On the expressive power of linear algebra on graphs

Floris Geerts
University of Antwerp, Antwerp, Belgium
floris.geerts@uantwerpen.be

November 14, 2018

Abstract

Most graph query languages are rooted in logic. By contrast, in this paper we consider graph query languages rooted in linear algebra. More specifically, we consider MATLANG, a matrix query language recently introduced, in which some basic linear algebra functionality is supported. We investigate the problem of characterising equivalence of graphs, represented by their adjacency matrices, for various fragments of MATLANG. A complete picture is painted of the impact of the linear algebra operations in MATLANG on their ability to distinguish graphs.

1 Introduction

Motivated by the importance of linear algebra for machine learning on big data [6, 7, 11, 46, 53] there is a current interest in languages that combine matrix operations with relational query languages in database systems [21, 34, 40, 41, 43]. Such hybrid languages raise many interesting questions from a database theoretical point of view. It seems natural, however, to first consider query languages for matrices alone. These are the focus of this paper.

More precisely, we continue the investigation of the expressive power of the matrix query language MATLANG, recently introduced as an analog for matrices of the relational algebra on relations [8]. Intuitively, queries in MATLANG are built-up by composing several linear algebra operations. The language MATLANG was shown to be subsumed by aggregate logic with only three non-numerical variables. Conversely, MATLANG can express all queries from graph databases to binary relations that can be expressed in first-order logic with three variables. The four-variable query asking if the graph contains a four-clique, however, is not expressible [8].

In this paper, we further zoom in on the expressive power of MATLANG on graphs. In particular, we investigate when two graphs are *equivalent* relative to some fragment of MATLANG. These fragments are defined by allowing only certain linear algebra operations in the queries and are denoted by $\mathsf{ML}(\mathscr{L})$, with $\mathscr{L}$ the list of allowed operations. A total of six (sensible) fragments are considered and $\mathsf{ML}(\mathscr{L})$-equivalence of graphs, i.e., their agreement on all sentences in $\mathsf{ML}(\mathscr{L})$ is characterised. Our results are as follows.

- For starters, we have the fragment $\mathsf{ML}(\cdot,\mathsf{tr})$ that allows for matrix multiplication $(\cdot)$ and trace (tr) computation (i.e., taking the sum of diagonal elements of a matrix). Equivalence of graphs relative to $\mathsf{ML}(\cdot,\mathsf{tr})$ coincides with being co-spectral, or equivalently, to having the same number of closed walks of any length (Section 5);

- Another fragment, $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$, allows for matrix multiplication, conjugate transposition $({}^*)$ and the introduction of the vector $\mathbb{1}$, consisting of all ones. Here, equivalence coincides with having the same number of (not necessarily closed) walks of any length (Section 6);

- When allowing both tr and $\mathbb{1}$, equivalence relative to $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$ coincides, not surprisingly, to having the same number of closed and non-closed walks of any length (Section 6);

- More interesting is the fragment $\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag})$, which also allows for the operation $\mathsf{diag}(\cdot)$ that turns a vector into a diagonal matrix with that vector on its diagonal. In this case, equivalence coincides with having a so-called common equitable partition, or equivalently, to C^2-equivalence. Here, C^2 denotes the two-variable fragment of C, the extension of first-order logic with counting (Section 7);

- The combination of tr with diag results in a stronger notion of equivalence: Graphs are equivalent relative to $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$ when they are C^2-equivalent *and* co-spectral (Section 7);

- Finally, equivalence relative to MATLANG is shown to correspond to C^3-equivalence, the three-variable fragment of C (Section 8). This is in agreement with the results from Brijder et al. [8] mentioned earlier.

We remark that each of these fragments can be extended with addition and scalar multiplication at no increase in distinguishing power. We exhibit examples *separating* all fragments.

The characterisations are shown in a pure algebraic way, without relying on simulations in logic. Underlying are reductions of $\mathsf{ML}(\mathscr{L})$-equivalence of graphs to *similarity notions* of their adjacency matrices. For example, it is known that two graphs G and H are C^2-equivalent if and only if they are fractionally isomorphic [50, 55, 56]. This means that the adjacency matrices A_G of G and A_H of H satisfy $A_G \cdot S = S \cdot A_H$ for some doubly stochastic matrix S. As another example, C^3-equivalence of graphs corresponds to $A_G \cdot O = O \cdot A_H$ for some orthogonal matrix O that is also an isomorphism between the cellular algebras of G and H [19]. We provide similar characterisations for all our matrix query language fragments. It is worth pointing out that beyond MATLANG, C^k-equivalence, for $k \geq 4$, can also be characterised in terms of solutions to linear problems [2, 28, 44].

Moreover, whenever possible, we also provide characterisations in terms of *spectral properties* of graphs. A wealth of results exists in spectral graph theory on what information can be obtained from the adjacency matrix, or from other matrices like the Laplacian, of a graph [9, 15, 25]. We rely quite a bit on known results in that area. Nevertheless, we believe that the connections made in this paper are of interest in their own right. They relate combinatorial and spectral graph invariants by means of query languages. We refer to work by Fürer [23, 24] for more examples of the power of graph invariants and to Dawar et al. [19] for connections between logic, combinatorial and spectral invariants.

Finally, although links to logics such as C^2 and C^3 are made, the connection between MATLANG, rank logics and fixed-point logics with counting, as studied in the context of the descriptive complexity of linear algebra [17, 16, 18, 26, 29, 33], is yet to be explored. Similarly for connections to logic-based graph query languages [1, 4].

2 Background

We denote the set of real numbers by $\mathbb{R}$; the set of complex numbers by $\mathbb{C}$. The set of $m \times n$-matrices over the real (resp., complex) numbers is denoted by $\mathbb{R}^{m\times n}$ (resp., $\mathbb{C}^{m\times n}$). Vectors are elements in $\mathbb{R}^{m\times 1}$ (or $\mathbb{C}^{m\times 1}$). The entries of an $m \times n$-matrix A are denoted by A_{ij}, for $i = 1,\dots,m$ and $j = 1,\dots,n$; entries of a vector v are denoted by v_i, for $i = 1,\dots,m$. We often identify $\mathbb{R}^{1\times 1}$ with $\mathbb{R}$; $\mathbb{C}^{1\times 1}$ with $\mathbb{C}$. The following classes of matrices are of interest in this paper: square matrices (elements in $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$), symmetric matrices (such that $A_{ij} = A_{ji}$ for all i and j), *doubly stochastic* matrices ($A_{ij} \in \mathbb{R}$, $A_{ij} \geq 0$, $\sum_{j=1}^{n} A_{ij} = 1$ and $\sum_{i=1}^{m} A_{ij} = 1$ for all i and j), *doubly quasi-stochastic* matrices ($A_{ij} \in \mathbb{R}$, $\sum_{j=1}^{n} A_{ij} = 1$ and $\sum_{i=1}^{m} A_{ij} = 1$ for all i and j), and *orthogonal* matrices ($O \in \mathbb{R}^{n\times n}$, $O^{\mathsf{t}} \cdot O = I = O \cdot O^{\mathsf{t}}$, where O^{t} denotes the transpose of O obtained by switching rows and columns, $\cdot$ denotes matrix multiplication, and I is the identity matrix in $\mathbb{R}^{n\times n}$).

We only need a couple of notions of linear algebra; we refer to the textbook by Axler [3] for more background. An *eigenvalue* of a matrix A is a scalar λ in $\mathbb{C}$ for which there is a non-zero vector v satisfying $A \cdot v = \lambda v$. Such a vector is called an *eigenvector* of A for eigenvalue λ. The *eigenspace* of an eigenvalue is the vector space obtained as the span of a maximal set of linear independent eigenvectors. Here, the *span* of a set of vectors just denotes the set of all linear combinations of vectors in that set. A set of vectors is linear independent if no vector in that set can be written as a linear combination of other vectors. The *dimension* of an eigenspace is the minimal number of eigenvectors that span the eigenspace.

We will only consider undirected graphs without self-loops. Let $G = (V,E)$ be such a graph with vertices $V = \{1,\dots,n\}$ and edges $E \subseteq V \times V$. The *order* of G is simply the number of vertices. Then, the *adjacency matrix* of a graph G of order n, denoted by A_G, is an $n \times n$-matrix whose entries $(A_G)_{ij}$ are set to 1 if and only if $(i,j) \in E$; all other entries are set to 0. It is a symmetric real matrix with zeroes on its diagonal. The *spectrum* of an undirected graph can be represented as

conjugate transposition ($\mathsf{op}(e) = e^*$)		
$e(\nu(X)) = A \in \mathbb{C}^{m\times n}$	$e(\nu(X))^* = A^* \in \mathbb{C}^{n\times m}$	$(A^*)_{ij} = A^*_{ji}$
one-vector ($\mathsf{op}(e) = \mathbb{1}(e)$)		
$e(\nu(X)) = A \in \mathbb{C}^{m\times n}$	$\mathbb{1}(e(\nu(X)) = \mathbb{1} \in \mathbb{C}^{m\times 1}$	$\mathbb{1}_i = 1$
diagonalization of a vector ($\mathsf{op}(e) = \mathsf{diag}(e)$)		
$e(\nu(X)) = A \in \mathbb{C}^{m\times 1}$	$\mathsf{diag}(e(\nu(X)) = \mathsf{diag}(A) \in \mathbb{C}^{m\times m}$	$\mathsf{diag}(A)_{ii} = A_i$, $\mathsf{diag}(A)_{ij} = 0,\ i \neq j$
matrix multiplication ($\mathsf{op}(e_1, e_2) = e_1 \cdot e_2$)		
$e_1(\nu(X)) = A \in \mathbb{C}^{m\times n}$ $e_2(\nu(X)) = B \in \mathbb{C}^{n\times o}$	$e_1(\nu(X)) \cdot e_2(\nu(X)) = C \in \mathbb{C}^{m\times o}$	$C_{ij} = \sum_{k=1}^{m} A_{ik} \times B_{kj}$
matrix addition ($\mathsf{op}(e_1, e_2) = e_1 + e_2$)		
$e_i(\nu(X)) = A^{(i)} \in \mathbb{C}^{m\times n}$	$e_1(\nu)(X) + e_2(\nu(X)) = B \in \mathbb{C}^{m\times n}$	$B_{ij} = A^{(1)}_{ij} + A^{(2)}_{ij}$
scalar multiplication ($\mathsf{op}(e) = c \times \mathsf{e}$, $c \in \mathbb{C}$)		
$e(\nu(X)) = A \in \mathbb{C}^{m\times n}$	$c \times e(\nu(X)) = B \in \mathbb{C}^{m\times n}$	$B_{ij} = c \times A_{ij}$
trace ($\mathsf{op}(e) = \mathsf{tr}(\mathsf{e})$)		
$e(\nu(X)) = A \in \mathbb{C}^{m\times m}$	$\mathsf{tr}(e(\nu(X)) = c \in \mathbb{C}$	$c = \sum_{i=1}^{m} A_{ii}$
$e(\nu(X)) = A \in \mathbb{C}^{m\times 1}$	$\mathsf{tr}(e(\nu(X)) = c \in \mathbb{C}$	$c = \sum_{i=1}^{m} A_i$
pointwise function application ($\mathsf{op}(e_1, \ldots, e_p) = \mathsf{apply}[f](e_1, \ldots, e_p)$), $f : \mathbb{C}^p \to \mathbb{C} \in \Omega$		
$e_i(\nu(X)) = A^{(i)} \in \mathbb{C}^{m\times n}$	$\mathsf{apply}[f]\big(e_1(\nu(X)), \ldots, e_p(\nu(X))\big) = B \in \mathbb{C}^{m\times n}$	$B_{ij} = f(A^{(1)}_{ij}, \ldots, A^{(p)}_{ij})$

Table 1: Linear algebra operations (supported in MATLANG [8]) and their semantics. In the first operation, for $A_{ji} \in \mathbb{C}$, A^*_{ji} denotes complex conjugation. In the last operation, $\Omega = \bigcup_{k>0} \Omega_k$, where Ω_k consists of functions $f : \mathbb{C}^k \to \mathbb{C}$.

$\mathsf{spec}(G) = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_p \\ m_1 & m_2 & \cdots & m_p \end{pmatrix}$, where $\lambda_1 < \lambda_2 < \cdots < \lambda_p$ are the distinct real eigenvalues of the adjacency matrix A_G of G, and where $m_1, m_2, \ldots, m_p$ denote the dimensions of the corresponding eigenspaces. Two graphs are said to be *co-spectral* if they have the same spectrum. We introduce other relevant notions throughout the paper.

3 Matrix query languages

As described in Brijder et al. [8], matrix query languages can be formalised as compositions of linear algebra operations. Intuitively, a linear algebra operation takes a number of matrices as input and returns another matrix. Examples of operations are matrix multiplication, conjugate transposition, computing the trace, just to name a few. By closing such operations under composition "matrix query languages" are formed. More specifically, for linear algebra operations $\mathsf{op}_1, \ldots, \mathsf{op}_k$ the corresponding matrix query language is denoted by $\mathsf{ML}(\mathsf{op}_1, \ldots, \mathsf{op}_k)$ and consists of expressions formed by the following grammar:

$$e := X \mid \mathsf{op}_1\big(e_1, \ldots, e_{p_1}\big) \mid \cdots \mid \mathsf{op}_k\big(e_1, \ldots, e_{p_k}\big),$$

where X denotes a *matrix variable* which serves to indicate the input to expressions and p_i denotes the number of inputs required by operation op_i. We focus on the case when only a single matrix variable X is present; the treatment of multiple variables is left for future work.

The semantics of an expression $e(X)$ in $\mathsf{ML}(\mathsf{op}_1, \ldots, \mathsf{op}_k)$ is defined inductively, relative to an *assignment* ν of X to a matrix $\nu(X) \in \mathbb{C}^{m\times n}$, for some dimensions m and n. We denote by $e\big(\nu(X)\big)$ the result of evaluating $e(X)$ on $\nu(X)$. As expected, we define $\mathsf{op}_i(e_1(X), \ldots, e_{p_i}(X))(\nu(X)) := \mathsf{op}_i\big(e_1(\nu(X)), \ldots, e_{p_i}(\nu(X))\big)$ for linear algebra operation op_i. In Table 1 we list the operations constituting the matrix query language MATLANG, introduced in Brijder et al. [8]. In the table we also show their semantics. We note that restrictions on the dimensions are in place to ensure that operations are well-defined. Using a simple type system one can formalise a notion of well-formed expressions which guarantees that the semantics of such expressions is well-defined [8]. We only consider well-formed expressions from here on.

REMARK 3.1. The list of operations in Table 1 differs slightly from the list presented in Brijder et al. [8]: We explicitly mention scalar multiplication ($\times$) and addition ($+$), and the trace operation (tr), all of which can be expressed in MATLANG. Hence, MATLANG and $\mathsf{ML}(\cdot, {}^*, \mathsf{tr}, \mathbb{1}, \mathsf{diag}, +, \times, \mathsf{apply}[f], f \in \Omega)$ are equivalent.

4 Expressive power

As mentioned in the introduction, we are interested in the expressive power of matrix query languages. In this paper, we consider sentences in these languages. We define an expression $e(X)$ in $\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_k)$ to be a *sentence* if $e(v(X))$ returns a 1×1-matrix for any assignment v of X. We note that the type system of MATLANG allows to check whether an expression in $\mathsf{ML}(\mathscr{L})$ is a sentence (see Brijder et al. [8] for more details). Having defined sentences, a notion of equivalence naturally follows.

DEFINITION 4.1. Two matrices A and B in $\mathbb{C}^{m \times n}$ are said to be $\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})$-*equivalent*, denoted by $A \equiv_{\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})} B$, if and only if $e(A) = e(B)$ for all sentences $e(X)$ in $\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})$.

In other words, equivalent matrices cannot be distinguished by sentences in the matrix query language under consideration. We aim to *characterise* equivalence for various matrix query languages. We will, however, not treat this problem in full generality and instead, to gain intuition, start by considering *adjacency matrices of undirected graphs*.

The corresponding notion of equivalence on graphs is defined, as expected:

DEFINITION 4.2. Two graphs G and H of the same order are said to be $\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})$-*equivalent*, denoted by $G \equiv_{\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})} H$, if and only if their adjacency matrices are $\mathsf{ML}(\mathsf{op}_1, \dots, \mathsf{op}_\mathsf{k})$-equivalent.

In the following sections we consider graph equivalence for various fragments, starting from simple fragments only supporting a couple of operations, up to the full MATLANG matrix query language. Most proofs are deferred to the appendix.

5 Expressive power of the matrix query language $\mathsf{ML}(\cdot, \mathsf{tr})$

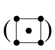
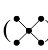
We start with the smallest nontrivial fragment (in terms of number of operations) in which sentences can be expressed: $\mathsf{ML}(\cdot, \mathsf{tr})$. An example sentence in this fragment is $\#\mathsf{cwalk}_k(X) := \mathsf{tr}(X^k)$, where X^k stands for the kth power of X, i.e., X multiplied k times with itself. When evaluated on an adjacency matrix A_G, $\#\mathsf{cwalk}_k(A_G)$ counts the number of *closed walks of length k* in G.

Indeed, the entries of the powers A_G^k of adjacency matrix A_G are known to correspond to the number of walks of length k in G. Recall that a *walk of length k* in a graph $G = (V, E)$ is a sequence $(v_0, v_1, \dots, v_k)$ of vertices of G such that consecutive vertices are adjacent in G, i.e., $(v_{i-1}, v_i) \in E$ for all $i = 1, \dots, k$. Furthermore, a *closed walk* is a walk that starts in and ends at the same vertex. Hence, $\#\mathsf{cwalk}_k(A_G) = \sum_i (A_G^k)_{ii}$ indeed counts closed walks of length k in G.

The following characterisations are known to hold.

PROPOSITION 5.1 ([9, 15]). *Let G and H be two graphs of the same order. The following are equivalent:*

- *G and H have the same total number of closed walks of length k, for all $k \geq 0$;*
- *$\mathsf{tr}(A_G^k) = \mathsf{tr}(A_H^k)$ for all $k \geq 0$;*
- *G and H are co-spectral; and*
- *there exists a real orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$.* □

EXAMPLE 5.2. The graphs G_1 () and H_1 () are the smallest pair (in terms of number of vertices) of non-isomorphic co-spectral graphs [12]. □

A characterisation of $\mathsf{ML}(\cdot, \mathsf{tr})$-equivalence now easily follows:

PROPOSITION 5.3. *For two graphs G and H of the same order, $G \equiv_{\mathsf{ML}(\cdot, \mathsf{tr})} H$ if and only if there exists a real orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ if and only if G and H have the same number of closed walks of any length.*

PROOF. By definition, if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$, then $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$. This holds in particular for the sentences $\#\mathsf{cwalk}_k(X) := \mathsf{tr}(X^k)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$, for $k \geq 1$. That is, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ implies that $\mathsf{tr}(A_G^k) = \mathsf{tr}(A_H^k)$ for all $k \geq 1$. Since G and H are of the same order and $A_G^0 = A_H^0 = I$ (by convention), $\mathsf{tr}(A_G^0) = \mathsf{tr}(A_H^0) = \mathsf{tr}(I) = n$. From the previous proposition it then follows that there exists an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$.

For the converse, assume that $A_G \cdot O = O \cdot A_H$ for some orthogonal matrix O. It can be shown by induction on the structure of expressions in $\mathsf{ML}(\cdot,\mathsf{tr})$ that $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$. The proof uses that O is an orthogonal matrix and that $\mathsf{tr}(P \cdot A \cdot P^{-1}) = \mathsf{tr}(A)$ for any matrix A and any invertible matrix P. □

From an expressiveness point of view, it tells that $\mathsf{ML}(\cdot,\mathsf{tr})$-equivalence of two graphs implies that their adjacency matrices share the same rank, characteristic polynomial, determinant, eigenvalues, and their algebraic multiplicities, geometric multiplicities of eigenvalues, just to name a few.

Given that the trace operation is a linear mapping, i.e., $\mathsf{tr}(c \times A + d \times B) = c \times \mathsf{tr}(A) + d \times \mathsf{tr}(B)$ for matrices A and B and complex numbers c and d, one would expect that matrix addition $(+)$ and scalar multiplication $(\times)$ can be added to $\mathsf{ML}(\cdot,\mathsf{tr})$ without an increase in expressiveness. Indeed, one can rewrite sentences in $\mathsf{ML}(\cdot,\mathsf{tr},+,\times)$ as a linear combination of sentences in $\mathsf{ML}(\cdot,\mathsf{tr})$. Combined with the linearity of $\mathsf{tr}(\cdot)$, Proposition 5.3 can be extended as follows.

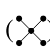
COROLLARY 5.4. *For two graphs G and H of the same order, we have that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times)} H$.* □

We can further strengthen Corollary 5.4 by allowing the application of *any* function $f : \mathbb{C}^p \to \mathbb{C}$ in Ω, *provided* that $\mathsf{apply}[f](e_1,\ldots,e_p)$ is only allowed when each e_i is a sentence. That is, we only allow pointwise function applications on scalars. The restriction of such function applications is denoted by $\mathsf{apply}_{\mathsf{s}}[f]$, for $f \in \Omega$. Indeed, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times)} H$ implies that $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},+,\times)$. Clearly, when $e_i(A_G) = e_i(A_H)$ for all $i = 1,\ldots,p$, $\mathsf{apply}_{\mathsf{s}}[f](e_1(A_G),\ldots,e_p(A_G)) = \mathsf{apply}_{\mathsf{s}}[f](e_1(A_H),\ldots,e_p(A_H))$.

COROLLARY 5.5. *For two graphs G and H of the same order, we have that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times)} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)} H$.* □

Finally, we can also add conjugate transposition $(^*)$ without increasing the expressive power, *provided* that we mildly restrict the class Ω of pointwise functions. More precisely, we assume that Ω is *closed under complex conjugation* in the sense that for every $f \in \Omega_k$ also the functions $\bar{f} : \mathbb{C}^k \to \mathbb{C} : (x_1,\ldots,x_k) \mapsto \overline{f(x_1,\ldots,x_k)}$ and $\underline{f} : \mathbb{C}^k \to \mathbb{C} : (x_1,\ldots,x_k) \mapsto f(\bar{x}_1,\ldots,\bar{x}_k)$ are in Ω_k, where $\bar{\ }$ denotes complex conjugation in $\mathbb{C}$. This assumption, together with standard properties of complex conjugation and conjugate transposition (in particular, $(A \cdot B)^* = B^* \cdot A^*$, $(A^*)^* = A$ and linearity) and using the fact that adjacency matrices of undirected graphs are symmetric, allows to rewrite expressions in $\mathsf{ML}(\cdot,^*,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)$ such that * is only applied on scalars. As a consequence, any expression in $\mathsf{ML}(\cdot,^*,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)$ is equivalent to an expression in $\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)$.

COROLLARY 5.6. *Let Ω be a class of pointwise functions that is closed under complex conjugation. Then, for two graphs G and H of the same order, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)} H$.* □

As a consequence, the graphs G_1 () and H_1 () from Example 5.2 cannot be distinguished by sentences in $\mathsf{ML}(\cdot,^*,\mathsf{tr},+,\times,\mathsf{apply}_{\mathsf{s}}[f],f\in\Omega)$.

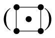
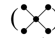
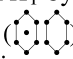
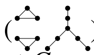
As we will see later, including any other operation from Table 1, such as $\mathbb{1}(\cdot)$, $\mathsf{diag}(\cdot)$ or pointwise function applications on vector or matrices, requires additional constraints on the orthogonal matrix O linking A_G with A_H.

6 The impact of the $\mathbb{1}(\cdot)$ operation

The $\mathbb{1}(\cdot)$ operation, which returns the all-ones vector $\mathbb{1}$, allows to extract other information from graphs than just the number of closed walks. Indeed, consider the sentences

$$\#\mathsf{walk}_k(X) := (\mathbb{1}(X))^* \cdot X^k \cdot \mathbb{1}(X) \text{ and } \#\mathsf{walk}'_k(X) := \mathsf{tr}(X^k \cdot \mathbb{1}(X)),$$

in $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$ and $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$, respectively. When applied on adjacency matrix A_G of a graph G, $\#\mathsf{walk}_k(A_G)$ (and also $\#\mathsf{walk}'_k(A_G)$) returns the *number of (not necessarily closed) walks* in G of length k. In relation to the previous section, co-spectral graphs do not necessarily have the same number of walks of any length. Similarly, graphs with the same number of walks of any length are not necessarily co-spectral.

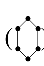
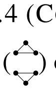
EXAMPLE 6.1. It can be verified that the co-spectral graphs G_1 () and H_1 () of Example 5.2 have 16 versus 20 walks of length 2, respectively. As a consequence, $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$ and $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$ can distinguish G_1 from H_1 by means of the sentences $\#\mathsf{walk}_2(X)$ and $\#\mathsf{walk}'_2(X)$, respectively. By contrast, the graphs G_2 () and H_2 () are not co-spectral, yet have the same number of walks of any length. It is easy to see that G_2 and H_2 are not co-spectral (apart from verifying that their spectra are different): H_2 has 12 closed walks of length 3 (because of the triangles), whereas G_2 has none. We argue below why they have the same number of walks. As a consequence, $\mathsf{ML}(\cdot,\mathsf{tr})$ (and thus also $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$) can distinguish G_2 and H_2. It follows from Proposition 6.6 below that these graphs cannot be distinguished by $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$. □

Graphs sharing the same number of walks of any length have been investigated before in spectral graph theory [13, 14, 31, 51]. To state a spectral characterisation, the so-called *main spectrum* of a graph needs to be considered. The main spectrum of a graph is the set of eigenvalues whose eigenspace is not orthogonal to the $\mathbb{1}$ vector. More formally, for an eigenvalue λ and corresponding eigenspace, represented by a matrix V whose columns are eigenvectors of λ that span the eigenspace, the *main angle* β_λ of λ's eigenspace is $\frac{1}{\sqrt{n}}\|V^{\mathsf{t}}\cdot\mathbb{1}\|_2$, where $\|\cdot\|_2$ is the Euclidean norm. Hence, main eigenvalues are those with a non-zero main angle. Two graphs are said to be *co-main* if they have the same set of main eigenvalues and corresponding main angles. Intuitively, the importance of the orthogonal projection on $\mathbb{1}$ stems from the observation that $\#\mathsf{walk}_k(A_G)$ can be expressed as $\sum_i \lambda_i^k \beta_{\lambda_i}^2$ where the λ_i's are eigenvalues of A_G. Clearly, only those eigenvalues λ_i for which $\beta_{\lambda_i} > 0$ matter when computing $\#\mathsf{walk}_k(A_G)$. This results in the following characterisation.

