On a Hierarchy of Spectral Invariants for Graphs

V. Arvind
The Institute of Mathematical Sciences (HBNI), Chennai, India
Chennai Mathematical Institute, India

Frank Fuhlbrück
Institut für Informatik, Humboldt-Universität zu Berlin, Germany

Johannes Köbler
Institut für Informatik, Humboldt-Universität zu Berlin, Germany

Oleg Verbitsky
Institut für Informatik, Humboldt-Universität zu Berlin, Germany

Abstract

We consider a hierarchy of graph invariants that naturally extends the spectral invariants defined by Fürer (Lin. Alg. Appl. 2010) based on the angles formed by the set of standard basis vectors and their projections onto eigenspaces of the adjacency matrix. We provide a purely combinatorial characterization of this hierarchy in terms of the walk counts. This allows us to give a complete answer to Fürer's question about the strength of his invariants in distinguishing non-isomorphic graphs in comparison to the 2-dimensional Weisfeiler-Leman algorithm, extending the recent work of Rattan and Seppelt (SODA 2023). As another application of the characterization, we prove that almost all graphs are determined up to isomorphism in terms of the spectrum and the angles, which is of interest in view of the long-standing open problem whether almost all graphs are determined by their eigenvalues alone. Finally, we describe the exact relationship between the hierarchy and the Weisfeiler-Leman algorithms for small dimensions, as also some other important spectral characteristics of a graph such as the generalized and the main spectra.

2012 ACM Subject Classification Theory of computation → Graph algorithms analysis; Mathematics of computing → Graph algorithms

Keywords and phrases Graph Isomorphism, spectra of graphs, combinatorial refinement, strongly regular graphs

Digital Object Identifier 10.4230/LIPIcs.STACS.2024.6

Related Version *Full Version*: `https://arxiv.org/abs/2310.04391` [1]

Supplementary Material *Software*: `https://gitlab.com/tcslib/tcsliblua/`
archived at `swh:1:rev:7e429da7fcf9bf8d93dfc50b54e823a47364cb4b`

Funding *Oleg Verbitsky*: Supported by DFG grant KO 1053/8–2. On leave from the IAPMM, Lviv, Ukraine.

1 Introduction

The spectrum of a graph is a remarkable graph invariant that has found numerous applications in computer science; e.g., [17, 20]. These applications are based on analyzing relevant information contained in the eigenvalues of a given graph. The maximum information possible is evidently obtained for graphs that are determined by their spectra up to isomorphism. This graph class, which is of direct relevance to the graph isomorphism problem, is often abbreviated as DS. Thus, a graph G is DS if every graph cospectral to G, i.e., with the same spectrum as G, is actually isomorphic to G. Though the problem of characterizing DS graphs has been intensively studied since the beginning of spectral graph theory (see [10] and

41st International Symposium on Theoretical Aspects of Computer Science (STACS 2024).
Editors: Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, and Daniel Lokshtanov;
Article No. 6; pp. 6:1–6:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

references therein), we are still far from a satisfactory solution. In particular, a long-standing open question [16, 25] is whether or not almost all graphs are DS. Here and in the rest of the paper, we say that *almost all* graphs have some property if a uniformly distributed random n-vertex graph[1] has this property with probability approaching 1 as n goes to infinity.

Somewhat surprisingly at first sight, the area is connected to a purely combinatorial approach to the graph isomorphism problem. In their seminal work, Weisfeiler and Leman [28] proposed and studied a method for distinguishing a graph G from another non-isomorphic graph by computing a sequence of canonical partitions of $V(G)^2$ into color classes. The final partition of $V(G)^2$ results in a *coherent configuration*, a concept which is studied in algebraic combinatorics [7] and plays an important role in isomorphism testing [2]. The method of [28] is now called the 2-dimensional Weisfeiler-Leman algorithm (2-WL). A similar approach based on partitioning $V(G)$ is known as *color refinement* and is often called 1-WL. Even this one-dimensional method is quite powerful as it suffices for identification of almost all graphs [3]. On the other hand, construction of graphs not identifiable by 2-WL is rather tricky. Particular examples are based on rare combinatorial objects. General constructions, like the far-reaching one in [6], give rather sporadic families of "hard" graphs. It turns out (see, e.g., [4, 15]) that if two graphs are indistinguishable by 2-WL, then they are cospectral. As a consequence, graphs not identifiable by 2-WL are examples of non-DS graphs.

Our overall goal in this paper is a systematic exploration of connections between spectral and combinatorial approaches to finding expressive graph invariants. A *graph invariant* $\mathcal{I}$ is a function of a graph such that $\mathcal{I}(G) = \mathcal{I}(H)$ whenever $G \cong H$. An invariant $\mathcal{I}$ is *stronger* than invariant $\mathcal{I}'$ if $\mathcal{I}(G)$ determines $\mathcal{I}'(G)$. That is, $\mathcal{I}(G) = \mathcal{I}(H)$ implies $\mathcal{I}'(G) = \mathcal{I}'(H)$. Equivalently, we sometimes say that $\mathcal{I}'$ is *weaker* than $\mathcal{I}$ and write $\mathcal{I}' \preceq \mathcal{I}$. A stronger invariant can be more effective in distinguishing non-isomorphic graphs: If $\mathcal{I}' \preceq \mathcal{I}$ and $\mathcal{I}'$ can distinguish non-isomorphic graphs G and H, i.e., $\mathcal{I}'(G) \neq \mathcal{I}'(H)$, then these graphs are distinguishable by $\mathcal{I}$ as well.

Let $\mathsf{Spec}(G)$ denote the spectrum of a graph G, and $\mathsf{WL}_2(G)$ denote the output of 2-WL on G (a formal definition is given in Section 2.2). The discussion above shows that $\mathsf{Spec} \preceq \mathsf{WL}_2$. This can be seen as evidence of limitations of Spec, as well as evidence of the power of WL_2.

A reasonable question is what can be achieved if $\mathsf{Spec}(G)$ is enhanced by other spectral characteristics of the adjacency matrix of G. One such line of research in spectral graph theory considers $\mathsf{Spec}(G)$ augmented with the multiset of all angles between the standard basis vectors and the eigenspaces of G. The parameters and properties of a graph G which are determined by its *e*igenvalues and *a*ngles are called *EA-reconstructible* and are thoroughly studied by Cvetković and co-authors; see [11, Ch. 4] and [10, Ch. 3].

A further natural step is to take into consideration the multisets of angles between the projections of the standard basis vectors onto eigenspaces. Fürer [15] uses this additional data to define two new graph invariants, namely, the *weak* and *strong spectral invariants*. We denote these spectral invariants by $\mathsf{weak\text{-}FSI}$ and $\mathsf{strong\text{-}FSI}$ respectively; formal definitions are in Section 3. Fürer shows that both $\mathsf{weak\text{-}FSI}$ and $\mathsf{strong\text{-}FSI}$ remain weaker than WL_2. That is,

$$\mathsf{weak\text{-}FSI} \preceq \mathsf{strong\text{-}FSI} \preceq \mathsf{WL}_2 \tag{1}$$

(note that $\mathsf{Spec} \preceq \mathsf{weak\text{-}FSI}$ by definition). An open problem posed in [15] is to determine which of the relations in (1) are strict. Rattan and Seppelt, in their recent paper [23], show that this small hierarchy does not entirely collapse by separating $\mathsf{weak\text{-}FSI}$ and WL_2. Hence,

[1] In the Erdős-Rényi $\mathcal{G}(n, 1/2)$ random graph model to be precise.

at least one of the two relations in (1) is strict. Fürer [15] conjectures that the first relation in (1) is strict and does not exclude that the last two invariants in (1) are equivalent, and our aim is to give precise answers to these questions.

In [23] the invariants weak-FSI and WL_2 are separated by introducing a new natural graph invariant $\mathsf{WL}_{3/2}$, whose strength lies between WL_1 and WL_2. The authors give an elegant algebraic characterization of $\mathsf{WL}_{3/2}$ using which they show that weak-FSI $\preceq \mathsf{WL}_{3/2}$. The final step in their analysis is an example of graph pair that separates $\mathsf{WL}_{3/2}$ and WL_2.

