces an isomorphism ϕ' from $C(G)$ to $C(H)$, which takes conducting paths in $C(G)$ to conducting paths in $C(H)$ and non-conducting ones to non-conducting ones.

If P is conducting, we can replace the subgraph $G[Y]$ with an isomorphic but different subgraph so that P becomes non-conducting in the cell graph $C(H)$ of the resulting graph H. Vice versa, if P is non-conducting, we can make such a replacement converting P to a conducting path.

By Lemma 2.5.1, CR does not distinguish between G and H. As another consequence of Lemma 2.5.1, $C(G) = C(H)$. Lemmas 2.5.1 and 2.1 imply that if there is an isomorphism ϕ from G to H, the induced isomorphism ϕ' from $C(G)$ to $C(H)$ is the identity map on $\mathcal{P}_G$, the vertex set of $C(G)$. Therefore, ϕ' takes P onto itself, which contradicts preservation of the conducting property. We conclude that $G \not\cong H$ and, hence, G is not amenable.

D. Suppose that $C(G)$ contains a uniform anisotropic cycle Q of length m. All vertices of Q have the same cardinality as cells; 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 vertices 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. We do replacement that changes $\tau(Q)$.

Similarly to Part C, we use Lemma 2.5.2 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 $XY_1 \ldots Y_l Z$ such that $|X| < |Y_1| = \ldots = |Y_l| > |Z|$ (for the case of a cycle, where $Z = X$, the argument is virtually the same). Like in the proof of Part C, the uniform anisotropic path $Y_1 \ldots Y_l$ determines a one-to-one correspondence between the sets Y_1 and Y_l. We make identification $Y_1 = Y_l = Y$ according to this correspondence.

Let $G[X, Y] = sK_{1,t}$ and $G[Z, Y] = aK_{1,b}$, 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 $|Y| = st = ab$. 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

$$|Y_x| = t, \quad |Y_z| = b, \quad Y_x \cap Y_{x'} = \emptyset, \quad \text{and} \quad Y_z \cap Y_{z'} = \emptyset \tag{6}$$

for any $x \neq x'$ in X and $z \neq z'$ in Z (the disjointness conditions can be replaced with the covering conditions $\bigcup_{x \in X} Y_x = \bigcup_{z \in Z} Y_z = Y$). We regard $\mathcal{Y}_G = \{Y_x\}_{x \in X} \cup \{Y_z\}_{z \in Z}$ as a hypergraph on the vertex set Y.

For an isomorphism ϕ from G to another graph H, let $\mathcal{Y}_H$ be determined similarly by the path $\phi(X)\phi(Y_1)\ldots\phi(Y_l)\phi(Z)$ (note that $\mathcal{Y}_H$ does not depend on ϕ by Lemma 2.1). Then obviously $\mathcal{Y}_H \cong \mathcal{Y}_G$.

Now, let $\mathcal{H} = \{Y_x\}_{x \in X} \cup \{Y_z\}_{z \in Z}$ be an arbitrary hypergraph on the vertex set Y satisfying the conditions (6). Given $\mathcal{H}$, we can replace the subgraph $G[X, Y_1]$ with an isomorphic but different bipartite graph so that $\mathcal{Y}_H \cong \mathcal{H}$ for the resulting graph H. When we take $\mathcal{H} \ncong \mathcal{Y}_G$, this will ensure that $H \ncong G$. Similarly to Parts C and D, Lemma 2.5.2 along with Lemma 2.1 implies that CR does not distinguish between G and H. Therefore, G cannot be amenable.

It remains to show that an appropriate choice of a hypergraph $\mathcal{H}$ is always available, that is, there are at least two non-isomorphic hypergraphs $\mathcal{H}_1$ and $\mathcal{H}_2$ on the vertex set Y with hyperedges denoted by Y_x, $x \in X$, and Y_z, $z \in Z$, satisfying the conditions (6) (if $t = b$, multiple hyperedges $Y_x = Y_z$ are allowed). Without loss of generality, suppose that $t \leq b$. In order to construct such $\mathcal{H}_1$ and $\mathcal{H}_2$, identify Y with the segment of integers $\{1, 2, \ldots, st\}$. Set $\{Y_x\}$ to be the partition of Y into s blocks of consecutive integers of length t in both $\mathcal{H}_1$ and $\mathcal{H}_2$. In $\mathcal{H}_1$, we set $\{Y_z\}$ to be the partition of Y into a blocks of consecutive integers of length b. In $\mathcal{H}_2$, we set $\{Y_z\}_{z \in Z} = \{\{j, j + a, j + 2a, \ldots, j + (b - 1)a\}\}_{j=1}^{a}$. The two hypergraphs are non-isomorphic because in $\mathcal{H}_1$ we have $Y_x \subset Y_z$ for some x and z, while no hyperedge includes another in $\mathcal{H}_2$.