PROPOSITION 6.2 (THEOREM 1.3.5 IN CVETKOVIĆ ET AL. [15]). *Two graphs G and H of the same order are co-main if and only if they have the same total number of walks of length k, for every $k \geq 0$.* □

Furthermore, the following proposition follows implicitly from the proof of Theorem 3 in van Dam et al. [58] (and is also shown in Theorem 1.2 in Dell et al. [20] in the context of distinguishing graphs by means of *homomorphism vectors*).

PROPOSITION 6.3. *Two graphs G and H of the same order have the same total number of walks of length k, for every $k \geq 0$, if and only if there is a doubly quasi-stochastic matrix Q such that $A_G\cdot Q = Q\cdot A_H$, i.e., $Q\cdot\mathbb{1} = \mathbb{1}$ and $Q^{\mathsf{t}}\cdot\mathbb{1} = \mathbb{1}$.*

EXAMPLE 6.4 (CONTINUATION OF EXAMPLE 6.1). Consider the subgraph G_3 () of G_2 and the subgraph H_3 () of H_2. We have that

$$A_{G_3}\cdot Q = \begin{bmatrix} 0&0&1&0&0&1\\ 0&0&0&1&1&0\\ 1&0&0&0&1&0\\ 0&1&0&0&0&1\\ 0&1&1&0&0&0\\ 1&0&0&1&0&0 \end{bmatrix}\cdot\begin{bmatrix} 0&0&\frac12&0&0&\frac12\\ 0&0&0&\frac12&\frac12&0\\ \frac12&0&0&0&\frac12&0\\ 0&\frac12&0&0&0&\frac12\\ 0&\frac12&\frac12&0&0&0\\ \frac12&0&0&\frac12&0&0 \end{bmatrix} = \begin{bmatrix} 0&0&\frac12&0&0&\frac12\\ 0&0&0&\frac12&\frac12&0\\ \frac12&0&0&0&\frac12&0\\ 0&\frac12&0&0&0&\frac12\\ 0&\frac12&\frac12&0&0&0\\ \frac12&0&0&\frac12&0&0 \end{bmatrix}\cdot\begin{bmatrix} 0&0&1&0&0&1\\ 0&0&0&1&1&0\\ 1&0&0&0&1&0\\ 0&1&0&0&0&1\\ 0&1&1&0&0&0\\ 1&0&0&1&0&0 \end{bmatrix} = Q\cdot A_{H_3}$$

and hence by Proposition 6.3, G_3 and H_3 have the same number of walks on any length. □

As it turns out, the value of the sentences $\#\mathsf{walk}_k(A_G)$ mentioned earlier, that count the number of walks of length k in G, fully determine the value of any sentence in $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$.

LEMMA 6.5. *Let G and H be two graphs of the same order. Then, $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathbb{1})} H$ if and only if $\#\mathsf{walk}_k(A_G) = \#\mathsf{walk}_k(A_H)$ for all $k \geq 1$.* □

The proof involves an analysis of expressions in $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$. We may thus conclude from Proposition 6.3 and Lemma 6.5 that:

PROPOSITION 6.6. *For two graphs G and H of the same order, $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ if and only if there exists a doubly quasi-stochastic matrix Q such that $A_G \cdot Q = Q \cdot A_H$ if and only if G and H have the same number of walks of any length.* □

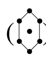
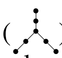
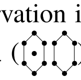
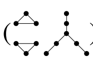
When it comes to $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$, we know from Proposition 5.1 and Theorem 5.3 that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ implies that G and H are co-spectral. Combined with Proposition 6.2 and the fact that the sentence $\#\mathsf{walk}'_k(X)$ count the number of walks of length k, we have that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ implies that G and H are co-spectral and co-main. The following is known about such graphs.

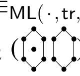
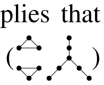
PROPOSITION 6.7 ([37, 58]). *Two co-spectral graphs G and H of the same order are co-main if and only if there exists an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ and $O \cdot \mathbb{1} = \mathbb{1}$.* □

In other words, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ implies the existence of an orthogonal matrix O such that $O \cdot \mathbb{1} = \mathbb{1}$ (i.e., O is also doubly quasi-stochastic) and $A_G \cdot O = O \cdot A_H$. An analysis of expressions in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$ shows that the converse also holds.

PROPOSITION 6.8. *For two graphs G and H of the same order, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ if and only if there exists an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ and $O \cdot \mathbb{1} = \mathbb{1}$ if and only if G and H have the same number of closed walks and the same number of walks of any length.* □

An alternative characterisation (also in van Dam et al. [58]) is that G and H are co-spectral and co-main if and only if both G and H *and* their complement graphs $\bar{G}$ and $\bar{H}$ are co-spectral. Here, the complement graph $\bar{G}$ of G is the graph with adjacency matrix given by $J - A_G - I$, where J is the all ones matrix; similarly for $\bar{H}$.

EXAMPLE 6.9 (CONTINUATION OF EXAMPLE 6.1). Consider the subgraph G_4 () of G_2 and the subgraph H_4 () of H_2. These are known to be the smallest non-isomorphic co-spectral graphs with co-spectral complements [30]. From Proposition 6.8 it then follows that G_4 and H_4 have the same number of walks of any length. Combined with our earlier observation in Example 6.4 that also G_3 and H_3 have this property, we may conclude that $G_2 = G_3 \cup G_4$ () and $H_2 = H_3 \cup H_4$ () have the same number of walks of any length, as anticipated in Example 6.1 □

We remark that as a consequence of Propositions 6.6 and 6.8, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ implies that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$. We already mentioned in Example 6.1 that the graphs G_2 () and H_2 () show that the converse does not hold.

As before, we observe that addition, scalar multiplication, conjugate transposition and pointwise function application on scalars can be included at no increase in expressiveness.

COROLLARY 6.10. *Let G and H be two graphs of the same order. Then,*

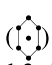
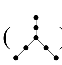
- $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ *if and only if* $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$*; and*
- $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ *if and only if* $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$*,*

where Ω is assumed to be closed under complex conjugation. □

7 The impact of the $\mathsf{diag}(\cdot)$ operation

We next consider the operation $\mathsf{diag}(\cdot)$ which takes a vector as input and returns a diagonal matrix with the input vector on its diagonal. The smallest fragments in which vectors (and sentences) can be defined are $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$ and $\mathsf{ML}(\cdot,^*,\mathbb{1})$. Therefore, in this section we consider equivalence with regards to $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$ and $\mathsf{ML}(\cdot,^*,\mathbb{1},\mathsf{diag})$.

Using $\mathsf{diag}(\cdot)$ we can again extract new information from graphs.

EXAMPLE 7.1. Consider graphs G_4 () and H_4 (). In G_4 we have vertices of degrees 0 and 2, and in H_4 vertices of degrees 1, 2 and 3. We will count the number of vertices of degree 3. To this aim consider the sentence $\#3\mathsf{degr}(X)$ given by

$$\left(\frac{1}{6}\right) \times \mathbb{1}(X)^* \cdot \big(\mathsf{diag}(X\cdot\mathbb{1}(X)) \cdot \mathsf{diag}(X\cdot\mathbb{1}(X) - \mathbb{1}(X)) \cdot \mathsf{diag}(X\cdot\mathbb{1}(X) - 2\times\mathbb{1}(X))\big) \cdot \mathbb{1}(X),$$

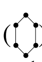
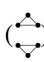
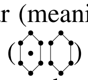
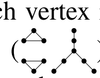
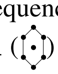
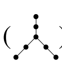
in which we, for convenience, allow addition and scalar multiplications. Each of the subexpressions $\mathsf{diag}(X \cdot \mathbb{1}(X) - d \times \mathbb{1}(X))$, for $d = 0, 1$ and 2, sets the diagonal entry corresponding to vertex v to 0 when v has degree d. By taking the product of these diagonal matrices, entries that are set to 0 will remain zero in the resulting diagonal matrix. This implies that the only non-zero diagonal entries are those corresponding to vertices of degree different from 0, 1 and 2. In other words, only for vertices of degree 3 the diagonal entries carry a non-zero value, i.e., value $3(3-1)(3-2)$. By appropriately rescaling by the factor $\frac{1}{6} = \frac{1}{3(3-1)(3-2)}$, the diagonal entries for the degree three vertices are set to 1, and then summed up. Hence, $\#3\mathsf{degr}(X)$ indeed counts the number vertices of degree three in G_4 and H_4. Since $\#3\mathsf{degr}(A_{G_4}) = [0] \neq [1] = \#3\mathsf{degr}(A_{H_4})$ we can distinguish these graphs. $\square$

The use of the diagonal matrices and their products as in our example sentence $\#3\mathsf{degr}(X)$ can be generalised to obtain information about so-called *iterated degrees* of vertices in graphs, e.g., to identify and/or count vertices that have a number of neighbours each of which have neighbours of specific degrees. Such iterated degree information is closely related to *equitable partitions* of graphs (see e.g., Scheinerman et al. [52]). We phrase our results in terms of such partitions instead of iterated degree sequences.

7.1 Equitable partitions

We show that the presence of $\mathsf{diag}(\cdot)$ allows to formulate a number of expressions, denoted by $\mathsf{eqpart}_i(X)$, for $i = 1, \dots, \ell$, that together extract the *coarsest equitable partition* from a given graph. Equitable partitions naturally arise as the result of the *colour refinement procedure* [5, 27, 59], also known as the 1-dimensional Weisfeiler-Lehman algorithm, used as a subroutine in graph isomorphism solvers. Furthermore, there is a close connection to the study of *fractional isomorphisms* of graphs [52, 55], already mentioned in the introduction. We recall: two graphs G and H are said to be fractional isomorphic if there exists a doubly stochastic matrix S such that $A_G \cdot S = S \cdot A_H$. Furthermore, a logical characterisation of graphs with a common equitable partition exists.

PROPOSITION 7.2 ([55], [35]). *Let G and H be two graphs of the same order. Then, G and H are fractional isomorphic if and only if G and H have a common equitable partition if and only if $G \equiv_{\mathsf{C}^2} H$.* $\square$

EXAMPLE 7.3. The matrix linking the adjacency matrices of G_3 () and H_3 () in Example 6.4 is in fact a doubly stochastic matrix (all its entries are either 0 or $\frac{1}{2}$). Hence, G_3 and H_3 have a common equitable partition, which in this case consists of a single part consisting of all vertices. This generally holds for graphs that are k-regular (meaning, each vertex is adjacent to k vertices) for the same k [50, 52]. By contrast, graphs G_2 () and H_2 () do not have a common equitable partition. Indeed, fractional isomorphic graphs must have the same degree sequence [52], which does not hold for G_2 and H_2. For the same reason, G_1 () and H_1 (), and G_4 () and H_4 () are not fractional isomorphic. $\square$

Formally, an *equitable partition* $\mathscr{V} = \{V_1, \dots, V_\ell\}$ of G is partition of the vertex set of G such that for all $i, j = 1, \dots, \ell$ and $v, v' \in V_i$, $\mathsf{deg}(v, V_j) = \mathsf{deg}(v', V_j)$. Here, $\mathsf{deg}(v, V_j)$ is the number of vertices in V_j that are adjacent to v. In other words, an equitable partition is such that the graph is regular within each part, and is bi-regular between any two different parts. Two graphs G and H are said to have a *common* equitable partition if there exists an equitable partition $\mathscr{V} = \{V_1, \dots, V_\ell\}$ of G and an equitable partition $\mathscr{W} = \{W_1, \dots, W_m\}$ of H such that (a) $\ell = m$; (b) the sizes of the parts agree, i.e., $|V_i| = |W_i|$ for each $i = 1, \dots, \ell$; and (c) $\mathsf{deg}(v, V_j) = \mathsf{deg}(w, W_j)$ for any $v \in V_i$ and $w \in W_i$ and any $i, j = 1, \dots, \ell$. A graph always has an equitable partition: simply treat each vertex as a part by its own. Most interesting is the *coarsest* equitable partition of a graph, i.e., the *unique* equitable partition for which any other equitable partition of the graph is a refinement thereof [52].

In the following, $\mathscr{L}$ can be either $\{\cdot, \mathsf{tr}, \mathbb{1}, \mathsf{diag}\}$ or $\{\cdot, ^*, \mathbb{1}, \mathsf{diag}\}$. Furthermore, we denote by $\mathscr{L}^+$ the extension of $\mathscr{L}$ with linear combinations (i.e., $+$ and $\times$), pointwise function applications on scalars (i.e., $\mathsf{apply_s}[f]$, $f \in \Omega$) and conjugate transposition (*). The corresponding matrix query languages are denoted by $\mathsf{ML}(\mathscr{L})$ and $\mathsf{ML}(\mathscr{L}^+)$, respectively.

We start by reducing the problem of $\mathsf{ML}(\mathscr{L})$-equivalence to $\mathsf{ML}(\mathscr{L}^+)$-equivalence.

LEMMA 7.4. *Let G and H be two graphs of the same order. Then, $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ if and only if $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$.* □

This lemma is verified by showing that expressions in $\mathsf{ML}(\mathscr{L}^+)$ can be seen as linear combinations of expressions in $\mathsf{ML}(\mathscr{L})$, in an analogous way as in the proof of Corollary 6.10. For example, it is clear that $\#3\mathsf{degr}(X)$ can be written as such a linear combination.

We next relate $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ and common equitable partitions of G and H.

PROPOSITION 7.5. *Let G and H be two graphs of the same order. Then, $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that G and H have a common equitable partition.*

PROOF. We show that the algorithm CGCR(A_G), described in Kersting et al. [39], which computes the coarsest equitable partition of a graph can be simulated by expressions in $\mathsf{ML}(\mathscr{L}^+)$. To describe a partition $\mathscr{V} = \{V_1, \dots, V_\ell\}$ of the vertex set of G we use *indicator vectors*. More precisely, we define $\mathbb{1}_{V_i}$ as the $n \times 1$-vector which has a "1" for those entries corresponding to vertices in V_i and has all its other entries set to "0". It is clear that we can also recover partitions from indicator vectors. The simulation of CGCR(A_G) results in a number of expressions, denoted by $\mathsf{eqpart}_i(X)$ for $i = 1, \dots, \ell$, in $\mathsf{ML}(\mathscr{L}^+)$ that *depend on* G and such that the set $\{\mathsf{eqpart}_i(A_G)\}$ consists of indicator vectors of the coarsest equitable partition of G. Since the algorithm CGCR(A_G) is phrased in linear algebra terms [39], its simulation follows easily. Underlying this simulation is the use of products of diagonal matrices as a means of taking conjunctions of indicator vectors, similar to the propagation of zeroes used in $\#3\mathsf{degr}(X)$. Details can be found in the appendix.

The expressions $\mathsf{eqpart}_i(X)$ are constructed based on G. Next, using our assumption $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$, we show that the vectors $\mathsf{eqpart}_i(A_H)$, for $i = 1, \dots, \ell$, also correspond to the coarsest equitable partition of H. This is done in a number of steps:

1. We verify that each $\mathsf{eqpart}_i(A_H)$ is also an indicator vector containing the same number of 1's as $\mathsf{eqpart}_i(A_G)$.

2. We verify that any distinct pair of indicator vectors in $\{\mathsf{eqpart}_i(A_H)\}$ have no common entry holding value "1"; This implies that the set $\{\mathsf{eqpart}_i(A_H)\}$ also represents a *partition*.

3. Finally, we verify that the set $\{\mathsf{eqpart}_i(A_H)\}$ corresponds to an *equitable* partition of H which, together with the partition corresponding to $\{\mathsf{eqpart}_i(A_G)\}$, witnesses that G and H have a *common* equitable partition. Since $\{\mathsf{eqpart}_i(A_G)\}$ is an equitable partition,

$$\mathsf{diag}(\mathsf{eqpart}_i(A_G)) \cdot A_G \cdot \mathsf{diag}(\mathsf{eqpart}_j(A_G)) = \mathsf{deg}(v, V_j) \times \mathsf{diag}(\mathsf{eqpart}_i(A_G)),$$

   for some $v \in V_i$. Here, $\mathscr{V} = \{V_1, \dots, V_\ell\}$ denotes the equitable partition corresponding to indicator vectors $\{\mathsf{eqpart}_i(A_G)\}$. Then, $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that $\mathsf{diag}(\mathsf{eqpart}_i(A_H)) \cdot A_H \cdot \mathsf{diag}(\mathsf{eqpart}_j(A_H))$ is the diagonal matrix $\mathsf{deg}(v, V_j) \times \mathsf{diag}(\mathsf{eqpart}_i(A_H))$. Hence, $\mathsf{deg}(w, W_j) = \mathsf{deg}(w', W_j)$, for any $w, w' \in W_i$, and furthermore, $\mathsf{deg}(v, V_j) = \mathsf{deg}(w, W_j)$. Here, $\mathscr{W} = \{W_1, \dots, W_\ell\}$ denotes the partition corresponding to $\{\mathsf{eqpart}_i(A_H)\}$.

All combined, we may conclude that G and H have indeed a common equitable partition. □

7.2 Characterisations

For $\mathsf{ML}(\cdot, {}^*, \mathbb{1}, \mathsf{diag}, +, \times, \mathsf{apply_s}[f], f \in \Omega)$ we also have the converse.

PROPOSITION 7.6. *Let G and H be two graphs of the same order. If G and H have a common equitable partition, then $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot, {}^*, \mathbb{1}, \mathsf{diag}, +, \times, \mathsf{apply_s}[f], f \in \Omega)$.*

PROOF. Let $\mathscr{V} = \{V_1, \dots, V_\ell\}$ and $\mathscr{W} = \{W_1, \dots, W_\ell\}$ be the common coarsest equitable partitions of G and H, respectively. Denote by $\{\mathbb{1}_{V_i}\}$ and $\{\mathbb{1}_{W_i}\}$, for $i = 1, \dots, \ell$, the corresponding indicator vectors. We know from Proposition 7.2 that there exists a doubly stochastic matrix S such that $A_G \cdot S = S \cdot A_H$. In fact, S can be assumed to have a *block structure* in which the only non-zero blocks are those relating $\mathbb{1}_{V_i}$ and $\mathbb{1}_{W_i}$ [52]. As a consequence, $\mathbb{1}_{V_i} = S \cdot \mathbb{1}_{W_i}$ and $\mathbb{1}_{V_i}^{\mathsf{t}} \cdot S = \mathbb{1}_{W_i}^{\mathsf{t}}$ for $i = 1, \dots, \ell$. The key insight in the proof is that when $e(A_G)$ is an $n \times 1$-vector, it can be written as a linear combination of $\mathbb{1}_{V_i}$'s, say $\sum a_i \times \mathbb{1}_{V_i}$. Moreover, also $e(A_H) = \sum a_i \times \mathbb{1}_{W_i}$. As a consequence,

$e(A_G) = S \cdot e(A_H)$ meaning that $e(A_G)$ is just a permutation of $e(A_H)$. For this to hold, it is essential that we work with equitable partitions common to G and H. For example, if $e(X) := X \cdot \mathbb{1}(X)$ then

$$e(A_G) = A_G \cdot \mathbb{1} = \sum_{i=1}^{\ell} A_G \cdot \mathbb{1}_{V_i} = \sum_{i,j=1}^{\ell} \mathsf{deg}(v_i, V_j) \times \mathbb{1}_{V_i} = \sum_{i,j=1}^{\ell} \mathsf{deg}(w_i, W_j) \times (S \cdot \mathbb{1}_{W_i}) = S \cdot e(A_H),$$

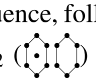
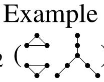
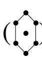
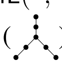
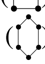
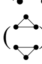
for some $v_i \in V_i$ and $w_i \in W_i$. The challenging case in the proof is when $e(X) := \mathsf{diag}(e'(X))$. Based on the decomposition of $n \times 1$-vectors and the block structure of S, we have

$$\mathsf{diag}(e'(A_G)) \cdot S = \sum_{i=1}^{\ell} a_i \times \mathsf{diag}(\mathbb{1}_{V_i}) \cdot S = \sum_{i=1}^{\ell} a_i \times (S \cdot \mathsf{diag}(\mathbb{1}_{W_i})) = S \cdot \mathsf{diag}(e'(A_H)),$$

which allows to prove that $A_G \cdot S = S \cdot A_H$ implies that $e(A_G) = e(A_H)$ for all sentences in our fragment. □

All combined, we obtain the following characterisation.

THEOREM 7.7. *Let G and H be two graphs of the same order. Then, $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},\mathsf{diag})} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ if and only if there is doubly stochastic matrix S such that $A_G \cdot S = S \cdot A_H$ if and only if $G \equiv_{\mathsf{C}^2} H$.* □

As a consequence, following Example 7.3, sentences in $\mathsf{ML}(\cdot,^*,\mathbb{1},\mathsf{diag})$ can distinguish G_1 () and H_1 (), G_2 () and H_2 (), G_4 () and H_4 (), but cannot distinguish G_3 () and H_3 ().

We next turn our attention to $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$- and $\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f\in\Omega)$-equivalence. Theorem 5.3 implies that G and H are co-spectral and we thus need to combine the existence of a common equitable partition with the existence of an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$. We remark that we cannot simply require O to be doubly stochastic as this would imply that O is a permutation matrix [1], which in turn would imply that G and H are isomorphic, contradicting that our fragments cannot go beyond C^3-equivalence, as we see later.

A characterisation is obtained inspired by a characterisation of simultaneous equivalence of the so-called 1-dimensional Weisfeiler-Lehman closure of adjacency matrices [54]. Let $\mathscr{V} = \{V_1, \ldots, V_\ell\}$ and $\mathscr{W} = \{W_1, \ldots, W_\ell\}$ be common equitable partitions of G and H. Following Thüne [54], we say that an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ is *compatible with $\mathscr{V}$ and $\mathscr{W}$* if O can be block partitioned into ℓ orthogonal matrices O_i of size $|V_i|$ such that $\mathbb{1}_{V_i} = O \cdot \mathbb{1}_{W_i}$, for all $i = 1, \ldots, \ell$. Given this notion, we have the following characterisation.

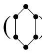
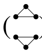
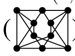
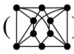
THEOREM 7.8. *Let G and H be graphs of the same order. Then the following holds: $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ if and only if G and H have a common equitable partition, say $\mathscr{V}$ and $\mathscr{W}$, and furthermore $A_G \cdot O = O \cdot A_H$ for some orthogonal matrix O that is compatible with $\mathscr{V}$ and $\mathscr{W}$.*

PROOF. If $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$, then for any k, $\mathsf{tr}(e(A_G)^k) = \mathsf{tr}(e(A_H)^k)$ for any expression $e(X)$ such that $e(A_G)$ (and thus also $e(A_H)$) is an $n \times n$-matrix. As argued in Thüne [54] this implies the existence of a single orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ and $e(A_G) \cdot O = O \cdot e(A_H)$. (The proof relies on Specht's Theorem which relates the existence of an orthogonal matrix *simultaneously* linking sets of matrices to trace equality conditions [36].) In particular, $\mathsf{diag}(\mathsf{eqpart}_i(A_G)) \cdot O = O \cdot \mathsf{diag}(\mathsf{eqpart}_i(A_H))$, for $i = 1, \ldots, \ell$, where $\mathsf{eqpart}_i(X)$ are the expressions computing the equitable partition given in the proof of Proposition 7.5. Lemma 6 in Thüne [54] shows that O must be compatible with the common equitable partitions represented by $\mathsf{eqpart}_i(A_G)$ and $\mathsf{eqpart}_i(A_H)$.

For the converse, we argue as in Proposition 7.6, using orthogonal matrices (which preserve the trace operation) instead of doubly stochastic matrices. □

Note that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$ implies $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},\mathsf{diag})} H$. The converse does not hold.

[1]This is an immediate consequence of the Birkhoff-von Neumann Theorem which states that any doubly stochastic matrix lies in the convex hull of permutation matrices [45].

EXAMPLE 7.9. Consider G_3 () and H_3 (). These graphs are fractional isomorphic but are not co-spectral. Hence, $G_3 \not\equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H_3$ since $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$-equivalence implies co-spectrality. On the other hand, G_5 () and H_5 () are co-spectral regular graphs [57], with co-spectral complements, which cannot be distinguished by $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$. □

A close inspection of the proofs of Proposition 7.6 and Theorem 7.8, shows that $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that for any expression $e(X)$ in $\mathsf{ML}(\mathscr{L}^+)$ such that $e(A_G)$ (and thus also $e(A_H)$) is an $n\times 1$-vector, $e(A_G)$ is a permutation of $e(A_H)$. Indeed, both can be written as linear combinations of indicator vectors, $e(A_G)$ in terms of $\mathbb{1}_{V_i}$'s and $e(A_H)$ in terms of $\mathbb{1}_{W_i}$'s, using the *same* coefficients. This implies that we can allow pointwise function applications *on vectors* and scalars, denoted by $\mathsf{apply}_{\mathsf{v}}[f]$, $f \in \Omega$, at no increase in expressiveness.

COROLLARY 7.10. *Let G and H be two graphs of the same order. We have that $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ if and only if $G \equiv_{\mathsf{ML}(\mathscr{L}^+\cup\{\mathsf{apply}_{\mathsf{v}}[f], f\in\Omega\})} H$.* □

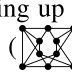
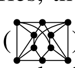
REMARK 7.11. An equitable partition can be defined *without* the $\mathsf{diag}(\cdot)$ operation, provided that function applications on *vectors* are allowed. Hence, the same story holds when first adding pointwise function applications on vectors to $\mathsf{ML}(\cdot,{}^*,\mathbb{1})$ and $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$, rather than first adding $\mathsf{diag}(\cdot)$ like we did in this section.

8 The impact of pointwise functions on matrices

We conclude by considering pointwise function applications on *matrices*, the only operation from Table 1 that we did not consider yet. As we will see shortly, *pointwise multiplication of matrices*, also known as the Schur-Hadamard product, is what results in an increase in expressive power. We denote the Schur-Hadamard product by the binary operator $\circ$, i.e., $(A\circ B)_{ij} = A_{ij}B_{ij}$ for matrices A and B.

