Supplementary material for "Tractable n-Metrics for Multiple Graphs"

A. Special case of orthogonal matrices

In this section, we discuss the special case, where the pairwise matching matrices are orthogonal. This will further illustrate why computing d_F is harder than computing $d_{\mathcal{G}}$. We consider the following assumption.

Assumption 1. *Ω is the set of real symmetric matrices, namely, $\Omega = \{A \in \mathbb{R}^{m\times m} : A = A^\top\}$. $\mathcal{P}$ is the set of orthogonal matrices, namely, $\mathcal{P} = \{P \in \mathbb{R}^{m\times m} : P^\top = P^{-1}\}$. $s(A, B, P) = \|\|AP - PB\|\|\ \forall A, B \in \Omega, P \in \mathcal{P}$, where $\|\|\cdot\|\|$ is the Frobenius norm or the operator norm, which are orthogonal invariant, and $d(A, B) = \min_{P\in\mathcal{P}} s(A, B, P)$.*

We now provide the main results of this section in the following theorems, and provide the detailed proofs in Appendix D-F.

Theorem 6. *Under Assumption 1, d_F induced by d, and $d_{\mathcal{G}}$ induced by s, are pseudo n-metrics.*

Theorem 7. *Let $\Lambda_{A_i} \in \mathbb{R}^m$ be the vector of eigenvalues of A_i, ordered from largest to smallest. Then, under Assumption 1,*

$$d_F(A_{1:n}) = \min_{\Lambda_C\in\mathbb{R}^m} \sum_{i=1}^{n} \|\Lambda_{A_i} - \Lambda_C\|. \tag{34}$$

Theorem 8. *Let $\Lambda_{A_i} \in \mathbb{R}^m$ be the vector of eigenvalues of A_i, ordered from largest to smallest. Then, under Assumption 1,*

$$d_{\mathcal{G}}(A_{1:n}) = \frac{1}{2} \sum_{i,j\in[n]} \|\Lambda_{A_i} - \Lambda_{A_j}\|. \tag{35}$$

Note that $d_F = d_{\mathcal{G}} = 0$ if and only if $A_{1:n}$ share the same spectrum.

The function d_F is related to the geometric median of the spectra of $A_{1:n}$. In order to write (35) as an optimization problem similar to d_F in (34), it is tempting to define $d_{\mathcal{G}}$ using s^2 instead of s, and take a square root. Let us call the resulting function $\bar{d}_{\mathcal{G}}$. A straightforward calculation allows us to write

$$(\bar{d}_{\mathcal{G}}(A_{1:n}))^2 = \frac{1}{2} \sum_{i,j\in[n]} \|\Lambda_{A_i} - \Lambda_{A_j}\|^2$$

$$= n^2 \left(\frac{1}{n} \sum_{i\in[n]} \left\|\Lambda_{A_i} - \frac{1}{n}\sum_{j\in[n]} \Lambda_{A_j}\right\|^2 \right) \equiv n^2 \mathrm{Var}(\Lambda_{A_{1:n}})$$

$$= n \min_{\Lambda_C\in\mathbb{R}^m} \frac{1}{2} \sum_{i\in[n]} \|\Lambda_{A_i} - \Lambda_C\|^2,$$

where we use $\mathrm{Var}(\Lambda_{A_{1:n}})$ to denote the geometric sample variance of the vectors $\{\Lambda_{A_i}\}$. This leads to a definition very close to (34), and a connection between $\bar{d}_{\mathcal{G}}$ and the geometric sample variance.

At this point it is important to note that sample variances can be computed exactly in $\mathcal{O}(n)$ steps involving only sums and products of numbers. Contrastingly, although there are fast approximation algorithms for the geometric median (Cohen et al., 2016), there are no procedures to compute it exactly in a finite number of simple algebraic operations (Bajaj, 1986; Cockayne & Melzak, 1969).

B. Proof of Theorem 1

In the following lemmas, we show that the Fermat distance function satisfies properties (6), (8), (9), and (10), and hence is a pseudo n-metric.

Lemma 7. *d_F is non-negative.*

Proof. If d is a pseudo metric, it is non-negative. Thus, (11) is the sum of non-negative functions, and hence also non-negative. □

Lemma 8. *d_F satisfies the self-identity property.*

Proof. If $A_1 = A_2 = \ldots = A_n$, then $d_F(A_{1:n}) = \min_B n \times d(A_1, B)$, which is zero if we choose $B = A_1 \in \Omega$, and (10) follows. □

Lemma 9. *d_F is symmetric.*

Proof. Property (8) simply follows from the commutative property of summation. □

Lemma 10. *d_F satisfies the generalized triangle inequality.*

Proof. Note that the following proof is a direct adaptation of the one in (Kiss et al., 2016), and is included for the sake of completeness. We show that the Fermat distance satisfies (9), i.e.,

$$d_F(A_{1:n}) \leqslant \sum_{i=1}^{n} d_F(A^i_{1:n,n+1}). \tag{36}$$

Consider $B_{1:n} \in \Omega$ such that,

$$d_F(A^i_{1:n,n+1}) = d(A_{n+1}, B_i) + \sum_{j\in[n]\backslash i} d(A_j, B_i). \tag{37}$$

Equation (37) implies that

$$\sum_{i=1}^{n} d_F(A^i_{1:n,n+1}) \geqslant \sum_{i=1}^{n}\sum_{j\in[n]\backslash i} d(A_j, B_i) \geqslant d(A_1, B_n) + d(A_2, B_n) + \sum_{i=2}^{n-1} (d(A_1, B_i) + d(A_{i+1}, B_i)). \tag{38}$$

Using triangle inequality, we have $d(A_1, B_n) + d(A_2, B_n) \geqslant d(A_1, A_2)$, and, $d(A_1, B_i) + d(A_{i+1}, B_i) \geqslant d(A_1, A_{i+1})$. Thus, from (38),

$$\sum_{i=1}^{n} d_F(A^i_{1:n,n+1}) \geqslant \sum_{i=2}^{n} d(A_1, A_i) = \sum_{i=1}^{n} d(A_1, A_i) \geqslant d_F(A_{1:n}),$$

where we used $d(A_1, A_1) = 0$ in the equality. The last inequality follows from Definition 7, and completes the proof. □

C. Proof of Theorem 2

In the following lemmas, we show that the $\mathcal{G}$-align distance function satisfies properties (6), (8), (9), and (10), and hence is a pseudo n-metric.