F. Suppose that $C(G)$ contains an anisotropic path $XY_1 \ldots Y_l$ where $|X| < |Y_1| = \ldots = |Y_l|$ and Y_l is heterogeneous. The uniform anisotropic path $Y_1 \ldots Y_l$ determines a one-to-one correspondence between Y_1 and Y_l, and we make identification $Y_1 = Y_l = Y$ accordingly to it. Consider an auxiliary graph A_G on the vertex set $X \cup Y$ where $A_G[X]$ is empty, $A_G[Y] = G[Y_l]$, and $A_G[X, Y] = G[X, Y_1]$.

For an isomorphism ϕ from G to another graph H, let A_H be determined similarly by the path $\phi(X)\phi(Y_1) \ldots \phi(Y_l)$. Then obviously $A_H \cong A_G$.

Like in Part C, we can replace $G[Y_l]$ (hence $A_G[Y]$) with an isomorphic but different graph so that $A_H \ncong A_G$ for the resulting graph H. This will imply that G and H are non-isomorphic while indistinguishable by CR and, therefore, that G is not amenable. All what we have to show is that at least two different isomorphism types of A_H can be obtained by such a replacement.

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. Consider a hypergraph $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ where $\mathcal{H}_1 = \{Y_x\}_{x \in X}$ has hyperedges $Y_x = N(x) \cap Y_1$ as in the proof of Part E, while $\mathcal{H}_2$ consists now of 2-element hyperedges corresponding to the edges of $G[Y_l]$. Similarly to Part E, we can change the isomorphism type of $\mathcal{H}$ by modifying the subgraph $G[X, Y_1]$. This results in different isomorphism types of A_H. ∎

It turns out that Conditions A–F are not only necessary for amenability but also sufficient.

Theorem 2.7. *A graph G is amenable if and only if it satisfies Conditions A–F.*

The necessity part of the theorem is given by Lemmas 2.2 and 2.6. The proof of the sufficiency consists of Lemmas 2.8 and 2.9 below, that reveal 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 all edges are anisotropic. Note that of a vertex of $C(G)$ has no incident anisotropic edges, it forms a single-vertex anisotropic component.

Lemma 2.8. *Suppose that a graph G satisfies Conditions A–F. For an anisotropic component A of $C(G)$, the following is true.*

1. *A is a tree with the following monotonicity property. Let R be a vertex of A of the minimum cardinality (as a cell). Let A_R the rooted directed tree obtained from A by rooting it at R. Then $|X| \leq |Y|$ for any directed edge (X, Y) of A_R.*

2. *A contains at most one heterogeneous vertex. If R is such a vertex, it has the minimum cardinality among the vertices of A.*

Proof. **1.** A cannot contain any uniform cycle by Condition D and any other cycle by Condition E. The monotonicity property follows from Condition E.

2. Assume that A contains more than one heterogeneous vertex. Consider two such vertices 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 Part 1 implies that there is j (possibly $j = 1, l$) such that $|Z_1| \geq \ldots \geq |Z_j| \leq \ldots \leq |Z_l|$. Since the path cannot be uniform by Condition C, at least one of the inequalities is strict. However, this contradicts Condition F.

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

Lemma 2.9. *Suppose that a graph G satisfies Conditions A and B, which allows us to consider the cell graph $C(G)$. Assume that every anisotropic component A of $C(G)$*

(i) *is a tree and*

(ii) *has at most one heterogeneous vertex.*

Then G is amenable.

Proof. Given a graph H indistinguishable from G by CR, we have to show that G and H are isomorphic.

Since G and H satify the condition (2) for any i, any coloring C^i stable on the disjoint union of G and H determines a one-to-one correspondence f between the cells of the stable partitions of G and H. As follows directly from (2), $|X| = |f(X)|$ for every cell X of G. The map f is an isomorphism from $C(G)$ to $C(H)$ and, moreover, for any cells X and Y of G

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

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

To show (a), consider a coloring $C = C^i$ stable on the disjoint union of G and H. Since C is stable on both G and H and X and $f(X)$ are cells of the corresponding partitions, both $G[X]$ and $H[f(X)]$ are regular. By Condition A, $G[X]$ is a unigraph. Since G and H have equal number of vertices, their non-isomorphism would mean that they have different degrees. This is impossible because then $X \cup f(X)$, a cell of C, would split in the next refinement step. The relation (b) follows from Condition B by a similar argument.

We now construct an isomorphism ϕ from G to H. By Lemma 2.1, 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 assumption, an anisotropic component A of the cell graph $C(G)$ contains at most one heterogeneous vertex. Denote it by R_A if it exists. Otherwise fix R_A to be an arbitrary vertex 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 assumption, 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. Then ϕ is extended to Y so that this is an isomorphism from $G[X, Y]$ to $H[f(X), f(Y)]$.

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. ∎

Lemmas 2.8 and 2.9 immediately imply the sufficiency part of Theorem 2.7. The proof of this theorem is therewith complete.

2.4 Examples and applications

Lemma 2.9 is a convenient tool for checking the amenability. For example, let us call a graph *discrete* if its coarsest equitable partition is discrete or, in other words, all vertices are colored differently in the stable coloring. The amenability of discrete graphs is a well-known fact. Note that Conditions A and B as well as Conditions (i) and (ii) in Lemma 2.9 are fulfilled for discrete graphs by trivial reasons.