EXAMPLE 8.1. We recall that in expression $\#3\mathsf{degr}(X)$ in Example 7.1, products of diagonal matrices resulted in the ability to zoom in on *vertices* that carry specific degree information. When diagonal matrices are concerned, the product of matrices coincides with pointwise multiplication of the *vectors* on the diagonals. Allowing pointwise multiplication on matrices has the same effect, but now on *edges* in graphs. As an example, suppose that we want to count the number of "triangle paths" in G, i.e., paths $(v_0,\ldots,v_k)$ of length k in G such that each edge (v_{i-1},v_i) on the path is part of a triangle. This can be done by expression

$$\#\Delta\mathsf{paths}_k(X) := \mathbb{1}(X)^*\cdot\big((\mathsf{apply}[f_{>0}](X^2\circ X)\big)^k\cdot\mathbb{1}(X),$$

where $f_{>0}(x)=1$ if $x\neq 0$ and $f_{>0}(x)=0$ otherwise [2]. Indeed, when evaluated on adjacency matrix A_G, $A_G^2\circ A_G$ extracts from A_G^2 only those entries corresponding to paths (u,v,w) of length 2 such that (u,w) is an edge as well, i.e., it identifies edges involved in triangles. Then, $\mathsf{apply}[f_{>0}](A_G^2\circ A_G)$ sets all non-zero entries to 1. By considering the kth power of this matrix and summing up all its entries, the number of triangle paths is obtained. It can be verified that for graphs G_5 () and H_5 (), $\#\Delta\mathsf{paths}_2(A_{G_5}) = [160] \neq [132] = \#\Delta\mathsf{paths}_2(A_{H_5})$ and hence, they can be distinguished when the Schur-Hadamard product is available. Recall that all previous fragments could not distinguish between these two graphs. □

In fact, in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$ we can compute the *coarsest stable edge colouring* of a graph $G=(V,E)$ which arises as the result of applying the edge colouring algorithm by Weisfeiler-Lehman [5, 10, 47, 59]. Initially, an edge colouring $\chi_0 : V\times V\to\{0,1,2\}$ is defined such that $\chi_0(v,v)=2$, $\chi_0(v,w)=1$ if $(v,w)\in E$, and $\chi_0(v,w)=0$ for $v\neq w$ and $(v,w)\notin E$. Such a colouring naturally induces a partitioning Π_{χ_0} of $V\times V$. A colouring $\chi : V\times V\to C$ for some set of colours C is called *stable* if and only if for any two pairs (v_1,v_2) and (v_1',v_2') in $V\times V$,

$$\chi(v_1,v_2)=\chi(v_1',v_2') \Leftrightarrow \mathsf{L}^2(v_1,v_2)=\mathsf{L}^2(v_1',v_2'),$$

[2] The use of $\mathsf{apply}[f_{>0}](\cdot)$ is just for convenience and can be simulated when evaluated on given instances using $\cdot$, $+$, $\times$ and $\circ$.

where for a pair $(v,v') \in V \times V$ and pairs (c,d) of colours in C,

$$\mathsf{L}^2(v,v') := \{(c,d,p^{c,d}_{v,v'}) \mid p^{c,d}_{v,v'} \neq 0\} \text{ and } p^{c,d}_{v,v'} := |\{v'' \in V \mid \chi(v,v'') = c, \chi(v'',v') = d\}|.$$

In other words, $\mathsf{L}^2(v,v')$ lists the number of triangles (v,v',v'') in which (v,v'') has colour c and (v'',v) has colour d, for each pair of colours. Such a stable edge colouring χ is called *coarsest* when the corresponding edge partition Π_χ is the coarsest stable edge partition. That is, Π_χ refines Π_{χ_0}, χ is stable and any other colouring satisfying these conditions results in a finer partition than Π_χ.

Two graphs $G=(V,E)$ and $H=(W,F)$ are said to be *indistinguishable by edge colouring*, denoted by $G \equiv_{\mathsf{2WL}} H$, if the following holds. Let $\Pi_{\chi_G} = \{E_1,\dots,E_\ell\}$ and $\Pi_{\chi_H} = \{F_1,\dots,F_\ell\}$ be the edge partitions corresponding to stable edge colourings χ_G and χ_H of H. Then, $G \equiv_{\mathsf{2WL}} H$ if there is a bijection $\iota : \Pi_{\chi_G} \to \Pi_{\chi_H}$ such that E_i and $F_{\iota(i)}$ have the same colour and the same number of entries carrying value 1.

In the seminal paper by Cai, Fürer and Immerman [10], the following was shown.

THEOREM 8.2. *Let G and H be two graphs of the same order. Then, $G \equiv_{\mathsf{2WL}} H$ if and only if $G \equiv_{\mathsf{C}^3} H$.* □

We have the following characterisation of $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$-equivalence.

THEOREM 8.3. *Let G and H be two graphs of the same order, then $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)} H$ if and only if $G \equiv_{\mathsf{C}^3} H$.*

PROOF. We only have space here to sketch the proof. The proof is not that different from the one used in the context of equitable partitions. Let $G=(V,E)$ and $H=(W,F)$ be two graphs. First, we simulate algorithm 2-STAB(A_G) [5], that computes the coarsest stable edge colouring, by expressions $\mathsf{stabcol}_i(X)$, for $i=1,\dots,\ell$, in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$. Each $\mathsf{stabcol}_i(A_G)$ is an *indicator matrix* representing the part of the partition Π of $V \times V$ corresponding to a specific colour. Based on well-known properties of these indicator matrices (they form standard basis of the *cellular* or *coherent algebra* associated with G [32]), we show that $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)} H$ implies that $\{\mathsf{stabcol}_i(A_H)\}$ also represent a partition of $W \times W$ corresponding to the coarsest stable colouring of H. Finally, G and H are shown to be indistinguishable by edge colouring, based on the partitions $\{\mathsf{stabcol}_i(A_G)\}$ and $\{\mathsf{stabcol}_i(A_H)\}$. Hence, by Theorem 8.2, $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)} H$ implies $G \equiv_{\mathsf{C}^3} H$.

For the converse, we use that $G \equiv_{\mathsf{2WL}} H$ implies that there exists an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$ and furthermore, the mapping $Y \mapsto O \cdot Y \cdot O^{\mathsf{t}}$ is an isomorphism between the cellular algebras of G and H. In particular, it commutes with the Schur-Hadamard product [22]. This is crucial to show that $e(A_G) = e(A_H)$ for all sentences $e(X) \in \mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$. More details can be found in the appendix. □

REMARK 8.4. We can do some simplification in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$. Indeed, the trace operator can be simulated by $\mathsf{tr}(e(X)) = \mathbb{1}(X)^* \cdot (e(X) \circ \mathsf{diag}(\mathbb{1}(X))) \cdot \mathbb{1}(X)$ and can hence be omitted. Moreover, $\mathsf{diag}(\cdot)$ can be replaced by a simpler operator, denoted by Id, which returns the identity matrix of the same dimensions as the input. Indeed, $\mathsf{diag}(e(X)) = (e(X) \cdot \mathbb{1}(X)^*) \circ \mathsf{Id}(X)$. We can thus work with $\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{Id},+,\times,\circ)$ instead.

REMARK 8.5. Similar to Corollary 7.10, we can allow *any* pointwise function application on matrices. This follows from the proof of Theorem 8.3 in which it is shown that for expressions $e_i(X)$, for $i=1,\dots,p$, such that each $e_i(A_G)$ (and thus also each $e_i(A_H)$) is an $n \times n$-matrix, $e_i(A_G) = \sum a_j^{(i)} \times \mathsf{stabcol}_j(A_G)$ and $e(A_H) = \sum a_j^{(i)} \times \mathsf{stabcol}_i(A_H)$, for scalars $a_j^{(i)} \in \mathbb{C}$. This implies that

$$\mathsf{apply}[f]\big(e_1(A_G),\dots,e_p(A_G)\big) = \sum f(a_j^{(1)},\dots,a_j^{(p)}) \times \mathsf{stabcol}_j(A_G),$$

and similarly,

$$\mathsf{apply}[f]\big(e_1(A_H),\dots,e_p(A_H)\big) = \sum f(a_j^{(1)},\dots,a_j^{(p)}) \times \mathsf{stabcol}_j(A_H).$$

As a consequence,

$$\mathsf{apply}[f]\big(e_1(A_G),\dots,e_p(A_G)\big) \cdot O = O \cdot \mathsf{apply}[f]\big(e_1(A_H),\dots,e_p(A_H)\big)$$

for the orthogonal matrix O in the proof of Theorem 8.3. This suffices to show that $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot, ^*, \mathsf{tr}, \mathbb{1}, \mathsf{diag}, +, \times, \mathsf{apply}[f], f \in \Omega)$, or in other words, for any sentence in MATLANG.

REMARK 8.6. The orthogonal matrix O in the proof of Theorem 8.3 can be taken to be compatible with the common equitable partitions of G and H, just as in Theorem 7.8. This follows from the fact that the diagonal indicator matrices $\mathsf{diag}(\mathsf{eqpart}_i(A_G))$ are part of the indicator matrices that constitute the basis of the cellular algebra of G [5].

9 Concluding Remarks

We have characterised $\mathsf{ML}(\mathscr{L})$-equivalence for undirected graphs and clearly identified what additional distinguishing power each of the operations has. That natural characterisations can be obtained once more attests that MATLANG is an adequate matrix language.

We conclude with some avenues for further investigation.

Although some of the results generalise to directed graphs (with asymmetric adjacency matrices), an extension to the case when queries can have multiple inputs seems do-able but challenging. The generalisation beyond graphs, i.e., for arbitrary matrices, is wide open.

Of interest may also be to connect $\mathsf{ML}(\mathscr{L})$-equivalence to fragments of first-order logic (without counting). A possible line of attack could be to work over the boolean semiring instead of over the complex numbers (see Grohe and Otto [28] for a similar approach).

We also note that MATLANG was extended in Brijder et al. [8] with an operator inv that computes the inverse of a matrix, if it exists, and returns the zero matrix otherwise. The extension, MATLANG + inv, was shown to be more expressive than MATLANG. For example, connectedness of graphs can be checked by a single sentence in MATLANG + inv. Of course, we here consider *equivalence* of graphs. Even when considering a "classical" logic like FO^3, the three-variable fragment of first-order logic, $G \equiv_{\mathsf{FO}^3} H$ implies that G is connected if and only if H is connected. Translated to our setting, for any fragment $\mathsf{ML}(\mathscr{L})$ in which $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ implies that the Laplacian $\mathsf{diag}(A_G \cdot \mathbb{1}) - A_G$ of G is co-spectral with the Laplacian of $\mathsf{diag}(A_H \cdot \mathbb{1}) - A_H$ of H, $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ implies that G is connected if and only if H is connected. It even implies that G and H must have the same number of connected components, as this is determined by the multiplicity of the eigenvalue 0 of the Laplacian [9].

Nevertheless, we can also consider equivalence of graphs relative to MATLANG + inv. We observe, however, that the inverse of a matrix can be computed using $+$ and $\times$, by the Cayley-Hamilton Theorem [3], given the coefficients of the characteristic polynomial of the adjacency matrix. These coefficients can be computed using $+$, $\times$ and tr. For fragments supporting $\cdot$, $+$, $\times$ and tr, the operator inv thus does not add distinguishing power. It is unclear what the impact is of inv for smaller fragments such as $\mathsf{ML}(\cdot, \mathbb{1})$ and $\mathsf{ML}(\cdot, ^*, \mathbb{1}, \mathsf{diag})$.

To relate our notion of equivalence more closely to the expressiveness questions studied in Brijder et al. [8], it may be interesting to investigate notions of *locality* of $\mathsf{ML}(\mathscr{L})$ expressions, as this underlies the inexpressibility of connectivity of MATLANG [42]. It would be nice if this can be achieved in purely algebraic terms, without relying on locality notions in logic.

Finally, MATLANG was also extended with an eigen operator which returns a matrix whose columns consist of eigenvectors spanning the eigenspaces [8]. Since the choice of eigenvectors is not unique, this results in a non-deterministic semantics. We leave it for future work to study the equivalence of graphs relative to *deterministic* fragments supporting the eigen operator, i.e., such that the result of expressions does not depend on the eigenvectors returned. As a starting point one could, for example, force determinism by considering a certain answer semantics. That is, if $e(X)$ is an expression using $\mathsf{eigen}(X)$, one can define $\mathsf{cert}(e(A_G)) := \bigcap_V e(A_G, V)$, where V ranges over all bases of the eigenspaces. Distinguishability with regards to such a certain answer semantics demands further investigation.

References

[1] Renzo Angles, Marcelo Arenas, Pablo Barceló, Aidan Hogan, Juan Reutter, and Domagoj Vrgoč. Foundations of modern query languages for graph databases. *ACM Comput. Surv.*,

50(5):68:1–68:40, 2017. http://doi.acm.org/10.1145/3104031.

[2] Albert Atserias and Elitza N. Maneva. Sherali-Adams relaxations and indistinguishability in counting logics. *SIAM J. Comput.*, 42(1):112–137, 2013. https://doi.org/10.1137/120867834.

[3] Sheldon Axler. *Linear Algebra Done Right*. Springer, third edition, 2015. https://doi.org/10.1007/978-3-319-11080-6.

[4] Pablo Barceló. Querying graph databases. In *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS, pages 175–188, 2013. http://doi.acm.org/10.1145/2463664.2465216.

[5] Oliver Bastert. *Stabilization procedures and applications*. PhD thesis, Technical University Munich, Germany, 2001. http://nbn-resolving.de/urn:nbn:de:bvb:91-diss2002070500045.

[6] Matthias Boehm, Michael W. Dusenberry, Deron Eriksson, Alexandre V. Evfimievski, Faraz Makari Manshadi, Niketan Pansare, Berthold Reinwald, Frederick R. Reiss, Prithviraj Sen, Arvind C. Surve, and Shirsih Tatikonda. SystemML: Declarative machine learning on Spark. *Proceedings of the VLDB Endowment*, 9(13):1425–1436, 2016. https://doi.org/10.14778/3007263.3007279.

[7] Matthias Boehm, Berthold Reinwald, Dylan Hutchison, Prithviraj Sen, Alexandre V. Evfimievski, and Niketan Pansare. On optimizing operator fusion plans for large-scale machine learning in SystemML. *Proceedings of the VLDB Endowment*, 11(12):1755–1768, 2018. https://doi.org/10.14778/3229863.3229865.

[8] Robert Brijder, Floris Geerts, Jan Van den Bussche, and Timmy Weerwag. On the expressive power of query languages for matrices. In *21st International Conference on Database Theory*, ICDT, pages 10:1–10:17, 2018. https://doi.org/10.4230/LIPIcs.ICDT.2018.10.

[9] Andries E. Brouwer and Willem H. Haemers. *Spectra of Graphs*. Universitext. Springer, 2012. https://doi.org/10.1007/978-1-4614-1939-6.

[10] Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992. https://doi.org/10.1007/BF01305232.

[11] Lingjiao Chen, Arun Kumar, Jeffrey Naughton, and Jignesh M. Patel. Towards linear algebra over normalized data. *Proceedings of the VLDB Endowment*, 10(11):1214–1225, 2017. https://doi.org/10.14778/3137628.3137633.

[12] Dragoš M. Cvetković. Graphs and their spectra. *Publikacije Elektrotehničkog fakulteta. Serija Matematika i fizika*, 354/356:1–50, 1971. http://www.jstor.org/stable/43667526.

[13] Dragoš M. Cvetković. The main part of the spectrum, divisors and switching of graphs. *Publ. Inst. Math. (Beograd) (N.S.)*, 23(37):31–38, 1978. http://elib.mi.sanu.ac.rs/files/journals/publ/43/6.pdf.

[14] Dragoš M. Cvetković, Peter Rowlinson, and Slobodan Simić. *Eigenspaces of Graphs*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1997. https://doi.org/10.1017/CBO9781139086547.

[15] Dragoš M. Cvetković, Peter Rowlinson, and Slobodan Simić. *An Introduction to the Theory of Graph Spectra*. London Mathematical Society Student Texts. Cambridge University Press, 2009. https://doi.org/10.1017/CBO9780511801518.

[16] Anuj Dawar. On the descriptive complexity of linear algebra. In *Proceedings of the 15th International Workshop on Logic, Language, Information and Computation*, WoLLIC, pages 17–25, 2008. http://dx.doi.org/10.1007/978-3-540-69937-8_2.

[17] Anuj Dawar, Martin Grohe, Bjarki Holm, and Bastian Laubner. Logics with rank operators. In *Proceedings of the 24th Annual IEEE Symposium on Logic In Computer Science*, LICS, pages 113–122, 2009. https://doi.org/10.1109/LICS.2009.24.

[18] Anuj Dawar and Bjarki Holm. Pebble games with algebraic rules. *Fund. Inform.*, 150(3-4):281–316, 2017. https://doi.org/10.3233/FI-2017-1471.

[19] Anuj Dawar, Simone Severini, and Octavio Zapata. Descriptive complexity of graph spectra. In *Proceedings of the 23rd International Workshop on Logic, Language, Information and Computation*, WoLLIC, pages 183–199, 2016. https://doi.org/10.1007/978-3-662-52921-8_12.

[20] Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets Weisfeiler and Lehman. In *45th International Colloquium on Automata, Languages, and Programming*, ICALP, pages 40:1–40:14, 2018. https://doi.org/10.4230/LIPIcs.ICALP.2018.40.

[21] Ahmed Elgohary, Matthias Boehm, Peter J. Haas, Frederick R. Reiss, and Berthold Reinwald. Compressed linear algebra for large-scale machine learning. *The VLDB Journal*, pages 1–26, 2017. https://doi.org/10.1007/s00778-017-0478-1".

[22] Shmuel Friedland. Coherent algebras and the graph isomorphism problem. *Discrete Applied Mathematics*, 25(1):73–98, 1989. https://doi.org/10.1016/0166-218X(89)90047-4.

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

[24] Martin Fürer. On the combinatorial power of the Weisfeiler-Lehman algorithm. In Dimitris Fotakis, Aris Pagourtzis, and Vangelis Th. Paschos, editors, *Algorithms and Complexity*, pages 260–271. Springer, 2017. https://doi.org/10.1007/978-3-319-57586-5_22.

[25] Chris Godsil and Gordon F. Royle. *Algebraic Graph Theory*, volume 207 of *Graduate Texts in Mathematics.* Springer, 2001. https://doi.org/10.1007/978-1-4613-0163-9.

[26] Erich Grädel and Wied Pakusa. Rank logic is dead, long live rank logic! In *24th EACSL Annual Conference on Computer Science Logic*, CSL, pages 390–404, 2015. https://doi.org/10.4230/LIPIcs.CSL.2015.390.

[27] Martin Grohe, Kristian Kersting, Martin Mladenov, and Erkal Selman. Dimension reduction via colour refinement. In *22th Annual European Symposium on Algorithms*, ESA, pages 505–516, 2014. https://doi.org/10.1007/978-3-662-44777-2_42.

[28] Martin Grohe and Martin Otto. Pebble games and linear equations. *The Journal of Symbolic Logic*, 80(3):797–844, 2015. http://doi.org/10.1017/jsl.2015.28.

[29] Martin Grohe and Wied Pakusa. Descriptive complexity of linear equation systems and applications to propositional proof complexity. In *32nd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS, pages 1–12, 2017. https://doi.org/10.1109/LICS.2017.8005081.

[30] Willem H. Haemers and Edward Spence. Enumeration of cospectral graphs. *European J. Combin.*, 25(2):199–211, 2004. https://doi.org/10.1016/S0195-6698(03)00100-8.

[31] Frank Harary and Allen J. Schwenk. The spectral approach to determining the number of walks in a graph. *Pacific J. Math.*, 80(2):443–449, 1979. https://projecteuclid.org:443/euclid.pjm/1102785717.

[32] Donald G. Higman. Coherent algebras. *Linear Algebra and its Applications*, 93:209–239, 1987. https://doi.org/10.1016/S0024-3795(87)90326-0.

[33] Bjarki Holm. *Descriptive Complexity of Linear Algebra*. PhD thesis, University of Cambridge, 2010.

[34] Dylan Hutchison, Bill Howe, and Dan Suciu. LaraDB: A minimalist kernel for linear and relational algebra computation. In *Proceedings of the 4th ACM SIGMOD Workshop on Algorithms and Systems for MapReduce and Beyond*, BeyondMR, pages 2:1–2:10, 2017. http://doi.acm.org/10.1145/3070607.3070608.

[35] Neil Immerman and Eric Lander. Describing graphs: A first-order approach to graph canonization. In Alan L. Selman, editor, *Complexity Theory Retrospective: In Honor of Juris Hartmanis on the Occasion of His Sixtieth Birthday*, pages 59–81. Springer, 1990. https://doi.org/10.1007/978-1-4612-4478-3_5.

[36] Naihuan Jing. Unitary and orthogonal equivalence of sets of matrices. *Linear Algebra and its Applications*, 481:235–242, 2015. https://doi.org/10.1016/j.laa.2015.04.036.

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

[38] Irving Kaplansky. *Linear Algebra and Geometry. A Second Course*. Chelsea Publishing Company, 1974.

[39] Kristian Kersting, Martin Mladenov, Roman Garnett, and Martin Grohe. Power iterated color refinement. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, AAAI'14, pages 1904–1910. AAAI Press, 2014. https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8377.

[40] Andreas Kunft, Alexander Alexandrov, Asterios Katsifodimos, and Volker Markl. Bridging the gap: Towards optimization across linear and relational algebra. In *Proceedings of the 3rd ACM SIGMOD Workshop on Algorithms and Systems for MapReduce and Beyond*, BeyondMR, pages 1:1–1:4, 2016. http://doi.acm.org/10.1145/2926534.2926540.

[41] Andreas Kunft, Asterios Katsifodimos, Sebastian Schelter, Tilmann Rabl, and Volker Markl. Blockjoin: Efficient matrix partitioning through joins. *Proceedings of the VLDB Endowment*, 10(13):2061–2072, 2017. https://doi.org/10.14778/3151106.3151110.

[42] Leonid Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004. https://doi.org/10.1007/978-3-662-07003-1.

[43] Shangyu Luo, Zekai J. Gao, Michael Gubanov, Luis L. Perez, and Christopher Jermaine. Scalable linear algebra on a relational database system. *SIGMOD Rec.*, 47(1):24–31, 2018. http://doi.acm.org/10.1145/3277006.3277013.

[44] Peter N. Malkin. Sherali–Adams relaxations of graph isomorphism polytopes. *Discrete Optimization*, 12:73 – 97, 2014. https://doi.org/10.1016/j.disopt.2014.01.004.

[45] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and its Applications*. Springer, 2011. https://doi.org/10.1007/978-0-387-68276-1.

[46] Hung Q. Ngo, XuanLong Nguyen, Dan Olteanu, and Maximilian Schleich. In-database factorized learning. In *Proceedings of the 11th Alberto Mendelzon International Workshop on Foundations of Data Management and the Web*, AMW, 2017. http://ceur-ws.org/Vol-1912/paper21.pdf.

[47] Martin Otto. *Bounded Variable Logics and Counting: A Study in Finite Models*, volume 9 of *Lecture Notes in Logic*. Cambridge University Press, 2017. https://doi.org/10.1017/9781316716878.

[48] Christopher J Pappacena. An upper bound for the length of a finite-dimensional algebra. *Journal of Algebra*, 197(2):535 – 545, 1997. https://doi.org/10.1006/jabr.1997.7140.

[49] Christian Pech. Coherent algebras. https://doi.org/10.13140/2.1.2856.2248, 2002.

[50] Motakuri V. Ramana, Edward R. Scheinerman, and Daniel Ullman. Fractional isomorphism of graphs. *Discrete Mathematics*, 132(1-3):247–265, 1994. https://doi.org/10.1016/0012-365X(94)90241-0.

[51] Peter Rowlinson. The main eigenvalues of a graph: A survey. *Applicable Analysis and Discrete Mathematics*, 1(2):455–471, 2007. http://www.jstor.org/stable/43666075.

[52] Edward R. Scheinerman and Daniel H. Ullman. *Fractional Graph Theory: a Rational Approach to the Theory of Graphs*. John Wiley & Sons, 1997. https://www.ams.jhu.edu/ers/wp-content/uploads/sites/2/2015/12/fgt.pdf.

[53] Maximilian Schleich, Dan Olteanu, and Radu Ciucanu. Learning linear regression models over factorized joins. In *Proceedings of the 2016 International Conference on Management of Data*, SIGMOD, pages 3–18, 2016. http://doi.acm.org/10.1145/2882903.2882939.

[54] Mario Thüne. *Eigenvalues of matrices and graphs*. PhD thesis, University of Leipzig, 2012.

[55] Gottfried Tinhofer. Graph isomorphism and theorems of Birkhoff type. *Computing*, 36(4):285–300, 1986. https://doi.org/10.1007/BF02240204.

[56] Gottfried Tinhofer. A note on compact graphs. *Discrete Applied Mathematics*, 30(2):253–264, 1991. https://doi.org/10.1016/0166-218X(91)90049-3.

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

[58] Erwin R. van Dam, Willem H. Haemers, and Jack H. Koolen. Cospectral graphs and the generalized adjacency matrix. *Linear Algebra and its Applications*, 423(1):33–41, 2007. https://doi.org/10.1016/j.laa.2006.07.017.

[59] Boris J. Weisfeiler and Andrei A. Lehman. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsiya*, 2(9):12–16, 1968. https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf.

A Simplifications

We start with some general observations about $\mathsf{ML}(\mathscr{L})$-equivalence. These observations are important for many of the proofs of the results presented in the paper.

First, since we work with adjacency matrices that are symmetric (recall, we only consider undirected graphs), we observe that conjugate transposition adds limited expressive power. Indeed, we can safely replace complex conjugation (*) by transposition ($^{\mathsf{t}}$). In fact, the only place where transposition is needed is to create the transpose of the "all ones" vector $\mathbb{1}$. For this purpose, we introduce a new operation, $\mathbb{1}^{\mathsf{t}}(\cdot)$, defined such that it returns the transpose of the operator $\mathbb{1}(\cdot)$.

We explicitly denote *complex conjugation on scalars* by the function $\bar{}: \mathbb{C} \to \mathbb{C}$. We note that this function can be regarded as conjugate transposition when applied to 1×1-matrices. The set Ω of pointwise functions is said to be *closed under complex conjugation* if for any $f:\mathbb{C}^k\to\mathbb{C}$ in Ω, also the function $\bar{f}:\mathbb{C}^k\to\mathbb{C}$, defined as $(x_1,\dots,x_k)\mapsto\overline{f(x_1,\dots,x_k)}$, is in Ω; Furthermore, also the function $\underline{f}:\mathbb{C}^k\to C$, defined as $(x_1,\dots,x_k)\mapsto f(\bar{x}_1,\dots,\bar{x}_k)$, is in Ω. Finally, we denote by $\underline{\bar{f}}:\mathbb{C}^k\to\mathbb{C}$ the function $(x_1,\dots,x_k)\mapsto\overline{f(\bar{x}_1,\dots,\bar{x}_k)}$. Clearly, when Ω is closed under complex conjugations, $\underline{\bar{f}}$ is in Ω as well.