Lemma 11. *$d_{\mathcal{G}}$ is non-negative.*

Proof. Since s is a P-score, it satisfies (2), i.e., $s \geqslant 0$, which implies $d_{\mathcal{G}} \geqslant 0$, since it is a sum of P-scores. □

Lemma 12. *$d_{\mathcal{G}}$ satisfies the self-identity property.*

Proof. If $A_1 = A_2 = \ldots = A_n$, then, if we choose $P \in S$ such that $P_{i,j} = I$ for all $i, j \in [n]$, we have $s(A_i, A_j, P_{i,j}) = 0$ by (3), for all $i, j \in [n]$. Therefore,

$$0 \leqslant d_{\mathcal{G}}(A_{1:n}) \leqslant \frac{1}{2} \sum_{i,j \in [n]} s(A_i, A_j, P_{i,j}) = 0.$$

□

Lemma 13. *$d_{\mathcal{G}}$ is symmetric.*

Proof. The definition, (12), involves summing $s(A_i, A_j, P_{i,j})$ over all pairs $i, j \in [n]$, which clearly makes $d_{\mathcal{G}}$ invariant to permuting $\{A_i\}$. □

Lemma 14. *$d_{\mathcal{G}}$ satisfies the generalized triangle inequality.*

Proof. We now show that $d_{\mathcal{G}}$ satisfies (9), i.e.,

$$d_{\mathcal{G}}(A_{1:n}) \leqslant \sum_{\ell=1}^{n} d_{\mathcal{G}}(A^i_{1:n,n+1}). \tag{39}$$

Let $P^* = \{P^*_{i,j}\} \in S$ be an optimal value for P in the optimization problem corresponding to the l.h.s of (39). Henceforth, just like Section 6, we use $s^*_{i,j} = s(A_i, A_j, P^*_{i,j})$ for all $i, j \in [n]$. Note that according to (3) and (4), we have $s^*_{i,i} = 0$, and $s^*_{i,j} = s^*_{j,i}$, respectively. From (14), we have,

$$d_{\mathcal{G}}(A_{1:n}) = \sum_{i,j \in [n],\ i<j} s(A_i, A_j, P^*_{i,j}) = \sum_{i,j \in [n],\ i<j} s^*_{i,j}. \tag{40}$$

Let $P^{k*} = \{P^{k*}_{i,j}\} \in S$ be an optimal value for P in the optimization problem associated to $d_{\mathcal{G}}(A^i_{1:n,n+1})$ on the r.h.s of (39). Henceforth, just like Section 6, we use $s^{\ell*}_{i,j} = s(A_i, A_j, P^{\ell*}_{i,j})$ for all $i, j \in [n+1], \ell \in [n] \backslash \{i, j\}$. Note that $s^{\ell*}_{i,i} = 0$, and $s^{\ell*}_{i,j} = s^{\ell*}_{j,i}$. From (14), we can write,

$$\sum_{\ell=1}^{n} d_{\mathcal{G}}(A^i_{1:n,n+1}) = \sum_{\ell=1}^{n} \sum_{\substack{i,j \in [n+1],\ i<j \\ \ell \notin \{i,j\}}} s^{\ell*}_{i,j}. \tag{41}$$

We will show that,

$$\sum_{i,j \in [n],\ i<j} s^*_{i,j} \leqslant \sum_{\ell=1}^{n} \sum_{\substack{i,j \in [n+1],\ i<j \\ \ell \notin \{i,j\}}} s^{\ell*}_{i,j}. \tag{42}$$

From the definition of $d_{\mathcal{G}}$ in Lemma 1,

$$\sum_{i,j \in [n],\ i<j} s^*_{i,j} \leqslant \sum_{i,j \in [n],\ i<j} s(A_i, A_j, \Gamma_{i,j}), \tag{43}$$

for any matrices $\{\Gamma_{i,j}\}_{i,j \in [n]}$ in S, where S satisfies Definition 8. Hence, from Lemma 2, we also know that

$$\sum_{i,j \in [n],\ i<j} s^*_{i,j} \leqslant \sum_{i,j \in [n],\ i<j} s(A_i, A_j, \Gamma_i \Gamma_j^{-1}), \tag{44}$$

for any invertible matrices $\{\Gamma_i\}_{i \in [n]}$ in $\mathcal{P}$.

Consider the following choice for Γ_i :

$$\Gamma_i = P^{i-1*}_{i,n+1}, \qquad 2 \leqslant i \leqslant n, \tag{45}$$

$$\Gamma_1 = P^{n*}_{1,n+1}. \tag{46}$$

Remark 5. *To simplify notation, we will just use $\Gamma_i = P^{i-1*}_{i,n+1}$ for all $i \in [n]$. It is assumed that when we writing $P^{\ell*}_{i,j}$, the index in superscript satisfies $\ell = 0 \Leftrightarrow \ell = n$.*

Note that since $P^{i-1*} \in S$, then $\Gamma_i = P^{i-1*}_{i,n+1}$ is invertible and belongs to $\mathcal{P}$. Using (45) to replace Γ_i and Γ_j in (44), and the fact that $(P^{j-1*}_{j,n+1})^{-1} = P^{j-1*}_{n+1,j}$, along with property (5) of the P-score s, we have

$$\begin{aligned}\sum_{\substack{i,j \in [n] \\ i<j}} s(A_i, A_j, \Gamma_i \Gamma_j^{-1}) &= \sum_{\substack{i,j \in [n] \\ i<j}} s(A_i, A_j, P^{i-1*}_{i,n+1} P^{j-1*}_{n+1,j}) \\ &\leqslant \sum_{\substack{i,j \in [n] \\ i<j}} s^{i-1*}_{i,n+1} + s^{j-1*}_{n+1,j}.\end{aligned}$$

We now show that

$$\sum_{\substack{i,j \in [n] \\ i<j}} s^{i-1*}_{i,n+1} + s^{j-1*}_{n+1,j} \leqslant \sum_{\ell=1}^{n} \sum_{\substack{i,j \in [n+1],\ i<j \\ \ell \notin \{i,j\}}} s^{\ell*}_{i,j}, \tag{47}$$

which will prove (42) and complete the proof of the generalized triangle inequality for $d_{\mathcal{G}}$.