Checking these four condition, we can also reprove the amenability of trees. Moreover, we can extend this result to the class of forests.

Corollary 2.10. *All forests are amenable graphs.*

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 forests of stars.

Let $C^*(G)$ be the version of the cell graph $C(G)$ where all empty edges are removed. Note that, 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 (which is Condition (i)).

To prove Condition (ii), suppose that an anisotropic component of $C(G)$ contains a path $X_0, X_1, \ldots, X_l$ connecting two heterogeneous vertices X_0 and X_l. Consider the subgraph $G[X_0] \cup G[X_0, X_1] \cup \ldots \cup G[X_{l-1}, X_l] \cup G[X_l]$. Since this graph has no vertex of degree 0 or 1, it must contain a cycle, a contradiction. ∎

The characterization of amenable graphs we obtained 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 [6]; see also [3]) 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 2.11. *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 $\mathcal{P}_G$ of the input graph G. Sorting the adjacency list of each vertex according to $\mathcal{P}_G$, we compute a list of entries $d_{X,Y}$ of the *degree refinement matrix* $(d_{X,Y})$, $X, Y \in \mathcal{P}_G$, where $d_{X,Y}$ is equal to the number of neighbors in Y of any vertex in X. Along with the numbers $|X|$ and $|Y|$, $d_{X,Y}$ allows us to determine whether or not each subgraph $G[X,Y]$ is one of the graphs listed in Condition B of Lemma 2.2. Similarly, $|X|$ and $d_{X,X}$ allows us to determine whether or not each subgraph $G[X]$ is one of the graphs listed in Condition A of this lemma. If Conditions A and B are fulfilled, therewith we also obtain the cell graph $C(G)$.

Using breadth-first search, we find all anisotropic components of $C(G)$ and, simulateneously, for each of them we check Conditions (i) and (ii) in Lemma 2.9. As follows from Theorem 2.7 along with Lemmas 2.8 and 2.9, any graph satisfying Conditions A and B is amenable if and only if it satisfies also Conditions (i) and (ii). ■

In the conclusion of this section, we consider 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 [12] discovered an intimate connection between color refinement and 2-variable first-order logic with counting quantifiers. This connection impies that amenability of a graph is equivalent to its definability in this logic. Thus, Corollary 2.11 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. Standard reductions lead to the following extension of this fact.

Corollary 2.12. *Let σ be a vocabulary consisting of binary relation symbols. Then the class of structures over σ definable in 2-variable first-order logic with counting quantifiers is recognizable in polynomial time.*

Finally, note that CR admits a natural extension to structures over any binary vocabulary σ (which is most obvious for vertex-colored graphs), and the Immerman-Lander results is preserved under this extension. Thus, Corollary 2.12 implies that there is an efficient way to check whether or not the generalized CR applies to a given input structure.

3 Compactness

3.1 Basic definitions and facts

Convex sets and extreme points. For $x, y \in \mathbb{R}^n$, a *convex combination* of x and y is any vector of the form $\alpha x + (1 - \alpha)y$ where $0 \le \alpha \le 1$. More generally, $\sum_{i=1}^{k} \alpha_i x_i$ is a convex combination of k points $x_1, \ldots, x_k \in \mathbb{R}^n$ if $\sum_{i=1}^{k} \alpha_i = 1$ and $\alpha_i \ge 0$ for all i. A set $S \subseteq \mathbb{R}^n$ is *convex* if, for every two points $x, y \in S$, S contains also any convex combination of these points, that is, the segment with endpoints x and y. The *convex hull* of a set $S \subseteq \mathbb{R}^n$, denoted by $\langle S \rangle$, is the inclusion-minimal convex set containing S. Equivalently, $\langle S \rangle$ is the set of all convex combinations of any finite number of points in S. A point $z \in S$ is called an *extreme point* of S if it cannot be represented as a convex combination of other points of S, that is, $z = \alpha x + (1 - \alpha)y$ with $0 \le \alpha \le 1$ implies $z = x = y$. We will denote the set of all extreme points of S by $\mathrm{Extr}\, S$. The Minkowski theorem says that, if a convex set S is bounded and closed, then $S = \langle \mathrm{Extr}\, S \rangle$.

Polytopes. Speaking of polytopes, we always mean *convex polytopes*. Such a polytope P can be defined as the intersection of a set of half-spaces of $\mathbb{R}^n$ or, algebraically, $P = \{\, x \in \mathbb{R}^n : Ax \le b \}$ where A is a real $m \times n$ matrix, x and b are supposed to be n-dimensional column vectors, and the inequality is understood row-wise. A matrix or a vector is *integral* if all its entries are integers. Given integral A and b, deciding if P is non-empty is exactly the linear programming problem, that can be solved in polynomial time, for example, by the famous ellipsoid method. A point $x \in P$ is called a *basic feasible solution* to the system $Ax \le b$ if $\mathrm{rank}\, A_x = n$, where the matrix A_x is obtained from A by removing those rows where the inequality is strict. It is known that $\mathrm{Extr}\, P$ consists exactly of basic feasible solutions to the underlying system of inequalities. This implies that $\mathrm{Extr}\, P$ is a finite set. If A and b are integral, this also implies that the bit length of any extreme point is bounded by a polynomial in the bit length of A and b. Though the set $\mathrm{Extr}\, P$ can be exponentially large in n, computing a single point in $\mathrm{Extr}\, P$ reduces in polynomial time to the linear programming problem.