We say that two expressions $e(X)$ and $e'(X)$ in some matrix query language fragments are *equivalent*, denoted by $e(X)\equiv e'(X)$, if $e(A)=e'(A)$ for all (adjacency) matrices A.

LEMMA A.1. *Let Ω be a class of pointwise functions that is closed under complex conjugation. Then, every expression $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ,\mathsf{apply}[f],f\in\Omega)$ is equivalent to an expression $e'(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},+,\times,\circ,\mathsf{apply}[f],f\in\Omega)$.*

PROOF. The proof is by induction on the structure of expressions $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ,\mathsf{apply}[f],f\in\Omega)$.

- (**base case**) $e(X):=X^*$. Clearly, $e(X)\equiv e'(X)$ for $e'(X):=X$. Indeed, for any adjacency matrix A, $e(A)^*=A^*=A=e'(A)$, due to A being symmetric and real.
- (**complex conjugate**) $e(X):=(e_1(X)^*)^*$. Then, $e(X)\equiv e'(X)$ for $e'(X):=e_1(X)$. Indeed, we recall that conjugate transposition is an involution, i.e., $(A^*)^*$ for any matrix A.

- **(multiplication)** $e(X) := (e_1(X)\cdot e_2(X))^*$. Then $e(X) \equiv e'(X)$ for $e'(X) := e_2(X)^* \cdot e_1(X)^*$. Indeed, for any two matrices A and B, $(A\cdot B)^* = B^*\cdot A^*$.
- **(ones vector)** $e(X) := \mathbb{1}(e_1(X))^*$. Then, $e(X) \equiv e'(X)$ for $e'(X) := (\mathbb{1}(e_1(X))^{\mathsf{t}}$.
- **(addition)** $e(X) := (e_1(X)+e_2(X))^*$. Then, $e(X) \equiv e'(X)$ for $e'(X) := e_1(X)^* + e_2(X)^*$. Indeed, for any two matrices A and B, $(A+B)^* = A^* + B^*$.
- **(scalar multiplication)** $e(X) := (c\times e_1(X))^*$. Then $e(X) \equiv e'(X)$ for $e'(X) := \overline{c}\times e_1(X)^*$. Indeed, for scalar c and matrix A, $(c\times A)^* = \overline{c}\times A^*$.
- **(Schur-Hadamard)** $e(X) := (e_1(X)\circ e_2(X))^*$. Then, $e(X) \equiv e'(X)$ where $e'(X) := e_1(X)^* \circ e_2(X)^*$. Indeed, for any two matrices A and B, $((A\circ B)^*)_{ij} = (A_{ji}B_{ji})^* = A^*_{ji}B^*_{ji} = (A^*\circ B^*)_{ij}$.
- **(diagonalisation)** $e(X) := (\mathsf{diag}(e_1(X))^*$. Then $e(X) \equiv e'(X)$ for $e'(X) := \mathsf{diag}((e_1(X)^*)^{\mathsf{t}})$.
- **(pointwise functions)** $e(X) := \big(\mathsf{apply}[f](e_1(X),\dots,e_k(X))\big)^*$. Then, $e(X) \equiv e'(X)$ for expression $e'(X) := \mathsf{apply}[\overline{f}](e_1(X)^*,\dots,e_k(X)^*)$.
- **(trace)** $e(X) := (\mathsf{tr}(e_1(X)))^*$. Clearly, $e(X) \equiv e'(X)$ for $e'(X) := \mathsf{tr}(e_1(X)^*)$.

All combined, this implies that we can push conjugate transpositions in $e(X)$ is equivalent to an expression $e'(X)$ that does not contain conjugate transposition, at the cost of introducting the transpose operation ($^{\mathsf{t}}$). Furthermore, we can assume that transposition only occurs on top of expressions of the form $\mathbb{1}(X)$. Indeed, we can use the same case analysis as above, this time applied on $e'(X)$ and by eliminating transposition rather than conjugate transposition. Complex conjugation and transposition indeed satisfy the same properties as used above (i.e., $(A^{\mathsf{t}})^{\mathsf{t}} = A$, $(A\cdot B)^{\mathsf{t}} = B^{\mathsf{t}}\cdot A^{\mathsf{t}}$, and so on). The only case where transposition cannot be eliminated is when it occurs in the form $(\mathbb{1}(X))^{\mathsf{t}}$, for which we introduced the operation $\mathbb{1}^{\mathsf{t}}(X)$. As a consequence, $e'(X)$ may be assumed to be an expression in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},+,\times,\circ,\mathsf{apply}[f],f\in\Omega)$. □

A second observation is that addition and scalar multiplication do not add expressive power. We leave out function applications for the moment; these will be discussed later on.

LEMMA A.2. *Every expression $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},+,\times,\circ)$ is equivalent to a linear combination of expressions in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},\circ)$. Furthermore, when $e(X) \in \mathsf{ML}(\mathscr{L},+,\times)$ for some $\mathscr{L} \subseteq \{\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},\circ\}$ then $e(X)$ is equivalent to a linear combination of expressions in $\mathsf{ML}(\mathscr{L})$*

PROOF. The lemma is shown by induction on the structure of expressions $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},+,\times,\circ)$.

- **(base case)** $e(X) := X$. Clearly, $e(X)$ is already in the desired form.
- **(multiplication)** $e(X) := e_1(X)\cdot e_2(X)$. By induction, $e_1(X) \equiv \sum_i a_i \times e_1^{(i)}(X)$ and $e_2(X) \equiv \sum_j b_j \times e_2^{(j)}(X)$. Hence, $e(X) \equiv \sum_{i,j}(a_ib_j)\times(e_1^{(i)}(X)\cdot e_2^{(j)}(X))$.
- **(ones vector)** $e(X) := \mathbb{1}(e_1(X))$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and hence, $e(X) \equiv \mathbb{1}(e_1^{(1)}(X))$.
- **(transposed ones vector)** $e(X) := \mathbb{1}^{\mathsf{t}}(e_1(X))$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and hence, $e(X) \equiv \mathbb{1}^{\mathsf{t}}(e_1^{(1)}(X))$.
- **(addition)** $e(X) := e_1(X)+e_2(X)$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and $e_2(X) \equiv \sum_j b_j \times e_2^{(j)}(X)$. Hence, $e(X) \equiv \sum_i a_i\times e_1^{(i)}(X) + \sum_j b_j\times e_2^{(j)}(X)$.
- **(scalar multiplication)** $e(X) := c\times e_1(X)$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and hence, $e(X) \equiv \sum_i (ca_i)\times e_1^{(i)}(X)$.
- **(Schur-Hadamard)** $e(X) := e_1(X)\circ e_2(X)$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and $e_2(X) \equiv \sum_j b_j\times e_2^{(j)}(X)$. Hence, $e(X) \equiv \sum_{i,j}(a_ib_j)\times(e_1^{(i)}(X)\circ e_2^{(j)}(X))$.
- **(diagonalisation)** $e(X) := \mathsf{diag}(e_1(X))$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and hence $e(X) \equiv \sum_i a_i \times \mathsf{diag}(e_1^{(i)}(X))$.
- **(trace)** $e(X) := \mathsf{tr}(e_1(X))$. By induction, $e_1(X) \equiv \sum_i a_i\times e_1^{(i)}(X)$ and hence we have that $e(X) \equiv \sum_i a_i\times\mathsf{tr}(e_1^{(i)}(X))$.

This concludes the proof. □

As an immediate consequence we have the following equivalence.

COROLLARY A.3. *Let $\mathscr{L} \subseteq \{\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag},\circ\}$. Then, $G \equiv_{\mathsf{ML}(\mathscr{L},+,\times)} H$ if and only if $G \equiv_{\mathsf{ML}(\mathscr{L})} H$.*

PROOF. Clearly, $G \equiv_{\mathsf{ML}(\mathscr{L},+,\times)} H$ implies $G \equiv_{\mathsf{ML}(\mathscr{L})} H$. For the converse, let $e(X)$ be a sentence in $\mathsf{ML}(\mathscr{L},+,\times)$. We know from the previous lemma that $e(X) \equiv \sum_i a_i \times e_i(X)$ for sentences $e_i(X)$ in $\mathsf{ML}(\mathscr{L})$. Since $G \equiv_{\mathsf{ML}(\mathscr{L})} H$, we have $e_i(A_G) = e_i(A_H)$. Hence, also $e(A_G) = \sum_i a_i \times e_i(A_G) = \sum_i a_i \times e_i(A_H) = e(A_H)$. □

When pointwise function applications are concerned, these do not add distinguishing power when only applied on sentences. Let $f : \mathbb{C}^k \to \mathbb{C}$ be any function in Ω. We denote by $\mathsf{apply_s}[f](e_1,\dots,e_k)$ the application of f on $e_1(X),\dots,e_k(X)$ when each $e_i(X)$ is a sentence. That is, we only allow pointwise function applications on scalars.

LEMMA A.4. *For any two graphs G and H of the same order, we have that $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ if and only if $G \equiv_{\mathsf{ML}(\mathscr{L}\cup\{\mathsf{apply}_s[f], f\in\Omega\})} H$.*

PROOF. Clearly, $G \equiv_{\mathsf{ML}(\mathscr{L}\cup\{\mathsf{apply}_s[f], f\in\Omega\})} H$ implies $G \equiv_{\mathsf{ML}(\mathscr{L})} H$. For the other direction, let $e(X)$ be a sentence in $\mathsf{ML}(\mathscr{L},\mathsf{apply_s}[f], f\in\Omega)$. We show that $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ implies $e(A_G) = e(A_H)$ by induction on the nesting depth of occurrences of $\mathsf{apply_s}[f]$ in $e(X)$ (we do not formalise the notion of nesting depth; this is defined as one would expect). Clearly, if the nesting depth is zero, $e(X) \in \mathsf{ML}(\mathscr{L})$ and we are done. Otherwise, suppose that $e(X) := \mathsf{apply_s}[f](e_1(X),\dots,e_k(X))$, where $e_i(X)$ are sentences in $\mathsf{ML}(\mathscr{L},\mathsf{apply_s}[f], f\in\Omega)$ in which $\mathsf{apply_s}[f]$ occurs at nesting depth at most ℓ. By induction, $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ implies that $e_i(A_G) = e_i(A_H)$ for all $i = 1,\dots,k$. Hence,

$$e(A_G) = f(e_1(A_G),\dots,e_k(A_G)) = f(e_1(A_H),\dots,e_k(A_H)) = e(A_H).$$

So, the induction hypothesis also holds for sentences of nesting depth $\ell+1$. □

In some cases, $\mathbb{1}^{\mathsf{t}}(\cdot)$ can be completely eliminated. An obvious case is when we consider fragments that do not contain $\mathbb{1}(\cdot)$ (see the proof of Lemma A.1). We also identify the following case.

LEMMA A.5. *We have that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag})} H$ if and only $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},,\mathsf{diag})} H$.*

PROOF. Clearly, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag})} H$ implies $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$. For the converse, it is easily verified by induction on expressions $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag})$ that

- If $e(A_G)$ is an $n\times n$-matrix, then $e(X) \equiv c \times f(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}(X,\mathbb{1})\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g(X)$;
- If $e(A_G)$ is an $n\times 1$-matrix, then $e(X) \equiv c \times f(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}(X)$;
- If $e(A_G)$ is a $1\times n$-matrix, then $e(X) \equiv c \times e_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g(X)$;
- If $e(A_G)$ is a 1×1-matrix, then $e(X) \equiv c \times e_{\mathsf{tr}}(X)$,

where $c\in\mathbb{C}$, $f(X)$ and $g(X)$ are expressions in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$ and $e_{\mathsf{tr}}(X)$ is an expression of the form

$$\prod_{i\in K}\mathsf{tr}(h_i(X)),$$

with $h_i(X)$ expressions in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$, for each $i\in K$. In all cases, $f(X)$, $g(X)$ are optional. Also, in the first case $\mathbb{1}(X)\cdot e_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)$ is optional and so are the expressions $e_{\mathsf{tr}}(X)$ in the other cases. In the case analyses below, we only detail cases in which all these optional parts are included.

(**base case**) $e := X$. We have that $e(X) = X$, which is clearly of the desired form.

(**multiplication**) $e(X) := e_1(X)\cdot e_2(X)$. We distinguish between the following cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- ($\mathbf{n\times n},\mathbf{n\times n}$): $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n\times n$. By induction, $e_1(X) \equiv c_1 \times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)$ and $e_2(X) \equiv c_2\times f_2(X)\cdot\mathbb{1}(X)\cdot e^{(2)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X)$. This implies that
$$e(X) \equiv (c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X)\cdot e^{(2)}_{\mathsf{tr}}(X,\mathbb{1})\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X),$$
which is, because $\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X)$ is equivalent to $e_{\mathsf{tr}}(X) := \mathsf{tr}(g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X))$, equivalent to
$$(c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot e_{\mathsf{tr}}(X)\cdot e^{(2)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X),$$
which is in the desired form.
- ($\mathbf{n\times n},\mathbf{n\times 1}$): $e_1(A_G)$ is of dimension $n\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction, $e_1(X) \equiv c_1\times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)$ and $e_2(X)\equiv c_2\times f_2(X)\cdot\mathbb{1}(X)\cdot e^{(2)}_{\mathsf{tr}}(X)$. Hence,
$$\begin{aligned} e(X) &\equiv (c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X)\cdot e^{(2)}_{\mathsf{tr}}(X)\\ &\equiv (c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e^{(1)}_{\mathsf{tr}}(X)\cdot e_{\mathsf{tr}}(X)\cdot e^{(2)}_{\mathsf{tr}}(X),\end{aligned}$$

where $e_{\mathsf{tr}}(X) := \mathsf{tr}(g_1(X)\cdot f_2(X)\cdot \mathbb{1}(X))$.

- **($\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{n}$)**: $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(X)\equiv c_1\times f_1(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(1)}(X)$ and $e_2(X)\equiv c_2\times e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X)$. Hence,
$$e(X)\equiv (c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X),$$
which is in the desired form.
- **($\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{1}$)**: $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension 1×1. By induction, $e_1(X)\equiv c_1\times f_1(X)\cdot\mathbb{1}\cdot e_{\mathsf{tr}}^{(1)}(X)$ and $e_2(X)\equiv c_2\times e_{\mathsf{tr}}^{(2)}(X)$. Hence,
$$e(X)\equiv (c_1c_2)\times f_1(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}^{(2)}(X),$$
which is already in the desired form.
- **($\mathbf{1}\times\mathbf{n},\mathbf{n}\times\mathbf{n}$)**: $e_1(A_G)$ is of dimension $1\times n$ and $e_2(A_G)$ is of dimension $n\times n$. By induction, $e_1(X)\equiv c_1\times e_{\mathsf{tr}}^{(1)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)$ and $e_2(X)\equiv c_2\times f_2(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X)$. As before, this implies that
$$e(X)\equiv (c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X)$$
which is equivalent to
$$e(X)\equiv (c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X),$$
where $e_{\mathsf{tr}}(X) := \mathsf{tr}(g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X))$.
- **($\mathbf{1}\times\mathbf{n},\mathbf{n}\times\mathbf{1}$)**: $e_1(A_G)$ is of dimension $1\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction, $e_1(X)\equiv c_1\times e_{\mathsf{tr}}^{(1)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)$ and $e_2(X)\equiv c_2\times f_2(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)$. Hence,
$$e(X)\equiv (c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}^{(2)}(X).$$
As before, let $e_{\mathsf{tr}}(X) := \mathsf{tr}(g_1(X)\cdot f_2(X)\cdot\mathbb{1}(X))$. Then, $e(X)$ is equivalent to
$$(c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}(X)\cdot e_{\mathsf{tr}}^{(2)}(X),$$
as desired.
- **($\mathbf{1}\times\mathbf{1},\mathbf{1}\times\mathbf{n}$)**: $e_1(A_G)$ is of dimension 1×1 and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(X)\equiv c_1\times e_{\mathsf{tr}}^{(1)}(X)$ and $e_2(X)\equiv c_2\times e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X)$. Hence,
$$e(X)\equiv (c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_2(X),$$
which is in the desired form.
- **($\mathbf{1}\times\mathbf{1},\mathbf{1}\times\mathbf{1}$)**: $e_1(A)$ and $e_2(A)$ are of dimension 1×1. By induction, $e_1(X)\equiv c_1\times e_{\mathsf{tr}}^{(1)}(X)$ and $e_2(X)\equiv c_2\times e_{\mathsf{tr}}^{(2)}(X)$. Clearly, this implies that $e(X)\equiv (c_1c_2)\times e_{\mathsf{tr}}^{(1)}(X)\cdot e_{\mathsf{tr}}^{(2)}(X)$ which is in the desired form.

(ones vector) $e(X) := \mathbb{1}(e_1(X))$. If $e_1(A_G)$ returns an $n\times n$-matrix or $n\times 1$-vector, then $e(X)$ is equivalent to $\mathbb{1}(X)$; if $e_1(A_G)$ returns a $1\times n$-vector or 1×1-matrix, then $e(X)$ is equivalent to $\mathsf{tr}(\mathbb{1}(e_1(X)))$.

(transposed ones vector) $e(X) := \mathbb{1}^{\mathsf{t}}(e_1(X))$. This is completely analogous to the previous case.

(trace) $e(X) := \mathsf{tr}(e_1(X))$. If $e_1(A_G)$ is a sentence, then $e(X)\equiv e_1(X)$. If $e_1(A_G)$ is an $n\times 1$-vector, by induction $e_1(X)\equiv c\times f_1(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}(X)$. Hence, $e(X)\equiv c\times\mathsf{tr}(f_1(X)\cdot\mathbb{1})\cdot e_{\mathsf{tr}}(X)$, which is the desired form. Finally, when $e_1(A_G)$ is an $n\times n$-matrix, by induction, $e_1(X)\equiv c\times f_1(X)\cdot\mathbb{1}(X)\cdot e_{\mathsf{tr}}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)$. We observe that
$$\begin{aligned}\mathsf{tr}(f_1(X)\cdot\mathbb{1}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)) &\equiv \mathbb{1}^{\mathsf{t}}(X)\cdot g_1(X)\cdot f_1(X)\cdot\mathbb{1}(X)\\ &\equiv \mathsf{tr}(g_1(X)\cdot f_1(X)\cdot\mathbb{1}(X)).\end{aligned}$$
Hence,
$$e(X)\equiv c\times\mathsf{tr}(g_1(X)\cdot f_1(X)\cdot\mathbb{1}(X))\cdot e_{\mathsf{tr}}(X).$$

(diagonalisation) $e(X) := \mathsf{diag}(e_1(X))$. Here, $e_1(X)$ can only be a 1×1-matrix or an $n\times 1$-vector. In both cases, $e_1(X)$ is an expression in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$. Hence, also $e(X)$ is an expression in this fragment. □

We note that we left out the operators Schur-Hadamard product in the previous Lemma. We reconsider the impact of $\mathbb{1}^{\mathsf{t}}(\cdot)$ for fragments containing this operators later.

We also note that when the Schur-Hadamard and $\mathbb{1}^{\mathsf{t}}(\cdot)$ operators are present, we can replace $\mathsf{diag}(\cdot)$ by a simple operation $\mathsf{Id}(\cdot)$ which returns the identity matrix in $\mathbb{R}^{n\times n}$ when given an $n\times n$-matrix as input. Furthermore, the trace operation can be derived from other operations as well.

LEMMA A.6. *Any expression* $e(X) \in \mathsf{ML}(\cdot, \mathsf{tr}, \mathbb{1}, \mathbb{1}^{\mathsf{t}}, \mathsf{diag}, \circ)$ *is equivalent to an expression* $e'(X) \in \mathsf{ML}(\cdot, \mathbb{1}, \mathbb{1}^{\mathsf{t}}, \mathsf{Id}, \circ)$.

PROOF. Clearly, any sub-expression in $e(X)$ of the form $\mathsf{tr}(e_1(X))$ can be replaced by $\mathbb{1}^{\mathsf{t}}(X) \cdot (e_1(X) \circ \mathsf{Id}(X))$ when $e_1(X)$ returns an $n \times n$-matrix. When $e_1(X)$ returns an $n \times 1$-vector, $\mathsf{tr}(e_1(X)) \equiv \mathbb{1}^{\mathsf{t}}(X) \cdot e_1(X)$. Furthermore, any sub-expression in $e(X)$ of the form $\mathsf{diag}(e_1(X))$ can be replaced by the expression $(e_1(X) \cdot \mathbb{1}^{\mathsf{t}}(X)) \circ \mathsf{Id}(X)$. So, indeed, $\mathsf{diag}(\cdot)$ and $\mathsf{tr}(\cdot)$ are not needed. ◻

As a final observation, we note that when considering $\mathsf{ML}(\mathscr{L})$-equivalence, we can arbitrarily permute rows (and their corresponding columns) of the input matrices. This is a consequence of the fact that all linear algebra operations considered are invariant under permutations (i.e., if A_G is a permutation of A_H, then $e(A_G)$ will be a permutation of $e(A_H)$). This property allows to simplify some of the proofs later.

We recall that a permutation matrix is a $0/1$-matrix that has exactly one "1" in each row and column. Permutation matrices are orthogonal, that is $P^{\mathsf{t}} \cdot P = P \cdot P^{\mathsf{t}} = I$. Furthermore, $P \cdot \mathbb{1} = \mathbb{1}$ and $P^{\mathsf{t}} \cdot \mathbb{1} = \mathbb{1}$.

LEMMA A.7. *For any sentence* $e(X)$ *in* $\mathsf{ML}(\cdot, \mathsf{tr}, \mathbb{1}, \mathbb{1}^{\mathsf{t}}, \mathsf{diag}, \circ, \mathsf{apply}[f], f \in \Omega)$ *and matrix* A*, we have that* $e(A) = e(P \cdot A \cdot P^{\mathsf{t}})$ *for any permutation matrix* P.

PROOF. The proof is an easy exercise, by induction on the structure of expressions. In particular, it suffices to verify the following induction hypotheses:

- if $e(A)$ returns an $n \times n$-matrix, then $e(P \cdot A \cdot P^{\mathsf{t}}) = P \cdot e(A) \cdot P^{\mathsf{t}}$;
- if $e(A)$ returns an $n \times 1$-vector, then $e(P \cdot A \cdot P^{\mathsf{t}}) = P \cdot e(A)$;
- if $e(A)$ returns a $1 \times n$-vector, then $e(P \cdot A \cdot P^{\mathsf{t}}) = e(A) \cdot P^{\mathsf{t}}$;
- if $e(A)$ returns a 1×1-matrix, then $e(P \cdot A \cdot P^{\mathsf{t}}) = e(A)$.

for an arbitrary permutation matrix P. ◻

The previous lemma implies that when showing $G \equiv_{\mathsf{ML}(\mathscr{L})} H$, we can reorder G and H arbitrarily.

COROLLARY A.8. *Let* P *and* Q *be two permutation matrices. Let* $e(X)$ *be a sentence in* $\mathsf{ML}(\mathscr{L})$*. Then,* $e(A_G) = e(A_H)$ *if and only if* $e(P \cdot A_G \cdot P^{\mathsf{t}}) = e(Q \cdot A_H \cdot Q^{\mathsf{t}})$.

PROOF. Indeed, from the previous lemma we can infer that $e(A_G) = e(P \cdot A_G \cdot P^{\mathsf{t}})$ and $e(A_H) = e(Q \cdot A_H \cdot Q^{\mathsf{t}})$. The corollary follows immediately. ◻

B Invariance under similarities

We next show that $e(A_G) = e(A_H)$ for sentences $e(X)$ in $\mathsf{ML}(\mathscr{L})$ when $A_G \cdot T = T \cdot A_H$ for some matrix T. As we have seen in the main part of the paper, different matrix query language fragments impose different constraints on the matrix T. We show how these constraints, starting from simple to more complex constraints, ensure equivalence relative to the language considered.

B.1 All fragments

All matrix query languages considered contain multiplication and can use an input variable X. Irregardless of what type of matrix T is used such that $A_G \cdot T = T \cdot A_H$ holds, one can verify that the following induction hypotheses hold in the base case (input variable X) and when expressions are combined using multiplication.

- if $e(A_G)$ returns an $n \times n$-matrix, then $e(A_G) \cdot T = T \cdot e(A_H)$;
- if $e(A_G)$ returns an $n \times 1$-vector, then $e(A_G) = T \cdot e(A_H)$;
- if $e(A_G)$ returns a $1 \times n$-vector, then $e(A_G) \cdot T = e(A_H)$; and finally,
- if $e(A_G)$ returns a 1×1-matrix, then $e(A_G) = e(A_H)$.

(†)

We first verify these hypotheses for the base case.

(**base case**, (†)) $e(X) := X$. Clearly, by assumption $e(A_G) \cdot T = A_G \cdot T = T \cdot A_H = T \cdot e(A_H)$.

We next verify the hypotheses (†) for an expression $e(X)$, assuming that they hold for any subexpression of $e(X)$.