To this end, let $I_1 = \{(i,j) \in [n]^2 : i < j,\ j-1 = i\}$, $I_2 = \{(i,j) \in [n]^2 : i = 1, j = n\}$, $I_3 = \{(i,j) \in [n]^2 : i < j,\ j-1 \neq i \text{ and } (i,j) \neq (1,n)\}$. We will make use of the following three inequalities, which follow directly from property (5) of the P-score s.

$$\sum_{(i,j)\in I_1} s_{i,n+1}^{i-1*} \leqslant \sum_{(i,j)\in I_1} s_{i,j}^{i-1*} + s_{j,n+1}^{i-1*}. \quad (48)$$

$$\sum_{(i,j)\in I_2} s_{n+1,j}^{j-1*} \leqslant \sum_{(i,j)\in I_2} s_{n+1,i}^{j-1*} + s_{i,j}^{j-1*}. \quad (49)$$

$$\sum_{(i,j)\in I_3} s_{i,n+1}^{i-1*} + s_{n+1,j}^{j-1*} \leqslant \sum_{(i,j)\in I_3} \Big(s_{i,j}^{i-1*} + s_{j,n+1}^{i-1*} + s_{n+1,i}^{j-1*} + s_{i,j}^{j-1*}\Big). \quad (50)$$

Since I_1, I_2 and I_3 are pairwise disjoint, we have

$$\sum_{i,j\in[n]} (\cdot) = \sum_{(i,j)\in I_1} (\cdot) + \sum_{(i,j)\in I_2} (\cdot) + \sum_{(i,j)\in I_3} (\cdot). \quad (51)$$

Using (48)-(50), and (51) we have

$$\sum_{i,j\in[n],\ i<j} s_{i,n+1}^{i-1*} + s_{n+1,j}^{j-1*} \leqslant \sum_{(i,j)\in I_1} s_{i,j}^{i-1*} + s_{j,n+1}^{i-1*} + s_{n+1,j}^{j-1*} + \sum_{(i,j)\in I_2} s_{i,n+1}^{i-1*} + s_{n+1,i}^{j-1*} + s_{i,j}^{j-1*} + \sum_{(i,j)\in I_3} s_{i,j}^{i-1*} + s_{j,n+1}^{i-1*} + s_{n+1,i}^{j-1*} + s_{i,j}^{j-1*}. \quad (52)$$

To complete the proof, we show that the r.h.s of (52) is less than, or equal to

$$\sum_{\ell=1}^{n} \sum_{\substack{i,j\in[n+1],\ i<j \\ \ell\notin\{i,j\}}} s_{i,j}^{\ell*}. \quad (53)$$

To establish this, we show that each term on the r.h.s of (52) is: (i) not repeated; and (ii) is included in (53).

Definition 11. *We call two P-scores, $s_{a_1,b_1}^{c_1*}$ and $s_{a_2,b_2}^{c_2*}$,* co-incident, *and denote it by $s_{a_1,b_1}^{c_1*} \sim s_{a_2,b_2}^{c_2*}$, if and only if* $c_1 = c_2$, and $\{a_1, b_1\} = \{a_2, b_2\}$.

Checking (i) amounts to verifying that there are no coincident terms on the r.h.s. of (52). Checking (ii) amounts to verifying that for each P-score $s_{a_1,b_1}^{c_1*}$ on the r.h.s. of (52), there exists a P-score $s_{a_2,b_2}^{c_2*}$ in (53) such that $s_{a_1,b_1}^{c_1*} \sim s_{a_2,b_2}^{c_2*}$.

Note that the r.h.s of (52) consists of three summations. To verify (i), we first compare the terms within each summation, and then compare the terms among different summations. Consider the first summation on the r.h.s of (52). We have $s_{i,j}^{i-1*} \nsim s_{j,n+1}^{i-1*}$ because $i \in [n]$ and therefore $i \neq n+1$. We have $s_{i,j}^{i-1*} \nsim s_{n+1,j}^{j-1*}$ because $i-1 \neq j-1$ in this case, since $i < j$. We can similarly infer that $s_{j,n+1}^{i-1*} \nsim s_{n+1,j}^{j-1*}$.

Now consider the second summation on the r.h.s of (52). Taking the definition of I_2 and (46) into account, we can rewrite this summation as,

$$s_{1,n+1}^{n*} + s_{n+1,1}^{n-1*} + s_{1,n}^{n-1*}. \quad (54)$$

Since $n \neq n-1$, we have $s_{1,n+1}^{n*} \nsim s_{n+1,1}^{n-1*}$, and $s_{1,n+1}^{n*} \nsim s_{1,n}^{n-1*}$. Also, since $n \neq n+1$ we have $s_{n+1,1}^{n-1*} \nsim s_{1,n}^{n-1*}$.

Finally, consider the third summation on the r.h.s of (52). Since $i < j$, by comparing the superscripts we immediately see that the first and second terms in the summation cannot be equal to either the third or the forth term. On the other hand, since $n+1 \neq i \in [n]$ and $n+1 \neq j \in [n]$, we have $s_{i,j}^{i-1*} \nsim s_{j,n+1}^{i-1*}$, and $s_{n+1,i}^{j-1*} \nsim s_{i,j}^{j-1*}$, respectively.

We proceed by showing that the summands are not coincident among three summations. We first make the following observations:

Observation 1: since in all summations $i, j \in [n]$, we have $i \neq n+1,\ j \neq n+1$, and therefore each term with $n+1$ in the subscript is not coincident with any term with $\{i, j\}$ in the subscript, e.g., on the r.h.s of (52), the first terms in the first and second summations cannot be coincident.

Observation 2: since I_1, I_2 and I_3 are pairwise disjoint, any two terms from different summations with the same indices cannot be coincident, e.g., on the r.h.s of (52), the third term in the second summation cannot be coincident with the third term in third summation.