A polytope is called *integral* if all its extreme points are integral.

Doubly stochastic matrices and the Birkhoff polytope. Let π be a permutation of the set $\{1, \ldots, n\}$. The corresponding *permutation matrix* $P_\pi = (p_{ij})$ is defined by $p_{ij} = 1$ if $\pi(i) = j$ and $p_{ij} = 0$ otherwise. An $n \times n$ real matrix $X = (x_{ij})$ is *doubly stochastic* if $\sum_{i=1}^{n} x_{ij} = 1$ for every j, $\sum_{j=1}^{n} x_{ij} = 1$ for every i, and $x_{ij} \ge 0$ for all i, j. The product and any convex combination of doubly stochastic matrices are themselves doubly stochastic matrices. Considered as a subset of $\mathbb{R}^{n^2}$, the set of all $n \times n$ doubly stochastic matrices is known as the Birkhoff polytope B_n. Note that every permutation matrix is an extreme point of B_n. The Birkhoff theorem says that B_n has no other extreme points. Equivalently, every doubly stochastic matrix is a convex combination of permutation matrices.

Fractional isomorphism. Without loss of generality, in this section we consider graphs with vertex set $\{1, \ldots, n\}$. Let $\mathrm{Iso}(G, H)$ consist of the permutation matrices P_π for all isomorphisms π from G to H. Furthermore, $\mathrm{Aut}(G) = \mathrm{Iso}(G, G)$.

Let A and B denote the adjacency matrices of graphs G and H respectively. Note that G and H are isomorphic iff there is a permutation matrix X such that

$$AX = XB. \tag{7}$$

In fact, (7) is true for a permutation matrix X exactly when $X \in \mathrm{Iso}(G, H)$.

G and H are called *fractionally isomorphic* if (7) is satisfied by a doubly stochastic matrix X. Such an X will be called a *fractional isomorphism* from G to H. Note that if X is a fractional isomorphism from G to H, then the transpose X^T is a fractional isomorphism from H to G. Furthermore, if X is a fractional isomorphism from G to H and Y is a fractional isomorphism from H to F, then the product XY is a fractional isomorphism from G to F. Thus, being fractionally isomorphic is an equivalence relation.

Analysing the relation of fractional isomorphism to color refinement, Ramana, Scheinerman, and Ullman [17] show a tight connection between fractional isomorphisms and equitable partitions. Let $V_1, \ldots, V_m$ be a partition of the set $\{1, \ldots, n\}$. For $i \le m$, let X_i be a matrix whose rows and columns are indexed by the elements of V_i. Then $X_1 \oplus \cdots \oplus X_m$ will denote the block-diagonal matrix with blocks $X_1, \ldots, X_m$. We will use the following fact, that can be considered an analog of Lemma 2.1 for fractional isomorphisms. For the reader's convenience we include a proof in Appendix A.

Lemma 3.1 (Ramana et al. [17]). *Let G be a graph on vertex set $\{1, \ldots, n\}$ and assume that the elements of the coarsest equitable partition $\mathcal{P}_G$ of G are intervals of consecutive integers. Let X be a fractional automorphism G, i.e., a doubly stochastic matrix commuting with the adjacency matrix of G. Then $X = X_1 \oplus \cdots \oplus X_m$, where the blocks $X_1, \ldots, X_m$ correspond to $\mathcal{P}_G$.*

Note that the assumtion of the lemma can be ensured for any graph by appropriately relabeling its vertices.

Compact graphs. Denote the set of all fractional isomorphisms from G to H by $S(G, H)$. Then $S(G, H)$ is a polytope in the Euclidean space $\mathbb{R}^{n^2}$. Any integral matrix X in $S(G, H)$ is actually a permutation matrix and, therefore, $X \in \mathrm{Iso}(G, H)$. Note also that any permutation matrix in $S(G, H)$ is an extreme point of this polytope. Thus, $\mathrm{Iso}(G, H) \subseteq \mathrm{Extr}\, S(G, H)$. In fact, $\mathrm{Iso}(G, H)$ coincides with the set of all integral extreme points of $S(G, H)$.

The set $S(G) = S(G, G)$ will be referred to as the polytope of *fractional automorphisms* of G. According to the discussion above, the integral extreme points of $S(G)$ are exactly the matrices in $\mathrm{Aut}(G)$, i.e., the permutation matrices corresponding to automorphisms of G. A graph G is *compact* [20] if $S(G)$ has no other extreme points, i.e., $\mathrm{Extr}\, S(G) = \mathrm{Aut}(G)$. Compactness of a graph 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 standard automorphisms of G, i.e., $S(G) = \langle \mathrm{Aut}(G) \rangle$.

Example 3.2.