(**multiplication**, (†)) $e(X) := e_1(X) \cdot e_2(X)$. We distinguish between the following cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- $(\mathbf{n}\times\mathbf{n},\mathbf{n}\times\mathbf{n})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n\times n$. By induction, $e_1(A_G)\cdot T = T\cdot e_1(A_H)$ and $e_2(A_G)\cdot T = T\cdot e_2(A_H)$. Hence,
$$e(A_G)\cdot T = e_1(A_G)\cdot e_2(A_G)\cdot T = e_1(A_G)\cdot T\cdot e_2(A_H) = T\cdot e_1(A_H)\cdot e_2(A_H) = T\cdot e(A_H).$$
- $(\mathbf{n}\times\mathbf{n},\mathbf{n}\times\mathbf{1})$: $e_1(A_G)$ is of dimension $n\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction, $e_1(A_G)\cdot T = T\cdot e_1(A_H)$ and $e_2(A_G) = T\cdot e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = e_1(A_G)\cdot T\cdot e_2(A_H) = T\cdot e_1(A_H)\cdot e_2(A_H) = T\cdot e(A_H).$$
- $(\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{n})$: $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(A_G) = T\cdot e_1(A_H)$ and $e_2(A_G)\cdot T = e_2(A_H)$. Hence,
$$e(A_G)\cdot T = e_1(A_G)\cdot e_2(A_G)\cdot T = e_1(A_G)\cdot e_2(A_H) = T\cdot e_1(A_H)\cdot e_2(A_H) = T\cdot e(A_H).$$
- $(\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{1})$: $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension 1×1. By induction, $e_1(A_G) = T\cdot e_1(A_H)$ and $e_2(A_G) = e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = e_1(A_G)\cdot e_2(A_H) = T\cdot e_1(A_H)\cdot e_2(A_H) = T\cdot e(A_H).$$
- $(\mathbf{1}\times\mathbf{n},\mathbf{n}\times\mathbf{n})$: $e_1(A_G)$ is of dimension $1\times n$ and $e_2(A_G)$ is of dimension $n\times n$. By induction, $e_1(A_G)\cdot T = e_1(A_H)$ and $e_2(A_G)\cdot T = T\cdot e_2(A_H)$. Hence,
$$e(A_G)\cdot T = e_1(A_G)\cdot e_2(A_G)\cdot T = e_1(A_H)\cdot T\cdot e_2(A_H) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$
- $(\mathbf{1}\times\mathbf{n},\mathbf{n}\times\mathbf{1})$: $e_1(A_G)$ is of dimension $1\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction, $e_1(A_G)\cdot T = e_1(A_H)$ and $e_2(A_G) = T\cdot e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = e_1(A_G)\cdot T\cdot e_2(A_H) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$
- $(\mathbf{1}\times\mathbf{1},\mathbf{1}\times\mathbf{n})$: $e_1(A_G)$ is of dimension 1×1 and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(A_G) = e_1(A_H)$ and $e_2(A_G)\cdot T = e_2(A_H)$. Hence,
$$e(A_G)\cdot T = e_1(A_G)\cdot e_2(A_G)\cdot T = e_1(A_G)\cdot e_2(A_H) = e_1(A_G)\cdot e_2(A_H) = e(A_H).$$
- $(\mathbf{1}\times\mathbf{1},\mathbf{1}\times\mathbf{1})$: $e_1(A)$ and $e_2(A)$ are of dimension 1×1. By induction, $e_1(A_G) = e_1(A_H)$ and $e_2(A_G) = e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$

B.2 Fragments containing $\mathsf{tr}(\cdot)$

For fragments containing the trace operation, the matrices T are restricted to orthogonal matrices. We verify that the hypotheses (†) still hold in this case. It suffices to verify the hypotheses for sentences $e(X) := \mathsf{tr}(e_1(X))$.

(**trace**, (†)) $e(X) := \mathsf{tr}(e_1(X))$. By induction, $e_1(A_G)\cdot T = T\cdot e_1(A_H)$ in case that $e_1(A_G)$ is an $n\times n$-matrix, and $e_1(A_G) = e_1(A_H)$ in case that $e_1(A_G)$ is a sentence. In the latter case, clearly also $e(A_G) = \mathsf{tr}(e_1(A_G)) = \mathsf{tr}(e_1(A_H)) = e(A_H)$. In the former case, we observe that

$$e(A_G) = \mathsf{tr}(e_1(A_G)) = \mathsf{tr}(T^{\mathsf{t}}\cdot e_1(A_G)\cdot T) = \mathsf{tr}(T^{\mathsf{t}}\cdot T\cdot e_1(A_H)) = \mathsf{tr}(I\cdot e_1(A_H)) = \mathsf{tr}(e_1(A_H)) = e(A_H).$$

We here crucially rely on the fact that T is an orthogonal matrix and thus $T^{\mathsf{t}}\cdot T = I$. In addition, we use that $\mathsf{tr}(P\cdot A\cdot P^{-1}) = \mathsf{tr}(A)$ for any matrix A and any invertible matrix P. We note that orthogonal matrices are invertible. We do not need to consider the case when $e_1(A_G)$ is an $n\times 1$-vector as this case only occurs when fragments also support $\mathbb{1}(\cdot)$ or $\mathbb{1}^{\mathsf{t}}(\cdot)$. Indeed, only these operations may cause the creation of vectors.

B.3 Fragments containing $\mathbb{1}(\cdot)$ and $\mathbb{1}^{\mathsf{t}}(\cdot)$

For fragments containing the $\mathbb{1}(\cdot)$ and $\mathbb{1}^{\mathsf{t}}(\cdot)$, the matrices T are restricted to matrices that satisfy $T\cdot\mathbb{1} = \mathbb{1}$ and $T^{\mathsf{t}}\cdot\mathbb{1} = \mathbb{1}$. We verify that the hypotheses (†) still hold in this case.

(**ones vector**, (†)) $e(X) := \mathbb{1}(e_1(X))$. We distinguish between the following cases, depending on the dimensions of $e_1(A_G)$.

- If $e_1(A_G)$ is an $n\times n$-matrix or $n\times 1$-vector, then $e(A_G) = e(A_H) = \mathbb{1}\in\mathbb{R}^{n\times 1}$. Clearly, $e(A_G) = \mathbb{1} = T\cdot\mathbb{1} = T\cdot e(A_H)$.

- if $e_1(A_G)$ is an $1\times n$-vector or sentence, then $e(A_G)=e(A_H)=[1]$ and thus these agree.

(**transposed ones vector**, (†)) $e(X) := \mathbb{1}^{\mathsf{t}}(e_1(X))$. We distinguish between the following cases, depending on the dimensions of $e_1(A_G)$.

- If $e_1(A_G)$ is an $n\times n$-matrix or $n\times 1$-vector, then $e(A_G)=e(A_H)=\mathbb{1}^{\mathsf{t}}\in\mathbb{R}^{1\times n}$. Clearly, $e(A_G)\cdot T=\mathbb{1}^{\mathsf{t}}\cdot T=(T^{\mathsf{t}}\cdot\mathbb{1})^{\mathsf{t}}=\mathbb{1}^{\mathsf{t}}=e(A_H)$.
- if $e_1(A_G)$ is an $1\times n$-vector or sentence, then $e(A_G)=e(A_H)=[1]$ and thus these agree.

(**trace operations (on vectors)**, (†)) $e(X) := \mathsf{tr}(e_1(X))$. We consider the case when $e_1(A_G)$ is an $n\times 1$-vector (the other cases have been considered before). By induction, $e_1(A_G)=T\cdot e_1(A_H)$. Hence,

$$e(A_G)=\mathsf{tr}(e_1(A_G))=\mathbb{1}^{\mathsf{t}}\cdot e_1(A_G)=\mathbb{1}^{\mathsf{t}}\cdot T\cdot e_1(A_H)=\mathbb{1}^{\mathsf{t}}\cdot e_1(A_H)=\mathsf{tr}(e_1(A_H))=e(A_H).$$

Again, the requirements $T\cdot\mathbb{1}=\mathbb{1}$ and $T^{\mathsf{t}}\cdot\mathbb{1}=\mathbb{1}$ are crucial here.

B.4 Fragments containing $\mathsf{diag}(\cdot)$

For fragments containing the $\mathsf{diag}(\cdot)$ operation, the matrices T are restricted to matrices that are *compatible* with the common coarsest equitable partitions of G and H. Let $\mathscr{V}=\{V_1,\dots,V_p\}$ and $\mathscr{W}=\{W_1,\dots,W_p\}$ be two such partitions in G and H, respectively. We represent these partitions by their corresponding indicator vectors $\mathbb{1}_{V_i}$ and $\mathbb{1}_{W_i}$, for $i=1,\dots,p$. To simplify the proof a bit, we rely on Lemma A.7 to permute A_G and A_H such that we can treat the vectors $\mathbb{1}_{V_i}$ and $\mathbb{1}_{W_i}$, for $i=1,\dots,p$, as the *same*. We henceforth refer to these indicator vectors by $\mathbb{1}_i$, for $i=1,\dots,\ell$. That these vectors indicate an equitable partition of G and H translates into

$$A_G\cdot\mathbb{1}_i=\sum_j \mathsf{deg}(v_j,V_i)\cdot\mathbb{1}_i=\sum_j \mathsf{deg}(w_j,W_i)\cdot\mathbb{1}_i=A_H\cdot\mathbb{1}_i,$$

where v_j is some vertex in V_j; w_j is some vertex in W_j. We let $d_{ji}:=\mathsf{deg}(v_j,V_i)=\mathsf{deg}(w_j,W_i)$ for $i,j=1,\dots,p$. The compatibility of T with the common coarsest equitable partition means that

$$\mathbb{1}_i=T\cdot\mathbb{1}_i \quad\text{and}\quad \mathbb{1}_i=T^{\mathsf{t}}\cdot\mathbb{1}_i,$$

for all $i=1,\dots,p$. Furthermore, compatibility also requires a diagonal block-structure of T, which can be expressed as requiring

$$\mathsf{diag}(\mathbb{1}_i)\cdot T=T\cdot\mathsf{diag}(\mathbb{1}_i),$$

for all $i=1,\dots,p$. We verify that the hypotheses (†) still hold in this case. To handle, however, the case $e(X):=\mathsf{diag}(e_1(X))$, we need some additional induction hypotheses:

- if $e(A_G)$ returns an $n\times n$-matrix, then $e(A_G)\cdot\mathbb{1}_i=\sum a_{ij}\times\mathbb{1}_j=e(A_H)\cdot\mathbb{1}_i$;
- if $e(A_G)$ returns an $n\times n$-matrix, then $\mathbb{1}_i^{\mathsf{t}}\cdot e(A_G)=\sum a_{ij}\times\mathbb{1}_j^{\mathsf{t}}=\mathbb{1}_i^{\mathsf{t}}\cdot e(A_H)$;
- if $e(A_G)$ returns an $n\times 1$-vector, then $e(A_G)=\sum a_i\times\mathbb{1}_i=e(A_H)$; and
- if $e(A_G)$ returns a $1\times n$-vector, then $e(A_G)=\sum a_i\times\mathbb{1}_i^{\mathsf{t}}=e(A_H)$.

(‡)

These hypotheses basically state that vectors (resp., transposed vectors) obtained from A_G and A_H can be written as the *same* linear combination of (resp. transposed) indicator vectors. We first verify the hypotheses (‡) and then show that the hypotheses (†) remain to holds for expressions containing $\mathsf{diag}(\cdot)$. For the hypotheses (‡) we do not need to consider when $e(A_G)$ is a sentence.

(**base cases**, (‡)) We have three base cases (a) $e(X):=X$; (b) $e(X):=\mathbb{1}(X)$; and (c) $e(X):=\mathbb{1}^{\mathsf{t}}(X)$. For case (a), we rely on the fact that $\mathbb{1}_i$, for $i=1,\dots,p$ denote the common equitable partitions $\mathscr{V}$ of G and $\mathscr{W}$ of H. Hence,

$$e(A_G)\cdot\mathbb{1}_j=A_G\cdot\mathbb{1}_j=\sum_{i=1}^{p} d_{ij}=A_H\cdot\mathbb{1}_j=e(A_H)\cdot\mathbb{1}_j,$$

for some $v_i\in V_i$ and $w_i\in W_i$. Since A_G and A_H are symmetric,

$$\mathbb{1}_j^{\mathsf{t}}\cdot e(A_G)=\mathbb{1}_j^{\mathsf{t}}\cdot A_G=(A_G\cdot\mathbb{1}_j)^{\mathsf{t}}=\sum_{i=1}^{p} d_{ij}\times\mathbb{1}_i^{\mathsf{t}}=(A_H\cdot\mathbb{1}_j)^{\mathsf{t}}=\mathbb{1}_j^{\mathsf{t}}\cdot A_H=\mathbb{1}_j^{\mathsf{t}}\cdot e(A_H).$$

For cases (b) and (c) we simply need that all $\mathbb{1}_i$ together, for $i=1,\dots,p$, form a partition, i.e., $\mathbb{1}=\sum_{i=1}^{p}\mathbb{1}_i$ and $\mathbb{1}^{\mathsf{t}}=\sum_{i=1}^{p}\mathbb{1}_i^{\mathsf{t}}$. Hence, $\mathbb{1}(A_G)=\mathbb{1}=\sum_{i=1}^{p}\mathbb{1}_i=\mathbb{1}=\mathbb{1}(A_H)$, and similarly for case (c), but using the transposed indicator vectors instead.

(**multiplication**, (‡)) $e(X):=e_1(X)\cdot e_2(X)$. We distinguish between a number of cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- ($\mathbf{n}\times\mathbf{n},\mathbf{n}\times\mathbf{n}$): $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n\times n$. By induction, $e_1(A_G)\cdot\mathbb{1}_i = \sum a_{ij}\times\mathbb{1}_j = e_1(A_H)\cdot\mathbb{1}_i$ and $e_2(A_G)\cdot\mathbb{1}_i = \sum b_{ij}\times\mathbb{1}_j = e_2(A_H)\cdot\mathbb{1}_i$. Hence,
$$e(A_G)\cdot\mathbb{1}_k = e_1(A_G)\cdot e_2(A_G)\cdot\mathbb{1}_k = \sum_j b_{kj}\times(e_1(A_G)\cdot\mathbb{1}_j) = \sum_{i,j} a_{ji}b_{kj}\times\mathbb{1}_i$$
$$= \sum_j b_{kj}\times(e_1(A_H)\cdot\mathbb{1}_j) = e_1(A_H)\cdot e_2(A_H)\cdot\mathbb{1}_k = e(A_H)\cdot\mathbb{1}_k.$$
In an entirely similar way one can verify that $\mathbb{1}_i^{\mathsf{t}}\cdot e(A_G) = \sum_j a_{ij}\times\mathbb{1}_j^{\mathsf{t}} = \mathbb{1}_i^{\mathsf{t}}\cdot e(A_H)$.
- ($\mathbf{n}\times\mathbf{n},\mathbf{n}\times\mathbf{1}$): $e_1(A_G)$ is of dimension $n\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction, $e_1(A_G)\cdot\mathbb{1}_i = \sum a_{ij}\times\mathbb{1}_j = e_1(A_H)\cdot\mathbb{1}_i$ and $e_2(A_G) = \sum b_i\times\mathbb{1}_i = e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = \sum_j b_j(\times e_1(A_G)\cdot\mathbb{1}_j) = \sum_{i,j} a_{ji}b_j\times\mathbb{1}_i$$
$$= \sum_j b_j\times(e_1(A_H)\cdot\mathbb{1}_j) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$
- ($\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{n}$): $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(A_G) = \sum a_i\times\mathbb{1}_i = e_1(A_H)$ and $e_2(A_G) = \sum b_i\times\mathbb{1}_i^{\mathsf{t}} = e_2(A_H)$. Hence,
$$e(A_G)\cdot\mathbb{1}_k = e_1(A_G)\cdot e_2(A_G)\cdot\mathbb{1}_k = \sum_{i,j} a_ib_j\times(\mathbb{1}_i\cdot\mathbb{1}_j^{\mathsf{t}}\cdot\mathbb{1}_k) = e_1(A_G)\cdot e_2(A_G)\cdot\mathbb{1}_k = e(A_H)\cdot\mathbb{1}_k.$$
We observe that $\mathbb{1}_i\cdot\mathbb{1}_j^{\mathsf{t}}\cdot\mathbb{1}_k = \delta_{jk}\|\mathbb{1}_j\|_1\times\mathbb{1}_i$, for Kronecker delta δ_{ij} and ℓ_1-norm $\|x\|_1 = \sum|x_i|$. Hence,
$$e(A_G)\cdot\mathbb{1}_k = \sum_{i,j} a_ib_j\delta_{jk}\|\mathbb{1}_j\|_1\times\mathbb{1}_i = e(A_H)\cdot\mathbb{1}_k,$$
showing that $e(A_G)$ and $e(A_H)$ can be expressed as the same linear combination.
- ($\mathbf{n}\times\mathbf{1},\mathbf{1}\times\mathbf{1}$): $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension 1×1. By induction, $e_1(A_G) = \sum a_i\times\mathbb{1}_i = e_1(A_H)$ and $e_2(A_G) = e_2(A_H) = [b]$ for some $b\in\mathbb{C}$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = \sum_i a_i\times(\mathbb{1}_i\cdot e_2(A_G)) = \sum_i (a_ib)\times\mathbb{1}_i$$
$$= \sum_i a_i\times(\mathbb{1}_i\cdot e_2(A_H)) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$
- ($\mathbf{1}\times\mathbf{n},\mathbf{n}\times\mathbf{n}$): $e_1(A_G)$ is of dimension $1\times n$ and $e_2(A_G)$ is of dimension $n\times n$. By induction, $e_1(A_G) = \sum a_i\times\mathbb{1}_i^{\mathsf{t}} = e_1(A_H)$ and $\mathbb{1}_i^{\mathsf{t}}\cdot e_2(A_G) = \sum_j b_{ij}\times\mathbb{1}_j^{\mathsf{t}} = \mathbb{1}_i^{\mathsf{t}}\cdot e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = \sum_i a_i\times(\mathbb{1}_i^{\mathsf{t}}\cdot e_2(A_G)) = \sum_{i,j}(a_ib_{ij})\times\mathbb{1}_j^{\mathsf{t}}$$
$$= \sum_i a_i\times(\mathbb{1}_i^{\mathsf{t}}\cdot e_2(A_H)) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$
- ($\mathbf{1}\times\mathbf{1},\mathbf{1}\times\mathbf{n}$): $e_1(A_G)$ is of dimension 1×1 and $e_2(A_G)$ is of dimension $1\times n$. By induction, $e_1(A_G) = e_1(A_H) = [a]$ for some $a\in\mathbb{C}$ and $e_2(A_G) = \sum b_i\times\mathbb{1}_i^{\mathsf{t}} = e_2(A_H)$. Hence,
$$e(A_G) = e_1(A_G)\cdot e_2(A_G) = \sum_i b_i\times(e_1(A_G)\cdot\mathbb{1}_i^{\mathsf{t}}) = \sum_i (ab_i)\times\mathbb{1}_i^{\mathsf{t}}$$
$$= \sum_i b_i\times(e_1(A_H)\cdot\mathbb{1}_i^{\mathsf{t}}) = e_1(A_H)\cdot e_2(A_H) = e(A_H).$$

(**ones vector**, (‡)) $e(X) := \mathbb{1}(e_1(X))$. We only need to consider the case when $e_1(A_G)$ is an $n\times n$-matrix or $n\times 1$-vector. In both cases, it suffices to observe that $\mathbb{1} = \sum\mathbb{1}_i$. Indeed,
$$e(A_G) = \mathbb{1} = \sum\mathbb{1}_i = \mathbb{1} = e(A_H).$$
(**transposed ones vector**, (‡)) $e(X) := \mathbb{1}^{\mathsf{t}}(e_1(X))$. This is analogous to the previous case, except that $\mathbb{1}^{\mathsf{t}} = \sum\mathbb{1}_i^{\mathsf{t}}$ is used instead.

At this point, we have verified the hypotheses (‡) for all cases, except for when $e(X) := \mathsf{diag}(e_1(X))$. We treat this case next.

(**diagonalisation**, (‡)) $e(X) := \mathsf{diag}(e_1(X))$ where $e_1(A_G)$ is an $n\times 1$-vector. By induction, $e_1(A_G) = \sum a_i\times\mathbb{1}_i = e_1(A_H)$. Hence,
$$e(A_G)\cdot\mathbb{1}_j = \sum_i a_i\times(\mathsf{diag}(\mathbb{1}_i)\cdot\mathbb{1}_j) = e(A_H)\cdot\mathbb{1}_j,$$
and furthermore, since $\sum_i a_i\times(\mathsf{diag}(\mathbb{1}_i)\cdot\mathbb{1}_j) = a_j\|\mathbb{1}_j\|_1\times\mathbb{1}_j$,
$$e(A_G)\cdot\mathbb{1}_j = a_j\|\mathbb{1}_j\|_1\times\mathbb{1}_j = e(A_H)\cdot\mathbb{1}_j.$$

In a similar way one can verify that

$$\mathbb{1}_j^{\mathsf{t}} \cdot e(A_G) = a_j \|\mathbb{1}_j\|_1 \times \mathbb{1}_j^{\mathsf{t}} = \mathbb{1}_j^{\mathsf{t}} \cdot e(A_H).$$

So the hypotheses (‡) hold. We recall that the hypotheses (‡) were introduced for showing that the hypotheses (†) still hold in the presence of $\mathsf{diag}(\cdot)$. We next verify that this is indeed the case.

(**diagonalisation**, (†)) Let $e(X) := \mathsf{diag}(e_1(X))$. We distinguish between two cases, depending on the dimension of $e_1(A_G)$. First, if $e_1(A_G)$ is a sentence then we know by induction that $e_1(A_G) = e_2(A_G$. Hence,

$$e(A_G) = \mathsf{diag}(e_1(A_G)) = e_1(A_G) = e_1(A_H) = \mathsf{diag}(e_1(A_H)) = e(A_H).$$

Next, if $e_1(A_G)$ is an $n \times 1$-vector we know, by induction using the hypotheses (‡), that $e_1(A_G) = \sum a_i \times \mathbb{1}_i = e_1(A_H)$. We thus have that

$$e(A_G) = \mathsf{diag}(e_1(A_G)) = \sum_i a_i \times \mathsf{diag}(\mathbb{1}_i) = \mathsf{diag}(e_1(A_H)) = e(A_H).$$

Furthermore, since T is compatible with the equitable partition, $\mathsf{diag}(\mathbb{1}_i) \cdot T = T \cdot \mathsf{diag}(\mathbb{1}_i)$. Hence,

$$e(A_G) \cdot T = \sum_i a_i \times (\mathsf{diag}(\mathbb{1}_i) \cdot T) = \sum_i a_i \times (T \cdot \mathsf{diag}(\mathbb{1}_i)) = T \cdot e(A_H).$$

So also the hypotheses (†) remain to hold in the presence of $\mathsf{diag}(\cdot)$.

B.5 Fragments containing the Schur-Hadamard product (∘)

For fragments containing the ∘ operation, the matrices T are restricted to algebraic isomorphisms of the Weisfeiler-Lehman closures of $\mathsf{WL}(A_G, I, J)$ and $\mathsf{WL}(A_H, I, J)$. In particular, if $\mathscr{E} = \{E_1, \dots, E_p\}$ and $\mathscr{F} = \{F_1, \dots, F_p\}$ be the standard bases of $\mathsf{WL}(A_G, I, J)$ and $\mathsf{WL}(A_H, I, J)$, respectively. That is, both $\mathscr{E}$ and $\mathscr{F}$ consist of pairwise disjoint 0/1-matrices that satisfy [5].

$$J = \sum_{i=1}^{p} E_i$$

$$I = \sum_{i \in K} E_i, \text{for some subset } K \text{ of } \{1, \dots, p\}$$

$$\text{for each } i: \quad E_i)^{\mathsf{t}} = E_j \text{ for some } j$$

$$\text{for any } i, j: \quad E_i \cdot E_j = \sum p_{i,j}^k \times E_k.$$

The constants $p_{i,j}^k$ are called the structure constants are the same for $\mathscr{E}$ and $\mathscr{F}$.

Furthermore, by reordering rows and columns of A_G and A_H, we may assume that E_i and F_i for $i \in K$ are the diagonal matrices in $\mathscr{E}$ and $\mathscr{F}$, respectively, that correspond to the common coarsest equitable partitions of G and H, respectively. We may assume that $E_i = F_i$, for $i \in K$, and denote by $\mathbb{1}_i$, for $i = 1, \dots, q$, the corresponding indicator vectors. That is, $E_i = F_i = \mathsf{diag}(\mathbb{1}_i)$. We require the orthogonal matrix T to satisfy

$$E_i \cdot T = T \cdot F_i \quad \text{for } i = 1, \dots, p.$$

We also recall some properties that of the standard bases [32], there exists a function $\pi : \{1, \dots, r\} \to \{1, \dots, p\} \times \{1, \dots, p\}$ such that

$$E_i \cdot \mathbb{1}_{\pi_2(i)} = \mathbb{1}_{\pi_1(i)}, \text{ and } E_i \cdot \mathbb{1}_j = \mathbb{0} \text{ for } j \neq \pi_2(i).$$

Similarly, $\mathbb{1}_{\pi_1(i)}^{\mathsf{t}} \cdot E_i = \mathbb{1}_{\pi_2(i)}^{\mathsf{t}}$ and $\mathbb{1}_j^{\mathsf{t}} \cdot E_i = \mathbb{0}^{\mathsf{t}}$ for $j \neq \pi_1(i)$, where $\mathbb{0}$ denotes the zero vector in $\mathbb{R}^{n \times 1}$. We have the same properties for the basis elements F_i, using the same function π. That is, $F_i \cdot \mathbb{1}_{\pi_2(i)} = \mathbb{1}_{\pi_1(i)}$ and $F_i \cdot \mathbb{1}_j = \mathbb{0}$ for $j \neq \pi_2(i)$. Furthermore, $\mathbb{1}_{\pi_1(i)}^{\mathsf{t}} \cdot F_i = \mathbb{1}_{\pi_2(i)}^{\mathsf{t}}$ and $\mathbb{1}_j^{\mathsf{t}} \cdot F_i = \mathbb{0}$ for $j \neq \pi_1(i)$. Furthermore, also $E_i \circ E_j = \delta_{ij} \times E_i$ for Kronecker delta δ_{ij}.

Given a matrix T as described above, it can be verified that $A_G \cdot T = T \cdot A_H$; as this basically follows from the fact that A_G and A_H can be written as the same linear combination of basis elements (see below); $J \cdot T = T \cdot J$ since J is the sum over all basis elements (this implies that T is doubly quasstochastic); and since

$$E_i \cdot T = T \cdot F_i \quad \text{for } i \in K$$

for the diagonal matrices, after reordering we may assume T to be block-structured according the equitable partitions induced by these diagonal elements. $\mathbb{1}(\cdot)$, $\mathbb{1}^{\mathsf{t}}$ and $\mathsf{diag}(\cdot)$. For the all ones vector and its transpose, we observe that $T \cdot \mathbb{1} = \mathbb{1}$, by the first condition on T. This suffices, combined with the orthogonality of T, to deal with fragments containing the ones operations. By contrast, for the diagonal operation we introduced the hypotheses (‡). We will need to verify these still holds when $e(X) := e_1(X) \circ e_2(X)$. Similarly, to show that sentences are preserved when $A_G \cdot T = T \cdot A_H$ we also need to verify hypotheses (†). To show these, we introduce yet another induction hypothesis:

- if $e(A_G)$ returns an $n \times n$-matrix, then $e(A_G) = \sum a_i \times E_i$ and $e(A_H) = \sum a_i \times F_i$. (§)

The key observation is that $e(A_G)$ can be represented in terms of the basis elements E_i and $e(A_H)$ in terms of the basis elements F_i, in such a way that the coefficients a_i in the linear combinations are the same. Before showing that hypothesis (§) holds, we verify that this hypothesis indeed suffices to conclude that also the hypotheses (†) and (‡) hold in the presence of the Schur-Hadamard product.

