Statistical-computational Gap in Multiple Gaussian Graph Alignment

Bertrand Even
Université Paris-Saclay
Laboratoire de mathématiques d'Orsay
Orsay, France

Luca Ganassali
Université Paris-Saclay, CNRS, Inria
Laboratoire de mathématiques d'Orsay
Orsay, France

Abstract

We investigate the existence of a statistical-computational gap in multiple Gaussian graph alignment. We first generalize a previously established informational threshold from Vassaux and Massoulié (2025) to regimes where the number of observed graphs p may also grow with the number of nodes n: when $p \leq O(n/\log(n))$, we recover the results from Vassaux and Massoulié (2025), and $p \geq \Omega(n/\log(n))$ corresponds to a regime where the problem is as difficult as aligning one single graph with some unknown "signal" graph. Moreover, when $\log p = \omega(\log n)$, the informational thresholds for partial and exact recovery no longer coincide, in contrast to the all-or-nothing phenomenon observed when $\log p = O(\log n)$. Then, we provide the first computational barrier in the low-degree framework for (multiple) Gaussian graph alignment. We prove that when the correlation ρ is less than 1, up to logarithmic terms, low degree non-trivial estimation fails. Our results suggest that the task of aligning p graphs in polynomial time is as hard as the problem of aligning two graphs in polynomial time, up to logarithmic factors. These results characterize the existence of a statistical-computational gap and provide another example in which polynomial-time algorithms cannot handle complex combinatorial bi-dimensional structures.

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

1 INTRODUCTION

Graph alignment. Graph alignment (also referred to as graph matching) is a fundamental statistical task which consists in finding an underlying correspondence between graphs, namely node relabelings which preserve most of the edges. This problem has emerged in the 2000's across diverse application domains, including network privacy (Narayanan and Shmatikov (2009)), computational biology (Singh et al. (2008)), computer vision (Conte et al. (2004)), and natural language processing (Haghighi et al. (2005)). A common feature across these applications is that the data naturally admits a graph-based representation: social networks, protein–protein interaction networks, three-dimensional images or meshes, and word or token embeddings.

Graph alignment has since been extensively investigated within the statistics and computer science communities, particularly in the two-graph setting, both for correlated Gaussian graphs Ganassali et al. (2022); Ganassali (2022); Wu et al. (2022); Ding and Li (2022) and for correlated Erdős–Rényi graphs Fan et al. (2020); Mao et al. (2021); Ganassali et al. (2024a); Mao et al. (2023).

More recently, the problem of aligning multiple correlated graphs has attracted growing interest in the literature (Vassaux and Massoulié (2025); Ameen and Hajek (2025)).

Set-up. Multiple Gaussian graph alignment can be formalized as follows. We observe $p \geq 2$ undirected weighted graphs $G_1, \dots, G_p$ on the set of edges $E = \binom{[n]}{2}$. We suppose that there exists $\pi^\star = (\pi_1^\star, \dots, \pi_p^\star)$ drawn uniformly on $(\mathcal{S}_n)^p$, where $\mathcal{S}_n$ stands for the set of permutations of $[n]$, such that the following holds. Conditionally on $\pi^\star$, the edge weights $(G_{je})_{j\in[p], e\in E}$ have $\mathcal{N}(0,1)$ marginal distribution, with covariance matrix defined for all $(j,e) \neq (j',e')$ by

$$\mathrm{Cov}(G_{je}, G_{j'e'}) = \rho \mathbf{1}_{\Pi_j^\star(e) = \Pi_{j'}^\star(e')}, \tag{1}$$

where we write, for $e = \{u, u'\}$,

$$\Pi_j^\star(e) = \left\{\pi_j^\star(u), \pi_j^\star(u')\right\} ,$$

and where $\rho \in [0,1]$ is the correlation parameter, that is the correlation between the Gaussian signal on aligned edges in (1). If $\rho = 0$, the graphs are independent. If $\rho = 1$, the graphs are identical, up to relabeling of the nodes.

Our goal is to recover the p-tuple of permutations $\pi^\star = \left(\pi_1^\star, \ldots, \pi_p^\star\right)$. Since the model is identifiable only up to a global permutation of $\pi_1^\star, \ldots, \pi_p^\star$, we are interested in controlling the proportion of misclassified points:

$$\mathrm{err}(\pi, \pi^\star) := \min_{\psi \in \mathcal{S}_n} \frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\left\{\psi(\pi_j(u)) \neq \pi_j^\star(u)\right\} . \tag{2}$$

Alternatively, the weighted graphs $G_1, \ldots, G_p$ can be sampled as follows. First draw H_0 a complete weighted graph with independent $\mathcal{N}(0,1)$ edges. Then, for all $j \in [p]$ and $e \in E$, set

$$G_{je} = \sqrt{\rho}\, H_{0\Pi_j^\star(e)} + \sqrt{1-\rho}\, Z_{je}, \tag{3}$$

where $H_{0\Pi_j^\star(e)}$ denotes the weight of edge $\Pi_j^\star(e)$ in the graph H_0, and where $(Z_{je})_{e \in E}$, $j \in [p]$ are i.i.d. copies of H_0.