1. Complete graphs are compact as a consequence of the Birkhoff theorem.
2. Trees and cycles are compact (Tinhofer [20]).
3. Matching graphs mK_2 are compact. This is a particular instance of a much more general result by Tinhofer [22]: If G is compact, then mG is compact for any m.
4. Tinhofer [22] observes that the class of compact graphs is closed under complementation. Indeed, if A is the adjacency matrix of G, then $\overline{G}$ has adjacency matrix $J - I - A$, where I denotes the identity matrix and J denotes the all-ones matrix. This readily implies that $S(\overline{G}) = S(G)$. In particular, empty and co-matching graphs are compact.
5. $C_3 + C_4$ is an example of a non-compact graph. Tinhofer [22] makes a much more general observation: A regular compact graph must be vertex-transitive. This easily follows from the fact that $\frac{1}{n}J \in S(G)$ for any regular graph G on n vertices.

The concept of a compact graph is motivated by a linear programming approach to Graph Isomorphism that is based on the following fact.

Proposition 3.3 (Tinhofer [22]). *If G is a compact graph, then the polytope $S(G, H)$, for any graph H, is either empty or integral.*

If $S(G, H) = \emptyset$, then $G \not\cong H$ (as these graphs are even not fractionally isomorphic). If $S(G, H)$ is integral, then $G \cong H$ (as any integral point in $S(G, H)$ corresponds to an isomorphism from G to H). Therefore, if we know that a graph G is compact, we can decide whether or not another graph H is isomorphic to G just by checking if the polytope $S(G, H)$ is not empty and, if it is not, by computing a vertex of $S(G, H)$ and checking if it is integral. All this can be done in polynomial time by linear programming.

There is no efficient characterization for the class of compact graphs. As noted by Tinhofer in [22], the only complexity upper bound for compactness is coNP, because testing if every vertex of a given polytope is integral is in coNP. Nevertheless, our main result on compact graphs, Theorem 3.4 below, shows a wide, efficiently recognizable range of applicability of Tinhofer's approach.

3.2 Statement of the result

Theorem 3.4. *Amenable graphs are compact.*

Theorem 3.4 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 [9, Corollary 1.6]. Indeed, while Babai, Erdös, and Selkow [2] proved that almost all graphs are discrete (and, moreover, the discrete partition is reachable in 2 refinement rounds), we already mentioned in Section 2.4 that all discrete graphs are amenable.

Furthermore, Theorem 3.4 reproves Tinhofer's result on compactness of trees.[2] Using Corollary 2.10, we can extend this result to forests. This extension is not straightforward as compact graphs are not closed under disjoint unions; see Example 3.2.5. In [21], 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 3.4, seems to be earlier never observed. Summarizing, we note the following result.

Corollary 3.5. *Discrete graphs, forests, and unigraphs are compact.*

3.3 Proof of Theorem 3.4

Given an amenable graph G and its fractional automorphism X, 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 2.7 and Lemma 2.8. The last lemma refers to anisotropic components of the cell graph $C(G)$. Given an anisotropic component A of $C(G)$, 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 facts that, according to Lemma 2.8, A is a tree with at most one heterogeneous vertex.

This scenario cannot be implemented directly by a simple reason: In order to run induction, we need that a subgraph of an amenable graph induced by a number of cells is also amenable, which is not always the case. In order to remove this complication, we introduce a definition generalizing amenable graphs for the purpose of applying induction.

Definition 3.6. Let G be a vertex-colored graph with the partition $V(G) = V_1 \cup \cdots \cup V_m$ of its vertex set into color classes. We say that G is *pseudo-amenable* if there is an amenable graph G' such that

1. $V(G) \subset V(G')$ and $V_1, \ldots, V_m$ are cells of G';
2. G is an induced subgraph of G' obtained by deleting the remaining cells of G'.

[2] The proof of Theorem 3.4 uses only compactness of complete graphs, matching graphs, and the 5-cycle.

The next definition is needed in order to extend the notion of compactness to pseudo-amenable graphs. From now on, without loss of generality we suppose that the vertices $1, \dots, n$ of a vertex-colored graph G are named so that every color class is an interval of consecutive integers.

Definition 3.7. For a vertex-colored graph G, *color-preserving automorphisms* are automorphisms that map each color class to itself. A *color-preserving fractional automorphism* of G is a fractional automorphism

$$X = X_1 \oplus \cdots \oplus X_m$$

such that the blocks $X_1, \dots, X_m$ of the block-diagonal doubly stochastic matrix X correspond to the color classes $V_1, \dots, V_m$ of G. More precisely, the rows and columns of X_i are indexed by the vertices in the set V_i, for each i.

Claim 1. *Every pseudo-amenable graph G is compact in the sense that every color-preserving fractional automorphism of G is a convex combination of color-preserving automorphisms of G.*

This claim implies the theorem because we can consider every amenable graph G as pseudo-amenable with the coarsest equitable partition $\mathcal{P}_G$ defining the color classes. By Lemma 3.1, all fractional automorphisms of G (which includes all automorphisms of G; see also Lemma 2.1) will be color-preserving.

