Graph Similarity Based on Matrix Norms

Timo Gervens
RWTH Aachen University, Germany

Martin Grohe
RWTH Aachen University, Germany

Abstract

Quantifying the similarity between two graphs is a fundamental algorithmic problem at the heart of many data analysis tasks for graph-based data. In this paper, we study the computational complexity of a family of similarity measures based on quantifying the mismatch between the two graphs, that is, the "symmetric difference" of the graphs under an optimal alignment of the vertices. An important example is similarity based on graph edit distance. While edit distance calculates the "global" mismatch, that is, the number of edges in the symmetric difference, our main focus is on "local" measures calculating the maximum mismatch per vertex.

Mathematically, our similarity measures are best expressed in terms of the adjacency matrices: the mismatch between graphs is expressed as the difference of their adjacency matrices (under an optimal alignment), and we measure it by applying some matrix norm. Roughly speaking, global measures like graph edit distance correspond to entrywise matrix norms like the Frobenius norm and local measures correspond to operator norms like the spectral norm.

We prove a number of strong NP-hardness and inapproximability results even for very restricted graph classes such as bounded-degree trees.

2012 ACM Subject Classification Mathematics of computing → Graph theory; Theory of computation → Complexity theory and logic

Keywords and phrases graph similarity, approximate graph isomorphism, graph matching

Digital Object Identifier 10.4230/LIPIcs.MFCS.2022.52

Related Version *Full Version*: `https://arxiv.org/abs/2207.00090` [11]

Funding *Timo Gervens*: This work was supported by the DFG grant GR 1492/20-1.

1 Introduction

Graphs are basic models ubiquitous in science and engineering. They are used to describe a diverse range of objects and processes from chemical compounds to social interactions. To understand and classify graph models, we need to compare graphs. Since data and models are not always guaranteed to be exact, it is essential to understand what makes two graphs similar or dissimilar, and to be able to compute similarity efficiently. There are many different approaches to similarity, for example, based on edit-distance (e.g. [3, 6, 27]), spectral similarity (e.g. [16, 29, 30]), optimal transport (e.g. [5, 23, 25]), or behavioral equivalence (e.g. [28, 31]). This is only natural, because the choice of a "good" similarity measure will usually depend on the application. While graph similarity has received considerable attention in application areas such as computer vision (see, for example, [7, 9]) and network science (see, for example, [8]), theoretical computer scientists have not explored similarity systematically; only specific "special cases" such as isomorphism [14] and bisimilarity [28] have been studied to great depth. Yet it seems worthwhile to develop a *theory of graph similarity* that compares different similarity measures, determines their algorithmic and semantic properties, and thus gives us a better understanding of their suitability for various kinds of applications. We see our paper as one contribution to such a theory.

Maybe the simplest graph similarity measure is based on graph edit distance: the *edit distance* between graphs G, H (for simplicity of the same order) is the minimum number of edges that need to be added and deleted from G to obtain a graph ismorphic to H.

In this paper, we study the computational complexity of a natural class of similarity measures that generalize the similarity derived from edit distance of graphs. In general, we view similarity as proximity with respect to some metric. A common way of converting a graph metric d into a similarity measure s is to let $s(G, H) \coloneqq \exp(-\beta \cdot d(G, H))$ for some constant $\beta > 0$. For our considerations the transformation between distance and similarity is irrelevant, so we focus directly on the metrics. In terms of computational complexity, computing similarity tends to be a hard algorithmic problem. It is known that computing edit distance, exactly or approximately, is NP-hard [3, 13, 18, 27] even on very restricted graph classes. In fact, the problem is closely related to the quadratic assignment problem [22, 26], which is notorious for being a very hard combinatorial optimization problem also for practical, heuristic approaches. Within the spectrum of similarities, the "limit" case of graph isomorphism shows that overall the complexity of graph similarity is far from trivial.

The metrics we study in this paper are based on minimizing the mismatch between two graphs. For graphs G, H with the same vertex set V, we define their *mismatch graph* $G - H$ to be the graph with vertex set V and edge set $E(G) \triangle E(H)$, the symmetric difference between the edge sets of the graphs. We assign *signs* to edges of the mismatch graph to indicate which graph they come from, say, a positive sign to the edges in $E(G) \setminus E(H)$ and a negative sign to the edges in $E(H) \setminus E(G)$. To quantify the mismatch between our graphs we introduce a *mismatch norm* μ on signed graphs that satisfies a few basic axioms such as subadditivity as well as invariance under permutations and under flipping the signs of all edges. Now for graphs G, H, not necessarily with the same vertex set, but for simplicity of the same order,[1] we define the distance $\mathrm{dist}_\mu(G, H)$ to be the minimum of $\mu(G^\pi - H)$, where π ranges over all bijective mappings from $V(G)$ to $V(H)$ and G^π is the image of G under π.

The simplest mismatch norm μ_{ed} just counts the number of edges of the mismatch graph $G - H$, ignoring their signs. Then the associated distance $\mathrm{dist}_{\mathrm{ed}}(G, H)$ is the edit distance between G and H. (Note that we write $\mathrm{dist}_{\mathrm{ed}}$ instead of the clunky $\mathrm{dist}_{\mu_{\mathrm{ed}}}$; we will do the same for other mismatch norms discussed here.) Another simple yet interesting mismatch norm is μ_{deg} measuring the maximum degree of the mismatch graph, again ignoring the signs of the edges. Then $\mathrm{dist}_{\mathrm{deg}}(G, H)$ measures how well we can align the two graphs in order to minimize the "local mismatch" at every node. Hence an alignment where at every vertex there is a mismatch of one edge yields a smaller $\mathrm{dist}_{\mathrm{deg}}$ than an alignment that is perfect at all nodes expect one, where it has a mismatch of, say, $n/2$, where n is the number of vertices. For edit distance it is the other way round. Depending on the application, one or the other may be preferable. Another well-known graph metric that can be described via the mismatch graph is Lovász's *cut distance* (see [19, Chapter 8] and Section 3 of this paper). And, last but not least, for the mismatch norm μ_{iso} defined to be 0 if the mismatch graph has no edges and 1 otherwise, $\mathrm{dist}_{\mathrm{iso}}(G, H)$ is 0 if G and H are isomorphic, and 1 otherwise, so computing the distance between two graphs amounts to deciding if they are isomorphic.

Mathematically, the framework of mismatch norms and the associated distances is best described in terms of the adjacency matrices of the graphs; the adjacency matrix A_{G-H} of the mismatch graph (viewed as a matrix with entries $0, +1, -1$ displaying the signs of the

[1] A general definition that also applies to graphs of distinct order can be found in Section 3, but for the hardness results we prove in this paper we can safely restrict our attention to graphs of the same order; this only makes the results stronger.

edges) is just the difference $A_G - A_H$. Then mismatch norms essentially are just matrix norms applied to to A_{G-H}. It turns out that "global norms" such as edit distance and direct generalizations correspond to entrywise matrix norms (obtained by applying a vector norm to the flattened matrix), and "local norms" such as μ_{deg} correspond to operator matrix norms (see Section 3). Cut distance corresponds to the cut norm of matrices. Instead of adjacency matrices, we can also consider the Laplacian matrices, exploiting that $L_{G-H} = L_G - L_H$, and obtain another interesting family of graph distances.