Considering the above observations, the number of pairs we need to compare reduces from $3 \times 7 + 3 \times 4 = 33$ (in (52)) pairs to only 13 pairs, whose distinction may not seem trivial. To be specific, Obs. 1, excludes 16 comparisons and Obs. 2 excludes 4 comparisons. We now rewrite the r.h.s of (52) as

$$\sum_{(i,j)\in I_1} s_{i,j}^{i-1*} + s_{j,n+1}^{i-1*} + s_{n+1,j}^{j-1*} + s_{1,n+1}^{n*} + s_{n+1,1}^{n-1*} + s_{1,n}^{n-1*} + \sum_{(i',j')\in I_3} s_{i',j'}^{i'-1*} + s_{j',n+1}^{i'-1*} + s_{n+1,i'}^{j'-1*} + s_{i',j'}^{j'-1*}. \quad (55)$$

In what follows, we discuss the non-trivial comparisons, and refer to the first, second and third summations in (55) as Σ_1, Σ_2, and Σ_3, respectively.

1. $s_{i,j}^{i-1*}$ in Σ_1 vs. $s_{1,n}^{n-1*}$ in Σ_2: for these two terms to be coincident we need $i = n$. We also need $\{n, j\} = \{1, n\}$, i.e., $j = 1$, which cannot be true, since in S_1 we have $i = j-1$ according to I_1.

2. $s_{i,j}^{i-1*}$ in Σ_1 vs. $s_{i',j'}^{j'-1*}$ in Σ_3: since $(i,j) \in I_1 = \{(i,j) \in [n]^2 : i < j,\ j-1 = i\}$, we have $j = i+1$. Thus, we can write the first term as $s_{i,i+1}^{i-1*}$. For the two terms to be coincident, their superscripts must be the

same so $i = j'$. On the other hand, for their subscripts to match, we need $j = i + 1 = i'$. The last two equalities imply that $i' = j' + 1$, which contradicts $(i', j') \in I_3$.

3. $s^{i-1*}_{j,n+1}$ in Σ_1 vs. $s^{n*}_{1,n+1}$ in Σ_2: for the superscripts to match, we need $i = 1$. We also need $j = 1$ for the equality of subscripts, which cannot be true since $i < j$.

4. $s^{i-1*}_{j,n+1}$ in Σ_1 vs. $s^{n-1*}_{n+1,1}$ in Σ_2: we need $i = n$ for the equality of superscripts, and $j = 1$ for the equality of subscripts, which cannot be true since $(i, j) \in I_1$, and therefore $i = j - 1$.

5. $s^{i-1*}_{j,n+1}$ in S_1 vs. $s^{j'-1*}_{n+1,i'}$ in S_3: we can write the first term as $s^{i-1}_{i+1,n+1}$. The equality of superscripts requires $i = j'$. The equality of subscripts requires $i' = i + 1$. Therefore, $i' = j' + 1$, which contradicts $(i', j') \in I_3$.

6. $s^{j-1*}_{n+1,j}$ in Σ_1 vs. $s^{n*}_{1,n+1}$ in Σ_2: the equality of superscripts requires $j = 1$, which is impossible since $j > i \in [n]$.

7. $s^{j-1*}_{n+1,j}$ in Σ_1 vs. $s^{n-1*}_{n+1,1}$ in Σ_2: for the equality of superscripts, we need $j = n$, in which case the subscripts will not match, since $\{n + 1, n\} \neq \{n + 1, 1\}$.

8. $s^{j-1*}_{n+1,j}$ in Σ_1 vs. $s^{i'-1*}_{j',n+1}$ in Σ_3: the equality of superscripts requires $i' = j$. The equality of the subscripts requires $j' = j$. The two equalities imply that $i' = j'$, which contradicts $i' < j'$.

9. $s^{j-1*}_{n+1,j}$ in Σ_1 vs. $s^{j'-1*}_{n+1,i'}$ in Σ_3: the equality of superscripts requires $j' = j$. The equality of the subscripts requires $i' = j$. The two equalities imply that $i' = j'$, which contradicts $i' < j'$.

10. $s^{n*}_{1,n+1}$ in Σ_2 vs. $s^{i'-1*}_{j',n+1}$ in Σ_3: for the equality of superscripts, we need $i' = 1$, and for the equality of subscripts, we need $j' = 1$. This contradicts $i' < j'$.

11. $s^{n*}_{1,n+1}$ in Σ_2 vs. $s^{j'-1*}_{n+1,i'}$ in the Σ_3: for equality of superscripts, we need $j' = 1$. For the equality of subscripts, we need $i' = 1$, which contradicts $i' \neq j'$.

12. $s^{n-1*}_{n+1,1}$ in Σ_2 vs. $s^{i'-1*}_{j',n+1}$ in Σ_3: for equality of superscripts, we need $i' = n$. For the equality of subscripts, we need $j' = 1$, which contradicts $i' < j'$.

13. $s^{n-1*}_{1,n}$ in Σ_2 vs. $s^{i'-1*}_{i',j'}$ in Σ_3: for the equality of superscripts, we need $i' = n$. This in turn requires $j' = 1$ for the equality of subscripts, which contradicts $i' < j'$.

What is left to show is (ii), i.e., that all terms in (55) are included in the summation in (53). To this aim, we will show that for each $s^c_{a,b}$ in (55), the indices $\{a, b, c\}$ satisfy

$$c \in [n], a, b \in [n + 1] \backslash \{c\} \text{ and } a \neq b, \tag{56}$$

which is enough to prove that either $s^c_{a,b}$ or $s^c_{b,a}$ exist in (53).

We first note that the superscripts in (55) are in $[n]$, see Remark 5. Moreover, all the subscripts in (55) are either 1, $n + 1$, or $i, j, i', j' \in [n]$. Thus, for any $s^c_{a,b}$ in (55), we have $a, b \in [n + 1]$. Also note that, for any $s^c_{a,b}$ in (55), we have $a \neq b$, since the definition of I_1, I_2 and I_3 implies that $i < j$, $i' < j'$ and $i, j, i', j' < n + 1$. Therefore, all we need to verify is that for any $s^c_{a,b}$ in (55), $a \neq c$ and $b \neq c$.