Our approach is different. First, we observe that the invariants weak-FSI and strong-FSI are part of a broader scheme, presented in Section 2.1, that leads to a potentially infinite hierarchy of graph invariants. We define the corresponding spectral hierarchy containing weak-FSI and strong-FSI on its lower levels in Section 3. Another level is taken by the aforementioned invariant EA. In Section 4 we characterize this hierarchy in terms of walk counts. A connection between spectral parameters and walk counts is actually well known (see an overview in Subsection 4.1). With a little extra effort we are able to show that this connection is tight; see Theorem 4. This yields a purely combinatorial characterization of the invariants EA, weak-FSI, and strong-FSI (Corollary 5), which also reveals some new relations. For example, we notice that weak-FSI determines the *generalized spectrum* of a graph (Theorem 11).

As another application of our combinatorial characterization, we prove that almost all graphs are determined up to isomorphism by weak-FSI, that is, by the eigenvalues and the angles formed by the standard basis vectors and their projections onto eigenspaces (Corollary 8). We find this interesting in the context of the open problem mentioned above: whether or not almost all graphs are DS.

We present the relations between the spectral and combinatorial invariants under consideration in Section 5; see the diagram in Fig. 1. In Section 6 we prove that this diagram is complete, that is, it shows all existent relations, and all of these relations are strict (perhaps up to higher levels of the hierarchy whose separation remains open). In particular, both relations in (1) are strict, which gives a complete answer to Fürer's questions. Another noteworthy separation is strong-FSI $\not\preceq \mathsf{WL}_{3/2}$ (Theorem 12). Curiously, the separating pair of graphs is the same that was used in [23], which yields more information now because we also use our characterization in Theorem 4.

The more involved separations are shown in Theorems 13, 16, and 18. The corresponding separating examples are not ad hoc. They are obtained by a quite general construction (Lemma 14). The construction is based on a considerable extension of the approach taken in [27] to separate various concepts related to 1-WL and the walk matrix of a graph (an important notion discussed in Section 4.2). Implementation of the construction requires vertex-colored strongly regular graphs with certain properties. The required colorings were found by a computer assisted search among members of the family of strongly regular graphs on 25 vertices.

Some proofs are missing due to the space constraints and can be found in the full version of the paper [1].

2 Preliminaries

2.1 From isomorphism-invariant colorings to isomorphism invariants

Let $\mathcal{C}$ be a set of colors and $\chi : V(G)^2 \to \mathcal{C}$ be a coloring of vertex pairs in a graph G. It is natural to see $\chi(x,x)$ as the color of a vertex x. We suppose that $\chi = \chi_G$ is defined for every graph G. That is, speaking of a coloring χ, we actually mean a map $G \mapsto \chi_G$. Such a coloring χ is *isomorphism invariant* if for every isomorphism f from a graph G to a graph H (the equality $G = H$ is not excluded) we have $\chi_G(x, y) = \chi_H(f(x), f(y))$ for all vertices x and y in G.

The simplest isomorphism-invariant colorings are the adjacency relation A and the identity relation I. That is, $A(x, y) = 1$ if x and y are adjacent and $A(x, y) = 0$ otherwise. For the identity relation, $I(x, y) = 1$ if $x = y$ and $I(x, y) = 0$ if $x \neq y$. Below, we will consider A and I also as the adjacency and the identity matrices. Other examples of isomorphism-invariant colorings are the distance $d(x, y)$ between two vertices x and y and the number of all such triangles in the graph that contain the vertices x and y. One can consider also more complex definitions like the triple $\chi(x, y) = (\deg x, d(x, y), \deg y)$, where $\deg x$ denotes the degree of x.

Given an isomorphism-invariant coloring χ, we can build on it to define various graph invariants. The simplest such examples are

$$\begin{aligned}
\mathcal{I}_1(G) &= \{\!\!\{\chi(x, y)\}\!\!\}_{x,y \in V(G)}\,, \\
\mathcal{I}_2(G) &= \Big(\{\!\!\{\chi(x, x)\}\!\!\}_{x \in V(G)}\,, \{\!\!\{\chi(x, y)\}\!\!\}_{x,y \in V(G)}\Big)\,, \\
\mathcal{I}_3(G) &= \{\!\!\{(\chi(x, x), \chi(x, y), \chi(y, y))\}\!\!\}_{x,y \in V(G)}\,, \\
\mathcal{I}_4(G) &= \Big\{\!\!\Big\{\Big(\chi(x, x), \{\!\!\{(\chi(x, y), \chi(y, y))\}\!\!\}_{y \in V(G)}\Big)\Big\}\!\!\Big\}_{x \in V(G)}\,,
\end{aligned}$$

where $\{\!\!\{\ldots\}\!\!\}$ denotes a multiset. Note that $\mathcal{I}_1(G) \preceq \mathcal{I}_2(G) \preceq \mathcal{I}_3(G) \preceq \mathcal{I}_4(G)$.

Given an isomorphism-invariant coloring χ, we can define a hierarchy of ever more complex graph invariants. We first inductively define a sequence of colorings $\chi_0, \chi_1, \chi_2, \ldots$ of single vertices by

$$\chi_0(x) = \chi(x, x) \text{ and } \chi_{r+1}(x) = \Big(\chi_r(x), \{\!\!\{(\chi(x, y), \chi_r(y))\}\!\!\}_{y \in V(G)}\Big)\,. \tag{2}$$

This definition is a natural extension of the well-known concept of *color refinement (CR)* to edge- and vertex-colored graphs. In broad outline, CR computes an isomorphism-invariant color of each vertex in an input graph and recognizes two graphs as non-isomorphic if one of the colors occurs in one graph more frequently than in the other. In the case of an uncolored undirected graph G, CR starts with a uniform coloring χ_0 of $V(G)$, that is, $\chi_0(x) = \chi_0(x')$ for all $x, x' \in V(G)$. In the $(r + 1)$-th round, the preceding coloring χ_r is refined to a new coloring χ_{r+1}. For each vertex x, its new color $\chi_{r+1}(x)$ consists of $\chi_r(x)$ and the multiset $\{\!\!\{\chi_r(y)\}\!\!\}_{y \in N(x)}$ of all colors occurring in the neighborhood $N(x)$ of x. In other words, CR counts how frequently each χ_r-color occurs among the vertices adjacent to x (or, equivalently, among the vertices non-adjacent to x). In an edge- and vertex-colored graph G, each vertex pair (x, y) is assigned a color, which we denote by $\chi(x, y)$. The edge colors must be taken into account while computing the refined color $\chi_{r+1}(x)$. In the colored case, CR first splits all vertices y into classes depending on $\chi(x, y)$ and then computes the frequencies of $\chi_r(y)$ within each class. This is exactly what (2) does.

Note that

$$\chi_1(x) = \Big(\chi(x, x), \{\!\!\{(\chi(x, y), \chi(y, y))\}\!\!\}_{y \in V(G)}\Big)\,. \tag{3}$$

In addition, we set

$$\chi_{1/2}(x) = \Big(\chi(x, x), \{\!\!\{\chi(x, y)\}\!\!\}_{y \in V(G)}\Big) \tag{4}$$

Now, we define

$$\chi^{(r)}(G) = \{\!\!\{\chi_r(x)\}\!\!\}_{x \in V(G)}\,. \tag{5}$$

For each r, the coloring χ_r is isomorphism invariant in the sense that $\chi_r(f(x)) = \chi_r(x)$ for any isomorphism of graphs f. This readily implies that $\chi^{(r)}$ is a graph invariant. As easily seen, $\chi^{(r)} \preceq \chi^{(s)}$ if $r \leq s$.