For every mismatch norm μ, we are interested in the problem $\textsc{Dist}_\mu$ of computing $\text{dist}_\mu(G, H)$ for two given graphs G, H. Note that this is a minimization problem where the feasible solutions are the bijective mappings between the vertex sets of the two graphs. It turns out that the problem is hard for most mismatch norms μ, in particular almost all the natural choices discussed above. The single exception is μ_{iso} related to the graph isomorphism problem. Furthermore, the hardness results usually hold even if the input graphs are restricted to be very simple, for example trees or bounded degree graphs. For edit distance and the related distance based on entrywise matrix norms this was already known [3,13]. Our focus in this paper is on operator norms. We prove a number of hardness results for different graph classes. One of the strongest ones is the following (see Theorem 19). Here $\textsc{Dist}_p$ denotes the distance measure derived from the ℓ_p-operator norm.

▶ **Theorem 1.** *For $1 \leq p \leq \infty$ there is a constant $c > 1$ such that, unless* P = NP*, $\textsc{Dist}_p$ has no factor-c approximation algorithm even if the input graphs are restricted to be trees of bounded degree.*

For details and additional results, we refer to Section 4 and 5.

Initially we aimed for a general hardness result that applies to all mismatch norms satisfying some additional natural conditions. However, we found that the hardness proofs, while following the same general strategies, usually have some intricacies exploiting special properties of the specific norms. Furthermore, for cut distance, none of these strategies seemed to work. Nevertheless, we were able to give a hardness proof for cut distance (Theorem 23) that is simply based on the hardness of computing the cut norm of the matrix [1]. This is remarkable in so far as usually the hard part of computing the distance is to find an optimal alignment π, whereas computing $\mu(G^\pi - H)$ is usually easy. For cut norm, it is even hard to compute $\mu(G^\pi - H)$ for a fixed alignment π.

Related Work

Graph similarity has mostly been studied in specific application areas, most importantly computer vision, data mining, and machine learning (see the references above). Of course not all similarity measures are based on mismatch. For example, metrics derived from vector embedding or graph kernels in machine learning (see [17]) provide a completely different approach (see [12] for a broader discussion). Of interest compared to our work (specifically for the ℓ_2-operator norm a.k.a. spectral norm) is the spectral approach proposed by Kolla, Koutis, Madan, and Sinop [16]. Intuitively, instead of the "difference" of two graphs that is described by our mismatch graphs, their approach is based on taking a "quotient".

The complexity of similarity, or "approximate graph isomorphism", or "robust graph isomorphism" has been studied in [2,3,13,15,16,18,27], mostly based on graph edit distance and small variations. Operator norms have not been considered in this context.

2 Preliminaries

We denote the class of real numbers by $\mathbb{R}$ and the nonnegative and positive reals by $\mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$, respectively. By $\mathbb{N}, \mathbb{N}_{>0}$ we denote the sets of nonnegative resp. positive integers. For every $n \in \mathbb{N}_{>0}$ we let $[n] := \{1, \ldots, n\}$.

We will consider matrices with real entries and with rows and columns indexed by arbitrary finite sets. Formally, for finite sets V, W, a $V \times W$ matrix is a function $A \colon V \times W \to \mathbb{R}$. A standard $m \times n$-matrix is just an $[m] \times [n]$-matrix. We denote the set of all $V \times W$ matrices by $\mathbb{R}^{V \times W}$, and we denote the entries of a matrix A by A_{vw}.

For a square matrix $A \in \mathbb{R}^{V \times V}$ and injective mapping $\pi : V \to W$, we let A^π be the $V^\pi \times V^\pi$-matrix with entries $A_{v^\pi w^\pi} := A_{vw}$. Note that we apply the mapping π from the right and denote the image of v under π by v^π. If $\rho : W \to X$ is another mapping, we denote the composition of π and ρ by $\pi\rho$. We typically use this notation for mappings between matrices and between graphs.

We use a standard graph theoretic notation. We denote the vertex set of a graph G by $V(G)$ and the edge set by $E(G)$. Graphs are finite, simple, and undirected, that is, $E(G) \subseteq \binom{V(G)}{2}$. We denote edges by vw instead of $\{v, w\}$. The *order* of a graph is $|G| := |V(G)|$. The *adjacency matrix* $A_G \in \{0,1\}^{V(G) \times V(G)}$ is defined in the usual way. We denote the class of all graphs by $\mathcal{G}$.

Let G be a graph with vertex set $V := V(G)$. For a mapping $\pi : V \to W$ we let G^π be the graph with vertex set $\{v^\pi \mid v \in V\}$ and edge set $\{v^\pi w^\pi \mid vw \in E(G) \text{ with } v^\pi \neq w^\pi\}$. Then $A_{G^\pi} = A_G^\pi$ if π is injective.

For graphs G, H, we denote the set of all injective mappings $\pi : V(G) \to V(H)$ by $\text{Inj}(G, H)$. Graphs G and H are *isomorphic* (we write $G \cong H$) if there is an $\pi \in \text{Inj}(G, H)$ such that $G^\pi = H$. We think of the mappings in $\text{Inj}(G, H)$, in particular if $|G| = |H|$ and they are bijective, as *alignments* between the graphs. Intuitively, to measure the distance between two graphs, we will align them in an optimal way to minimize the mismatch.

3 Graph Metrics Based on Mismatches

A *graph metric* is a function $\delta : \mathcal{G} \times \mathcal{G} \to \mathbb{R}_{\geq 0}$ such that

(GM0) $\delta(G, H) = \delta(G', H')$ for all G, G', H, H' such that $G \cong G'$ and $H \cong H'$;
(GM1) $\delta(G, G) = 0$ for all G;
(GM2) $\delta(G, H) = \delta(H, G)$ for all G, H;
(GM3) $\delta(F, H) \leq \delta(F, G) + \delta(G, H)$ for all F, G, H.

Note that we do not require $\delta(G, H) > 0$ for all $G \neq H$, not even for $G \not\cong H$, so strictly speaking this is just a *pseudometric*. We are interested in the complexity of the following problem:

DIST_δ
Instance: Graphs G, H, $p, q \in \mathbb{N}_{>0}$
Problem: Decide if $\delta(G, H) \geq \frac{p}{q}$

A *signed graph* is a weighted graph with edge weights $-1, +1$, and for every edge e of a signed graph we denote its *sign* by $\text{sg}(e)$. For a signed graph Δ, we let $E_+(\Delta) := \{e \in E(\Delta) \mid \text{sg}(e) = +1\}$ and $E_-(\Delta) := \{e \in E(\Delta) \mid \text{sg}(e) = -1\}$. *Isomorphisms* of signed graphs must preserve signs. A signed graph Δ is a subgraph of a signed graph Γ (we write $\Delta \subseteq \Gamma$) if $V(\Delta) \subseteq V(\Gamma)$, $E_+(\Delta) \subseteq E_+(\Gamma)$, and $E_-(\Delta) \subseteq E_-(\Gamma)$. We denote the class of all signed graphs by $\mathcal{S}$.