In the sequel we prove the claim, essentially by induction on the number of color classes. For a pseudo-amenable graph G we define its cell graph $C(G)$ on the set of color classes exactly as $C(G)$ is defined on the coarsest equitable partitions of amenable graphs. Now, it makes sense to talk of anisotropic components of $C(G)$ and, hence, of anisotropic components of the pseudo-amenable graph G.

We first consider the case when G consists of a single anisotropic component. By Lemma 2.8, 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 $V_1, \dots, V_m$ of G so that V_1 is the unique heterogeneous cell if such exists; otherwise V_1 is chosen among the cells of minimum cardinality. Moreover, we can suppose that, for each $i \le m$, the cells $V_1, \dots, V_i$ 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 Lemma 2.2. All of them are known to be compact; see Example 3.2.1–4. 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 compactness of $G = G[V_1 \cup \cdots \cup V_m]$.

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 anisotripic link to the remaning part H of G. Let $C \in \{V_1, \dots, V_{m-1}\}$ be the unique cell such that $\{C, D\}$ is an anisotropic edge of $C(G)$. To be specific, suppose that $G[C, D] = sK_{1,t}$. If $G[C, D]$ is a co-constallation, 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 color-preserving fractional isomorphisms of the graph (cf. Example 3.2.4). By the monotonicity property in Lemma 2.8.1, $|C| = s$ 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 color-preserving 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

$$X' \oplus Y = \sum_{P' \oplus P \in \mathrm{Aut}(H)} \alpha_{P',P} \, P' \oplus P, \tag{8}$$

where $P' \oplus P$ are permutation matrices corresponding to automorphisms of the graph H, such that the permutation matrix block P denotes automorphism's action on 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 $A_{C,D}$ denote the $s \times st$ submatrix corresponding to the block row-indexed by C and column-indexed by D. Likewise, $A_{D,C}$ denotes the $st \times s$ submatrix with rows indexed by D and columns by C. Since X is a fractional automorphism of G, we have

$$XA = AX. \tag{9}$$

Recall that Y and Z are blocks of X corresponding to color classes C and D. Looking at the corner fragments of the matrices in the left and the right hand sides of (9), 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

$$\begin{aligned} Y A_{C,D} &= A_{C,D} \, Z, & (10) \\ A_{D,C} \, Y &= Z \, A_{D,C}. & (11) \end{aligned}$$

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 (10) and (11).

Claim 2. *Each block $Z^{(k,\ell)}$ in Z is of the form*

$$Z^{(k,\ell)} = y_{k,\ell} W^{(k,\ell)}, \tag{12}$$

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 (10) 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 (11) 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 (12) 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 (8)), 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:

$$W_P^{(k,\ell)} = \begin{cases} W^{(k,\ell)} & \text{if } p_{k\ell} = 1, \\ 0 & \text{if } p_{k\ell} = 0. \end{cases} \tag{13}$$

Equations (8) and (12) imply that

$$X = X' \oplus Y \oplus Z = \sum_{P' \oplus P \in \mathrm{Aut}(H)} \alpha_{P',P}\, P' \oplus P \oplus W_P. \tag{14}$$

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 (8).

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 (13), 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 (13): 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 (14), 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 3. *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, $1 \leq i \leq k$, the permutation $\pi = (\pi_1, \ldots, \pi_k)$ is a color-preserving automorphism of the graph G. ∎

As follows from Lemma 3.1, 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

$$I \oplus \cdots \oplus I \oplus X_i \oplus I \oplus \cdots \oplus I = \sum_{\pi \in \mathrm{Aut}(G_i)} \alpha_{i,\pi} \, I \oplus \cdots \oplus I \oplus P_\pi \oplus I \oplus \cdots \oplus I, \quad (15)$$

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

$$\begin{aligned} X &= X_1 \oplus \cdots \oplus X_k \\ &= (X_1 \oplus I \oplus \cdots \oplus I) \cdot (I \oplus X_2 \oplus \cdots \oplus I) \cdot \cdots \cdot (I \oplus \cdots \oplus I \oplus X_k). \end{aligned}$$

Substituting for $I \oplus \cdots \oplus I \oplus X_i \oplus I \oplus \cdots \oplus I$ from Equation (15) 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 3, all the terms $P_{\pi_1} \oplus \cdots \oplus P_{\pi_k}$ correspond to automorphisms of G. Therefore, G is compact.

The proof of Claim 1 and, hence, of Theorem 3.4 is complete.

4 A color-refinement hierarchy between the isomorphism and the automorphism problems

Let $u \in V(G)$ and $v \in V(H)$. *Individualization* of u and v assigns the same new color to each of these vertices, making them distinguished from the rest of G and H. Tinhofer [22] proved that, if G is compact, then whether or not G and H are isomorphic can be recognized by the following algorithm.

1. Run Color Refinement 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, decide non-isomorphism. Otherwise,
   (a) if all color classes are singletons within G and H, check if all pairs of equally colored vertices in G and H are in the same (non)adjacency relation in both graphs;
   (b) if there is a color class with at least two vertices in both G and H, select arbitrary $u \in V(G)$ and $v \in V(H)$ in this class, individualize them, and repeat Step 1.