We start with the first summation, where the first term is $s^{i-1*}_{i,j}$. Clearly $i \neq i - 1$ and $j \neq i - 1$, from the definition of I_1. In the second term, $s^{i-1*}_{j,n+1}$, $j \neq i - 1$, from the definition of I_1, and $i - 1 \neq n + 1$, because otherwise $i = n + 2 \notin [n]$. In the third term, $s^{j-1*}_{n+1,j}$, we have $n + 1 \neq j \in [n]$. Moreover, clearly $j \neq j - 1$.

For any term $s^c_{a,b}$, in the second summation, we clearly see in (55) that $a \neq c$ and $b \neq c$.

We now consider the last summation in (55). In the first term, $s^{i'-1*}_{i',j'}$, clearly $i' \neq i' - 1$. Moreover, $i' - 1 < i' < j'$, since $(i', j') \in I_3$. In the second term, $s^{i'-1*}_{j',n+1}$, $j' \neq i' - 1$, because since $i' - 1 < i' < j'$. Moreover, $n + 1 \neq i' - 1$ because otherwise $i' = n + 2 \notin [n]$. In the third term, $s^{j'-1*}_{n+1,i'}$, we have $n + 1 \neq j' - 1$ because otherwise $j' = n + 2 \notin [n]$. On the other hand, $i' \neq j' - 1$ since $(i', j') \in I_3$. In the fourth term, $s^{j'-1*}_{i',j'}$, we have $i' \neq j' - 1$ since $(i', j') \in I_3$. Also, clearly $j' \neq j' - 1$. $\square$

D. Proof of Theorem 6

To show that d_F is a pseudo n-metric, it suffices to show that d is a pseudometric, and evoke Theorem 1. To show that d is a pseudometric, we can evoke Theorem 3 in (Bento & Ioannidis, 2018).

To show that $d_{\mathcal{G}}$ is a pseudo n-metric, it suffices to show that s is a P-score, and evoke Theorem 2. Clearly, s is non-negative, and also $s(A, A, I) = 0$. Recall that, if P is orthogonal then, for any matrix M, we have $|||PM||| = |||MP||| = |||M|||$. Thus,

$$\begin{aligned} s(A, B, P) &= |||AP - PB||| = |||P^{-1}(AP - PB)P^{-1}||| \\ &= |||P^{-1}A - BP^{-1}||| = s(B, A, P^{-1}). \end{aligned}$$

Finally, for any $P, P' \in \mathcal{P}$,

$$\begin{aligned} & s(A, B, PP') = |||APP' - PP'B||| = \\ & |||APP' - PCP' + PCP' - PP'B||| \leqslant \\ & |||APP' - PCP'||| + |||PCP' - PP'B||| = \\ & |||AP - PC||| + |||CP' - P'B||| = \\ & s(A, C, P) + s(C, B, P'). \end{aligned}$$

E. Proof of Theorem 7

The proof uses the following lemmas by (Hoffman et al., 1953) and (Bento & Ioannidis, 2018).

Lemma 15. *For any matrix $M \in \mathbb{R}^{m\times m}$, and any orthogonal matrix $P \in \mathbb{R}^{m\times m}$, we have that $|||PM||| = |||MP||| = |||M|||$.*

Lemma 16. *Let $|||\cdot|||$ be the Frobenius norm. If A and B are Hermitian matrices with eigenvalues $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_m$ and $b_1 \leqslant b_2 \leqslant \ldots \leqslant b_m$ then*

$$|||A - B||| \geqslant \sqrt{\sum_{i\in[m]} (a_i - b_i)^2}. \tag{57}$$

Lemma 17. *Let $|||\cdot|||$ be the operator 2-norm. If A and B are Hermitian matrices with eigenvalues $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_m$ and $b_1 \leqslant b_2 \leqslant \ldots \leqslant b_m$ then*

$$|||A - B||| \geqslant \max_{i\in[m]} |a_i - b_i|. \tag{58}$$

We also need the following result.

Corollary 1. *If $a \in \mathbb{R}^m$, with $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_m$, $b \in \mathbb{R}^m$, with $b_1 \leqslant b_2 \leqslant \cdots \leqslant b_m$, and $P \in \mathbb{R}^{m\times m}$ is a permutation matrix, then*

$$\|a - b\| \leqslant \|a - Pb\|. \tag{59}$$

Proof. This follows directly from Lemma 16 and Lemma 17 by letting A and B be diagonal matrices with a and Pb in the diagonal, respectively. □

We now proceed with the proof of Theorem 7. Let $A_i = U_i \text{diag}(\Lambda_{A_i}) U_i^{-1}$ and $C = V \text{diag}(\Lambda_C) V^{-1}$ be the eigendecomposition of the real and symmetric matrices A_i and C, respectively. The eigenvalues in the vectors Λ_{A_i} and Λ_C are ordered in increasing order, and U_i and V are orthonormal matrices. Using Lemma 15, we have that

$$\begin{aligned}
&|||A_i P_i - P_i C||| = |||(A_i - P_i C (P_i)^{-1}) P_i||| \quad (60)\\
&= |||A_i - P_i C (P_i)^{-1}|||\\
&= |||U_i(\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_i C (P_i)^{-1} U_i) U_i^{-1}|||\\
&= |||\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_i C (P_i)^{-1} U_i||| \geqslant \|\Lambda_{A_i} - \Lambda_C\|,
\end{aligned}$$

where the last inequality follows from Lemma 16 or Lemma 17 (depending on the norm).

It follows from (60) that $d_F(A_{1:n}) \geqslant \min_{\Lambda_C \in \mathbb{R}^m : (\Lambda_C)_i \leqslant (\Lambda_C)_{i+1}} \sum_{i=1}^n \|\Lambda_{A_i} - \Lambda_C\| = \min_{\Lambda_C\in\mathbb{R}^m} \sum_{i=1}^n \|\Lambda_{A_i} - \Lambda_C\|$, where the last equality follows from Corollary 1.

Finally, notice that, by the equalities in (60), we have

$$\begin{aligned}
&d_F(A_{1:n}) = \min_{P\in\mathcal{P}^n, C\in\Omega} \sum_{i=1}^n |||\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_i C (P_i)^{-1} U_i|||\\
&\leqslant |||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_C)|||, \quad (61)
\end{aligned}$$

where the inequality follows from upper bounding $\min_{C\in\Omega}(\cdot)$ with the particular choice of $C = P_i^\top U_i \text{diag}(\Lambda_C) U_i^\top P_i \in \Omega$.