For every $\Delta \in \mathcal{S}$, we let $-\Delta$ be the signed graph obtained from Δ by flipping the signs of all edges. We define the *sum* of $\Delta, \Gamma \in \mathcal{S}$ to be the signed graph $\Delta+\Gamma$ with vertex set $V(\Delta+\Gamma) = V(\Delta) \cup V(\Gamma)$ and signed edge sets $E_+(\Delta+\Gamma) = \big(E_+(\Delta) \cup E_+(\Gamma)\big) \setminus \big(E_-(\Delta) \cup E_-(\Gamma)\big)$ and $E_-(\Delta+\Gamma) = \big(E_-(\Delta) \cup E_-(\Gamma)\big) \setminus \big(E_+(\Delta) \cup E_+(\Gamma)\big)$. The adjacency matrix A_Δ of a signed graph Δ displays the signs of the edges, so $A_\Delta \in \{0,1,-1\}^{V(\Delta)\times V(\Delta)}$ with $(A_\Delta)_{vw} = \operatorname{sg}(vw)$ if $vw \in E(\Delta)$ and $(A_\Delta)_{vw} = 0$ otherwise. Note that $A_{-\Delta} = -A_\Delta$ for all $\Delta \in \mathcal{S}$ and $A_{\Delta+\Gamma} = A_\Delta + A_\Gamma$ for all $\Delta, \Gamma \in \mathcal{S}$ with $V(\Delta) = V(\Gamma)$ and $E_+(\Delta) \cap E_+(\Gamma) = \emptyset$ and $E_-(\Delta) \cap E_-(\Gamma) = \emptyset$.

The *mismatch graph* of two graphs G, H is the signed graph $G - H$ with vertex set $V(G-H) := V(G) \cup V(H)$ and signed edge set $E_+(G-H) := E(G) \setminus E(H)$, $E_-(G-H) := E(H) \setminus E(G)$. Note that if $V(G) = V(H)$ then for the adjacency matrices we have $A_{G-H} = A_G - A_H$. Observe that every signed graph Δ is the mismatch graph of the graphs $\Delta_+ := (V(\Delta), E_+(\Delta))$ and $\Delta_- := (V(\Delta), E_-(\Delta))$.

A *mismatch norm* is a function $\mu : \mathcal{S} \to \mathbb{R}_{\geq 0}$ satisfying the following conditions:

(MN0) $\mu(\Delta) = \mu(\Gamma)$ for all $\Delta, \Gamma \in \mathcal{S}$ such that $\Delta \cong \Gamma$;

(MN1) $\mu(\Delta) = 0$ for all $\Delta \in \mathcal{S}$ with $E(\Delta) = \emptyset$;

(MN2) $\mu(\Delta) = \mu(-\Delta)$ for all Δ;

(MN3) $\mu(\Delta+\Gamma) \leq \mu(\Delta)+\mu(\Gamma)$ for all $\Delta, \Gamma \in \mathcal{S}$ with $V(\Delta) = V(\Gamma)$ and $E_+(\Delta) \cap E_+(\Gamma) = \emptyset$ and $E_-(\Delta) \cap E_-(\Gamma) = \emptyset$.

(MN4) $\mu(\Delta) = \mu(\Gamma)$ for all $\Delta, \Gamma \in \mathcal{S}$ with $E_+(\Delta) = E_+(\Gamma)$ and $E_-(\Delta) = E_-(\Gamma)$;

For every mismatch norm μ, we define $\operatorname{dist}_\mu : \mathcal{G} \times \mathcal{G} \to \mathbb{R}_{\geq 0}$ by

$$\operatorname{dist}_\mu(G,H) := \begin{cases} \min_{\pi \in \operatorname{Inj}(G,H)} \mu(G^\pi - H) & \text{if } |G| \leq |H|, \\ \min_{\pi \in \operatorname{Inj}(H,G)} \mu(G - H^\pi) & \text{if } |H| < |G|. \end{cases}$$

We write $\textsc{Dist}_\mu$ instead of $\textsc{Dist}_{\operatorname{dist}_\mu}$ to denote the algorithmic problem of computing dist_μ for two graphs G, H.

▶ **Lemma 2.** *For every mismatch norm μ the function dist_μ is a graph metric.*

Proof. Conditions (GM0), (GM1), (GM2) follow from (MN0), (MN1), (MN2), respectively. To prove (GM3), let F, G, H be graphs. Padding the graphs with isolated vertices, by (MM4) we may assume that $|F| = |G| = |H|$. By (MN0) we may further assume that $V(F) = V(G) = V(H) := V$. Choose $\pi, \rho \in \operatorname{Inj}(V,V)$ such that $\operatorname{dist}_\mu(F,G) = \mu(F^\pi - G)$ and $\operatorname{dist}_\mu(G,H) = \mu(G^\rho - H)$.

Then by (MN0) we have

$$\mu(F^{\pi\rho} - G^\rho) = \mu\big((F^\pi - G)^\rho\big) = \mu(F^\pi - G) = \operatorname{dist}_\mu(F,G).$$

Now consider the two mismatch graphs $\Delta := F^{\pi\rho} - G^\rho$ and $\Gamma := G^\rho - H$. We have $E_+(\Delta) = E(F^{\pi\rho}) \setminus E(G^\rho)$ and $E_+(\Gamma) = E(G^\rho) \setminus E(H)$. Thus $E_+(\Delta) \cap E(G^\rho) = \emptyset$ and $E_+(\Gamma) \subseteq E(G^\rho)$, which implies $E_+(\Delta) \cap E_+(\Gamma) = \emptyset$. Similarly, $E_-(\Delta) \cap E_-(\Gamma) = \emptyset$. Moreover, $\Delta + \Gamma = F^{\pi\rho} - H$, because

$$A_{\Delta+\Gamma} = A_\Delta + A_\Gamma = (A_F^{\pi\rho} - A_G^\rho) + (A_G^\rho - A_H) = A_F^{\pi\rho} - A_H = A_{F^{\pi\rho}-H}.$$

Thus by (MN3),

$$\operatorname{dist}_\mu(F,H) \leq \mu(F^{\pi\rho} - H) = \mu(\Delta+\Gamma) \leq \mu(\Delta) + \mu(\Gamma) = \operatorname{dist}_\mu(F,G) + \operatorname{dist}_\mu(G,H).$$

This proves (GM3). ◀

▶ Remark 3. None of the five conditions (MN0)–(MN4) on a mismatch norm μ can be dropped if we want to guarantee that dist_μ is a graph metric, but of course we could replace them by other conditions. While (MN0)–(MN3) are very natural and directly correspond to conditions (GM0)–(GM3) for graph metrics, condition (MN4) is may be less so. We chose (MN4) as the simplest condition that allows us to compare graphs of different sizes.