(**Schur-Hadamard, (‡)**) $e(X) := e_1(X) \circ e_2(X)$. We distinguish between a couple of cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- $(\mathbf{n} \times \mathbf{n}, \mathbf{n} \times \mathbf{n})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n \times n$. By induction (using (§), $e_1(A_G) = \sum a_j \times E_j$, $e_1(A_H) = \sum a_j \times F_j$, and $e_2(A_G) = \sum b_j \times E_j$ and $e_2(A_H) = \sum b_j \times F_j$. Hence,

$$\begin{aligned} e(A_G) \cdot \mathbb{1}_k &= (e_1(A_G) \circ e_2(A_G)) \cdot \mathbb{1}_k = (\sum a_i b_j \times (E_i \circ E_j)) \cdot \mathbb{1}_k \\ &= \sum a_i b_i \times (E_i \cdot \mathbb{1}_k) = \sum_{i,k=\pi_2(i)} a_i b_i \times \mathbb{1}_k, \end{aligned}$$

  and similarly,

$$\begin{aligned} e(A_H) \cdot \mathbb{1}_k &= (e_1(A_H) \circ e_2(A_H)) \cdot \mathbb{1}_k = (\sum a_i b_j \times (F_i \circ F_j)) \cdot \mathbb{1}_k \\ &= \sum a_i b_i \times (F_i \cdot \mathbb{1}_k) = \sum_{i,k=\pi_2(i)} a_i b_i \times \mathbb{1}_k. \end{aligned}$$

  We can similarly show that $\mathbb{1}_k^{\mathsf{t}} \cdot e(A_G) = \mathbb{1}_k^{\mathsf{t}} \cdot e(A_H) = \sum c_i \times \mathbb{1}_i^{\mathsf{t}}$ for some coefficients $c_i \in C$.
- $(\mathbf{n} \times \mathbf{1}, \mathbf{n} \times \mathbf{1})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n \times 1$. By induction, $e_1(A_G) = \sum a_j \times \mathbb{1}_j = e_1(A_H)$ and $e_2(A_G) = \sum b_j \times \mathbb{1}_j = e_2(A_H)$. Hence,

$$e(A_G) = e_1(A_G) \circ e_2(A_G) = \sum a_i b_j \times (\mathbb{1}_i \circ \mathbb{1}_j) = \sum a_i b_i \times \mathbb{1}_i$$

  and similarly,

$$e(A_H) = e_1(A_H) \circ e_2(A_H) = \sum a_i b_j \times (\mathbb{1}_i \circ \mathbb{1}_j) = \sum a_i b_i \times \mathbb{1}_i.$$

- $(\mathbf{1} \times \mathbf{n}, \mathbf{1} \times \mathbf{n})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $1 \times n$. This case is completely analogous to the previous one.
- $(\mathbf{1} \times \mathbf{1}, \mathbf{1} \times \mathbf{1})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $1 \times n$. By induction, $e_1(A_G) = e_1(A_H)$ and $e_2(A_G) = e_2(A_H)$. Clearly, $e(A_G) = e(A_H)$.

We may thus conclude that, assuming the validity of hypothesis (§), the hypotheses (‡) hold in the presence of the Schur-Hadamard product. We next show the hypotheses (†) also remain to hold. We only need to verify that the hypotheses hold when $e(X) := e_1(X) \circ e_2(X)$.

(**Schur-Hadamard, (†)**). $e(X) := e_1(X) \circ e_2(X)$. We distinguish between a couple of cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- $(\mathbf{n} \times \mathbf{n}, \mathbf{n} \times \mathbf{n})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n \times n$. By induction (using (§)), $e_1(A_G) = \sum a_j \times E_j$, $e_1(A_H) = \sum a_j \times F_j$, and $e_2(A_G) = \sum b_j \times E_j$ and $e_2(A_H) = \sum b_j \times F_j$. Hence,

$$\begin{aligned} e(A_G) \cdot T &= (e_1(A_G) \circ e_2(A_G)) \cdot T = (\sum a_i b_j \times (E_i \circ E_j)) \cdot T \\ &= (\sum a_i b_i \times E_i) \cdot T = \sum a_i b_i \times (E_i \cdot T) \\ &= \sum a_i b_i \times (T \cdot F_i) = T \cdot (\sum a_i b_j \times (F_i \circ F_j)) = T \cdot (e_1(A_H) \circ e_2(A_H)) = T \cdot e(A_H). \end{aligned}$$

- $(\mathbf{n} \times \mathbf{1}, \mathbf{n} \times \mathbf{1})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n \times 1$. We know from our previous analysis of the Schur-Hadamard product for the hypotheses (‡))

$$e(A_G) = \sum a_i b_i \times \mathbb{1}_i = e(A_H).$$

  We recall that $T \cdot \mathbb{1}_i = \mathbb{1}_i$. Hence, $e(A_G) = T \cdot e(A_H)$.
- $(\mathbf{1} \times \mathbf{n}, \mathbf{1} \times \mathbf{n})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension $1 \times n$. This case is completely analogous to the previous case, but using transposed indicator vectors instead.
- $(\mathbf{1} \times \mathbf{1}, \mathbf{1} \times \mathbf{1})$: $e_1(A_G)$ and $e_2(A_G)$ are of dimension 1×1. By induction, $e_1(A_G) = e_1(A_H)$ and $e_2(A_G) = e_2(A_H)$. Clearly, this implies that $e(A_G) = e(A_H)$.

So, under the assumption that hypothesis (§) holds, we have shown that hypotheses (†) and (‡) still hold. We now finally verify hypothesis (§).

We start by considering the base case.

(**base case,** (§)) $e(X) := X$. We know that $A_G = \sum a_i \times E_i$ and $A_H = \sum b_i \times F_i$. Moreover, by assumption, we have

$$A_G \cdot T = \sum a_i \times (E_i \cdot T) = \sum a_i \times (T \cdot F_i) = T \cdot (\sum a_i \times F_i) = T \cdot A_H.$$

This implies that $T \cdot (\sum a_i \times F_i) = T \cdot (\sum b_i \times F_i)$ and by orthogonality of T, $\sum a_i \times F_i = \sum b_i \times F_i$.

Next, we verify hypothesis (§) for expression $e(X)$, assuming that the hypothesis holds for any sub-expression of $e(X)$. We note that it suffices to consider cases that return a $n \times n$-matrix.

(**multiplication,** (§)) $e(X) := e_1(X) \cdot e_2(X)$. We distinguish between a number of cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$. We only need cases that generate an $n \times n$-matrix.

- ($\mathbf{n \times n, n \times n}$): $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n \times n$. By induction, $e_1(A_G) = \sum a_i \times E_i$ and $e_1(A_H) = \sum a_i \times F_i$, and $e_2(A_G) = \sum b_i \times E_i$ and $e_2(A_H) = \sum b_i \times F_i$. Hence,

$$e(A_G) = e_1(A_G) \cdot e_2(A_G) = \sum_{i,j} a_i b_i \times (E_i \cdot E_j) = \sum_{i,j,k} a_i b_i p_{ij}^k \times E_k$$

  and

$$e(A_H) = e_1(A_H) \cdot e_2(A_H) = \sum_{i,j} a_i b_i \times (F_i \cdot F_j) = \sum_{i,j,k} a_i b_i p_{ij}^k \times F_k.$$

- ($\mathbf{n \times 1, 1 \times n}$): $e_1(A_G)$ is of dimension $n \times 1$ and $e_2(A_G)$ is of dimension $1 \times n$. By induction (using (‡)), $e_1(A_G) = \sum a_i \times \mathbb{1}_i = e_1(A_H)$ and $e_2(A_G) = \sum b_i \times \mathbb{1}_i^{\mathsf{t}} = e_2(A_H)$. Hence,

$$e(A_G) = e_1(A_G) \cdot e_2(A_G) = \sum_{i,j} a_i b_i \times (\mathbb{1}_i \cdot \mathbb{1}_j^{\mathsf{t}}) = \sum_{i,j} a_i b_j \times (E_i \cdot \mathbb{1} \cdot \mathbb{1}^{\mathsf{t}} \cdot E_j)$$
$$= \sum_{i,j,k} a_i b_j \times (E_i \cdot E_k \cdot E_j) = \sum_{i,j,k,\ell,m} a_i b_j p_{ik}^{\ell} p_{\ell j}^{m} \times E_m$$

  and similarly for $e(A_H)$. Here we use that $J = \mathbb{1} \cdot \mathbb{1}^{\mathsf{t}} = \sum E_i = \sum F_i$.

(**Identity,** (§)) $e(X) := \mathsf{Id}(X)$. Clearly,

$$e(A_G) = I = \sum_i E_i = \sum_i F_i = I = e(A_H).$$

(**Schur-Hadamard,** (§)) $e(X) := e_1(X) \circ e_2(X)$. We only need to consider the case that $e_1(A_G)$ and $e_2(A_G)$ are $n \times n$-matrices. By induction, $e_1(A_G) = \sum a_i \times E_i$ and $e_1(A_H) = \sum a_i \times F_i$, and $e_2(A_G) = \sum b_i \times E_i$ and $e_2(A_H) = \sum b_i \times F_i$. Hence,

$$e(A_G) = e_1(A_G) \circ e_2(A_G) = \sum_{i,j} a_i b_i \times (E_i \circ E_j) = \sum_i a_i b_i E_i$$

and

$$e(A_H) = e_1(A_H) \cdot e_2(A_H) = \sum_{i,j} a_i b_i \times (F_i \circ F_j) = \sum_i a_i b_i F_i.$$

We may thus conclude that hypothesis (§) holds.

C Proofs of Section 5

C.1 Proof of Proposition 5.3

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ if and only if $A_G \cdot O = O \cdot A_H$ for an orthogonal matrix O.

$\boxed{\Rightarrow}$ By definition, if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$, then $e(A_G) = e(A_H)$ for any sentence $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$. This holds in particular for the sentences $\#\mathsf{cwalk}_k(X) := \mathsf{tr}(X^k)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$, for $k \geq 1$. That is, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ implies that $\mathsf{tr}(A_G^k) = \mathsf{tr}(A_H^k)$ for all $k \geq 1$. Since G and H are of the same order and $A_G^0 = A_H^0 = I$ (by convention), $\mathsf{tr}(A_G^0) = \mathsf{tr}(A_H^0) = \mathsf{tr}(I) = n$. Hence, $\mathsf{tr}(A_G^k) = \mathsf{tr}(A_H^k)$ for all $k \geq 0$. From Proposition 5.1 it then follows that there exists an orthogonal matrix O such that $A_G \cdot O = O \cdot A_H$.

$\boxed{\Leftarrow}$ For the converse, assume that $A_G \cdot O = O \cdot A_H$ for an orthogonal matrix O. We already showed in Sections B.1 and B.2 that this indeed implies that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr})$.

C.2 Proof of Corollary 5.4

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times)} H$. This is an immediate consequence of the more general Corollary A.3 that states that addition and scalar multiplication do not add distinguishing power.

C.3 Proof of Corollary 5.5

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr})} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$. This is an immediate consequence of Corollary 5.4 and the more general Lemma A.4.

C.4 Proof of Corollary 5.6

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ if and only if $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply_s}[f],f\in\Omega,^*)} H$. Inspecting the proof of Lemma A.1 tells us that, in the absence of $\mathbb{1}(\cdot)$ and $\mathsf{diag}(\cdot)$, conjugate transposition can be completed eliminated, provided that Ω is closed under complex conjugation. hence, any expression in $\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply_s}[f],f\in\Omega,^*)$ is equivalent to an expression in $\mathsf{ML}(\cdot,\mathsf{tr},+,\times,\mathsf{apply_s}[f],$ $f\in\Omega)$.

D Proofs of Section 6

D.1 Proof of Proposition 6.6

We show that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ if and only if $A_G\cdot Q = Q\cdot A_H$ for a doubly quasi-stochastic matrix Q. The proof presented here is different (it is a more direct proof) than the one sketched in the paper.

$\boxed{\Rightarrow}$ Suppose that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$. By definition, this implies that $e(A_G) = e(A_H)$ for any sentence $e(X)\in\mathsf{ML}(\cdot,^*,\mathbb{1})$. In particular, $\#\mathsf{walk}_k(A_G) = \#\mathsf{walk}_k(A_H)$ for all k. This in turn is equivalent, by Proposition 6.3, to the existence of doubly quasi-stochastic matrix Q such that $A_G\cdot Q = Q\cdot A_H$.

$\boxed{\Leftarrow}$ For the converse, assume that $A_G\cdot Q = Q\cdot A_H$ for a doubly quasi-stochastic matrix Q. We already showed in Sections B.1 and B.3 that this indeed implies that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathbb{1})$.

D.2 Proof of Proposition 6.8

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ if and only if $A_G\cdot O = O\cdot A_H$ for a doubly quasi-stochastic orthogonal matrix O.

$\boxed{\Rightarrow}$ Suppose that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$. This implies that $\#\mathsf{cwalk}_k(A_G) = \#\mathsf{cwalk}_k(A_H)$ and $\#\mathsf{walk}'_k(A_G) = \#\mathsf{walk}'_k(A_H)$ for any k. Hence, G and H have the same number of closed walks of any length, and hence are co-spectral by Proposition 5.1. Moreover, G and H also must have the same number of walks of any length, and hence are co-main by Proposition 6.2. From Proposition 6.7 it then follows that there exists an orthogonal matrix O satisfying $O\cdot\mathbb{1} = \mathbb{1}$ and such that $A_G\cdot O = O\cdot A_H$.

$\boxed{\Leftarrow}$ For the converse, assume that $A_G\cdot O = O\cdot A_H$ for a doubly quasi-stochastic orthogonal matrix O. We already showed in Sections B.1, B.2 and B.3 that this indeed implies that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$.

D.3 Proof of Corollary 6.10

We first show that $\mathsf{ML}(\cdot,^*,\mathbb{1})$-equivalence and $\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega)$-equivalence coincide. Clearly, $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$ implies $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$. We thus focus on the other direction.

An immediate consequence of the more general Corollary A.3 is that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ implies that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times)} H$. Furthermore, an immediate consequence of the more general Lemma A.4 is that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times)} H$ implies $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$. In fact, we even know from Lemma A.1 that every such sentence is equivalent to an expression that only uses $\mathbb{1}^{\mathsf{t}}(X)$, provided that Ω is closed under complex conjugation.

We next show that $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$-equivalence and $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega,^*)$-equivalence coincide. Clearly, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega,^*)} H$ implies $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$. We thus focus on the other direction.

Lemma A.1 implies that any sentence $e(X)\in\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},^*)$ is equivalent to an expression $e'(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}})$. Lemma A.5 further implies that $e'(X)$ can be assumed to an expression in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})$. Hence, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1})} H$ implies $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},^*)} H$. An immediate consequence of the more general Corollary A.3 is that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},^*)} H$ implies $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},+,\times,^*)} H$, Furthermore, an immediate consequence of the more general Lemma A.4 is that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},+,\times,^*)} H$, implies $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},+,\times,\mathsf{apply_s}[f],f\in\Omega,^*)} H$.

D.4 Proof of Lemma 6.5

We show that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ if and only if $\#\mathsf{walk}_k(A_G) = \#\mathsf{walk}_k(A_H)$ for all $k \geq 0$. Clearly, if $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ holds, then $G \equiv_{\mathsf{ML}(\cdot,^*,\mathbb{1})} H$ for all $k \geq 0$. After all, the expressions $\#\mathsf{walk}_k(X)$ are sentences in $\mathsf{ML}(\cdot,^*,\mathbb{1})$.

The show the converse, we analyse the structure of expressions $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathbb{1})$. In the following, $w(x,y)$ is a word over variables x and y. We write $w(X,J)$ when every occurrence of x in $w(x,y)$ is replaced by matrix variable X, every occurrence y is replaced by J (the all ones matrix which is a shorthand notation for $\mathbb{1}(X)\cdot\mathbb{1}^{\mathsf{t}}(X)$), and concatenation of variables in $w(x,y)$ is interpreted as matrix multiplication. By Lemma A.1 we may assume that $e(X)$ is an expression in $\mathsf{ML}(\cdot,\mathbb{1},\mathbb{1}^{\mathsf{t}})$.

The following induction hypotheses underly the proof.

- if $e(A_G)$ is an $n\times n$-matrix, then $e(X)\equiv c\times w(X,J)$, for a scalar $c\in\mathbb{C}$ and some word $w(x,y)$;
- if $e(A_G)$ is an $1\times n$-matrix, then
$$e(X)\equiv c\times w(X,J)\cdot\mathbb{1}(X)\cdot\prod_{k\in K}\#\mathsf{walk}_k(X)$$
for scalar $c\in\mathbb{C}$, word $w(x,y)$ and multiset of non-zero natural numbers K;
- similarly, if $e(A_G)$ is a $1\times n$ vector, then
$$e(X)\equiv c\times\prod_{k\in K}\#\mathsf{walk}_k(X)\cdot\mathbb{1}^{\mathsf{t}}(X)\cdot w(X,J),$$
and finally,
- if $e(A_G)$ is a 1×1-matrix, then
$$e(X)\equiv c\times\mathbb{1}^{\mathsf{t}}(X)\cdot w(X,J)\cdot\mathbb{1}(X)\equiv c'\times\prod_{k\in K}\#\mathsf{walk}_k(X),\tag{1}$$
for scalars $c,c'\in\mathbb{C}$, word $w(x,y)$ and multiset K of non-zero natural numbers.

It is the last case that is of interest here. Indeed, it states that any sentence $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathbb{1})$ is equivalent to an expression of the form $e'(X)=c'\times\prod_{k\in K}\#\mathsf{walk}_k(X)$. Hence, when $\#\mathsf{walk}_k(A_G)=\#\mathsf{walk}_k(A_H)$ for any $k>0$, we have that $e(A_G)=e'(A_G)=e'(A_H)=e(A_H)$, as desired.

We start by showing the second equivalence in equation (1). It suffices to observe that $J^\ell=n^{\ell-1}J$. As a consequence,

$$\begin{aligned}
e(X) &\equiv c\times\mathbb{1}^{\mathsf{t}}(X)\cdot w(X,J)\cdot\mathbb{1}(X)\\
&= c\times\mathbb{1}^{\mathsf{t}}(X)\cdot X^{k_1}\cdot J^{\ell_1}\cdot X^{k_2}\cdot J^{\ell_2}\cdots J^{\ell_{p-1}}\cdot X^{k_p}\cdot\mathbb{1}(X)\quad(\text{assuming } w(x,y)=x^{k_1}y^{\ell_1}\cdots y^{\ell_{p-1}}x^{k_p})\\
&= cn^{\ell_1+\cdots+\ell_{p-1}-p}\times\mathbb{1}^{\mathsf{t}}(X)\cdot X^{k_1}\cdot J\cdot X^{k_2}\cdots J\cdot X^{k_p}\cdot\mathbb{1}(X)\\
&= cn^{\ell_1+\cdots+\ell_{p-1}-p}\times\big(\mathbb{1}^{\mathsf{t}}(X)\cdot X^{k_1}\cdot\mathbb{1}(X)\big)\cdot\big(\mathbb{1}^{\mathsf{t}}(X)\cdot X^{k_2}\cdot\mathbb{1}(X)\big)\cdot\cdots\cdot\big(\mathbb{1}^{\mathsf{t}}(X)^*\cdot X^{k_p}\cdot\mathbb{1}(X)\big)\\
&= c'\times\#\mathsf{walk}_{k_1}(X)\cdot\#\mathsf{walk}_{k_2}(X)\cdot\cdots\cdot\#\mathsf{walk}_{k_p}(X),
\end{aligned}$$

for $c'=cn^{\ell_1+\cdots+\ell_{p-1}-p}$.

It remains to verify that hypotheses by induction on the structure of expressions $e(X)$ in $\mathsf{ML}(\cdot,^*,\mathbb{1})$.

(**base case**) $e:=X$. We have that $e(X)=X$, which is clearly of the desired form.

(**multiplication**) $e(X):=e_1(X)\cdot e_2(X)$. We distinguish between the following cases, depending on the dimensions of $e_1(A_G)$ and $e_2(A_G)$.

- ($\mathbf{n\times n,n\times n}$): $e_1(A_G)$ and $e_2(A_G)$ are of dimension $n\times n$. By induction $e_1(X)\equiv c_1\times w_1(X,J)$ and $e_2(X)\equiv c_2\times w_2(X,J)$. Hence, $e(X)\equiv c_1c_2\times w_1(X,J)\cdot w_2(X,J)=c\times w(X,J)$, for $c=c_1c_2$ and $w(x,y)$ the concatenation of $w_1(x,y)$ and $w_2(x,y)$.
- ($\mathbf{n\times n,n\times 1}$): By induction, we have $e_1(X)\equiv c_1\times w_1(X,J)$ and $e_2(X)\equiv c_2\times w_2(X,J)\cdot\mathbb{1}(X)\cdot\prod_{k\in K}\#\mathsf{walk}_k(X)$. Hence,
$$e(X)\equiv c_1c_2\times w_1(X,J)\cdot w_2(X,J)\cdot\mathbb{1}(X)\prod_{k\in K}\#\mathsf{walk}_k(X),$$
which can clearly be written in the form $c\times w(X,J)\cdot\mathbb{1}(X)\cdot\prod_{k\in K}\#\mathsf{walk}_k(X)$, for $c=c_1c_2$ and $w(x,y)=w_1(x,y)w_2(x,y)$.
- ($\mathbf{n\times 1,1\times n}$): $e_1(A_G)$ is of dimension $n\times 1$ and $e_2(A_G)$ is of dimension $1\times n$. $e_1(A_G)$ is of dimension $n\times n$ and $e_2(A_G)$ is of dimension $n\times 1$. By induction we have that $e_1(X)\equiv$

$c_1 \times w_1(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_1} \#\mathsf{walk}_k(X)$ and $e_2(X) \equiv c_2 \times \prod_{k \in K_2} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_2(X,J)$. Hence,

$$e(X) \equiv c_1 c_2 \times w_1(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_2(X,J).$$

Since each $\#\mathsf{walk}_k(X) = \mathbb{1}(X)^* \cdot X^k \cdot \mathbb{1}(X)$, we can rewrite this expression back in the form $c \times w(X,J)$, as desired.

- **($\mathbf{n \times 1, 1 \times 1}$)**: $e_1(A_G)$ is of dimension $n \times 1$ and $e_2(A_G)$ is of dimension 1×1. By induction $e_1(X) \equiv c_1 \times w_1(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_1} \#\mathsf{walk}_k(X)$ and $e_2(X) \equiv c_2 \times \prod_{k \in K_2} \#\mathsf{walk}_k(X)$. Hence,
$$e(X) \equiv c_1 c_2 w_1(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X),$$
which can be rewritten as $c \times w_1(X,J) \cdot \prod_{k \in K} \#\mathsf{walk}_k(X)$ for $c = c_1 c_2$ and K the multiset union of K_1 and K_2.
- **($\mathbf{1 \times n, n \times n}$)**: $e_1(A_G)$ is of dimension $1 \times n$ and $e_2(A_G)$ is of dimension $n \times n$. By induction, $e_1(X) \equiv c_1 \times \prod_{k \in K} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_1(X,J)$ and $e_2(X) \equiv c_2 \times w_2(X,J)$. Hence,
$$e(X) \equiv c_1 c_2 \times \prod_{k \in K} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_1(X,J) \cdot w_2(X,J),$$
which can clearly be written in the form $c \times \prod_{k \in K} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w(X,J)$ for $c = c_1 c_2$ and $w(x,y) = w_1(x,y) w_2(x,y)$.
- **($\mathbf{1 \times n, n \times 1}$)**: $e_1(A_G)$ is of dimension $1 \times n$ and $e_2(A_G)$ is of dimension $n \times 1$. By induction, $e_1(X) \equiv c_1 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_1(X,J)$ and $e_2(X) \equiv c_2 \times w_2(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X)$. Hence,
$$\begin{aligned} e(X) &\equiv c_1 c_2 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \mathbb{1}(X)^* \cdot w_1(X,J) \cdot w_2(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X) \\ &\equiv c_1 c_2 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \mathbb{1}(X)^* \cdot w(X,J) \cdot \mathbb{1}(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X) \\ &\equiv c_1 c_2 c_3 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \prod_{k \in K_3} \#\mathsf{walk}_k(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X) \\ &\equiv c \times \prod_{k \in K} \#\mathsf{walk}_k(X), \end{aligned}$$
for $c = c_1 c_2 c_3$ and K the multiset union of K_1, K_2 and K_3. In the second equivalence we use our earlier observation that $\mathbb{1}(X)^* \cdot w(X,J) \cdot \mathbb{1}(X) = c_3 \times \prod_{k \in K_3} \#\mathsf{walk}_k(X)$ for some $c_3 \in \mathbb{C}$ and multiset K_3.
- **($\mathbf{1 \times 1, 1 \times n}$)**: $e_1(A_G)$ is of dimension 1×1 and $e_2(A_G)$ is of dimension $1 \times n$. By induction, $e_1(X) \equiv c_1 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X)$ and $e_2(X) \equiv c_2 \times \prod_{k \in K_2} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_2(X,J)$. Hence,
$$e(X) \equiv c_1 c_2 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X) \cdot \prod_{k \in K_2} \#\mathsf{walk}_k(X) \cdot \mathbb{1}^{\mathsf{t}}(X) \cdot w_2(X,J),$$
which is clearly in the desired form.
- **($\mathbf{1 \times 1, 1 \times 1}$)**: $e_1(A)$ and $e_2(A)$ are of dimension 1×1. By induction, we have $e_1(X) \equiv c_1 \times \prod_{k \in K_1} \#\mathsf{walk}_k(X)$ and $e_2(X) \equiv c_2 \times \prod_{k \in K_2} \#\mathsf{walk}_k(X)$. Hence,
$$e(X) \equiv c \times \prod_{k \in K} \#\mathsf{walk}_k(X),$$
for $c = c_1 c_2$ and K the multiset union of K_1 and K_2.

(**ones vector**) $e(X) := \mathbb{1}(e_1(X))$. If $e_1(A_G)$ returns an $n \times n$-matrix or $n \times 1$-vector, then $e(X)$ is equivalent to $\mathbb{1}(X)$; if $e_1(A_G)$ returns a $1 \times n$-vector or 1×1-matrix, then $e(X)$ is equivalent to $\mathbb{1}^{\mathsf{t}}(X) \cdot \mathbb{1}(X)$, which are all expressions of the desired form.