2.2 First two dimensions of combinatorial refinement

We now give formal descriptions of the isomorphism tests 1-WL and 2-WL already introduced in Section 1. Note that 1-WL is an alternative name for CR. In what follows we will apply the 1-WL procedure to *vertex-colored* graphs. Given a vertex-colored graph G, we consider a coloring χ of $V(G)^2$ defined by $\chi(x, y) = A(x, y)$, i.e., according to the adjacency relation, for $x \neq y$ and by setting $\chi(x, x)$ to be the color of a vertex x. On an input G, 1-WL iteratively computes the vertex colorings χ_r according to (2). After performing n iterations, where n is the number of vertices in G, 1-WL outputs $\mathsf{WL}_1(G) = \chi^{(n)}(G)$ as defined by (5). Two graphs G and H are recognized as non-isomorphic if $\mathsf{WL}_1(G) \neq \mathsf{WL}_1(H)$.

2-WL can be similarly formulated, except that it computes colorings of vertex pairs. If an input graph G is uncolored, then 2-WL begins with an initial coloring χ_0 of $V(G)^2$ defined by $\chi_0(x, y) = A(x, y)$ if $x \neq y$ and by $\chi_0(x, x) = 2$ for every vertex x of G. If G is a vertex-colored graph, then $\chi_0(x, y)$ must include also the colors of x and y. Furthermore,

$$\chi_{r+1}(x, y) = \Big(\chi_r(x, y),\ \{\!\!\{(\chi_r(x, z),\ \chi_r(z, y))\}\!\!\}_{z \in V(G)}\Big).$$

Thus, the new color of a pair (x, y) can be seen as a kind of “superposition” of the old color pairs observable along all extensions of (x, y) to a triangle xzy. Finally, 2-WL outputs the multiset $\mathsf{WL}_2(G) = \{\!\!\{\chi_{n^2}(x, y)\}\!\!\}_{x, y \in V(G)}$.

3 A hierarchy of spectral invariants

Speaking of an n-vertex graph G, we will assume that $V(G) = \{1, 2, \ldots, n\}$. Let

$$\mu_1 < \mu_2 < \ldots < \mu_m \tag{6}$$

be all pairwise distinct eigenvalues of the adjacency matrix A of G. Let $\mathsf{Spec}(G)$ denote the spectrum of G, i.e., the multiset of all eigenvalues where each μ_i occurs with its multiplicity. As mentioned before, $\mathsf{Spec}(G)$ is a well-studied graph invariant with numerous applications in computer science.

Let E_i be the eigenspace of μ_i. Recall that E_i consists of all eigenvectors of μ_i, i.e., $E_i = \{v \in \mathbb{R}^n : Av = \mu_i v\}$. Let P_i be the matrix of the orthogonal projection of $\mathbb{R}^n$ onto E_i. Note that $P_i^2 = P_i = P_i^\top$. For $1 \leq x, y \leq n$, the matrix entry $P_i(x, y)$ can be seen a color of the vertex pair (x, y). This coloring is isomorphism invariant.

Throughout the paper, we use the following notational convention for compact representations of sequences.

▶ **Notation 1.** *For an indexed set $\{a_i\}$ with index i ranging through the interval of integers $s, s+1, \ldots, t-1, t$ we set $a_* = (a_s, \ldots, a_t)$.*

In particular,

$$P_*(x, y) = (P_1(x, y), \ldots, P_m(x, y)).$$

Since the order on the index set is determined by (6), P_* is also an isomorphism-invariant coloring. Following the general framework in Section 2.1, the coloring P_* determines the sequence of graph invariants $P_*^{(0)}, P_*^{(1/2)}, P_*^{(1)}, P_*^{(2)}, \ldots$. In particular, by (2)–(5) we have

$$\begin{aligned} P_*^{(0)}(G) &= \{\!\{\ P_*(x,x)\ \}\!\}_{1\le x\le n}\,, && (7)\\ P_*^{(1/2)}(G) &= \left\{\!\!\left\{\ \left(\ P_*(x,x),\ \{\!\{P_*(x,y)\}\!\}_{1\le y\le n}\ \right)\ \right\}\!\!\right\}_{1\le x\le n}\,, \\ P_*^{(1)}(G) &= \left\{\!\!\left\{\ \left(\ P_*(x,x),\ \{\!\{(P_*(x,y),P_*(y,y))\}\!\}_{1\le y\le n}\ \right)\ \right\}\!\!\right\}_{1\le x\le n}\,. \end{aligned}$$

Fürer [15] introduces the *weak* and *strong spectral invariants*. Using our notation, Fürer's spectral invariants (FSI) can be defined as follows:

$$\begin{aligned} \mathsf{weak\text{-}FSI}(G) &= \left(\mathsf{Spec}(G), P_*^{(1/2)}(G)\right) \text{ and} && (8)\\ \mathsf{strong\text{-}FSI}(G) &= \left(\mathsf{Spec}(G), P_*^{(1)}(G)\right). && (9) \end{aligned}$$

The entries of the projection matrices P_i have a well-known geometric meaning [10, 12]. For $1 \le x \le n$, the standard basis vector e_x of $\mathbb{R}^n$ has 1 in the position x and 0 elsewhere. The *angle* $\alpha_{i,x}$ of a graph G is defined to be the cosine of the angle between e_x and the eigenspace E_i, i.e., the angle between e_x and its projection $P_i\mathrm{e}_x$ onto E_i. We have the equality

$$P_i(x,x) = \alpha_{i,x}^2. \tag{10}$$

Indeed, let $\langle u, v\rangle$ denote the scalar product of two vectors $u, v \in \mathbb{R}^n$. Then

$$P_i(x,x) = \langle \mathrm{e}_x, P_i\mathrm{e}_x\rangle = \|\mathrm{e}_x\|\|P_i\mathrm{e}_x\|\alpha_{i,x} = \alpha_{i,x}^2.$$

Furthermore, let $\alpha_{i,xy}$ be the cosine of the angle between the projections $P_i\mathrm{e}_x$ and $P_i\mathrm{e}_y$ of the standard basis vector e_x and e_y onto E_i. If e_x or e_y is orthogonal to E_i, i.e., $\alpha_{i,x} = 0$ or $\alpha_{i,y} = 0$, then the angle is undefined and we set $\alpha_{i,xy} = 0$ in this case. In particular, $\alpha_{i,xx} = 0$ if $\alpha_{i,x} = 0$ while $\alpha_{i,xx} = 1$ if $\alpha_{i,x} \ne 0$. Equality (10) generalizes to

$$P_i(x,y) = \alpha_{i,x}\alpha_{i,y}\alpha_{i,xy}. \tag{11}$$

Indeed,

$$P_i(x,y) = \langle \mathrm{e}_x, P_i\mathrm{e}_y\rangle = \langle \mathrm{e}_x, P_i^2\mathrm{e}_y\rangle = \langle P_i\mathrm{e}_x, P_i\mathrm{e}_y\rangle = \|P_i\mathrm{e}_x\|\|P_i\mathrm{e}_y\|\alpha_{i,xy} = \alpha_{i,x}\alpha_{i,y}\alpha_{i,xy}.$$

Now, define a coloring α_i by $\alpha_i(x,x) = \alpha_{i,x}$ and $\alpha_i(x,y) = \alpha_{i,xy}$ for $x \ne y$. This coloring is isomorphism invariant basically because an isomorphism is represented by a permutation matrix, which is the transformation matrix of an isometry of $\mathbb{R}^n$. Using Notation 1,

$$\alpha_*(x,y) = (\alpha_1(x,y), \dots, \alpha_m(x,y))\,,$$

where α_* is an isomorphism-invariant coloring as well. The corresponding graph invariants $\alpha_*^{(r)}$ are closely related to the invariants $P_*^{(r)}$. More precisely, we say that two graph invariants $\mathcal{I}$ and $\mathcal{I}'$ are *equivalent* and write $\mathcal{I} \equiv \mathcal{I}'$ if $\mathcal{I}' \preceq \mathcal{I}$ and $\mathcal{I} \preceq \mathcal{I}'$.

► **Lemma 2.** $P_*^{(r)} \equiv \alpha_*^{(r)}$ *for every integer* $r \ge 0$.

Motivated by the equivalence $P_*^{(0)} \equiv \alpha_*^{(0)}$, we define the graph invariant EA similar to (8)–(9) as

$$\mathsf{EA}(G) = \left(\mathsf{Spec}(G), P_*^{(0)}(G)\right), \tag{12}$$

where the abbreviation EA stands for *E*igenvalues and *A*ngles and corresponds to the known concept [11, 10] mentioned in the introduction.