Having said this, it is worth noting that (MN0), (MN1) and (MN3) imply (MN4) for graphs Δ, Γ with $|\Delta| = |\Gamma|$. For graphs Δ, Γ with $|\Delta| < |\Gamma|$ they only imply $\mu(\Delta) \geq \mu(\Gamma)$. Thus as long as we only compare graphs of the same order, (MN4) is not needed. In particular, since our hardness results always apply to graphs of the same order, (MN4) is inessential for the rest of the paper.

However, it is possible to replace (MN4) by other natural conditions. For example, Lovász's metric based on a scaled cut norm [20] does not satisfy (MN4) and instead uses a blowup of graphs to a common size to compare graphs of different sizes.

Let us now consider a few examples of mismatch norms.

▶ **Example 4** (Isomorphism). The mapping $\iota \coloneqq \mathcal{S} \to \mathbb{R}_{\geq 0}$ defined by $\iota(\Delta) \coloneqq 0$ if $E(\Delta) = \emptyset$ and $\iota(\Delta) \coloneqq 1$ otherwise is a mismatch norm. Under the metric dist_ι, all nonisomorphic graphs have distance 1 (and isomorphic graphs have distance 0, as they have under all graph metrics).

▶ **Example 5** (Matrix Norms). Recall that a matrix (pseudo) norm $\|\cdot\|$ associates with every matrix A (say, with real entries) an nonnegative real $\|A\|$ in such a way that $\|N\| = 0$ for matrices N with only 0-entries, $\|aA\| = |a| \cdot \|A\|$ for all matrices A and reals $a \in \mathbb{R}$, and $\|A + B\| \leq \|A\| + \|B\|$ for all matrices A, B of the same dimensions.

Actually, we are only interested in square matrices here. We call a matrix norm $\|\cdot\|$ *permutation invariant* if for all $A \in \mathbb{R}^{V\times V}$ and all bijective $\pi : V \to V$ we have $\|A\| = \|A^\pi\|$. It is easy to see that for every permutation invariant matrix norm $\|\cdot\|$, the mapping $\mu_{\|\cdot\|} : \mathcal{S} \to \mathbb{R}_{\geq 0}$ defined by $\mu_{\|\cdot\|}(\Delta) \coloneqq \|A_\Delta\|$ satisfies (MN0)–(MN3).

We call a permutation invariant matrix norm $\|\cdot\|$ *paddable* if it is invariant under extending matrices by zero entries, that is, $\|A\| = \|A'\|$ for all $A \in \mathbb{R}^{V\times V}$, $A' \in \mathbb{R}^{V'\times V'}$ such that $V' \supseteq V$, $A_{vw} = A'_{vw}$ for all $v, w \in V$, and $A'_{vw} = 0$ if $v \in V' \setminus V$ or $w \in V' \setminus V$. A paddable matrix norm also satisfies (MN4).

The following common matrix norms are paddable (and thus by definition also invariant). Let $1 \leq p \leq \infty$ and $A \in \mathbb{R}^{V\times V}$.

1. *Entrywise p-norm:* $\|A\|_{(p)} \coloneqq \left(\sum_{v,w\in V} |A_{vw}|^p\right)^{\frac{1}{p}}$. The best-known special case is the *Frobenius norm* $\|\cdot\|_F \coloneqq \|\cdot\|_{(2)}$.
2. *ℓ_p-operator norm:* $\|A\|_p \coloneqq \sup_{\boldsymbol{x}\in\mathbb{R}^V\setminus\{\boldsymbol{0}\}} \frac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}$, where the vector p-norm is defined by $\|\boldsymbol{a}\|_p \coloneqq \left(\sum_{v\in V} a_v^p\right)^{\frac{1}{p}}$. In particular, $\|A\|_2$ is known as the *spectral norm.*
3. *Absolute ℓ_p-operator norm:* $\|A\|_{|p|} \coloneqq \|\mathrm{abs}(A)\|_p$, where abs takes entrywise absolute values. For the mismatch norm, taking entrywise absolute values means that we ignore the signs of the edges.
4. *Cut norm:* $\|A\|_\square \coloneqq \max_{S,T\subseteq V} \left|\sum_{v\in S,w\in T} A_{vw}\right|$.

▶ **Example 6** (Laplacian Matrices). Recall that the *Laplacian matrix* of a weighted graph G with vertex set $V \coloneqq V(G)$ is the $V \times V$ matrix L_G with off-diagonal entries $(L_G)_{vw}$ being the negative weight of the edge $vw \in E(G)$ if there is such an edge and 0 otherwise and diagonal entries $(L_G)_{vv}$ being the sum of the weights of all edges incident with v. For an unweighted graph we have $L_G = D_G - A_G$, where D_G is the diagonal matrix with the vertex degrees as diagonal entries.

Observe that for a signed graph $\Delta = G - H$ we have $L_\Delta = L_G - L_H$.

It is easy to see that for every paddable matrix norm $\|\cdot\|$, the function $\mu^L_{\|\cdot\|} : \mathcal{S} \to \mathbb{R}_{\geq 0}$ defined by $\mu^L_{\|\cdot\|}(\Delta) := \|L_\Delta\|$ is a mismatch norm.

Clearly, DIST_μ is not a hard problem for every mismatch norm. For example, DIST_ν is trivial for the trivial mismatch norm ν defined by $\nu(\Delta) := 0$ for all Δ, and DIST_ι for ι from Example 4 is equivalent to the graph isomorphism problem and hence in quasipolynomial time [4].

However, for most natural matrix norms the associated graph distance problem is NP-hard. In particular, for every $p \in \mathbb{R}_{>0}$, this holds for the metric $\text{dist}_{(p)}$ based on the entrywise p-norm $\|\cdot\|_{(p)}$.

▶ **Theorem 7** ([13]). *For $p \in \mathbb{R}_{>0}$, $\text{DIST}_{(p)}$ is NP-hard even if restricted to trees or bounded-degree graphs.*

The proof in [13] is only given for the Frobenius norm, that is, $\text{DIST}_{(2)}$, but it actually applies to all p. In the rest of the paper, we study the complexity of DIST_p, $\text{DIST}_{|p|}$ and $\text{DIST}_\square$.

4 Complexity for Operator Norms

In this section we investigate the complexity of DIST_p and $\text{DIST}_{|p|}$ for $1 \leq p \leq \infty$. However, as $\mu_{|p|}$ would only be a special case within the upcoming proofs, we omit to mention it explicitly. We also omit to specify the possible values for p.

Given graphs G and H, an alignment from G to H, and a node $v \in V(G)$ we refer to the nodes whose adjacency is not preserved by π as the *π-mismatches of v*. If π is clear from the context we might omit it. We call $G^\pi - H$ the *mismatch graph of π*.

For all ℓ_p-operator norms the value of $\mu_p(G^\pi - H)$ strongly depends on the maximum degree in the mismatch graph of π. We capture this property with the following definition.