Note that, in general, Tinhofer's algorithm is always correct on non-isomorphic G and H but can be wrong on isomorphic inputs. We call G a *Tinhofer graph* if the algorithm works correctly on G and every H for every choice of vertices to be individualized.

Note that, if G is a Tinhofer graph, then the algorithm above can be used to find a canonical labeling of G. In particular, this applies to all compact graphs, and Theorem 3.4 gives us also the following fact.

Corollary 4.1. *The class of amenable graphs admits a polynomial-time canonical labeling algorithm.*

Let $A \subseteq Aut(G)$ be a group of automorphisms of a graph G. The partition of G into the orbits of A is called an *orbit partition*. Note that any orbit partition of G is equitable. The converse is in general not true. However, Godsil [9, Corollary 1.3] proved that any equitable partition of a compact graph is an orbit partition. We call graphs with this property *Godsil graphs*. The aforementioned Tinhofer's result can be refined as follows.

Lemma 4.2. *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, $\phi_i = \phi$. 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 $\phi_{i+1} = \phi_i \circ \alpha$. ∎

Note that the orbit partition of a graph G (with respect to the entire automorphism group $Aut(G)$) is a refinement of the coarsest equitable partition of G. We call G *refinable* if the two partition coincide. Any Godsil graph is refinable by definition.

Lemma 4.3. *Any Tinhofer graph is refinable.*

Proof. Suppose that G is not refinable. Then G has vertices u and v that are in the same element of the coarsest equitable partition but in different orbits. The latter 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 non-isomorphism. ∎

Summarizing Theorem 3.4, Lemmas 4.2 and 4.3, and [9, Corollary 1.3], we obtain the following chain of inclusions:

$$\textsf{Amenable} \subseteq \textsf{Compact} \subseteq \textsf{Godsil} \subseteq \textsf{Tinhofer} \subseteq \textsf{Refinable}. \tag{16}$$

Note that, while Amenable is the class of graphs for which color refinement is powerful enough to solve the isomorphism problem, Refinable consists of the graphs on which color refinement succeeds in solving the *automorphism problem*.

Note that discrete graphs are asymmetric. Godsil [9, Corollary 1.6] shows that the asymmetric compact graphs are exactly the discrete graphs. In fact, the entire hierarchy (16) collapses to the class of Discrete graphs when it is restricted to asymmetric graphs. Indeed, the asymmetric amenable graphs include Discrete and the class of asymmetric refinable graphs obviously coincides with Discrete.

References

[1] A. Atserias and E. N. Maneva. Sherali-adams relaxations and indistinguishability in counting logics. *SIAM J. Comput.*, 42(1):112–137, 2013.

[2] L. Babai, P. Erdös, and S. M. Selkow. Random graph isomorphism. *SIAM J. Comput.*, 9(3):628–635, 1980.

[3] C. Berkholz, P. Bonsma, and M. Grohe. Tight lower and upper bounds for the complexity of canonical colour refinement. In H. L. Bodlaender and G. F. Italiano, editors, *Algorithms - ESA 2013 - 21st Annual European Symposium, Sophia Antipolis, France, September 2-4, 2013. Proceedings*, volume 8125 of *Lecture Notes in Computer Science*, pages 145–156. Springer, 2013.

[4] A. Borri, T. Calamoneri, and R. Petreschi. Recognition of unigraphs through superposition of graphs. *J. Graph Algorithms Appl.*, 15(3):323–343, 2011.

[5] 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.

[6] A. Cardon and M. Crochemore. Partitioning a graph in $O(|A| \log_2 |V|)$. *Theor. Comput. Sci.*, 19:85–98, 1982.

[7] S. Evdokimov, M. Karpinski, and I. N. Ponomarenko. Compact cellular algebras and permutation groups. *Discrete Mathematics*, 197-198:247–267, 1999.

[8] S. Evdokimov, I. N. Ponomarenko, and G. Tinhofer. Forestal algebras and algebraic forests (on a new class of weakly compact graphs). *Discrete Mathematics*, 225(1-3):149–172, 2000.

[9] C. Godsil. Compact graphs and equitable partitions. *Linear Algebra and its Applications*, 255(1–3):259 – 266, 1997.

[10] M. Grohe, K. Kersting, M. Mladenov, and E. Selman. Dimension reduction via colour refinement. In A. S. Schulz and D. Wagner, editors, *Algorithms - ESA 2014 - 22th Annual European Symposium, Wroclaw, Poland, September 8-10, 2014. Proceedings*, volume 8737 of *Lecture Notes in Computer Science*, pages 505–516. Springer, 2014.

[11] M. Grohe and M. Otto. Pebble games and linear equations. In P. Cégielski and A. Durand, editors, *Computer Science Logic (CSL'12) - 26th International Workshop/21st Annual Conference of the EACSL*, volume 16 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 289–304, Dagstuhl, Germany, 2012. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[12] N. Immerman and E. Lander. Describing graphs: A first-order approach to graph canonization. In A. Selman, editor, *Complexity Theory Retrospective*, pages 59–81. Springer, 1990.

[13] R. Johnson. Simple separable graphs. *Pac. J. Math.*, 56:143–158, 1975.