4 Characterization of the spectral invariants by walk counts

A walk of length k (or k-walk) from a vertex x to a vertex y is a sequence of vertices $x = x_0, x_1, \dots, x_k = y$ such that every two successive vertices x_i, x_{i+1} are adjacent. Let $w_k(x, y)$ denote the number of walks of length k from x to y in a graph. Obviously, w_k is an isomorphism-invariant coloring in the sense of Section 2.1. In accordance with Notation 1, we also consider the isomorphism-invariant coloring w_* defined by

$$w_*(x, y) = (w_0(x, y), w_1(x, y), \dots, w_{n-1}(x, y)),$$

where n, as usually, denotes the number of vertices in a graph. Note that for each y, the matrix $(w_k(x, y))_{1 \le x \le n,\, 0 \le k \le n-1}$ determines the value of $w_k(x, y)$ for every x and for every arbitrarily large k.

It is well known that the walk counts are expressible in terms of spectral parameters of a graph; see, e.g., [10]. On the other hand, it is also well known that the numbers of closed k-walks in a graph determine the graph spectrum. We give a brief overview of these facts in Subsection 4.1. In Subsection 4.2 we make further use of this connection between walks and spectra. We are able to characterize the spectral invariants defined in Section 3 using solely the walk numbers, that is, in purely combinatorial terms without involving any linear algebra.

4.1 Linear-algebraic background and known relations

Since the adjacency matrix A of a graph G is symmetric, the eigenspaces E_i are pairwise orthogonal, and hence $P_i P_j = O$ for $i \neq j$, where O denotes the zero matrix. The spectral theorem for symmetric matrices says in essence that

$$\mathbb{R}^n = E_1 \oplus \dots \oplus E_m,$$

that is, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of A. This decomposition implies that

$$I = P_1 + \dots + P_m. \tag{13}$$

From Equality (13) it is easy to derive the spectral decomposition

$$A = \mu_1 P_1 + \dots + \mu_m P_m.$$

Raising both sides of this equality to the k-th power and taking into account that $P_i^2 = P_i$ and $P_i P_j = O$ for $i \neq j$, we conclude that

$$A^k = \mu_1^k P_1 + \dots + \mu_m^k P_m.$$

Since $w_k(x, y) = A^k(x, y)$, we get

$$w_k(x, y) = \mu_1^k P_1(x, y) + \dots + \mu_m^k P_m(x, y). \tag{14}$$

Let $c_k(G) = \sum_{x \in V(G)} w_k(x, x)$ denote the total number of closed k-walks in a graph G. In particular, $c_0(G) = n$ and $c_1(G) = 0$.

▶ **Lemma 3** (folklore). Spec G = Spec H *if and only if* $c_k(G) = c_k(H)$ *for* $k = 0, 1, \dots, n$.

4.2 General characterization and its consequences

▶ **Theorem 4.** $(\mathsf{Spec}, P_*^{(r)}) \equiv w_*^{(r)}$ *for every* $r = 0, 1/2, 1, 2, \ldots$.

Before proving Theorem 4, we describe some of its consequences.

▶ **Corollary 5.**
1. $\mathsf{EA} \equiv w_*^{(0)}$.
2. weak-FSI $\equiv w_*^{(1/2)}$.
3. strong-FSI $\equiv w_*^{(1)}$.

Parts 2 and 3 of Corollary 5, which are particular cases of Theorem 4 for $r = 1/2$ and $r = 1$ respectively, provide a characterization of both Fürer's spectral invariants. We now comment on Part 1. By Definition (12), this part is the special case of Theorem 4 for $r = 0$. Note that $w_*^{(0)}(G) = w_*^{(0)}(H)$ if and only if the graphs G and H are *closed-walk-equivalent* in the sense that there is a bijection $f : V(G) \to V(H)$ such that $w_k(x, x) = w_k(f(x), f(x))$ for all $x \in V(G)$ and all k. On the other hand, let us say that G and H are *EA-equivalent* if these graphs have the same eigenvalues and angles, i.e., $\mathsf{EA}(G) = \mathsf{EA}(H)$. As seen from the summary in Subsection 4.1, there are well-known connections between the closed walk numbers and the eigenvalues and angles. Part 1 of Corollary 5 pinpoints the fact that the two equivalence concepts actually coincide.

Corollary 5 reveals connections of spectral invariants to other graph invariants studied in the literature, which we introduce now.

The *walk matrix* W of a graph G is indexed by vertices $1 \le x \le n$ and the length parameter $0 \le k \le n-1$ and defined by

$$W(x, k) = \sum_{y=1}^{n} w_k(x, y). \tag{15}$$

That is, $W(x, k)$ is the total number of k-walks starting from the vertex x. Let $WL_1^k(G, x)$ denote the color assigned by 1-WL to a vertex x of G after the k-th refinement round. For a vertex x of G, let G_x denote the version of G with x *individualized*. This means that G_x is a vertex-colored graph where x has a special unique color while the other vertices are colored uniformly. We now state two well-known facts.

▶ **Lemma 6.**
1. $W(x, k)$ *is determined by* $WL_1^k(G, x)$*;*
2. $w_k(x, y)$ *is determined by* $WL_1^k(G_x, y)$*;*

Part 1 of Lemma 6 is proved in algebraic terms in [22, Theorem 2]. Another proof, involving logical concepts, is provided in [13, Lemma 4] and a direct combinatorial proof can be found in [27, Lemma 8]. Part 2 is a straightforward extension of Part 1.

In addition to the graph invariants WL_1 and WL_2 introduced in Section 2.2, we define

$$\mathsf{WL}_{3/2}(G) = \{\!\{\mathsf{WL}_1(G_x)\}\!\}_{x \in V(G)}\,.$$

This yields a chain of graph invariants WL_d for $d \in \{1, 3/2, 2\}$, where $\mathsf{WL}_c \preceq \mathsf{WL}_d$ if $c \le d$. The walk matrix naturally gives us a graph invariant, which we denote by WM and define as $\mathsf{WM}(G) = \{\!\{W(x, *)\}\!\}_{x \in V(G)}$ where $W(x, *) = \big(W(x, 0), W(x, 1), \ldots, W(x, n-1)\big)$. In other words, $\mathsf{WM}(G)$ is the multiset of the rows of the walk matrix of G. Part 1 of Lemma 6 readily implies that $\mathsf{WM} \preceq \mathsf{WL}_1$. Thus,

$$\mathsf{WM} \preceq \mathsf{WL}_1 \preceq \mathsf{WL}_{3/2} \preceq \mathsf{WL}_2.$$

The second relation in the following corollary is the recent result in [23] already described in Section 1.

▶ **Corollary 7.** $\mathsf{WM} \preceq \mathsf{weak\text{-}FSI} \preceq \mathsf{WL}_{3/2}$.

Proof. By Part 2 of Corollary 5, it is enough to show that

$$\mathsf{WM} \preceq w_*^{(1/2)} \preceq \mathsf{WL}_{3/2}.$$

The former relation follows directly from the definitions of WM and $w_*^{(1/2)}$. Indeed, $w_*^{(1/2)}(G)$ comprises the multiset $\{\!\{w_*(x,y)\}\!\}_{y\in V(G)}$ for each vertex x of G; see (4). This multiset allows us to calculate the sum (15) for each k. The latter relation follows from the definitions of $w_*^{(1/2)}$ and $\mathsf{WL}_{3/2}$ by Part 2 of Lemma 6. ◀

The relationship between the graph invariants discussed above is summarized in Figure 1 below, which also puts these invariants in a somewhat broader context.

The next consequence of Theorem 4 is interesting in view of the long-standing open question whether almost all graphs are determined up to isomorphism by their spectrum [16, 25].