▶ **Definition 8.** *Let G, H be graphs of the same order and π an alignment from G to H. The* π-mismatch count (π-MC) *of a node $v \in V(G)$ is defined as:*

$$\text{MC}(v, \pi) := |\{w \in V(G) \mid w \text{ is a } \pi\text{-mismatch of } v\}|.$$

We use MC for nodes in H analogously. The *maximum mismatch count (π-*MMC*)* of π is defined as:

$$\text{MMC}(\pi) := \max_{v \in V(G)} \text{MC}(v, \pi).$$

Again we might drop the π if it is clear from the context. Note that the MMC corresponds to the maximum degree in the mismatch graph and that we use a slightly abbreviated notation in which we assume the graphs are given by the alignment.

The ℓ_1-operator norm and the ℓ_∞-operator norm measure exactly the maximum mismatch count. Due to the relation between the ℓ_p-operator norms we can derive an upper bound for μ_p. The proof of Lemma 9 can be found in the full version [11].

▶ **Lemma 9.** *Let G, H be graphs of the same order and π an alignment from G to H. Then*

$$\mu_1(G^\pi - H) = \mu_\infty(G^\pi - H) = \text{MMC}(\pi),$$
$$\mu_p(G^\pi - H) \leq \text{MMC}(\pi).$$

Next, we observe that μ_p is fully determined by the connected component of the mismatch graph with the highest mismatch norm. The proof of Lemma 10 can be found in the full version [11].

▶ **Lemma 10.** *Let G, H be graphs of the same order, π an alignment from G to H, and $\mathcal{C}$ the set of all connected components in $G^\pi - H$. Then*

$$\mu_p(G^\pi - H) = \max_{C \in \mathcal{C}} \mu_p(C).$$

For the sake of readability, we introduce a function as abbreviation for our upcoming bounds.

▶ **Definition 11.** *For all $1 \leq p \leq \infty$ we define the function* bound_p *as follows:*

$$\text{bound}_p(c) \coloneqq \max\left(c^{1/p}, c^{1-1/p}\right).$$

In particular, $\text{bound}_2(c) = \sqrt{c}$. Now we derive our lower bound. The proof of Lemma 12 can be found in the full version [11].

▶ **Lemma 12.** *Let G, H be graphs of the same order, π an alignment from G to H, and $v \in V(G)$.*
If v has at least c mismatches, then

$$\mu_p(G^\pi - H) \geq \text{bound}_p(c).$$

If $G^\pi - H$ is a star, then

$$\mu_p(G^\pi - H) = \text{bound}_p(\text{MMC}(\pi)).$$

While μ_1 and μ_∞ simply measure the MMC, μ_2 also considers the connectedness of the mismatches around the node with the highest MC. As Lemma 9 and Lemma 12 tell us, μ_2 ranges between the $\sqrt{\text{MMC}}$ and the MMC. In fact, the lower bound is tight for stars and the upper bound is tight for regular graphs. This is intuitive as these are the extreme cases in which no mismatch can be removed/added without decreasing/increasing the MMC, respectively. Other ℓ_p-operator norms interpolate between μ_2 and μ_1 / μ_∞ in terms of how much they value the connectedness within the mismatch component of the node with the highest MC.

The lower bound for $\mu_p(G^\pi - H)$ gives us lower bound for $\text{dist}_p(G, H)$.

▶ **Lemma 13.** *Let G, H be graphs of the same order and π an alignment from G to H. If all alignments from G to H have a node with at least c mismatches, then*

$$\text{dist}_p(G, H) \geq \text{bound}_p(c).$$

Proof. This follows directly from the first claim of Lemma 12. ◀

The following upper bound might seem to have very restrictive conditions but is actually used in several hardness proofs.

▶ **Lemma 14.** *Let G, H be graphs of the same order and π an alignment from G to H. If the mismatch graph of π consists only of stars, then*

$$\text{dist}_p(G, H) \leq \mu_p(G^\pi - H) = \text{bound}_p(\text{MMC}(\pi)).$$

Proof. This follows directly from Lemma 10 and the second claim of Lemma 12. ◀

The last tool we need to prove the hardness is that we can distinguish two alignments by their MMC as long as the mismatch graph of the alignment with the lower MMC consists only of stars. The proof of Lemma 15 can be found in the full version [11].

▶ **Lemma 15.** *For all $c, d \in \mathbb{N}$ with $c < d$ it holds that* $\text{bound}_p(c) < \text{bound}_p(d)$.

The graph isomorphism problem becomes solvable in polynomial time if restricted to graphs of bounded degree [21]. In contrast to this, DIST_F is NP-hard even under this restriction [13]. We show that the same applies to DIST_p.

The reduction in the hardness proof works for any mismatch norm which can, said intuitively, distinguish the mismatch norm of the 1-regular graph of order n from any other n-nodes mismatch graph in which every node has at least one -1 edge but at least one node has an additional -1 edge and $+1$ edge. In particular, the construction also works for $\text{DIST}_{(p)}$. However, it does not work for the cut-distance, for which we independently prove the hardness in Section 6.

▶ **Theorem 16.** *DIST_p and $\text{DIST}_{|p|}$ are NP-hard for $1 \leq p \leq \infty$ even if both graphs have bounded degree.*

Proof. The proof is done by reduction from the NP-hard Hamiltonian Cycle problem in 3-regular graphs (HAM-CYCLE) [10]. Given a 3-regular graph G of order n as an instance of HAM-CYCLE, the reduction uses the n-nodes cycle C_n and G as inputs for DIST_p. We claim G has a Hamiltonian cycle if and only if $\text{dist}_p(C_n, G) \leq \text{bound}_p(1)$.

Assume that G has a Hamiltonian cycle. Then there exists a bijection $\pi : V(C_n) \to V(G)$ that aligns the cycle C_n perfectly with the Hamiltonian cycle in G. Each node in G has three neighbors, two of which are matched correctly by π as they are part of the Hamiltonian cycle. Hence, each node has a π-mismatch count of 1 and obviously the MMC of π is 1 as well. According to Lemma 14, we get $\text{dist}_p(C_n, G) \leq \text{bound}_p(1)$.

Conversely, assume that G has no Hamiltonian cycle. Then, for any alignment π' from C_n to G, there exists at least one edge vw in C_n that is not mapped to an edge in G. Hence, only one of the three edges incident to $\pi'(v)$ in G can be matched correctly. In total, v has at least one mismatch from C_n to G and two mismatches from G to C_n, which implies $\text{MC}(v, \pi') \geq 3$. Using Lemma 13, we get $\text{dist}_p(C_n, G) \geq \text{bound}_p(3)$. And then $\text{dist}_p(C_n, G) > \text{bound}_p(1)$ according to Lemma 15. ◀