(**transposed ones vector**) $e(X) := \mathbb{1}^{\mathsf{t}}(e_1(X))$. This is completely analogous to the previous case.

E Proofs of Section 7

E.1 Proof of Lemma 7.4

We show that $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ if and only if $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$. It suffices to show that $G \equiv_{\mathsf{ML}(\mathscr{L})} H$ implies $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$. This follows from the more general Corollary A.3, Lemma A.4 and Lemma A.5.

Algorithm 1: Computing coarsest equitable partition.

input : Adjacency matrix A
output: Indicator vectors corresponding to coarsest equitable partition

1 Let $B^{(0)} := \mathbb{1}$;
2 Let $i = 1$;
3 **while** $i \leq n$ **do**
4   Let $M^{(i)} := A \cdot B^{(i-1)}$;
5   Let $\mathscr{V}^{(i)} := \{V_1^{(i)}, \ldots, V_{p_i}^{(i)}\}$ a partition such that $k, \ell \in V_j^{(i)}$ if and only if $M_{k*}^{(i)} = M_{\ell *}^{(i)}$;
6   Let $B^{(i)} := [\mathbb{1}_{V_1^{(i)}}, \ldots, \mathbb{1}_{V_{p_i}^{(i)}}]$;
7   Let $i = i + 1$;
8 **end**
9 Return $B^{(n)}$.

E.2 Proof of Proposition 7.5

We show that $\mathsf{ML}(\mathscr{L}^+)$ has sufficient power to compute the coarsest equitable partition of a given graph G. To see this, we implement the algorithm GDCR for finding this partition given in [39]. We recall this algorithm (in a slightly different form than presented in Kersting et al. [39]) in Algorithm 1.

In a nutshell, the algorithm starts by creating a partition consisting of a single part containing all vertices, represented by the indicator vector $\mathbb{1}$ (line 1). Then, in the ith step, the current partition is represented by p_{i-1} indicator vectors $\mathbb{1}_{V_1^{(i-1)}}, \ldots, \mathbb{1}_{V_{p_{i-1}}^{(i-1)}}$ which constitute the columns of matrix $B^{(i-1)}$. The refinement of this partition is then computed in two steps. First, the matrix $M^{(i)} := A \cdot B^{(i-1)}$ (line 4) is computed; Second, each $\mathbb{1}_{V_j^{(i-1)}}$ is refined by putting vertices k and ℓ in the same part if and only if they have the same rows in $M^{(i)}$, i.e., when $M_{k*}^{(i)} = M_{\ell *}^{(i)}$ holds (line 5). The corresponding partition $\mathscr{V}^{(i)}$ is then represented again by indicator vectors and stored as columns of $B^{(i)}$ (line 6). This is repeated until no further refinement of the partition is obtained.

We next detail that we can indeed simulate a run of the algorithm using expressions in $\mathsf{ML}(\mathscr{L}^+)$. We run the algorithm on A_G. Initially, on line 1, we simulate $B^{(0)}$ by the expression $b^{(0)}(X) := \mathbb{1}(X)$. Then, suppose by induction that we have p_{i-1} expressions $b_1^{(i-1)}(X), \ldots, b_{p_{i-1}}^{(i-1)}(X)$ such that the indicator vectors in the partition of the vertex set of V, i.e., those in $B^{(i-1)}$, are given by $b_1^{(i-1)}(A_G), \ldots, b_{p_{i-1}}^{(i-1)}(A_G)$. We next show how the ith iteration is simulated. We first compute the p_{i-1} vectors stored in the columns of $M^{(i)}$. More precisely, we consider

$$m_j^{(i)}(X) := X \cdot b_j^{(i-1)}(X), \quad j = 1, \ldots, p_{i-1}.$$

To compute the refined partition in $\mathscr{V}^{(i)}$, we need to inspect all $m_j^{(i)}(A_G)$ and partition the vertices (rows) according to the values in the matrix $M^{(i)} = [m_1^{(i)}(A_G), \ldots, m_{p_{i-1}}^{(i)}(A_G)]$. That is, two vertices belong to the same part when their rows in $M^{(i)}$ are equal, as explained above.

It is in this test that the $\mathsf{diag}(\cdot)$ operation plays a crucial rule. Let $D_j^{(i)}$ be the set of values occurring in $m_j^{(i)}(A_G)$, for $j = 1, \ldots, p_i$. We assume for convenience that 0 does not occur in $D_j^{(i)}$; if it does we add 1 by all values and work with those incremented values instead. We can compute an indicator vector that identifies the rows in $m_j^{(i)}(A_G)$ that hold a value $c \in D_j^{(i)}$, as follows:

$$\mathbb{1}_{=c}^{(i),j}(X) = \frac{1}{\prod_{c' \neq c}(c - c')} \times \prod_{c' \in D_j^{(i)}, c' \neq c} \mathsf{diag}\big(m_j^{(i)}(X) - c' \times \mathbb{1}(X)\big)$$

Then, when considering

$$\mathbb{1}_{=c_1,\ldots,c_{p_{i-1}}}^{(i)}(X) = \mathsf{diag}(\mathbb{1}_{=c_1}^{(i),1}(X)) \cdot \cdots \cdot \mathsf{diag}(\mathbb{1}_{=c_{p_{i-1}}}^{(i),p_{i-1}}(X)) \cdot \mathbb{1}(X)$$

we obtain an indicator vector identifying all rows in $M^{(i)}$ that hold the value combination $(c_1, \ldots, c_{p_{i-1}})$. Suppose that are p_i distinct combinations that return a non-zero indicator vector. We denote by

$b_1^{(i)}(X),\ldots,b_{p_i}^{(i)}(X)$ the corresponding expressions of the form $\mathbb{1}^{(i)}_{=c_1,\ldots,c_{p_{i-1}}}(X)$. Clearly, these represent the refined partition. This process is then repeated until no refinement takes place. At the end of the algorithm, we let

$$\mathsf{eqpart}_i(X) := b_i^{(n)}(X),$$

for $i=1,\ldots,p$.

We remark that these expressions only work for A_G, as their definitions rely on the values occurring in the matrices $M^{(i)}$ computed along the way. Indeed, we used these values to extract the indicator vectors and also to find the number of expression needed.

Recall that we want to show that G and H have a common equitable partition. We use our assumption that $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ holds to show that the vectors $\mathsf{eqpart}_i(A_H)$, for $i=1,\ldots,p$, also correspond to an equitable partition of H. This is done in a number of steps.

1. For each $i=1,\ldots,p$, we first check whether $\mathsf{eqpart}_i(A_H)$ is also a binary vector containing the same number of 1's as $\mathsf{eqpart}_i(A_G)$. We note that, by construction, $\mathsf{eqpart}_i(A_H)$ is a real vector. Consider the sentence

$$\mathsf{binary\_diag}(X) := \mathbb{1}^{\mathsf{t}}(X)\cdot\big((X\cdot X - X)\cdot(X\cdot X - X)\big)\cdot\mathbb{1}(X).$$

When evaluated on a diagonal real matrix Δ, $\mathsf{binary\_diag}(\Delta)=[0]$ if and only if Δ is a binary diagonal matrix. Indeed, if Δ is a binary diagonal matrix, then $\Delta\cdot\Delta=\Delta$, $\Delta\cdot\Delta-\Delta=Z$, and hence $\mathsf{binary\_diag}(\Delta)=\mathbb{1}^{\mathsf{t}}\cdot Z\cdot Z\cdot\mathbb{1}=[0]$. Conversely, suppose that $\mathsf{binary\_diag}(\Delta)=[0]$. We observe that $(\Delta\cdot\Delta-\Delta)\cdot(\Delta\cdot\Delta-\Delta)$ is a diagonal matrix with squares on its diagonal. Hence, $\mathsf{binary\_diag}(\Delta)=[0]$ implies that every element on the diagonal of $\Delta\cdot\Delta-\Delta$ must be zero. Because we work with diagonal real matrices, this implies that Δ is binary. Indeed, every element on its diagonal must satisfy the equation $x^2-x=0$, implying that $x=0$ or $x=1$. We also note that $\mathsf{binary\_diag}(X)$ can be expressed without $\mathbb{1}^{\mathsf{t}}(X)$ when $\mathsf{tr}(\cdot)$ is present.

   Clearly, $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that

$$\mathsf{binary\_diag}(\mathsf{diag}(\mathsf{eqpart}_i(A_G))) = [0] = \mathsf{binary\_diag}(\mathsf{diag}(\mathsf{eqpart}_i(A_H))),$$

   for all $i=1,\ldots,p$. Hence, the matrices $\mathsf{diag}(\mathsf{eqpart}_i(A_H))$ are indeed binary. In addition, we must also have that $\mathbb{1}^{\mathsf{t}}(X)\cdot\mathsf{eqpart}_i(X)$ must return the same number when evaluated on A_G and A_H. We may thus conclude that $\mathsf{eqpart}_i(A_H)$ also has the same number of entries set to 1 as $\mathsf{eqpart}_i(A_G)$.

2. We next verify that all $\mathsf{eqpart}_i(A_H)$ together form a partition of the vertex set of H. This is done by observing that for binary diagonal matrices Δ_1 and Δ_2, $\Delta_1\cdot\Delta_2$ holds on its diagonal the conjunction of the binary vectors on the diagonals of Δ_1 and Δ_2, respectively. If we want to test that all positions in which Δ_1 and Δ_2 carry value 1 are different, $\Delta_1\cdot\Delta_2$ should be the zero matrix Z. Consider the sentence

$$\mathsf{zerotest\_diag}(X) := \mathbb{1}^{\mathsf{t}}(X)\cdot X\cdot X\cdot\mathbb{1}(X).$$

   It is clear that for real diagonal matrices Δ, $\mathsf{zerotest\_diag}(\Delta)=[0]$ if and only if $\Delta=Z$. We have that $G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that for $i\neq j$,

$$[0] = \mathsf{zerotest\_diag}(\mathsf{diag}(\mathsf{eqpart}_i(A_G))\cdot\mathsf{diag}(\mathsf{eqpart}_j(A_G))) = \mathsf{zerotest\_diag}(\mathsf{diag}(\mathsf{eqpart}_i(A_H))\cdot\mathsf{diag}(\mathsf{eqpart}_j(A_H))).$$

   Hence, the indicator vectors $\mathsf{eqpart}_i(A_H)$, for $i=1,\ldots,p$, are all pairwise disjoint, and based on the fact that $\mathsf{eqpart}_i(A_G)$ are a partition, and $\mathsf{eqpart}_i(A_H)$ and $\mathsf{eqpart}_i(A_G)$ contain the same number of ones, this implies that also $\mathsf{eqpart}_i(A_H)$ correspond to a partition of the vertex set of H.

3. We know that, since the partition $\mathscr{V}=\{V_1,\ldots,V_p\}$ corresponding to the indicator vectors $\mathsf{eqpart}_i(A_G)$ is an equitable partition of G, that

$$\mathsf{diag}(\mathsf{eqpart}_i(A_G))\cdot A_G\cdot\mathsf{diag}(\mathsf{eqpart}_j(A_G))\cdot\mathbb{1} = \deg(v,V_j)\times\mathsf{eqpart}_i(A_G),$$

   where v is any vertex in V_i, the part corresponding to the indicator vector $\mathsf{diag}(\mathsf{eqpart}_i(A_G))$. We can evaluate the expression

$$\mathsf{diag}\big(\mathsf{diag}(\mathsf{eqpart}_i(X))\cdot X\cdot\mathsf{diag}(\mathsf{eqpart}_j(X))\cdot\mathbb{1}(X) - \deg(v,V_j)\times\mathsf{eqpart}_i(X)\big)$$

on A_H, and check again whether the obtained diagonal matrix is the zero matrix, using the sentence $\mathsf{zerotest\_diag}(X)$. This must be the case when $G \equiv_{\mathsf{ML}(\mathcal{L}^+)} H$ holds. As a consequence, $\mathsf{eqpart}_i(A_H)$ is also an equitable partition with the same parameters as the equitable partition of $\mathsf{eqpart}_i(A_G)$.

Hence G and H have indeed a common equitable partition.

E.3 Proof of Proposition 7.6

We show that if G and H have a common equitable partition, then $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f \in \Omega)$. If G and H have a common equitable partition, then Proposition 7.2 tells that there exists a doubly stochastic matrix S such that $A_G \cdot S = S \cdot A_H$. As observed in [52], after rearranging rows (and corresponding columns) of input matrices A_G and A_H, the matrix S may be assumed to be block diagonal. That is, when $\mathscr{V} = \{V_1,\ldots,V_p\}$ and $\mathscr{W} = \{W_1,\ldots,W_p\}$ denote the common coarsest equitable partitions of G and H, respectively, after reordering one can assume that the indicator vectors of these partitions, i.e., $\mathbb{1}_{V_i}$ and $\mathbb{1}_{W_i}$ are the same. More precisely, S can be taken to be

$$\begin{bmatrix} \frac{1}{n_1}J_{n_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{n_2}J_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{n_p}J_{n_p} \end{bmatrix}$$

where J_{n_i} is the all-ones matrix of size $n_i \times n_i$ where n_i is the size of the part corresponding to $\mathbb{1}_{V_i} = \mathbb{1}_{W_i}$. Clearly, this implies that S is compatible with the equitable partitions of G and H. We already showed in Sections B.1, B.3 and B.4 that this indeed implies that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag})$. In view of Corollary A.3 and Lemma A.4, also $e(A_G) = e(A_H)$ holds for sentences in $\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f \in \Omega)$.

E.4 Proof of Theorem 7.7

Clearly, when $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag})} H$ then $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag},+,\times,\mathsf{apply_s}[f],f\in\Omega)} H$, as shown in Corollary A.3 and Lemma A.4. Then, proposition 7.5 tells that G and H must have a common equitable partition and hence, by Proposition 7.2, this implies that there exists a doubly stochastic matrix S such that $A_G \cdot S = S \cdot A_H$. Proposition 7.6 implies that $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathbb{1},\mathsf{diag})} H$. So all three statements in the theorem are equivalent.

E.5 Proof of Theorem 7.8

We show that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$ if and only if there exists an orthogonal doubly quasi-stochastic matrix O that is compatible with the common coarsest equitable partitions of G and H and such that $A_G \cdot O = O \cdot A_H$. Given such a matrix, we have argued in Sections B.1, B.2, B.3 and B.4 that this indeed implies that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$. We therefore focus on the other direction.

The existence of the orthogonal matrix O is shown using Specht's Theorem (see e.g., [36]), which we recall next. Let $\mathscr{A} = \{A_1,\ldots,A_p\}$ and $\mathscr{B} = \{B_1,\ldots,B_p\}$ two sets of complex matrices that are closed under complex conjugation. These sets are called simultaneously unitary equivalent if there exists a unitary matrix U such that $A_i \cdot U = U \cdot B_i$, for $i = 1,\ldots,p$. Here, a unitary matrix U is such that $U^* \cdot U = U \cdot U^* = I$; it is the complex analogue of a real orthogonal matrix. Specht's Theorem provides a means of checking simultaneous unitary equivalence. Indeed, $\mathscr{A}$ and $\mathscr{B}$ are simultaneously unitary equivalent if and only if

$$\mathsf{tr}(w(A_1,\ldots,A_p)) = \mathsf{tr}(w(B_1,\ldots,B_p)),$$

for all words $w(x_1,\ldots,x_p)$ over the alphabet $\{x_1,\ldots,x_p\}$. When each x_i is instantiated with a matrix A_i, we interpret concatenation in w as matrix multiplication. As a note aside, only words of length at most $\mathscr{O}(pn\sqrt{\frac{2(pn)^2}{4(pn-1)}+\frac{1}{4}}+\frac{pn-4}{2})$ are needed [48]. Specht's Theorem also holds when $\mathscr{A}$ and $\mathscr{B}$ are real matrices and similarity is expressed in terms of orthogonal matrices [36]. The required condition is that $\mathscr{A}$ and $\mathscr{B}$ are closed under transposition.

We know that $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$ implies that G and H have a common equitable partition, described in terms of the indicator vectors $\mathsf{eqpart}_i(A_G)$, for $i = 1,\dots,p$, and $\mathsf{eqpart}_i(A_H)$, for $i = 1,\dots,p$, respectively. Consider the following sets of real symmetric matrices: $\mathscr{A}_G := \{A_G, J\} \cup \{\mathsf{diag}(\mathsf{eqpart}_i(A_G) \mid i = 1,\dots,p\}$ and $\mathscr{A}_H := \{A_H, J\} \cup \{\mathsf{diag}(\mathsf{eqpart}_i(A_H) \mid i = 1,\dots,p\}$. We observe that these sets are closed under transposition. By the real counterpart of Specht's Theorem we can check whether there exists an orthogonal matrix O such that

$$A_G \cdot O = O \cdot A_H$$
$$J \cdot O = O \cdot J$$
$$\mathsf{diag}(\mathsf{eqpart}_i(A_G)) \cdot O = O \cdot \mathsf{diag}(\mathsf{eqpart}_i(A_H)),$$

for $i = 1,\dots,p$ in terms of trace identities. These identities can be expressed by sentences in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$. Indeed, we just need to consider sentences

$$e_w(X) := \mathsf{tr}(w(X, \mathbb{1}^{\mathsf{t}}(X) \cdot \mathbb{1}(X), \mathsf{diag}(\mathsf{eqpart}_1(X)), \dots, \mathsf{diag}(\mathsf{eqpart}_p(X)))),$$

where $w(x, j, b_1, \dots, b_p)$ is a word over variables $x, j, b_1, \dots, b_p$. As before when $x \mapsto X$, $j \to \mathbb{1}(X) \cdot \mathbb{1}^{\mathsf{t}}(X)$, and $b_i \mapsto \mathsf{diag}(\mathsf{eqpart}_i(X))$, for $i = 1,\dots,p$, and interpreting concatenation as matrix multiplication, $e_w(X)$ is a sentence in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathbb{1}^{\mathsf{t}},\mathsf{diag})$. Lemma A.5 implies that $e_w(X)$ is equivalent to a sentence in $e'_w(X)$ in $\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})$.

Hence, $G \equiv_{\mathsf{ML}(\cdot,\mathsf{tr},\mathbb{1},\mathsf{diag})} H$ implies

$$\begin{aligned} e'_w(A_G) &\equiv \mathsf{tr}(w(A_G, J, \mathsf{diag}(\mathsf{eqpart}_1(A_G)), \dots, \mathsf{diag}(\mathsf{eqpart}_p(A_G)))) \\ &= e'_w(A_H) \equiv \mathsf{tr}(w(A_H, J, \mathsf{diag}(\mathsf{eqpart}_1(A_H)), \dots, \mathsf{diag}(\mathsf{eqpart}_p(A_H)))). \end{aligned}$$

Since this equality holds for any such word (and thus any sentence $e_w(X)$ or $e'_w(X)$), Specht's Theorem guarantees the existence of the orthogonal matrix O. We verify that O can indeed be chosen to be a doubly quasi-stochastic (orthogonal) matrix that is compatible with the coarsest equitable partitions of G and H.

Since $J \cdot O = O \cdot J$, we observe that $\mathbb{1} \cdot (\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1}) = (O \cdot \mathbb{1}) \cdot (\mathbb{1}^{\mathsf{t}} \cdot \mathbb{1})$ and also $\mathbb{1} \cdot (\mathbb{1}^{\mathsf{t}} \cdot O^{\mathsf{t}} \cdot \mathbb{1}) = (O^{\mathsf{t}} \cdot \mathbb{1}) \cdot (\mathbb{1}^{\mathsf{t}} \cdot \mathbb{1})$. Since $\mathbb{1}^{\mathsf{t}} \cdot \mathbb{1} = n$, $(\mathbb{1}^{\mathsf{t}} \cdot O^{\mathsf{t}} \cdot \mathbb{1}) = (\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1})^{\mathsf{t}}$ and $\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1}$ is a real number, $\mathbb{1}^{\mathsf{t}} \cdot O^{\mathsf{t}} \cdot \mathbb{1} = \mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1}$. Furthermore,

$$n = \mathbb{1}^{\mathsf{t}} \cdot I \cdot \mathbb{1} = \mathbb{1}^{\mathsf{t}} \cdot O \cdot O^{\mathsf{t}} \cdot \mathbb{1} = n \frac{(\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1})^2}{n^2} = \frac{(\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1})^2}{n},$$

or $\mathbb{1}^{\mathsf{t}} \cdot O \cdot \mathbb{1} = n$. In other words, $O \cdot \mathbb{1} = \mathbb{1}$ and $O^{\mathsf{t}} \cdot \mathbb{1} = \mathbb{1}$.

The block diagonal structure of O (after possibly reordering rows and columns of A_G and A_H) follows from the constraints $\mathsf{diag}(\mathsf{eqpart}_i(A_G)) \cdot O = O \cdot \mathsf{diag}(\mathsf{eqpart}_i(A_H))$, for $i = 1,\dots,p$. Indeed, in view of the common equitable partition, we can indeed reorder rows and columns such that the indicator vectors of the partitions coincide. That is, we can use $\mathbb{1}_i = \mathsf{eqpart}_i(A_G) = \mathsf{eqpart}_i(A_H)$ for $i = 1,\dots,p$ to indicate both partitions in G and H, respectively.

It is routine exercise to show that $\mathsf{diag}(\mathsf{eqpart}_i(A_G)) \cdot O = O \cdot \mathsf{diag}(\mathsf{eqpart}_i(A_H))$, for $i = 1,\dots,p$, imply a diagonal block structure (see also Lemma 6 in Thüne [54]). We repeat the argument here, for completeness. Indeed, define $\Delta_1 = \mathsf{diag}(\mathbb{1}_1)$, $\Delta_2 = I - \Delta_1$ and $O_{ij} = \Delta_i \cdot O \cdot \Delta_j$ for $i, j = 1, 2$. We have $O^{\mathsf{t}} \cdot E_1 \cdot O = E_1 = O \cdot E_1 \cdot O^{\mathsf{t}}$. This implies that O_{12} is the zero matrix and $(O_{11})^{\mathsf{t}} \cdot O_{11} = I$ and also O_{21} is the zero matrix and $(O_{22})^{\mathsf{t}} \cdot O_{22} = I$. Repeating the argument on O_{22} using $\mathsf{diag}(\mathbb{1}_2)$, and so forth results in the block diagonal structure. Combined $O \cdot \mathbb{1} = \mathbb{1}$ and $O^{\mathsf{t}} \cdot \mathbb{1} = \mathbb{1}$, this implies that $O \cdot \mathbb{1}_i = \mathbb{1}_i$ and $O^{\mathsf{t}} \cdot \mathbb{1}_i = \mathbb{1}_i$. When no reordering on the input matrices is applied, these conditions correspond to $\mathbb{1}_{V_i} = O \cdot \mathbb{1}_{W_i}$ and $O^{\mathsf{t}} \cdot \mathbb{1}_{W_i}$, for $i = 1,\dots,p$.

E.6 Proof of Corollary 7.10

Inspecting the proofs of Proposition 7.6 and Theorem 7.8, and more specifically by considering the analysis carried out in Section B.4, we see that the induction hypotheses imply that for expressions $e(X) \in \mathsf{ML}(\mathscr{L}^+)$, when $e(A_G)$ is an $n \times 1$-vector (or a $1 \times n$-vector), then $e(A_G)$ is a permutation of $e(A_H)$. (after reordering they can be assumed to be the same vectors.) This follows from the observation that any vector is a linear combination of indicator vectors of the common equitable partition, these indicator vectors are permutations of each other, and the coefficients in the linear combination can be take the same for $e(A_G)$ and $e(A_H)$.

If we allow function application on vectors, consider $f \in \Omega$ and $e(X) := \mathsf{apply}_{\mathsf{v}}[f](e_1(X), \dots, e_p(X))$, where either all $e_i(A_G)$'s are 1×1-matrices, or all are $n \times 1$, or $1 \times n$-vectors. Clearly,

Algorithm 2: Computing stable edge colouring.

input : Adjacency matrix A, initial colouring $\chi : V \times V \to \{0,1,2\}$
output: Coarsest stable edge colouring χ.

1 Let $\chi := \chi_0$;
2 Let $C := \{0,1,2\}$;
3 **repeat**
4   **for** $(v,v') \in V \times V$ **do**
5   Compute $\mathsf{L}^2(v,v')$ relative to χ;
6   **end**
7   Replace C by a minimal set of new colours C' and define $\chi' : V \times V \to C'$ such that
8   **for** *pairs* (v_1,v_2), (v_1',v_2') *in* $V \times V$ **do**
9         $\chi'(v_1,v_2) = \chi'(v_1',v_2') \Leftrightarrow \mathsf{L}^2(v_1,v_2) = \mathsf{L}^2(v_1',v_2')$
10   **end**
11   Let $C := C'$;
12   Let $\chi := \chi'$;
13 **until** $|C|$ does not change;

$G \equiv_{\mathsf{ML}(\mathscr{L}^+)} H$ implies that $e_i(A_G) = e_i(A_H)$ for sentences, and we just observed that in case of vectors, the $e_i(A_G)$'s are related by the same permutation to the $e_i(A_H)$'s. As a consequence, also $e(A_H)$ is related as such. It now suffices to recall Corollary A.8 which states that we can allow arbitrary permutations when showing equivalence. We may thus assume that $e(A_G) = e(A_H)$. By further induction, this implies that pointwise function application on vectors does not add distinguishing power.

F Proofs of Section 8

F.1 Proof of Theorem 8.3

We show that $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$ has sufficient power to compute the coarsest stable edge colouring of a given graph G. To see this, we implement the algorithm 2-STAB [5], that computes the coarsest stable edge colouring, by expressions $\mathsf{stabcol}_i(X)$, for $i = 1,\dots,\ell$, in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$. Each $\mathsf{stabcol}_i(A_G)$ is an *indicator matrix* representing the part of the partition Π of $V \times V$ induced by the stable edge colouring. In other words, these indicator matrices carry a 1 in entries corresponding pairs $(v,v') \in V \times V$ in the same part in the partition and carry 0's everywhere else. We recall the algorithm 2-STAB in Algorithm 2.