▶ **Corollary 8.** *Almost all graphs are determined up to isomorphism by* weak-FSI*.*

Proof. It is known that if the walk matrix is non-singular, then it determines the adjacency matrix [19]. Moreover, the walk matrix of a random graph is non-singular with high probability [21]. As a consequence, almost all graphs are determined by the graph invariant WM; see [19, Th. 7.2]. The same is true also for weak-FSI because weak-FSI is stronger than WM by Corollary 7. ◀

Since $\mathsf{Spec} \preceq \mathsf{EA} \preceq \mathsf{weak\text{-}FSI}$, a natural further question is whether Corollary 8 can be improved to the identifiability of almost all graphs by the graph invariant EA, that is, by using the eigenvalues and the angles between each standard basis vector and its projections onto eigenspaces but not between the projections themselves. If true, this would be yet closer to the aforementioned open problem.

4.3 Proof of Theorem 4

The case of $r = 0$. We have to prove that

$$(\mathsf{Spec}, \{\!\{P_*(x,x)\}\!\}_x) \equiv \{\!\{w_*(x,x)\}\!\}_x\,. \tag{16}$$

The "$\succeq$" part immediately follows from Equality (14). In the other direction, the relation $\mathsf{Spec} \preceq \{\!\{w_*(x,x)\}\!\}_x$ is a direct consequence of Lemma 3. To complete the proof of (16), we show that for each vertex x, the sequence $P_*(x,x)$ can be obtained from the sequences $w_*(x,x)$ and μ_*. To this end, put $y = x$ in Equality (14), obtaining

$$\mu_1^k P_1(x,x) + \cdots + \mu_m^k P_m(x,x) = w_k(x,x). \tag{17}$$

This equality makes sense also for $k = 0$. In this case, it reads

$$P_1(x,x) + \cdots + P_m(x,x) = 1, \tag{18}$$

which is true by Equality (13) (if $\mu_i = 0$, we need to use the convention $0^0 = 1$). Consider Equalities (17) for $k = 0, 1, \ldots, m-1$ as a system of m linear equations for m unknowns $P_1(x,x), \ldots, P_m(x,x)$. The coefficients of this system are powers of the m pairwise distinct eigenvalues. They form a Vandermonde matrix. Therefore, the system is uniquely solvable, and the sequence $P_*(x,x)$ is determined.

The case of $r = 1/2$. Now we have to prove that

$$\left(\mathsf{Spec}, \left\{\!\!\left\{ \left(P_*(x,x), \{\!\!\{ P_*(x,y) \}\!\!\}_y \right) \right\}\!\!\right\}_x \right) \equiv \left\{\!\!\left\{ \left(w_*(x,x), \{\!\!\{ w_*(x,y) \}\!\!\}_y \right) \right\}\!\!\right\}_x . \tag{19}$$

The "$\succeq$" part is again an immediate consequence of Equality (14).

Let us prove the part "$\preceq$". That is, we have to show that for a given graph, the left hand side of (19) can be obtained from the right hand side. The spectrum is, as already observed in the case of $r = 0$, determined by the multiset $\{\!\!\{ w_*(x,x) \}\!\!\}_x$, which is easy to obtain from the right hand side of (19). Thus, in what follows we can assume that the sequence $\mu_1, \ldots, \mu_m$ of distinct eigenvalues is known. For each vertex x, we have to compute the sequence $P_*(x,x)$ and the multiset $\{\!\!\{ P_*(x,y) \}\!\!\}_y$. The former task is solvable exactly as in the case of $r = 0$, and we focus on the latter task. In addition to x, we now fix also y and consider Equalities (14) for $k = 0, 1, \ldots, m-1$. Here, the equality for $k = 0$ is actually a particular instance of Equality (13), that is, this is (18) if $y = x$ and

$$P_1(x,y) + \cdots + P_m(x,y) = 0$$

if $y \neq x$. As in the case of $r = 0$, for m unknowns $P_1(x,y), \ldots, P_m(x,y)$ we obtain a system of m linear equation whose coefficients form a non-singular Vandermonde matrix. Hence, we can determine the sequence $P_*(x,y)$, completing the proof of (19).

The case of $r \geq 1$. We proceed as above using induction. To facilitate the notation, we set $\pi(x,y) = P_*(x,y)$ and $\omega(x,y) = w_*(x,y)$. We write $a \hookleftarrow b$ to say that a is obtainable from b. As we have already seen,

$$\omega(x,y) \hookleftarrow \mathsf{Spec},\ \pi(x,y) \tag{20}$$

by Equality (14) and

$$\pi(x,y) \hookleftarrow \mathsf{Spec},\ \omega(x,y) \tag{21}$$

by the Vandermonde matrix argument. The vertex colorings π_r and ω_r are defined as in (2). For every r,

$$\mathsf{Spec} \hookleftarrow \{\!\!\{ \omega_0(x) \}\!\!\}_x \hookleftarrow \{\!\!\{ \omega_r(x) \}\!\!\}_x . \tag{22}$$

The former relation is, as already observed above, a consequence of Lemma 3, while the latter relation follows directly from the definition of ω_r. Therefore, in order to prove that

$$(\mathsf{Spec}, \{\!\!\{ \pi_r(x) \}\!\!\}_x) \preceq \{\!\!\{ \omega_r(x) \}\!\!\}_x ,$$

it suffices to prove for each x that

$$\pi_r(x) \hookleftarrow \mathsf{Spec},\ \omega_r(x). \tag{23}$$

In order to prove that

$$\{\!\!\{ \omega_r(x) \}\!\!\}_x \preceq (\mathsf{Spec}, \{\!\!\{ \pi_r(x) \}\!\!\}_x) ,$$

it suffices to prove for each x that

$$\omega_r(x) \hookleftarrow \mathsf{Spec},\ \pi_r(x). \tag{24}$$

We prove (23) and (24) by induction on r.

In the base case we have $r = 0$. The relation (24) for $r = 0$ follows from the relation (20) for $y = x$. Similarly, the relation (23) for $r = 0$ follows from the relation (21) for $y = x$.

For the induction step, suppose that $r \geq 1$. Consider first (23). Recall that

$$\pi_r(x) = \left(\pi_{r-1}(x), \{\!\{(\pi(x,y), \pi_{r-1}(y))\}\!\}_y\right).$$

By the induction hypothesis,

$$\pi_{r-1}(x) \hookleftarrow \mathsf{Spec},\ \omega_{r-1}(x) \hookleftarrow \mathsf{Spec},\ \omega_r(x).$$

The latter relation follows from the fact that $\omega_{r-1}(x)$ is a part of $\omega_r(x)$. Another part of $\omega_r(x)$ gives us the multiset $\{\!\{(\omega(x,y), \omega_{r-1}(y))\}\!\}_y$. Therefore, it suffices to argue that, for each y,

$$(\pi(x,y), \pi_{r-1}(y)) \hookleftarrow \mathsf{Spec},\ (\omega(x,y), \omega_{r-1}(y)).$$

Indeed, $\pi(x,y)$ is determined by (21), and $\pi_{r-1}(y)$ is determined by the induction hypothesis.

The argument for (24) is virtually the same, with the roles of π and ω interchanged. In place of (21), we have to refer to (20). The proof is complete.

5 The Hierarchy

Figure 1 shows the invariants from Section 4.2 and the relations between them as a part of a more general picture involving also some other spectral invariants studied in the literature, which we introduce in the next two subsections.

Moreover, we define a new invariant $w_*^{(\bullet)}$ which is, in a sense, the limit of the sequence of invariants $w_*^{(r)}$ for $r = 1, 2, \ldots$. The definition is rather general. As we already mentioned, the formal framework of Section 2.1 is analogous to the concept of color refinement. Given an isomorphism-invariant coloring χ, we therefore can along with graph invariants $\chi^{(r)}$ define the stable version $\chi^{(\bullet)}$. One possibility to do this is to set $\chi^{(\bullet)}(G) = \chi^{(n)}(G)$, where n is the number of vertices in G. Note that $\chi^{(r)} \preceq \chi^{(\bullet)}$ for every r. For $\chi = \omega_*$, we obtain

$$w_*^{(r)} \preceq w_*^{(r+1)} \preceq w_*^{(\bullet)} \preceq \mathsf{WL}_2.$$

To see the last relation above, we first recall the well-known fact that $w_*(x,y)$ is determined by the color assigned to the vertex pair (x,y) by 2-WL. Furthermore, since $w_*^{(\bullet)}(G)$ is obtained from G endowed with the coloring w_* by running the version of 1-WL for edge-colored graphs, the outcome can be simulated by 2-WL.