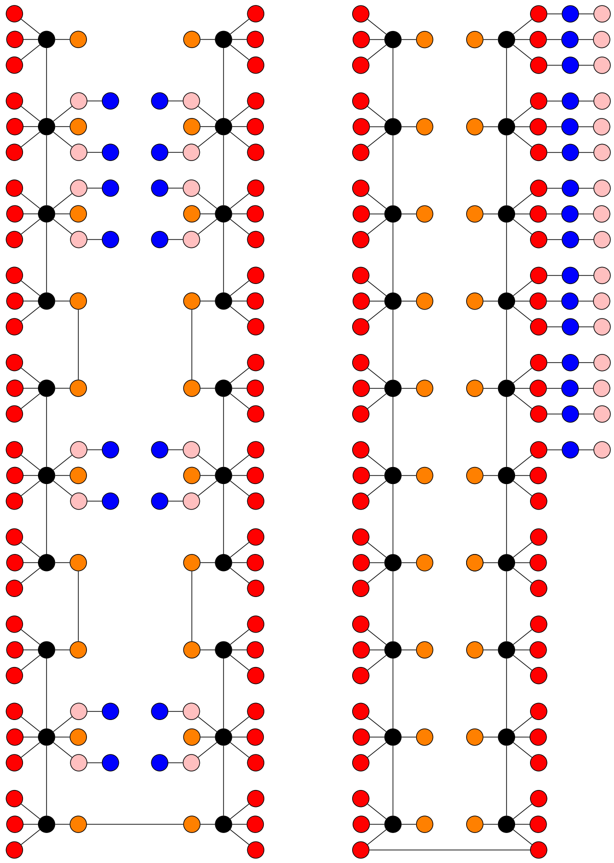
Next, we modify the construction to get an even stronger NP-hardness result. The proof can be found in the full version [11].

▶ **Theorem 17.** *DIST_p and $\text{DIST}_{|p|}$ are NP-hard for $1 \leq p \leq \infty$ even if restricted to a path and a graph of maximum degree 3.*

Similar to bounded degree input graphs, restricting graph isomorphism to trees allows it to be solved in polynomial time [24] but DIST_F is NP-hard for trees [13]. We show that DIST_p remains NP-hard for trees even when applying the bounded degree restriction simultaneously.

▶ **Theorem 18.** *DIST_p and $\text{DIST}_{|p|}$ are NP-hard for $1 \leq p \leq \infty$ even if restricted to bounded-degree trees.*

Proof. The proof is done by reduction from the NP-hard THREE-PARTITION problem [10] that is defined as follows. Given the integers A and $a_1, \dots, a_{3m}$ in unary representation, such that $\sum_{i=1}^{3m} a_i = mA$ and $A/4 < a_i < A/2$ for $1 \leq i \leq 3m$, decide whether there exists a



(a) T_1.   **(b)** T_2.

Figure 1 Example of the construction in the proof of Theorem 18 where m=2, A=10, a_1=a_2=4 and a_3=a_4=a_5=a_6=3. Best viewed in color.

partition of $a_1, \ldots, a_{3m}$ into m groups of size 3 such that the elements in each group sum up to exactly A. For technical reasons, we restrict the reduction to $A \geq 8$. However, the ignored cases are trivial. Precisely, for $A \in \{6, 7\}$ the answer is always YES and for $A \in [5]$ there exist no valid instances.

Given an instance of THREE-PARTITION with $A \geq 8$, we construct two trees T_1 and T_2 such that the given instance has answer YES, if and only if $\text{dist}_p(T_1, T_2) \leq \text{bound}_p(2)$. Figure 1 shows an example of T_1 and T_2 for $m = 2$ and $A = 10$. For illustrative reasons we assign each node a color during the construction. The colors are used in the example and we will refer to certain nodes by their color later in the proof. However, they do not restrict the possible alignments in any way.

Initialize T_1 as the disjoint union of paths $p_1^1, \ldots, p_{3m}^1$ such that p_i^1 has a_i black nodes; initialize T_2 as the disjoint union of paths $p_1^2, \ldots, p_m^2$ consisting of A black nodes each. In the following we refer to one endpoint of p_i^k as $e_1(p_i^k)$ and the other endpoint as $e_2(p_i^k)$. We attach three new red leaves and one new orange leaf to each black node in both T_1 and T_2. Next we modify the graphs into trees by connecting the paths. For $1 \leq i \leq 3m - 1$ we add an edge between the orange leaf adjacent to $e_1(p_i^1)$ and the orange leaf adjacent to $e_2(p_{i+1}^1)$. For $1 \leq i \leq m - 1$ we add an edge between one of the red leaves adjacent to $e_1(p_i^2)$ and one of the red leaves adjacent to $e_2(p_{i+1}^2)$.

Next we attach two new pink leaves to each inner (non-endpoint) path node in T_1 and attach a new blue leaf to each pink node. Then we add the same number of blue nodes to T_2 and connect each blue node to one of the red nodes with degree 1. Finally, we attach a new pink leaf to each blue node. Note that both T_1 and T_2 are trees with bounded degree. Precisely, the highest degree in T_1 is 8 and 6 in T_2 independent of the problem instance.

Intuitively, the construction ensures that every inner path node has already 2 mismatches just because of the degree difference. If there is no partition, at least one path in T_1 cannot be mapped contiguously into a path in T_2 which raises the MC of some inner path node to at least 3. Simultaneously, the construction ensures that there is an alignment for which the mismatch graph consists only of stars with maximum degree 2 if there is an alignment.

The formal continuation of this proof can be found in the full version [11]. ◀

5 Approximability for Operator Norms

In this section we investigate the approximability of Dist_p and $\text{Dist}_{|p|}$ for $1 \leq p \leq \infty$. Again, we omit specifying the possible values for p and mentioning $\text{Dist}_{|p|}$ explicitly as the proofs work the same for it. We consider the following two possibilities to measure the error of an approximation algorithm for a minimization problem. An algorithm $\mathcal{A}$ has *multiplicative error* $\alpha > 1$, if for any instance $\mathcal{I}$ of the problem with an optimum $\text{OPT}(\mathcal{I})$, $\mathcal{A}$ outputs a solution with value $\mathcal{A}(\mathcal{I})$ such that $\text{OPT}(\mathcal{I}) \leq \mathcal{A}(\mathcal{I}) \leq \alpha\text{OPT}(\mathcal{I})$. In this case we call $\mathcal{A}$ an *α-approximation algorithm*. An algorithm $\mathcal{B}$ has *additive error* $\varepsilon > 0$, if $\text{OPT}(\mathcal{I}) \leq \mathcal{B}(\mathcal{I}) \leq \text{OPT}(\mathcal{I}) + \varepsilon$ for any instance $\mathcal{I}$.

Approximating Dist_p with multiplicative error is at least as hard as the graph isomorphism problem (GI) since such an approximation algorithm $\mathcal{A}$ could be used to decide GI considering that $\mathcal{A}(G, H) = 0$ if and only if G is isomorphic to H.

Furthermore, we can deduce thresholds under which the α-approximation is NP-hard using the gap between the decision values of the reduction in each hardness proof from Section 4.