Since $|||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_C)|||_{\text{Frobenius}} = \|\Lambda_{A_i} - \Lambda_C\|_{\text{Eucledian}}$ and $|||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_C)|||_{\text{operator}} = \|\Lambda_{A_i} - \Lambda_C\|_{\infty\text{-norm}}$, the proof follows.

F. Proof of Theorem 8

Let $A_i = U_i \text{diag}(\Lambda_{A_i}) U_i^{-1}$ be the eigendecomposition of the real and symmetric matrix A_i. The eigenvalues in the vector Λ_{A_i} are ordered in increasing order, and U_i is an orthonormal matrix. Using Lemma 15, we get

$$\begin{aligned}
&|||A_i P_{i,j} - P_{i,j} A_j||| = |||(A_i - P_{i,j} A_j (P_{i,j})^{-1}) P_{i,j}||| \quad (62)\\
&= |||A_i - P_{i,j} A_j (P_{i,j})^{-1}|||\\
&= |||U_i(\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_{i,j} A_j (P_{i,j})^{-1} U_i) U_i^{-1}|||\\
&= |||\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_{i,j} A_j (P_{i,j})^{-1} U_i||| \geqslant \|\Lambda_{A_i} - \Lambda_{A_j}\|,
\end{aligned}$$

where the last inequality follows from Lemma 16 or Lemma 17 (depending on the norm).

From (62) we have $d_{\mathcal{G}}(A_{1:n}) \geqslant \frac{1}{2}\sum_{i,j\in[n]} \|\Lambda_{A_i} - \Lambda_{A_j}\|$.

At the same time, $d_{\mathcal{G}}(A_{1:n}) =$

$$\begin{aligned}
&\min_{P\in S} \frac{1}{2} \sum_{i,j\in[n]} |||\text{diag}(\Lambda_{A_i}) - U_i^{-1} P_{i,j} A_j (P_{i,j})^{-1} U_i|||\\
&\leqslant |||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_{A_j})|||, \quad (63)
\end{aligned}$$

where the inequality follows from upper bounding $\min_{P\in S}(\cdot)$ by choosing $P = \{P_{i,j}\}_{i,j\in[n]}$ such that $P_{i,j} = U_i U_j^{-1}$, which by Lemma 2 implies that $P \in S$.

Since $|||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_{A_j})|||_{\text{Frobenius}} = \|\Lambda_{A_i} - \Lambda_{A_j}\|_{\text{Eucledian}}$ and $|||\text{diag}(\Lambda_{A_i}) - \text{diag}(\Lambda_{A_j})|||_{\text{operator}} = \|\Lambda_{A_i} - \Lambda_{A_j}\|_{\infty\text{-norm}}$, the proof follows.

G. Proof of Theorem 3

We first show that (17) is well defined. Let $A_i' \in [A_i]$. Since d satisfies the triangle inequality we have

$$\begin{aligned}
&d_F'([A']_{1:n}) = d_F(A'_{1:n}) = \min_{B\in\Omega} \sum_{i\in[n]} d(A_i', B)\\
&\leqslant \min_{B\in\Omega} \sum_{i\in[n]} d(A_i', A_i) + d(A_i, B) = \min_{B\in\Omega} \sum_{i\in[n]} d(A_i, B)\\
&= d_F(A_{1:n}) = d_F'([A]_{1:n}),
\end{aligned}$$

where in the last equality we used $d(A_i', A_i) = 0$, since $A_i' \in [A_i]$. Similarly, we can show that $d_F'([A]_{1:n}) \leqslant d_F'([A']_{1:n})$. It follows that $d_F'([A]_{1:n}) = d_F'([A']_{1:n})$, and hence (17) is well defined.

We now prove that d'_F satisfies (7). Recall that, by Theorem 1, d_F is a pseudo n-metric. If $[A_1] = \cdots = [A_n]$, then

$$d'_F([A]_{1:n}) = d'_F([A_1], \ldots, [A_1]) = d_F(A_1, \ldots, A_1) = 0,$$

since, d_F is a pseudometric, and hence satisfies the property of self-identity (10).

On the other hand, if $d'_F([A]_{1:n}) = d_F(A_{1:n}) = 0$, then there exists $B \in \Omega$, such that $d(A_i, B) = 0$ for all $i \in [n]$. Since d is non-negative and symmetric, and also satisfies the triangle inequality, it follows that

$$\begin{aligned} 0 \leqslant d(A_i, A_j) &\leqslant d(A_i, B) + d(B, A_j) \\ &= d(A_i, B) + d(A_j, B) = 0. \end{aligned}$$

Hence, $[A_i] = [A_j]$ for all $i, j \in [n]$.

H. Proof of Theorem 4

In the proof, we let S_2 denote the set S in definition (13) for the distance d on two graphs and we let S_n denote the set S in definition (13) for the distance $d_{\mathcal{G}}$ on n graphs.

We first verify that (18) is well defined. Let $A'_i \in [A_i]$. Let $\{I, P^*_i, (P^*_i)^{-1}\} \in S_2$ be such that

$$\begin{aligned} d_{\mathcal{G}_2}(A_i, A'_i) \equiv \frac{1}{2}(s(A_i, A_i, I) + s(A'_i, A'_i, I) + \\ s(A'_i, A_i, P^*_i) + s(A_i, A'_i, (P^*_i)^{-1})) = 0. \end{aligned}$$

Since s is a P-score, $s(A'_i, A_i, P^*_i) = 0$. For any $\tilde{P} = \{\tilde{P}_{i,j}\}_{i,j \in [n]} \in S$ we have $\{P^*_i \tilde{P}_{i,j} (P^*_j)^{-1}\}_{i,j \in [n]} \in S$. Thus,

$$\begin{aligned} d'_{\mathcal{G}}([A']_{1:n}) = d_{\mathcal{G}}(A'_{1:n}) &= \min_{P \in S} \frac{1}{2} \sum_{i,j \in [n]} s(A'_i, A'_j, P_{i,j}) \\ &\leqslant \frac{1}{2} \sum_{i,j \in [n]} s(A'_i, A'_j, P^*_i \tilde{P}_{i,j} (P^*_j)^{-1}). \end{aligned}$$