5.1 Main eigenvalues and angles

Let j denote the all-ones vector (the dimension should be clear from the context). Using our usual notation, suppose that G has m distinct eigenvalues $\mu_1, \ldots, \mu_m$, and let $E_1, \ldots, E_m$ be the corresponding eigenspaces of G. Consider the angle between E_i and j and denote its cosine by β_i. If $\beta_i \neq 0$, then the corresponding eigenvalue μ_i is called a *main eigenvalue*, and then the positive number β_i is called a *main angle*. Let $\nu_1, \ldots, \nu_{m'}$ be the sequence of all main eigenvalues in the ascending order and $\theta_1, \ldots, \theta_{m'}$ be the sequence of the main angles in the corresponding order. We define a graph invariant main-EA by

$$\mathsf{main\text{-}EA}(G) = (\nu_*, \theta_*).$$

$$
\begin{tikzcd}
\mathsf{WL}_2 \arrow[r] \arrow[dd] & w_*^{(\bullet)} \arrow[r] & \cdots \arrow[r] & w_*^{(r)} \arrow[d] & (r \geq 1) \\
 & & & \vdots \arrow[d] & \\
\mathsf{WL}_{3/2} \arrow[d] \arrow[drrr] & & & w_*^{(1)} \arrow[d] \arrow[r, leftrightarrow] & \mathsf{strong\text{-}FSI} \\
\mathsf{WL}_1 \arrow[d] & & & w_*^{(1/2)} \arrow[d] \arrow[dl] \arrow[dlll] \arrow[r, leftrightarrow] & \mathsf{weak\text{-}FSI} \\
\mathsf{WM} \arrow[d] & & \mathsf{gen\text{-}Spec} \arrow[dll] \arrow[dr] & w_*^{(0)} \arrow[d] \arrow[r, leftrightarrow] & \mathsf{EA} \\
\mathsf{main\text{-}EA} & & & \mathsf{Spec} &
\end{tikzcd}
$$

Figure 1 Relations between graph invariants. An arrow $\mathcal{I} \to \mathcal{I}'$ means $\mathcal{I}' \preceq \mathcal{I}$.

A characterization of main-EA in terms of walk numbers is known. Let

$$w_k(G) = \sum_{x \in V(G)} w_k(x)$$

be the total number of k-walks in G. By W_G we will denote the corresponding generating function, that is, the formal series

$$W_G(z) = \sum_{k=0}^{\infty} w_k(G) z^k.$$

▶ **Proposition 9** (folklore, e.g. [24]). *Let G and H be graphs with n vertices. Then* main-EA$(G) =$ main-EA(H) *if and only if $w_k(G) = w_k(H)$ for $k = 1, \ldots, n-1$.*

As a direct consequence of Proposition 9, we get the relation main-EA $\preceq$ WM.

5.2 The generalized spectrum

Another important spectral invariant of G is the spectrum of the complement graph $\overline{G}$. The equalities Spec $G =$ Spec H and, simultaneously, Spec $\overline{G} =$ Spec $\overline{H}$ are equivalent to the condition that the graphs G and H have the same *generalized spectrum*. For the definition of this concept and its various characterizations we refer the reader to [18] and [26, Th. 3]. We define the graph invariant gen-Spec by

$$\mathsf{gen\text{-}Spec}(G) = \left(\mathsf{Spec}\, G, \mathsf{Spec}\, \overline{G}\right).$$

We note that

$$\mathsf{main\text{-}EA} \preceq \mathsf{gen\text{-}Spec}\,. \tag{25}$$

This follows from Proposition 9 and the following result in [8]. Let P_G denote the characteristic polynomial of a graph G.

▶ **Proposition 10** (Cvetković [8]). $W_G(z) = \frac{1}{z}\left((-1)^n \frac{P_{\overline{G}}(-1/z-1)}{P_G(1/z)} - 1\right)$.

Alternatively, (25) can be obtained by using [26, Th. 3].

To complete the diagram in Figure 1, it remains to prove the following relation.

▶ **Theorem 11.** gen-Spec $\preceq$ weak-FSI.

Proof. The spectrum of G, hence also the characteristic polynomial P_G, is determined by weak-FSI(G) just by definition. We have to show that Spec $\overline{G}$ or, equivalently, $P_{\overline{G}}$ is also determined. By Proposition 10, $P_{\overline{G}}$ is obtainable from P_G and W_G. Using Part 2 of Corollary 5, it remains to notice that W_G is determined by $w_*^{(1/2)}(G)$. Indeed,

$$w_k(G) = \sum_x w_k(x) = \sum_x \sum_y w_k(x, y),$$

where the right hand side is obtainable from the multiset $\left\{\!\!\left\{ \{\!\!\{ w_k(x,y) \}\!\!\}_y \right\}\!\!\right\}_x$, which is a part of $w_*^{(1/2)}(G)$. ◀

6 Separations

Fürer [15] poses the open problem of determining which of the relations in the chain

$$\text{weak-FSI} \preceq \text{strong-FSI} \preceq \text{WL}_2 \tag{26}$$

are strict. As mentioned before, Rattan and Seppelt [23] show that this chain does not entirely collapse. They separate weak-FSI and WL$_2$ by proving that weak-FSI $\preceq$ WL$_{3/2}$ and separating WL$_{3/2}$ and WL$_2$. Hence, at least one of the two relations in (26) is strict. Fürer [15] conjectures that the first relation in (26) is strict and does not exclude the possibility that the last two invariants in (26) are equivalent. We settle this by showing that, in fact, both relations in (26) are strict. We actually prove much more: up to one question remaining open, the diagram shown in Figure 1 is exact in the sense that all present arrows are non-reversible and that any two invariants not connected by arrows are provably incomparable. The only remaining question concerns the chain $w_*^{(\bullet)} \to \cdots \to w_*^{(r)} \to \cdots \to w_*^{(1)}$; see Problem 17 stated below and an approach to its solution in Theorem 18.

We now present a minimal set of separations from which all other separations follow. For compatibility with Figure 1, we write $\mathcal{I} \nrightarrow \mathcal{I}'$ to negate $\mathcal{I}' \preceq \mathcal{I}$.

WL$_1$ $\nrightarrow$ Spec. To show this, we have to present two WL$_1$-equivalent but not cospectral graphs. The simplest pair of WL$_1$-equivalent graphs, $2C_3$ and C_6, works. Indeed, the eigenvalues of C_n are $2\cos\frac{2\pi k}{n}$ for $k = 0, 1, \ldots, n-1$, and the spectrum of the disjoint union of graphs is the union of their spectra; see, e.g., [10, Example 1.1.4 and Theorem 2.1.1].

EA $\nrightarrow$ main-EA. It is known [9] that among trees with up to 20 vertices there is a single pair of non-isomorphic trees, with 19 vertices, with the same eigenvalues and angles. Using Proposition 9, a direct computation shows that these two trees are not main-EA-equivalent.

gen-Spec $\nrightarrow$ WM. The smallest, with respect to the number of vertices and the number of edges, pair of generalized cospectral graphs consists of 7-vertex graphs G and H where $G = C_6 \cup K_1$ and H is obtained from the 3-star $K_{1,3}$ by subdividing each edge; see [16, Fig. 4]. Since G has an isolated vertex and H does not, these graphs are not WM-equivalent.

gen-Spec $\not\to$ EA. The same pair of graphs G and H works. They are not EA-equivalent because connectedness of a graph is determined by its spectrum and angles [10, Th. 3.3.3]. Another reason for this is that every graph with at most 9 vertices is determined by its spectrum and angles up to isomorphism [9].

strong-FSI $\not\to$ WL$_1$. An even stronger fact, namely $w_*^{(2)} \not\to \mathsf{WL}_1$ is proved as Theorem 13 below. The separation implies that strong-FSI $\not\to$ WL$_2$, answering one part of Fürer's question. Another consequence is also the separation WM $\not\to$ WL$_1$, which follows as well from [27, Theorem 3].