▶ **Theorem 19.** *For $1 \leq p \leq \infty$ and any $\varepsilon > 0$, unless $\mathsf{P} = \mathsf{NP}$, there is no polynomial time approximation algorithm for Dist_p or $\text{Dist}_{|p|}$ with a multiplicative error guarantee of*

1. $\text{bound}_p(3) - \varepsilon$, *even if both input graphs have bounded degree,*
2. $\text{bound}_p(2) - \varepsilon$, *even if one input graph is a path and the other one has bounded degree,*
3. $\frac{\text{bound}_p(3)}{\text{bound}_p(2)} - \varepsilon$, *even if both input graphs are trees with bounded degree.*

Proof. We recall the proof of Theorem 16. If G has a Hamiltonian cycle, then $\text{dist}_p(C_n, G) \leq \text{bound}_p(1) = 1$. Otherwise $\text{dist}_p(C_n, G) \geq \text{bound}_p(3)$. Assume there is a polynomial time approximation algorithm $\mathcal{A}$ with a multiplicative error guarantee of $\text{bound}_p(3) - \varepsilon$ for $\varepsilon > 0$. Then we can distinguish the two cases by checking whether $\mathcal{A}(C_n, G) < \text{bound}_p(3)$ and therefore decide Ham-Cycle on 3-regular graphs in polynomial time. The same argument can be used for the other bounds using the proofs of Theorem 17 and Theorem 18, respectively. ◀

In particular, this implies that there is no polynomial time approximation scheme (PTAS) for Dist_p or $\text{Dist}_{|p|}$ under the respective restrictions.

Next we show the additive approximation hardness by scaling up the gap between the two decision values of the reduction in the proof of Theorem 16. For this we replace each edge with a colored gadget and then modify the graph so that an optimal alignment has to be color-preserving. The proof of Theorem 20 can be found in the full version [11].

▶ **Theorem 20.** *For $1 \leq p \leq \infty$ there is no polynomial time approximation algorithm for Dist_p with any constant additive error guarantee unless $\mathsf{P} = \mathsf{NP}$.*

However, the approximation of $\textsc{Dist}_p$ becomes trivial once we restrict the input to graphs of bounded degree, although $\textsc{Dist}_p$ stays NP-hard under this restriction. The proof of Theorem 21 can be found in the full version [11].

▶ **Theorem 21.** *For $1 \leq p \leq \infty$, if one graph has maximum degree d, then there is a polynomial time approximation algorithm for $\textsc{Dist}_p$ and $\textsc{Dist}_{|p|}$ with*
1. *a constant additive error guarantee $2d$,*
2. *a constant multiplicative error guarantee $1 + 2d$.*

6 Complexity for Cut Norm

Finally, we show the hardness for $\textsc{Dist}_\square$ which corresponds to the cut distance $\hat{\delta}_\square$ (see [19, Chapter 8]. For any signed graph G and $V \subseteq V(G)$ the induced subgraph $G[V]$ is the signed graph with vertex set V, $E_+(G[V]) = \{vw \in E_+(G) \mid v, w \in V)\}$, and $E_-(G[V]) = \{vw \in E_-(G) \mid v, w \in V)\}$.

▶ **Lemma 22.** *Let Δ be a signed graph and $W \subseteq V(\Delta)$. Then $\mu_\square(\Delta) \geq \mu_\square(\Delta[W])$.*

Proof. Let $V := V(\Delta)$, $A := A_\Delta$, $B := A_{\Delta[W]}$ and $S', T' := \operatorname{argmax}_{S,T \subseteq W} \left|\sum_{v \in S, w \in T} B_{vw}\right|$. Then

$$\|\Delta[W]\|_\square = \left|\sum_{v \in S', w \in T'} B_{vw}\right| = \left|\sum_{v \in S', w \in T'} A_{vw}\right| \leq \max_{S,T \subseteq V} \left|\sum_{v \in S, w \in T} A_{vw}\right| = \|\Delta\|_\square . \qquad ◀$$

Our hardness proof for $\textsc{Dist}_\square$ is based on the hardness of computing the cut norm. Intuitively, the construction enforces a specific alignment by modifying the nodes degrees.

▶ **Theorem 23.** *The problem $\textsc{Dist}_\square$ is* NP*-hard.*

Proof. The proof is done by reduction from the NP-hard $\textsc{Max-Cut}$ problem on unweighted graphs [10]. First, we recall how Alan and Naor [1] construct a matrix A for any graph G so that $\|A\|_\square = \mathit{MAXCUT}(G)$. Orient G in an arbitrary manner, let $V(G) = \{v_1, v_2, \ldots, v_n\}$ and $E(G) = \{e_1, e_2, \ldots, e_m\}$. Then A is the $2m \times n$ matrix defined as follows. For each $1 \leq i \leq m$, if e_i is oriented from v_j to v_k, then $A_{2i-1,j} = A_{2i,k} = 1$ and $A_{2i-1,k} = A_{2i,j} = -1$. The rest of the entries in A are all 0.

Next, we observe that the matrix

$$B = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$$

has the property $\|B\|_\square = 2\,\|A\|_\square = 2 \cdot \mathit{MAXCUT}(G)$ since the cut norm is invariant under transposition and the two submatrices A, A^T have no common rows or columns in B.

We interpret B as the adjacency matrix of a signed graph Δ' and construct the two unsigned graphs $F' := (V(\Delta), E_+(\Delta))$, $H' := (V(\Delta), E_-(\Delta))$. Then $\mu_\square(F' - H') = \mu_\square(\Delta') = \|B\|_\square = 2 \cdot \mathit{MAXCUT}(G)$. Next, we modify F', H' into the graphs F, H by adding $i \cdot \left(\left\lceil \frac{n^2}{4} \right\rceil + n\right)$ leaves to node v_i for $1 \leq i \leq n$. The reduction follows from the claim that $\operatorname{dist}_\square(F, H) = 2 \cdot \mathit{MAXCUT}(G)$, which we prove in the following.

Let π be an alignment that maps v_i to v_i for $1 \leq i \leq n$ and each leaf in F to a leaf in H so that its adjacency is preserved; let Δ be the mismatch graph of π. Then $E(\Delta) = E(\Delta')$ and therefore $\mu_\square(F^\pi - H) = 2 \cdot \mathit{MAXCUT}(G)$. We conclude $\operatorname{dist}_\square(F, H) \leq 2 \cdot \mathit{MAXCUT}(G)$.

It remains to show that no alignment can lead to a lower mismatch norm. First, let σ be any alignment that maps v_i to v_i for $1 \leq i \leq n$; let Λ be the mismatch graph of σ. Then $\Lambda[\{v_1, \ldots, v_n\}] = \Delta'$ and we get $\mu_\square(\Lambda) \geq 2 \cdot MAXCUT(G)$ from Lemma 22.