By property (5) and the fact that $s(A'_i, A_i, P^*_i) = s(A_i, A'_i, (P^*_i)^{-1}) = 0$ for all $i \in [n]$, we can write

$$\begin{aligned} \frac{1}{2} \sum_{i,j \in [n]} s(A'_i, A'_j, P^*_i \tilde{P}_{i,j} (P^*_j)^{-1}) \leqslant \frac{1}{2} \sum_{i,j \in [n]} \Big(s(A'_i, A_i, P^*_i) \\ + s(A_i, A_j, \tilde{P}_{i,j}) + s(A_j, A'_j, (P^*_j)^{-1} \Big) = s(A_i, A_j, \tilde{P}_{i,j}). \end{aligned}$$

Taking the minimum of the r.h.s. of the above expression over $\tilde{P}$ we get $d'_{\mathcal{G}}([A']_{1:n}) \leqslant d_{\mathcal{G}}(A_{1:n}) = d'_{\mathcal{G}}([A]_{1:n})$. Similarly, we can prove $d'_{\mathcal{G}}([A]_{1:n}) \leqslant d'_{\mathcal{G}}([A']_{1:n})$. It follows that $d'_{\mathcal{G}}([A]_{1:n}) = d'_{\mathcal{G}}([A']_{1:n})$, and hence (18) is well defined.

Now we show that $d'_{\mathcal{G}}$ satisfies (7). Recall that, by Thm. 2, $d_{\mathcal{G}}$ is a pseudo n-metric. If $[A_1] = \cdots = [A_n]$, then

$$d'_{G}([A]_{1:n}) = d'_{\mathcal{G}}([A_1], \ldots, [A_1]) = d_{\mathcal{G}}(A_1, \ldots, A_1) = 0,$$

since, $d_{\mathcal{G}}$ is a pseudometric, and hence satisfies the property of self-identity (10).

On the other hand, if $d'_{G}([A]_{1:n}) = d_{\mathcal{G}}(A_{1:n}) = 0$, then, for any $i, j \in [n]$, we have that $s(A_i, A_j, P_{i,j}) = 0$ for some $P_{i,j}$, and hence $d(A_i, A_j) = 0$. This implies that $[A_i] = [A_j]$ for all $i, j \in [n]$.

I. Proof of Theorem 5

The following lemma will be used later.

Lemma 18. *Let $\Gamma_i \in \mathbb{R}^{m \times m}$, $\|\|\Gamma_i\|\|_2 \leqslant 1$ for all $i \in [n]$. Let $\mathbf{P} \in \mathbb{R}^{nm \times nm}$ have n^2 blocks such that the (i, j)th block is $\Gamma_i \Gamma_j^\top$ if $i \neq j$, and I otherwise. We have that $\mathbf{P} \succeq 0$, and that $\|\|\mathbf{P}\|\|_* \leqslant mn$.*

Proof. Let us first prove that $\mathbf{P} \succeq 0$. Let $\mathbf{v} \in \mathbb{R}^{nm}$ have n blocks, the ith block being $v_i \in \mathbb{R}^m$. Since $\|\|\Gamma_i \Gamma_i^\top\|\|_2 \leqslant \|\|\Gamma_i\|\|_2 \|\|\Gamma_i^\top\|\|_2 \leqslant 1$, we have that $\|\Gamma_i^\top v_i\|_2^2 = \|v_i^\top \Gamma_i \Gamma_i^\top v_i\|_2 \leqslant \|v_i\|_2^2$ for all $i \in [n]$. Therefore, we have $\mathbf{v}^\top \mathbf{P} \mathbf{v} = \|\sum_{i \in [n]} \Gamma_i^\top v_i\|_2^2 + \sum_{i \in [n]} \|v_i\|_2^2 - \sum_{i \in [n]} \|\Gamma_i^\top v_i\|_2^2 \geqslant 0$, for any $\mathbf{v}$, which implies that $\mathbf{P} \succeq 0$. We now prove that $\|\|\mathbf{P}\|\|_* \leqslant mn$. Let σ_r and λ_r be the rth singular value and rth eigenvalue of $\mathbf{P}$ respectively. Since $\mathbf{P}$ is real-symmetric and positive semi-definite, we have that $\|\|\mathbf{P}\|\|_* = \sum_r \sigma_r = \sum_r |\lambda_r| = \sum_r \lambda_r = \text{tr}(\mathbf{P}) = mn$. ☐

Proof of Theorem 5.

(Non-negativity): Since s is a modified P-score, it satisfies (30), i.e., $s \geqslant 0$, which implies $d_{sc\mathcal{G}} \geqslant 0$, since the objective function on the r.h.s of (29) is a sum of modified P-scores.

(Self-identity): If $A_1 = A_2 = \ldots = A_n$, then, if we choose $P_{i,j} = I$ for all $i, j \in [n]$, we have $s(A_i, A_j, P_{i,j}) = 0$ by (31), for all $i, j \in [n]$. Note that from the definition of $d_{sc\mathcal{G}}$, we are assuming that $I \in \mathcal{P}$. Furthermore, $\mathbf{P}$ defined using these $P_{i,j}$'s satisfies $\|\|\mathbf{P}\|\|_* \leqslant mn$. Therefore, this choice of $P_{i,j}$'s satisfies the constraints in the minimization problem in the definition of $d_{sc\mathcal{G}}(A_{1:n})$. Therefore, $d_{sc\mathcal{G}}(A_{1:n})$ is upper-bounded by 0, which along with its non-negativity leads to $d_{sc\mathcal{G}}(A_{1:n}) = 0$.

(Symmetry): The optimization problem in (29), involves summing $s(A_i, A_j, P_{i,j})$ over all pairs $i, j \in [n]$. Thus, permuting the matrices $\{A_i\}$ is the same as solving (29) with $P_{i,j}$ replaced by $P_{\sigma(i),\sigma(j)}$ for some permutation σ. Thus, all that we need to show is that $\mathbf{P} \succeq 0$ if and only if $\mathbf{P}' \succeq 0$, where $\mathbf{P}'$ is just like $\mathbf{P}$ but with its blocks' indexes permuted. To see this, note that the eigenvalues of a matrix M do not change if M is then permuted under some permutation matrix T.