[14] K. Kersting, M. Mladenov, R. Garnett, and M. Grohe. Power iterated color refinement. In C. E. Brodley and P. Stone, editors, *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pages 1904–1910. AAAI Press, 2014.

[15] M. Koren. Pairs of sequences with a unique realization by bipartite graphs. *Journal of Combinatorial Theory, Series B*, 21(3):224 – 234, 1976.

[16] A. Krebs and O. Verbitsky. Universal covers, color refinement, and two-variable logic with counting quantifiers: Lower bounds for the depth. *E-print, arxiv.org/abs/1407.3175*, 2014.

[17] M. V. Ramana, E. R. Scheinerman, and D. Ullman. Fractional isomorphism of graphs. *Discrete Mathematics*, 132(1-3):247–265, 1994.

[18] E. R. Scheinerman and D. H. Ullman. *Fractional graph theory. A rational approach to the theory of graphs.* Wiley: John Wiley & Sons, 1997.

[19] 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.

[20] G. Tinhofer. Graph isomorphism and theorems of birkhoff type. *Computing*, 36:285–300, 1986.

[21] G. Tinhofer. Strong tree-cographs are birkhoff graphs. *Discrete Applied Mathematics*, 22(3):275–288, 1989.

[22] G. Tinhofer. A note on compact graphs. *Discrete Applied Mathematics*, 30(2-3):253–264, 1991.

[23] G. Tinhofer and M. Klin. Algebraic combinatorics in mathematical chemistry. Methods and algorithms III. Graph invariants and stabilization methods, 1999.

[24] R. Tyshkevich. Decomposition of graphical sequences and unigraphs. *Discrete Mathematics*, 220(1-3):201–238, 2000.

[25] G. Valiente. *Algorithms on Trees and Graphs.* Springer, 2002.

A Proof of Lemma 3.1

Various proofs of this fact are available in [9, 10, 17]. Our exposition is close to [18].

Let $X = (x_{ij})$ be any fractional automorphism of G. We define a directed graph D_X with vertex set $V(G) = \{1, \ldots, n\}$ and edge set

$$E = \{(i, j) : X_{ij} \neq 0\}.$$

Let $V = S_1 \cup S_2 \cup \cdots \cup S_t$ be the partition of V into the strongly connected components S_i of the digraph D_X. We first show that there can be no directed edges between the strongly connected components. We can assume, without loss of generality, that the components are topologically sorted: if (u, v) is a directed edge in D_X with $u \in S_i$ and $v \in S_j$ then $i \leq j$. Hence, for $S = S_i \cup S_{i+1} \cup \ldots S_t$ there are no directed edges (u, v) with $u \in S$ and $v \notin S$. It implies that for each $i \in S$ the row sum $\sum_{j \in S} X_{ij} = 1$. Hence all the row sums of the submatrix indexed by S on both rows and columns are 1. Since X is doubly stochastic this forces each column sum of this submatrix is also 1. More precisely, $\sum_{i \in S} X_{ij} = 1$ for $j \in S$. Hence there are *no* edges between S and $\overline{S}$ for each such S. Consequently, there are no edges between the strongly connected components S_i of D_X.

Therefore, relabeling $V(G)$ so that $S_1, S_2, \ldots, S_t$ become intervals of consecutive integers, we can bring X to a block diagonal form

$$X = X_1 \oplus X_2 \oplus \cdots \oplus X_t \tag{17}$$

where each X_i is a doubly stochastic matrix. Note that the underlying directed graph D_{X_i} induced by S_i is strongly connected, which means that each matrix X_i is *irreducible.*

Now, we claim that the strongly connected components S_i form an equitable partition of the graph G. Since X is a fractional automorphism we have $AX = XA$ where A is the adjacency matrix of the relabeled graph G. Then for any pair of components S_i and S_j we have $A_{ij}X_j = X_iA_{ij}$, where A_{ij} denotes the $(i, j)^{th}$ block indexed by S_i on the rows and S_j on columns. Let u denote the all 1's vector of dimension $|S_j|$. Multiplying by u to the right of both sides we obtain

$$A_{ij}u = A_{ij}X_ju = X_iA_{ij}u,$$

since $X_ju = u$. Hence $A_{ij}u$ is an eigenvector of X_j for eigenvalue 1. However, the matrix X_j is is nonnegative and irreducible, hence by the Perron-Frobenius theorem the maximum eigenvalue (which is 1 for a stochastic matrix) has a 1 dimensional eigenspace. Since u is an eigenvector, it follows that $A_{ij}u$ is its scalar multiple. This means that $A_{ij}u = d_{ij}u$, where d_{ij} is the degree of every vertex in S_i in the bipartite graph $G[S_i, S_j]$ if $i \neq j$ and in the subgraph $G[S_i]$ if $i = j$. We conclude that each $G[S_i, S_j]$ is biregular and each $G[S_i]$ is regular.

Thus, the S_i's do yield an equitable partition. Since this partition is a refinement of the coarsest equitable partition $\mathcal{P}_G$, (17) is a block diagonal form of X also with respect to $\mathcal{P}_G$.