Conversely, let ρ be any alignment that maps v_i to v_j for $i \neq j$ and $i, j \leq n$; let Γ be the mismatch graph of ρ. The number of leaves adjacent to v_i in F and to v_j in H differs at least by $\left\lceil \frac{n^2}{4} \right\rceil + n$. Without restriction we can assume there are at least $l := \left\lceil \frac{n^2}{4} \right\rceil$ leaves $w_1, \ldots, w_l$ adjacent to v_i in F that are mismatched by ρ. Let $S := \{v_i^\rho, w_1^\rho, \ldots, w_l^\rho\}$. Then $\Gamma[S]$ has exactly l edges all of which have the same sign. It is easy to see that $\mu_\square(\Gamma[S]) = 2l$ which implies $\mu_\square(\Gamma) \geq 2l$ according to Lemma 22. We chose l so that $2l \geq \frac{n^2}{2} \geq 2 \cdot MAXCUT(G)$. After considering all alignments, we get $\text{dist}_\square(F, G) \geq 2 \cdot MAXCUT(G)$. This proves our claim and therefore concludes the reduction. ◀

7 Concluding Remarks

We study the computational complexity of a class of graph metrics based on mismatch norms, or equivalently, matrix norms applied to the difference of the adjacency matrices of the input graphs under an optimal alignment between the vertex sets. We find that computing the distance between graphs under these metrics (at least for the standard, natural matrix norms) is NP-hard, often already on simple input graphs such as trees. This was essentially known for entrywise matrix norms. We prove it for operator norms and also for the cut norm.

We leave it open to find (natural) general conditions on a mismatch norm such that the corresponding distance problem becomes hard. Maybe more importantly, we leave it open to find meaningful tractable relaxations of the distance measures.

Measuring similarity via mismatch norms is only one approach. There are several other, fundamentally different ways to measure similarity. We are convinced that graph similarity deserves a systematic and general theory that compares the different approaches and studies their semantic as well as algorithmic properties. Our paper is one contribution to such a theory.

References

1. Noga Alon and Assaf Naor. Approximating the cut-norm via grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803, April 2006. doi:10.1137/S0097539704441629.
2. S. Arora, A. Frieze, and H. Kaplan. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. *Mathematical Programming*, 92(1):1–36, 2002.
3. V. Arvind, J. Köbler, S. Kuhnert, and Y. Vasudev. Approximate graph isomorphism. In *Proceedings of the 37th International Symposium on Mathematical Foundations of Computer Science (MFCS 2012)*, volume 7464, pages 100–111, 2012.
4. László Babai. Graph isomorphism in quasipolynomial time [extended abstract]. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 684–697. ACM, 2016. doi:10.1145/2897518.2897542.
5. Amélie Barbe, Marc Sebban, Paulo Gonçalves, Pierre Borgnat, and Rémi Gribonval. Graph diffusion wasserstein distances. In Frank Hutter, Kristian Kersting, Jefrey Lijffijt, and Isabel Valera, editors, *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2020, Ghent, Belgium, September 14-18, 2020, Proceedings, Part II*, volume 12458 of *Lecture Notes in Computer Science*, pages 577–592. Springer, 2020. doi:10.1007/978-3-030-67661-2_34.
6. H. Bunke. On a relation between graph edit distance and maximum common subgraph. *Pattern Recognition Letters*, 18:689–694, 1997.

7 D. Conte, P. Foggia, C. Sansone, and M. Vento. Thirty years of graph matching in pattern recognition. *International journal of pattern recognition and artificial intelligence*, 18(03):265–298, May 2004.

8 F. Emmert-Streib, M. Dehmer, and Y. Shi. Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, 346:180–197, 2016.

9 P. Foggia, G. Percannella, and M. Vento. Graph matching and learning in pattern recognition in the last 10 years. *International Journal of Pattern Recognition and Artificial Intelligence*, 28(1), 2014.

10 M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, 1979.

11 Timo Gervens and Martin Grohe. Graph similarity based on matrix norms, 2022. doi:10.48550/ARXIV.2207.00090.

12 M. Grohe. word2vec, node2vec, graph2vec, x2vec: Towards a theory of vector embeddings of structured data. In D. Suciu, Y. Tao, and Z. Wei, editors, *Proceedings of the 39th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 1–16, 2020. doi:10.1145/3375395.3387641.

13 M. Grohe, G. Rattan, and G. J. Woeginger. Graph similarity and approximate isomorphism. In *43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)*, volume 117, pages 20:1–20:16, 2018.

14 Martin Grohe and Pascal Schweitzer. The graph isomorphism problem. *Communications of the ACM*, 63(11):128–134, 2020. doi:10.1145/3372123.

15 P. Keldenich. Random robust graph isomorphism. Master's thesis, Department of Computer Science, RWTH Aachen University, 2015.

16 A. Kolla, I. Koutis, V. Madan, and A. K. Sinop. Spectrally Robust Graph Isomorphism. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107, pages 84:1–84:13, 2018.

17 N.M. Kriege, F.D Johansson, and C. Morris. A survey on graph kernels. *ArXiv*, arXiv:1903.11835 [cs.LG], 2019.

18 C.-L. Lin. Hardness of approximating graph transformation problem. In D.-T. Du and X.-S. Zhang, editors, *Proceedings of the 5th International Symposium on Algorithms and Computation*, volume 834 of *Lecture Notes in Computer Science*, pages 74–82. Springer, 1994.

19 L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society, 2012.

20 László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012. URL: http://www.ams.org/bookstore-getitem/item=COLL-60.

21 E. M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of Computer and System Sciences*, 25(1):42–65, 1982.

22 K. Makarychev, R. Manokaran, and M. Sviridenko. Maximum quadratic assignment problem: Reduction from maximum label cover and lp-based approximation algorithm. *ACM Transactions on Algorithms*, 10(4):18, 2014.

23 Hermina Petric Maretic, Mireille El Gheche, Giovanni Chierchia, and Pascal Frossard. GOT: an optimal transport framework for graph comparison. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 13876–13887, 2019. URL: https://proceedings.neurips.cc/paper/2019/hash/fdd5b16fc8134339089ef25b3cf0e588-Abstract.html.

24 D. W. Matula. Subtree isomorphism in $o(n^{5/2})$. In B. Alspach, P. Hell, and D.J. Miller, editors, *Algorithmic Aspects of Combinatorics*, volume 2, pages 91–106. Elsevier, 1978.

25 Facundo Mémoli. Gromov-wasserstein distances and the metric approach to object matching. *Found. Comput. Math.*, 11(4):417–487, 2011. doi:10.1007/s10208-011-9093-5.

26 V. Nagarajan and M. Sviridenko. On the maximum quadratic assignment problem. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 516–524, 2009.

27 R. O'Donnell, J. Wright, C. Wu, and Y. Zhou. Hardness of robust graph isomorphism, lasserre gaps, and asymmetry of random graphs. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1659–1677, 2014.

28 D. Sangiorgi. *Introduction to Bisimulation and Coinduction*. Cambridge University Press, 2012.

29 A. Tsitsulin, D. Mottin, P. Karras, A. Bronstein, and E. Müller. Netlsd: Hearing the shape of a graph. In Y. Guo and F. Farooq, editors, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2347–2356, 2018.

30 S. Umeyama. An eigendecomposition approach to weighted graph matching problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):695–703, 1988.

31 R.J. van Glabbeek. The linear time - branching time spectrum I. The semantics of concrete, sequential processes. In J.A. Bergstra, A. Ponse, and S.A. Smolka, editors, *Handbook of Process Algebra*, chapter 1, pages 3–99. Elsevier, 2001.