(Generalized triangle inequality): We will follow exactly the same argument as in the proof of the generalized triangle inequality for Theorem 2, which is provided in Appendix C.

The only modification is in equation (44), and in a couple of steps afterwards.

Equation (44) should be replaced with

$$\sum_{i\neq j} s(A_i, A_j, P^*_{i,j}) \leqslant \sum_{i\neq j} s(A_i, A_j, \Gamma_i\Gamma_j^\top), \tag{64}$$

where $\{\Gamma_i\}_{i\in[n]}$ are matrices in $\mathcal{P}$. This inequality holds because $P_{i,j}$ defined by $P_{i,j} = \Gamma_i\Gamma_j^\top\ \forall i \neq j$, and $P_{i,i} = I \forall i$, satisfies the constraints in (28), and hence the r.h.s. of (64) upper bounds the optimal objective value for (28). Indeed, since $\Gamma_i \in \mathcal{P}$, and since, by assumption, $\mathcal{P}$ is closed under multiplication and transposition, it follows that $\Gamma_i\Gamma_j^\top \in \mathcal{P}$. Furthermore, if we define $\mathbf{P}$ to have as the (i,j)th block, $i \neq j$, $\Gamma_i\Gamma_j^\top$, and have as the (i,i)th block the identity I, then, by Lemma 18, we know that $\mathbf{P} \geq 0$.

Starting from (64), we use (33) and (32) from the modified P-score properties and obtain

$$\sum_{i\neq j} s(A_i, A_j, \Gamma_i\Gamma_j^\top) \leqslant \sum_{i\neq j} s(A_i, A_{n+1}, \Gamma_i) +$$
$$\sum_{i\neq j} s(A_{n+1}, A_j, \Gamma_j^\top) = s(A_i, A_{n+1}, \Gamma_i) \tag{65}$$
$$+ \sum_{i\neq j} s(A_j, A_{n+1}, \Gamma_j). \tag{66}$$

The rest of the proof follows by choosing Γ_i has in (45) and (46), and noting that the new definition of $s^*_{i,j}$ and $s^{\ell*}_{i,j}$ satisfies the same properties as in the proof of Theorem 2. In particular, we have that $s^*_{i,j} = s^*_{j,i}$ and $s^{\ell*}_{i,j} = s^{\ell*}_{j,i}$, because $\mathbf{P}$ in (29) is symmetric, and because we are assuming that (32) holds.

J. Distribution of AQ and AC for the alignment experiment

K. Distribution of clustering errors for the clustering experiment

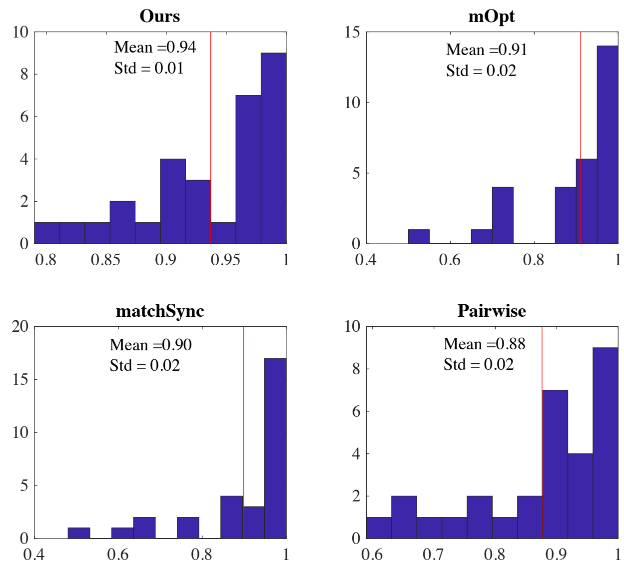

Figure 2. Distribution of alignment quality (AQ) for the 30 tests in Section 8.1.

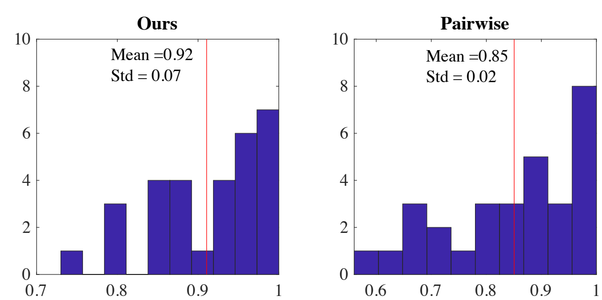

Figure 3. Distribution of alignment consistency (AC) for the 30 tests in Section 8.1. Note that, by construction, *mOpt* and *matchSync* always have AC = 1.

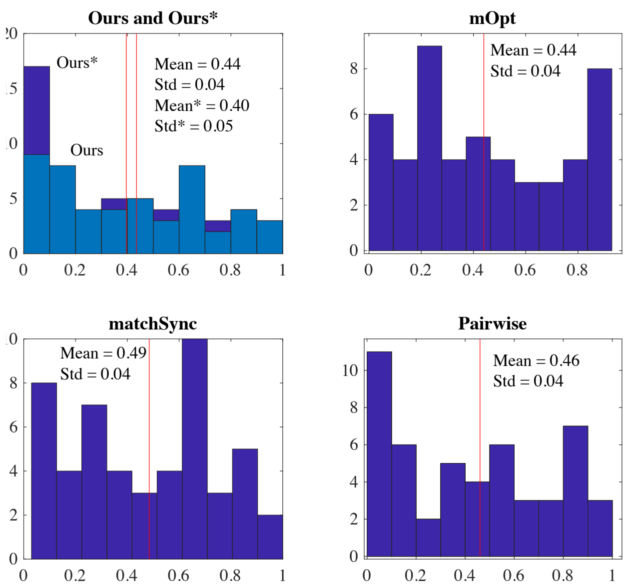

Figure 4. Distribution of errors for clustering for the 50 tests in Section 8.2. Recall that the error is the fraction of misclassified graphs times the number of clusters, which is 2 in our case. A random guess gives an average clustering error of 1.