Initially, the algorithm starts with the initial colouring $\chi_0 : V \times V \to \{0,1,2\}$ (line 1) defined such that $\chi_0(v,v) = 2$, $\chi_0(v,w) = 1$ if $(v,w) \in E$, and $\chi_0(v,w) = 0$ for $v \neq w$ and $(v,w) \notin E$. We simulate this colouring by considering expressions $\mathsf{colpart}_2^{(0)}(X) := \mathsf{diag}(\mathbb{1}(X))$; $\mathsf{colpart}_1^{(0)}(X) := X$; and $\mathsf{colpart}_0^{(0)}(X) := \mathbb{1}(X) \cdot \mathbb{1}^{\mathsf{t}}(X) - X - \mathsf{diag}(\mathbb{1}(X))$. Clearly, $\mathsf{colpart}_0^{(0)}(A_G)$, $\mathsf{colpart}_1^{(0)}(A_G)$, and $\mathsf{colpart}_2^{(0)}(A_G)$ represent the initial partition Π_{χ_0} corresponding to χ_0. The initial set of colours C is equal to $\{0,1,2\}$ (line 2).

In consequent steps, the algorithm refines the current colouring $\chi : V \times V \to C$, based on $\mathsf{L}^2(v,v')$. We recall that for a pair $(v,v') \in V \times V$ and pairs (c,d) of colours in C,

$$\mathsf{L}^2(v,v') := \{(c,d,p_{v,v'}^{c,d}) \mid p_{v,v'}^{c,d} \neq 0\}, \text{ where } p_{v,v'}^{c,d} := |\{v'' \in V \mid \chi(v,v'') = c, \chi(v'',v') = d\}|.$$

These are computed on line 5 for any pair of vertices $(v,v') \in V \times V$. The refinement $\chi' : V \times V \to C'$ uses an updated set of colours C and the corresponding partition $\Pi_{\chi'}$ is a refinement of the partition Π_χ.

Suppose, by induction that, after iteration i, the current set of colours C consists of ℓ_i colours and the colouring is $\chi : V \times V \to C$. Assume that we have ℓ_i expressions $\mathsf{colpart}_c^{(i)}(X)$, one for each $c \in C$, such that $\mathsf{colpart}_c^{(i)}(A_G)$ are indicator matrices representing the edge partition Π_χ corresponding to χ. Given these, we construct expressions for the refined partition computed in the next iteration.

More precisely, we consider for each pair of colours (c,d) in C, the expression

$$P_{c,d}(X) := \mathsf{colpart}_c^{(i)}(X)\cdot\mathsf{colpart}_d^{(i)}(X)$$

which, on input A_G, results in a matrix whose entry corresponding to vertices v and v' holds the value $p_{v,v'}^{c,d}$; these are needed for $\mathsf{L}^2(v,v')$.

Let $\mathscr{P}_{c,d}$ be the set of values in $P_{c,d}(A_G)$. For each non-zero value p in $\mathscr{P}_{c,d}$, we extract an indicator matrix indicating the positions in $P_{c,d}(A_G)$ that hold value p. This can be done using an expression $\mathsf{ind}_{c,d,p}(X)$ which works in a similar way as $\#3\mathsf{deg}(X)$ in Example 7.1, but uses the Schur-Hadamard product instead of products of diagonal matrices. More precisely,

$$\mathsf{ind}_{c,d,p}(X) := \frac{1}{p\prod_{p_i}(p-p_i)} X\circ(X-p_1\times\mathbb{1}(X)\cdot\mathbb{1}^{\mathsf{t}}(X))\circ\cdots\circ(X-p_r\times\mathbb{1}(X)\cdot\mathbb{1}^{\mathsf{t}}(X))$$

where $p_1,\dots,p_r$ are all values in $\mathscr{P}_{c,d}$ different from p. In fact, the construction of indicator matrices pointing out the entries in a matrix that hold a specific value is referred to as the Schur-Wielandt Principle [49].

We now look at $\mathsf{L}^2(v,v')$ for all v,v' in V. If we denote by C' the new list of colours (one for each unique $\mathsf{L}(v,v')$), then a colour c in C' corresponds to a set of s triples $(c_1,d_2,p_{v,v'}^{c_1,d_1}),\dots,(c_s,d_s,p_{v,v'}^{c_s,d_s})$ for some pair v, v' of vertices. The Schur-Hadamard product plays a key part in finding those vertex pairs having colour c. Indeed, we consider the expression

$$\mathsf{colpart}_c^{(i+1)}(X) := \mathsf{ind}_{c_1,d_2,p_{v,v'}^{c_1,d_1}}(X)\circ\cdots\circ\mathsf{ind}_{c_s,d_s,p_{v,v'}^{c_s,d_s}}(X).$$

On input A_G, $\mathsf{colpart}_c^{(i+1)}(A_G)$ returns an indicator matrix in which the entries holding a 1 correspond precisely to the pairs $(v'',v''')\in V\times V$ such that $c=\mathsf{L}^2(v'',v''')=\mathsf{L}^2(v,v')$. In other words, $\mathsf{colpart}_c^{(i+1)}(A_G)$ represent the refined edge partition corresponding to the new colour c. We do this for every new colour. Clearly, all $\mathsf{colpart}_c^{(i+1)}(A_G)$ together represent the refined partition $\Pi_{\chi'}$ corresponding to $\chi': V\times V\to C'$ as computed on line 9.

We continue in this way until the colouring stabilises. i.e., no further colours are needed. We denote by $\mathsf{stabcol}_i(X)$ the final expressions obtained. We have that $\mathsf{stabcol}_i(A_G)$ are indicator matrices of the partition induced by the stable edge colouring.

The construction of the expressions $\mathsf{stabcol}_i(X)$ depend on A_G; at first sight, correctness is only guaranteed for this adjacency matrix. Nevertheless, when $G\equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$, we next show that $\mathsf{stabcol}_i(A_H)$ also corresponds to a partition corresponding to the stable edge colouring on H. To check this, we rely on an equivalent characterisation of partitions induced by stable edge colourings. More precisely, $\{\mathsf{stabcol}_i(X_H)\}$ represents a stable edge colouring of H if and only if they form the standard basis of the cellular algebra of H. Here, the cellular algebra of H, sometimes also referred to as the *Weisfeiler-Lehman closure* $\mathsf{WL}(A_H,I,J)$ of A_H, I and J, is the smallest algebra containing A_H, I and J, and which is closed under the Schur-Hadamard product [5, 32]. More precisely, $\mathsf{stabcol}_i(A_H)$ is the partition of a stable edge colouring χ_H of H if and only if

$$J=\sum_{i=1}^{\ell}\mathsf{stabcol}_i(A_H)$$

$$I=\sum_{i\in K}\mathsf{stabcol}_i(A_H), \text{for some subset } K \text{ of } \{1,\dots,\ell\}$$

$$\text{for each } i\text{:}\quad (\mathsf{stabcol}_i(A_H))^{\mathsf{t}}=\mathsf{stabcol}_j(A_H) \text{ for some } j$$

$$\text{for any } i,j\text{:}\quad \mathsf{stabcol}_i(A_H)\cdot\mathsf{stabcol}_j(A_H)=\sum p_{i,j}^k\times\mathsf{stabcol}_k(A_H),$$

where $p_{(i,j)}^k := p_{(w,w')}^{(i,j)}$ for some $(w,w')\in W^2$ such that $\chi(w,w')=k$, where χ_H is a stable edge colouring on H.. We know that $\mathsf{stabcol}_i(A_G)$ satisfies these conditions and they thus represent the standard basis of $\mathsf{WL}(A_G,I,J)$. We use $G\equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ to show that $\mathsf{stabcol}_i(A_H)$ also satisfies these conditions. At the same time we derive the necessary ingredients to show that G and H cannot be distinguished by edge-colouring. This is done in a number of steps.

1. For each $i=1,\dots,\ell$, we first check whether $\mathsf{stabcol}_i(A_H)$ is also a binary matrix. We observe, based on inspecting the defining expressions, that $\mathsf{stabcol}_i(A_H)$ is a real matrix. If we consider the sentence

$$\mathsf{binary}(X) := \mathbb{1}^{\mathsf{t}}(X)\cdot\big((X\circ X-X)\circ(X\circ X-X)\big)\cdot\mathbb{1}(X),$$

then this sentence will return $[0]$, when given a real matrix as input A, if and only if the input matrix is a binary matrix. Indeed, for a binary matrix A, $A\circ A = A$ and hence $A\circ A - A = Z$. Since $Z\circ Z = Z$, $\mathsf{binary}(A) = \mathbb{1}^{\mathsf{t}}\cdot Z\cdot \mathbb{1} = [0]$. For the converse, assume that $\mathsf{binary}(A) = [0]$. We observe that each entry in $(A\circ A - A)\circ(A\circ A - A)$ is non-negative value. Indeed, all entries are squares of real numbers. This implies that $A\circ A - A = Z$. Clearly, this implies that A can only contain 0 or 1 as entries, since these are the only real values satisfying $x^2 - x = 0$.

Hence, when $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ this implies

$$\mathsf{binary}(\mathsf{stabcol}_i(A_G)) = [0] = \mathsf{binary}(\mathsf{stabcol}_i(A_H)),$$

for all $i = 1,\dots \ell$. So indeed, $\mathsf{stabcol}_i(A_H)$ is a binary matrix.

2. Next, we check that the matrices $\mathsf{stabcol}_i(A_H)$, for $i \in K$, are diagonal matrices. We will use the following zero test:

$$\mathsf{zerotest}(X) := \mathbb{1}^{\mathsf{t}}(X)\cdot(X\circ X)\cdot\mathbb{1}(X).$$

A similar reasoning as before shows that $\mathsf{zerotest}(A) = [0]$ for a real matrix A if and only if $A = Z$. If we consider

$$\mathsf{diagtest}(X) := \mathsf{zerotest}(X\circ(\mathbb{1}(X)\cdot\mathbb{1}^{\mathsf{t}}(X) - \mathsf{diag}(\mathbb{1}(X)))),$$

then $\mathsf{diagtest}(A) = [0]$ if and only if A is a diagonal matrix. We have that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies

$$\mathsf{diagtest}(\mathsf{stabcol}_i(A_G)) = [0] = \mathsf{diagtest}(\mathsf{stabcol}_i(A_H)),$$

for all $i \in K$. Hence, $\mathsf{stabcol}_i(A_H)$ for $i \in K$ are indeed diagonal matrices.

3. We next check that $\mathsf{stabcol}_i(A_G)$ and $\mathsf{stabcol}_i(A_H)$ have the same number of ones. This will be important when showing indistinguishability of G and H by edge-colouring. Clearly, $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies

$$\mathsf{tr}(\mathsf{stabcol}_i(A_G)\cdot\mathbb{1}(A_G)) = \mathsf{tr}(\mathsf{stabcol}_i(A_H)\cdot\mathbb{1}(A_H)),$$

for $i = 1,\dots,\ell$, which implies that the number of ones agree.

4. We also verify that all $\mathsf{stabcol}_i(A_H)$ are disjoint, i.e., all the occurrences of 1 are in different positions. This, together with the fact that $\sum_{i=1}^{\ell}\mathsf{stabcol}_i(A_G) = J$, $\sum_{i\in K}\mathsf{stabcol}_i(A_G) = I$ and that $\mathsf{stabcol}_i(A_G)$ and $\mathsf{stabcol}_i(A_H)$ have the same number of ones, this implies that $\sum_{i=1}^{\ell}\mathsf{stabcol}_i(A_H) = J$ and $\sum_{i\in K}\mathsf{stabcol}_i(A_H) = I$. To check disjointness, we uses sentences

$$\mathsf{disjoint}_{i,j}(X) := \mathsf{zerotest}(\mathsf{stabcol}_i(X)\circ\mathsf{stabcol}_j(X)),$$

for $i \neq j$. We have that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies $\mathsf{disjoint}_{i,j}(A_G) = [0] = \mathsf{disjoint}_{i,j}(A_H)$ for all $i \neq j$, establishing the pairwise disjointness of the $\mathsf{stabcol}_i(A_H)$'s.

5. Next, we proceed in a similar way to show that, when $(\mathsf{stabcol}_i(A_G))^{\mathsf{t}} = \mathsf{stabcol}_j(A_G)$ holds, also $(\mathsf{stabcol}_i(A_H))^{\mathsf{t}} = \mathsf{stabcol}_j(A_H)$ holds. To this aim, we consider the sentence

$$\mathsf{is\_transpose}_{i,j}(X) := \mathsf{zerotest}((\mathsf{stabcol}_i(X))^{\mathsf{t}} - \mathsf{stabcol}_j(X)),$$

where (i,j) are such that $\mathsf{stabcol}_j(A_G)$ is the transpose of $\mathsf{stabcol}_j(A_G)$. Then, clearly, we have that $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies $\mathsf{is\_transpose}_{i,j}(A_G) = [0] = \mathsf{is\_transpose}_{i,j}(A_H)$ for all valid pairs (i,j). So, we may again conclude that $(\mathsf{stabcol}_i(A_H))^{\mathsf{t}} = \mathsf{stabcol}_j(A_H)$ when $(\mathsf{stabcol}_i(A_G))^{\mathsf{t}} = \mathsf{stabcol}_j(A_G)$ holds.

6. We also verify that $(\mathsf{stabcol}_i(A_H)\cdot\mathsf{stabcol}_j(A_H))\circ\mathsf{stabcol}_k(A_H) = p_{i,j}^k\times\mathsf{stabcol}_k(A_H)$ holds. This is again done in the same way as before, i.e., we consider the sentences

$$\mathsf{structconst}_{i,j,k} := \mathsf{zerotest}\big((\mathsf{stabcol}_i(X)\cdot\mathsf{stabcol}_j(X))\circ\mathsf{stabcol}_k(X) - p_{i,j}^k\times\mathsf{stabcol}_k(X)\big).$$

Then, $G \equiv_{\mathsf{ML}(\cdot,^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies that $\mathsf{structconst}_{i,j,k}(A_G) = [0] = \mathsf{structconst}_{i,j,k}(A_H)$, for all $i,j,k = 1,\dots,\ell$.

7. Finally, we know that $A_G = \sum_{i\in L}\mathsf{stabcol}_i(A_G)$ for some subset $L \subseteq \{1,\dots,\ell\}$. Indeed, A_G belongs to $\mathsf{WL}(A_G,I,J)$ and thus $A_G = \sum_{i=1}^{\ell} a_i\times\mathsf{stabcol}_i(A_G)$. Since we have that $\mathsf{stabcol}_i(A_G)$ and $\mathsf{stabcol}_j(A_G)$, for $i \neq j$, are disjoint and binary matrices, and also A_G is a binary matrix, the coefficients a_i must be either 0 or 1. If we let L be the index set such that $a_i = 1$, we indeed have that $A_G = \sum_{i\in L}\mathsf{stabcol}_i(A_G)$. We next show that $A_H = \sum_{i\in L}\mathsf{stabcol}_i(A_H)$ as well. We know already that $A_H = \sum b_i\times\mathsf{stabcol}_i(A_H)$ and for the same reasons as above,

$A_H = \sum_{i\in L'} \mathsf{stabcol}_i(A_H)$ for some subset $L' \subseteq \{1,\dots,\ell\}$. We verify that $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ implies that $L = L'$. Consider the sentence

$$\mathsf{included}_i(X) := \mathbb{1}^{\mathsf{t}}(X)\cdot(X \circ \mathsf{stabcol}_i(X))\cdot \mathbb{1}(X).$$

We have that $\mathsf{included}_i(A_G) = [0]$ if and only if $i \in L$. Similarly, $\mathsf{included}_i(A_H) = [0]$ if and only if $i \in L'$. Suppose, for the sake of contradiction that $i \in L' \setminus L$ (the case $i \in L \setminus L'$ is analogous). Then, $\mathsf{included}_i(A_G) \neq [0] = \mathsf{included}_i(A_H)$, contradicting that $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},+,\times,\circ)} H$ holds. We may thus conclude that $L = L'$.

All combined, $\mathsf{stabcol}_j(A_H)$, for $i = 1,\dots,\ell$, satisfy the conditions to be the standard basis of $\mathsf{WL}(A_H, I, J)$, and hence these indicator matrices represent a stable edge colouring of H. We also have that $G \equiv_{\mathsf{2WL}} H$. To see this, it suffices to show that $\mathsf{stabcol}_j(A_G)$ and $\mathsf{stabcol}_j(A_H)$ represent a part in their corresponding stable partitions that have the same colour. Indeed, since both $\mathsf{stabcol}_j(A_G)$ and $\mathsf{stabcol}_j(A_H)$ contain the same number of pairs, this implies that $G \equiv_{\mathsf{2WL}} H$.

Let $\chi_G : V \times V \to C$ and $\chi_H : W \times W \to C$ the stable edge-colourings corresponding to the indicator matrices $\mathsf{stabcol}_i(A_G)$ and $\mathsf{stabcol}_i(A_H)$. We know the colours of pairs of vertices in $\mathsf{stabcol}_i(A_G)$, as the expressions $\mathsf{stabcol}_i(X)$ were constructed by running 2-STAB on A_G. For $\mathsf{stabcol}_i(A_H)$, we only know that they correspond to indicator matrices of a stable edge-colouring. We argue that χ_G and χ_H can be taken the same edge-colouring function. This is done by induction on the number of steps of 2-STAB on A_G and A_H.

More precisely, we show that for any $i = 1,\dots,\ell$, $\chi_H(w,w') = \chi_G(v,v')$ for any pair $(w,w') \in W^2$ and pair $(v,v') \in V^2$ such that $\mathsf{stabcol}_i(A_H)$ and $\mathsf{stabcol}_i(A_G)$ has a 1 in the positions corresponding to these pairs of vertices.

Initially, we know that when $\mathsf{stabcol}_i(A_G)$ is a diagonal matrix if and only if $\mathsf{stabcol}_i(A_H)$ is a diagonal matrix. Similarly, we have shown that $A_G = \sum_{i\in L} \mathsf{stabcol}_i(A_G)$ and $A_H = \sum_{i\in L} \mathsf{stabcol}_i(A_H)$. Since $J = \sum_{i=1}^{\ell} \mathsf{stabcol}_i(A_G)$ and $J = \sum_{i=1}^{\ell} \mathsf{stabcol}_i(A_H)$, also $J - A_G - I = \sum_{i\in M} \mathsf{stabcol}_i(A_G)$ and $J - A_H - I = \sum_{i\in M} \mathsf{stabcol}_i(A_H)$. This implies that the initial colourings of G and H will assign the same colour to corresponding indicator matrices. Indeed, recall that the initial colouring is fully determined by self-loops, edges in the graph, and edges in the complement graph.

Suppose, by induction, that $\mathsf{stabcol}_j(A_G)$ and $\mathsf{stabcol}_j(A_H)$ are coloured the same, when the algorithm 2-STAB starts a refinement step. The recipe for refinement was explained before and is fully determined by $\mathsf{L}^2(v,v')$ for A_G, and $\mathsf{L}^2(w,w')$ for A_H. It suffices to show that these are the same for pairs of vertices in $\mathsf{stabcol}_j(A_G)$ and $\mathsf{stabcol}_j(A_H)$, respectively. This in turn pours down to showing that for each pair of colours (c,d) in C (the current colours in the colouring function of G and H), $p^{c,d}_{v,v'} = p^{c,d}_{w,w'}$. For a pair (v,v') in the partition $\mathsf{stabcol}_k(A_G)$, let L_c and L_d be the index sets of partitions coloured by c and d, respectively. We know from before that

$$\Big(\sum_{i\in L_c, j\in L_d} \mathsf{stabcol}_i(A_G)\cdot \mathsf{stabcol}_j(A_G)\Big)\circ \mathsf{stabcol}_k(A_G) = \Big(\sum_{i\in L_c, j\in L_d} p^k_{i,j}\Big)\times \mathsf{stabcol}_k(A_G),$$

and that $p^{c,d}_{v,v'} = \sum_{i\in L_c, j\in L_d} p^k_{i,j}$. We know that $\mathsf{stabcol}_i(A_H)$ share the same structure constants with $\mathsf{stabcol}_i(A_G)$, hence, $p^{c,d}_{w,w'} = \sum_{i\in L_c, j\in L_d} p^k_{i,j} = p^{c,d}_{v,v'}$. This holds for every colour pair (c,d) and vertex pairs (v,v') and (w,w') in $\mathsf{stabcol}_k(A_G)$ and $\mathsf{stabcol}_k(A_H)$, respectively. As a consequence, $\mathsf{L}^2(v,v') = \mathsf{L}^2(w,w')$ for every such pairs of vertices. Hence, after refinement the updated colour functions χ'_G and χ'_H will be the same.

To conclude, we have that $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)} H$ implies that for every colour c, there is unique, by stability, k such that $\mathsf{stabcol}_k(A_G)$ and $\mathsf{stabcol}_k(A_H)$ have colour c. Since these carry the same number of ones (as shown earlier), G and H cannot be distinguished by edge-colouring and thus $G \equiv_{\mathsf{WL}} H$. As a consequence, by Theorem 8.2, $G \equiv_{\mathsf{2WL}} H$ if and only if $G \equiv_{\mathsf{C}^3} H$. Hence, $G \equiv_{\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)} H$ implies $G \equiv_{\mathsf{C}^3} H$.

For the converse, it is known that $G \equiv_{\mathsf{C}^3} H$ implies that there exists an isomorphism between the Weisfeiler-Lehman closures $\mathsf{WL}(A_G, I, J)$ and $\mathsf{WL}(A_H, I, J)$ [19]. This can be shown by using that $G \equiv_{\mathsf{C}^3} H$ and $G \equiv_{\mathsf{2WL}} H$ are equivalent [10], and that the latter implies that there is a so-called algebraic isomorphism between the coherent *configurations* underlying the coherent algebras $\mathsf{WL}(A_G, I, J)$ and $\mathsf{WL}(A_H, I, J)$ [49]. The latter is known to be equivalent to the existence of a unitary matrix U such that $\mathsf{stabcol}_i(A_G)\cdot U = U \cdot \mathsf{stabcol}_i(A_H)$ [49]. Since the indicator matrices are real matrices, closed under transposition, U may be assumed to be an orthogonal matrix O (see e.g., Jing [36] or Theorem 65, Section 2.6. in Kaplansky [38]).

We can also directly infer from $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$-equivalence that there exists an orthogonal matrix O satisfying $\mathsf{stabcol}_i(A_G)\cdot O = O\cdot\mathsf{stabcol}_i(A_H)$. This is shown in precisely the same way as in the proof of Theorem 7.8, by relying the Specht's Theorem. The two sets, that are closed under transposition, are now $\mathscr{A} := \{\mathsf{stabcol}_i(A_G) \mid i = 1,\dots,\ell\}$ and $\mathscr{B} := \{\mathsf{stabcol}_i(A_H) \mid i = 1,\dots,\ell\}$. We remark that when $\mathsf{stabcol}_i(A_G)\cdot O = O\cdot\mathsf{stabcol}_i(A_H)$ implies $A_G\cdot O = O\cdot A_H$, since $A_G = \sum i \in L\mathsf{stabcol}_i(A_G)$ and $A_H = \sum_{i\in L}\mathsf{stabcol}_i(A_H)$. Furthermore, $J\cdot O = O\cdot J$, because J is the sum over all indicator matrices. We may thus assume that $O\cdot\mathbb{1} = \mathbb{1}$ (and thus also $O^{\mathsf{t}}\cdot\mathbb{1} = \mathbb{1}$), as argued in the proof of Theorem 7.8. Moreover, we may reorder the input matrices A_G and A_H such that O has a block structure compatible with the common coarsest equitable partitions of G and H, induced by the diagonal basis elements $\mathsf{stabcol}_i(A_G)$ and $\mathsf{stabcol}_i(A_H)$, for $i \in K$.

Given such an orthogonal matrix O, we have already argued in Section B.5 that this suffices to ensure that $e(A_G) = e(A_H)$ for all sentences $e(X)$ in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$.

F.2 Functions on matrices

in Section B.5 we have shown that for expressions $e(X)$ in $\mathsf{ML}(\cdot,{}^*,\mathsf{tr},\mathbb{1},\mathsf{diag},+,\times,\circ)$, if $A_G\cdot O = O\cdot A_H$ for an orthogonal matrix which is an (algebraic) isomorphism between $\mathsf{WL}(A_G,I,J)$ and $\mathsf{WL}(A_H,I,J)$, then $e(A_G)$ and $e(A_H)$ (when they return an $n\times n$-matrix) can be written as a linear combination of elements in the standard bases of the coherent algebras, and more importantly, these are linear combinations with the same coefficients. That is, $e(A_G) = \sum a_i \times E_i$ and $e(A_H) = \sum a_i \times F_i$, for standard bases $\mathscr{E} = \{E_1,\dots,E_\ell\}$ and $\mathscr{F} = \{F_1,\dots,F_\ell\}$.

We next show that the induction hypotheses (†), (‡), and (§) remain to hold when pointwise function applications on matrices are considered.

(**pointwise functions**, (†, §)) $e(X) := \mathsf{apply}[f](e_1(X),\dots,e_p(X))$. We consider the case when all $e_i(X)$ return $n\times n$-matrices. By induction, using (§), we have that $e_i(A_G) = \sum_j a_j^{(i)} \times E_j$ and $e_i(A_H) = \sum_j a^{(i)} \times F_j$. Since all E_i's are pairwise disjoint (have entries with 1 in different positions), we have that

$$e(A_G) = \mathsf{apply}[f](e_1(A_G),\dots,e_p(A_G)) = \sum_j f(a_j^{(1)},\dots,e_j^{(p)}) \times E_i,$$

and similarly,

$$e(A_H) = \mathsf{apply}[f](e_1(A_H),\dots,e_p(A_H)) = \sum_j f(a_j^{(1)},\dots,e_j^{(p)}) \times F_i.$$

From $E_i\cdot O = O\cdot F_i$, it then follows that $e(A_G)\cdot O = O\cdot e(A_H)$ and the hypotheses (†) remains satisfied. The expressions above also show that hypothesis (§) is satisfied.

(**pointwise functions**, (‡)) $e(X) := \mathsf{apply}[f](e_1(X),\dots,e_p(X))$. We need to show that when $e(A_G)$ is an $n\times n$-matrix, then $e(A_G)\cdot\mathbb{1}_i = \sum a_{ij}\cdot\mathbb{1}_j = e(A_H)\cdot\mathbb{1}_i$. As observed above, we have that $e(A_G) = \sum a_i \times E_i$ and $e(A_H) = \sum a_i \times F_i$. This follows from properties of the standard basis, described earlier. That is, $E_i\cdot\mathbb{1}_i$ is either the zero vector or a specific vector $\mathbb{1}_j$; similarly for $F_i\cdot\mathbb{1}_i$. This implies that hypotheses (‡) remain satisfied.