WL$_1$ $\not\to$ WL$_{3/2}$. Consider C_6 and $2C_3$.

WL$_{3/2}$ $\not\to$ strong-FSI. This is Theorem 12 below. As a consequence, weak-FSI $\not\to$ strong-FSI, answering the other part of Fürer's question. Another consequence is the separation WL$_{3/2}$ $\not\to$ WL$_2$ shown in [23].

$w_*^{(\bullet)}$ $\not\to$ WL$_2$. This is Theorem 16 below. It considerably strengthens the negative answer to Fürer's question by showing that the whole hierarchy of the invariants $w_*^{(r)}$ for all r is strictly weaker than WL$_2$.

We state and prove the three results announced above in the rest of this section.

6.1 WL$_{3/2}$ is not stronger than strong-FSI

▶ **Theorem 12.** strong-FSI $\not\preceq$ WL$_{3/2}$.

Proof. The separation WL$_2$ $\not\preceq$ WL$_{3/2}$ in [23] is shown by constructing a pair of WL$_{3/2}$-equivalent graphs G and H as follows. Consider two copies of $C_6 * K_1$, where $*$ is the join of graphs. Denote the vertices of degree 6 by u_1 and u_2. Consider also two copies of $(2C_3) * K_1$, denoting their vertices of degree 6 by v_1 and v_2. The graph G is obtained by adding four edges forming the cycle $u_1u_2v_1v_2$, and the graph H is obtained by adding the cycle $u_1v_1u_2v_2$. In [23] it is observed that G and H are distinguishable by 2-WL. We now strengthen this observation to show that strong-FSI$(G) \neq$ strong-FSI(H).

The vertices u_1, u_2, v_1, v_2 will be referred to as Q-vertices. The other vertices are split into two classes, H-vertices and T-vertices, depending on whether they belong to a hexagonal or a triangular part. A vertex x is a Q-vertex exactly when $w_2(x,x) = \deg x = 8$. The H- and the T-vertices are distinguishable by the condition $w_3(x,x) = 6$ for a T-vertex and $w_3(x,x) = 4$ for an H-vertex. Consider an arbitrary T-vertex x in G. Note that from x there is at least one 3-walk to each of the twelve T-vertices y. If we consider a T-vertex x in H, then from x there are 3-walks only to six T-vertices y. This implies that $w_*^{(1)}(G) \neq w_*^{(1)}(H)$. We conclude by Part 3 of Corollary 5 that G and H are not strong-FSI-equivalent. ◀

6.2 $w_*^{(2)}$ is not stronger than WL$_1$

▶ **Theorem 13.** WL$_1$ $\not\preceq$ $w_*^{(2)}$.

The proof requires a substantial extension of the approach in [27] to separate various WL$_1$- and WM-based concepts.

Construction. Suppose that we have a graph A with m designated vertices $a_1, \ldots, a_m$ and a graph B with m designated vertices $b_1, \ldots, b_m$, which will be referred to as *port vertices*. In each of the graphs, the port vertices are colored by different colors. Specifically, a_i and b_i are colored by the same color i. The resulting partially colored graphs are denoted by A' and B'. We construct a graph $G(A', B')$ with no colored vertices as follows. $G(A', B')$

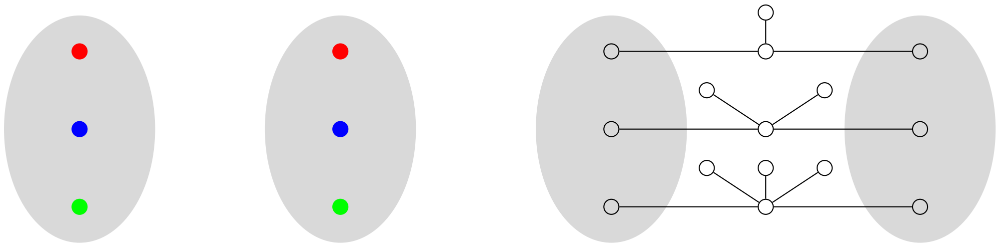

Figure 2 Construction of $G(A', B')$.

consists of the vertex-disjoint union of A and B and a number of new vertices of two sorts, namely *connecting* and *pendant vertices*. For each i, there is a connecting vertex c_i adjacent to a_i and b_i. Moreover, for each i there are i pendant vertices $p_{i,1}, \dots, p_{i,i}$ of degree 1 all adjacent to c_i. An example of the construction for $m = 3$ is shown in Figure 2.

The main lemma. The crux of the proof is the following lemma. Recall that a *strongly regular graph* with parameters (n, d, λ, μ) is an n-vertex d-regular graph where every two adjacent vertices have λ common neighbors, and every two non-adjacent vertices have μ common neighbors. Extending the notation used in Section 4.2, for a graph G we set $\mathit{WL}_1^r(G) = \{\!\{ \mathit{WL}_1^r(G, x) \}\!\}_{x \in V(G)}$.

▶ **Lemma 14.** *Let A and B be strongly regular graphs with the same parameters (in particular, A and B can be isomorphic). Let A' and B' be their partially colored versions such that each color occurs in A', as well as in B', at most once. Assume that $\mathit{WL}_1^0(A') = \mathit{WL}_1^0(B')$, which means the the sets of the colors occurring in A' and B' are equal and, therefore, we can construct the uncolored graph $G = G(A', B')$. Consider also $H = G(A', A')$ constructed from two vertex-disjoint copies of A'.*

1. *If $\mathit{WL}_1(A') \neq \mathit{WL}_1(B')$, then $\mathit{WL}_1(G) \neq \mathit{WL}_1(H)$. In words: if color refinement distinguishes A' and B', then it distinguishes also G and H.*
2. *If $\mathit{WL}_1^r(A') = \mathit{WL}_1^r(B')$ for some $r \geq 1$ (i.e., r rounds of color refinement do not suffice for distinguishing A' and B'), then $w_*^{(r-1)}(G) = w_*^{(r-1)}(H)$.*

The rest of the proof. We can separate WL_1 and $w_*^{(2)}$ by finding partially colored strongly regular graphs A' and B' as in Lemma 14 with $r = 3$. Let $\mathrm{SRG}(n, d, \lambda, \mu)$ denote the set of strongly regular graphs with parameters (n, d, λ, μ). Two suitable colorings A' and B' exist for a graph in the set SRG(25,12,5,6) of *Paulus graphs*, namely for the graph $P_{25.12}$ in Brouwer's collection [5]. These colorings are described in the full version of the paper [1]. They were found by computer search using the Lua package TCSLibLua [14].

Note that $P_{25.12}$ is one of the two Latin square graphs in SRG(25,12,5,6). The other Latin square graph in SRG(25,12,5,6) is $P_{25.15}$, which is the Paley graph on 25 vertices. This is the only vertex-transitive graph in this family. Curiously, it is not suitable for our purposes. Moreover, it seems that strongly regular graphs with less than 25 vertices do not admit appropriate colorings even for $r = 2$.

6.3 Separation of the $w_*^{(r)}$ hierarchy from WL_2

▶ **Lemma 15.** *Let A and B be (possibly isomorphic) strongly regular graphs with the same parameters. Let A' and B' be their versions, each containing a single individualized vertex. Let $G = G(A', B')$ and $H = G(A', A')$.*

1. $w_*^{(r)}(G) = w_*^{(r)}(H)$ *for all r and, therefore,* $w_*^{(\bullet)}(G) = w_*^{(\bullet)}(H)$.
2. *If $WL_2(A') \neq WL_2(B')$, then $WL_2(G) \neq WL_2(H)$.*

Proof. For Part 1, note that $\mathsf{WL}_1(A') = \mathsf{WL}_1(B')$. Indeed, the distinguishability of A' and B' by 1-WL would imply the distinguishability of A and B by 2-WL, contradicting the assumption that these strongly regular graphs have the same parameters. Thus, $WL_1^r(A') = WL_1^r(B')$ for all $r \geq 1$, and we can apply the second part of Lemma 14. Part 2 is easy. ◀

▶ **Theorem 16.** $\mathsf{WL}_2 \not\preceq w_*^{(\bullet)}$.

Proof. We apply Lemma 15, where A is the Shrikhande graph and B the 4×4 rook's graph, both strongly regular graphs with parameters $(16, 6, 2, 2)$. Since both graphs are vertex-transitive, A' and B' are uniquely defined. The neighborhood of the individualized vertex induces C_6 in the Shrikhande graph and $2C_3$ in the 4×4 rook's graph. Therefore, $\mathsf{WL}_2(A') \neq \mathsf{WL}_2(B')$. ◀

▶ **Open Problem 17.** *We leave open the question whether the hierarchy*

$$w_*^{(1)} \preceq w_*^{(2)} \preceq w_*^{(3)} \preceq w_*^{(4)} \preceq \cdots \preceq w_*^{(\bullet)} \tag{27}$$

is strict or at least does not collapse to some level. While we know that each $w_*^{(r)}$ is strictly weaker than WL_2, it remains open whether $w_*^{(r)}$ can be stronger than WL_1 for some large r. A negative answer will follow from Lemma 14 if there is an infinite sequence of partially colored strongly regular graphs A'_r and B'_r for $r = 1, 2, 3 \ldots$, where the underlying graphs are equal or have the same parameters, such that A'_r and B'_r are distinguished by 1-WL, but requiring at least r refinement rounds.

Suppose that strongly regular graphs A and B have the same parameters and their partially colored versions A' and B' are distinguished by 1-WL exactly in the $(r+1)$-th round. By Lemma 14, $G(A', B')$ and $G(A', A')$ are $w_*^{(r-1)}$-equivalent and, therefore, this pair of graphs is a good candidate for separation of $w_*^{(r-1)}$ from $w_*^{(r)}$. This approach works indeed pretty well.

▶ **Theorem 18.** *The hierarchy (27) is strict up to the 4-th level, that is, the first three relations in (27) are strict.*

References

1. V. Arvind, Frank Fuhlbrück, Johannes Köbler, and Oleg Verbitsky. On a hierarchy of spectral invariants for graphs. Technical report, `arxiv.org/abs/2310.04391`, 2023.
2. László Babai. Graph isomorphism in quasipolynomial time. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC'16)*, pages 684–697, 2016. `doi:10.1145/2897518.2897542`.
3. László Babai, Paul Erdős, and Stanley M. Selkow. Random graph isomorphism. *SIAM Journal on Computing*, 9(3):628–635, 1980. `doi:10.1137/0209047`.
4. Amir Rahnamai Barghi and Ilia Ponomarenko. Non-isomorphic graphs with cospectral symmetric powers. *Electron. J. Comb.*, 16(1), 2009. `doi:10.37236/209`.
5. Andries E. Brouwer. Paulus-Rozenfeld graphs. URL: `https://www.win.tue.nl/~aeb/drg/graphs/Paulus.html`.

6 Jin-yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identifications. *Combinatorica*, 12(4):389–410, 1992. `doi:10.1007/BF01305232`.

7 Gang Chen and Ilia Ponomarenko. *Coherent configurations*. Wuhan: Central China Normal University Press, 2019.

8 Dragoš Cvetković. The main part of the spectrum, divisors and switching of graphs. *Publ. Inst. Math., Nouv. Sér.*, 23:31–38, 1978. URL: `http://elib.mi.sanu.ac.rs/files/journals/publ/43/6.pdf`.

9 Dragoš Cvetković and Mirko Lepović. Cospectral graphs with the same angles and with a minimal number of vertices. *Publ. Elektroteh. Fak., Univ. Beogr., Ser. Mat.*, 8:88–102, 1997. URL: `http://www.jstor.org/stable/43666387`.

10 Dragoš Cvetković, Peter Rowlinson, and Slobodan Simić. *An introduction to the theory of graph spectra*, volume 75 of *Lond. Math. Soc. Stud. Texts*. Cambridge: Cambridge University Press, 2010. `doi:10.1017/CBO9780511801518`.

11 Dragoš Cvetković, Peter Rowlinson, and Slobodan Simic. *Eigenspaces of Graphs*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1997. `doi:10.1017/CBO9781139086547`.

12 Dragoš Cvetković, Peter Rowlinson, and Slobodan Simić. A study of eigenspaces of graphs. *Linear Algebra and its Applications*, 182:45–66, 1993. `doi:10.1016/0024-3795(93)90491-6`.

13 Z. Dvořák. On recognizing graphs by numbers of homomorphisms. *Journal of Graph Theory*, 64(4):330–342, 2010. `doi:10.1002/jgt.20461`.

14 Frank Fuhlbrück. TCSLibLua. GitHub repository, 2023. URL: `https://gitlab.com/tcslib/tcsliblua/-/tree/7e429da7fcf9bf8d93dfc50b54e823a47364cb4b`.

15 Martin Fürer. On the power of combinatorial and spectral invariants. *Linear Algebra and its Applications*, 432(9):2373–2380, 2010. `doi:10.1016/j.laa.2009.07.019`.

16 Willem H. Haemers and Edward Spence. Enumeration of cospectral graphs. *European Journal of Combinatorics*, 25(2):199–211, 2004. `doi:10.1016/S0195-6698(03)00100-8`.

17 Shlomo Hoory, Nathan Linial, and Avi Widgerson. Expander graphs and their applications. *Bull. Am. Math. Soc., New Ser.*, 43(4):439–561, 2006. `doi:10.1090/S0273-0979-06-01126-8`.

18 Charles R. Johnson and Morris Newman. A note on cospectral graphs. *Journal of Combinatorial Theory, Series B*, 28(1):96–103, 1980. `doi:10.1016/0095-8956(80)90058-1`.

19 Fenjin Liu and Johannes Siemons. Unlocking the walk matrix of a graph. *J. Algebr. Comb.*, 55(3):663–690, 2022. `doi:10.1007/s10801-021-01065-3`.

20 Bojan Mohar and Svatopluk Poljak. Eigenvalues in combinatorial optimization. In *Combinatorial and graph-theoretical problems in linear algebra*, pages 107–151. New York: Springer-Verlag, 1993. `doi:10.1007/978-1-4613-8354-3`.

21 Sean O'Rourke and Behrouz Touri. On a conjecture of Godsil concerning controllable random graphs. *SIAM J. Control. Optim.*, 54(6):3347–3378, 2016. `doi:10.1137/15M1049622`.

22 David L. Powers and Mohammad M. Sulaiman. The walk partition and colorations of a graph. *Linear Algebra Appl.*, 48:145–159, 1982. `doi:10.1016/0024-3795(82)90104-5`.

23 Gaurav Rattan and Tim Seppelt. Weisfeiler-Leman and graph spectra. In *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms (SODA'23)*, pages 2268–2285. SIAM, 2023. `doi:10.1137/1.9781611977554.ch87`.

24 Sirolf. Characterization of k-walk-equivalent graphs, 2018. URL: `https://mathoverflow.net/questions/304106/characterization-of-k-walk-equivalent-graphs`.

25 Edwin R. van Dam and Willem H. Haemers. Which graphs are determined by their spectrum? *Linear Algebra and its Applications*, 373:241–272, 2003. `doi:10.1016/S0024-3795(03)00483-X`.

26 E.R. van Dam, W.H. Haemers, and J.H. Koolen. Cospectral graphs and the generalized adjacency matrix. *Linear Algebra and its Applications*, 423(1):33–41, 2007. `doi:10.1016/j.laa.2006.07.017`.

27 Oleg Verbitsky and Maksim Zhukovskii. Canonization of a random graph by two matrix-vector multiplications. In *31st Annual European Symposium on Algorithms (ESA 2023)*, volume 274 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 100:1–100:13, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.ESA.2023.100`.

28 B.Yu. Weisfeiler and A.A. Leman. The reduction of a graph to canonical form and the algebra which appears therein. *NTI, Ser. 2*, 9:12–16, 1968. English translation is available at `https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf`.