We say that an estimator $\widehat{\pi} : \mathbb{R}^{E \times p} \to (\mathcal{S}_n)^p$ of $\pi^\star$ achieves:

- *perfect recovery* if $\mathrm{err}\left(\widehat{\pi}, \pi^\star\right) = 0$ with high probability[1] when $n \to \infty$,
- *partial recovery* if $\mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right] \to 0$ when $n \to \infty$,
- *non-trivial recovery* if $\mathbb{E}[\mathrm{err}\left(\widehat{\pi}, \pi^\star\right)]$ does not converge to 1 when $n \to \infty$.

Note that, since partial recovery does not require any specific rate of convergence of $\mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right]$, it is a weaker notion than exact recovery. Also note that if $\pi \in (\mathcal{S}_n)^p$ is taken uniformly at random, then $\mathbb{E}[\mathrm{err}\left(\pi, \pi^\star\right)] \to 1$ when $n \to \infty$. This motivates the 'non-trivial recovery' terminology.

Related work. The problem of aligning two Gaussian graphs has been studied both for characterizing the information-theoretic (IT) threshold of the problem (Ganassali (2022); Wu et al. (2022)) and for finding polynomial-time algorithms able to recover the single hidden permutation $(\pi_1^\star)^{-1} \circ \pi_2^*$ (Ganassali et al. (2022); Fan et al. (2020); Ding and Li (2022)). For this problem, it is well known that the IT threshold for non-trivial recovery is exactly

$$\rho = \sqrt{\frac{4\log(n)}{n}} \tag{4}$$

and that, above this threshold, perfect recovery of $\pi^\star$ is feasible via maximum likelihood estimation.

However, this maximum likelihood estimator is not computable in polynomial time (see Pardalos et al. (1993) and Makarychev et al. (2014)). The state-of-the-art polynomial time procedure for aligning two correlated Gaussian graphs, developed by Ding and Li (2022) requires a non-vanishing correlation ρ. This suggests the existence of a statistical-computational gap for the problem of aligning two correlated Gaussian graphs.

More recently, alignment of multiple correlated Gaussian graphs has been studied by Vassaux and Massoulié (2025) and Ameen and Hajek (2025). In a regime where the number of graphs p is a fixed constant and where n tends to infinity, Ameen and Hajek (2025) studied the related detection problem, namely distinguishing the null hypothesis under which the p graphs are p independent Gaussian graphs, from the alternative where they are correlated according to the model (1). The authors derived a sufficient condition

$$\rho \geq \sqrt{\frac{8\log(n)}{(n-1)p}}$$

for reliable detection, as well as a sufficient condition

$$\rho \leq \sqrt{\left(\frac{4}{p-1} - \varepsilon\right)\frac{\log(n)}{n}}$$

for impossibility of detection.

Then, the IT threshold for non-trivial recovery is proven by Vassaux and Massoulié (2025) to be exactly

$$\rho = \sqrt{\frac{8\log(n)}{np}},$$

and here again, above this threshold, perfect recovery of $\pi^\star$ is possible with high probability. In this regime with fixed p and $n \to \infty$, an *all-or-nothing phenomenon* occurs: either non-trivial recovery is impossible, or perfect recovery is achievable.

The case where the number of observed graphs p can also grow with the number of nodes n remains open, and raises the following question:

> *(Q1). When the number of observed graphs p grows with the number of nodes n, what is the*

[1] an event $\mathcal{E}$ holds *with high probability* if $\mathbb{P}(\mathcal{E}) \to 1$ when $n \to \infty$.

information-theoretic threshold for perfect/exact recovery, and does an all-or-nothing phenomenon always occur?

Evidence for a statistical–computational gap for Erdős-Rényi graph alignment when $p = 2$ has been established by recent results (Ding et al. (2023); Ganassali et al. (2024b); Li (2025)). However, to the best of our knowledge, the existence of a statistical–computational gap for (multiple) Gaussian graph alignment has not been studied. We formalize this in the following question:

(Q2). Can we provide evidence for the existence of a statistical–computational gap for multiple Gaussian graph alignment?

The low-degree model of computation. In many high dimensional statistical problems, such as sparse PCA Hopkins et al. (2017), planted clique Barak et al. (2019), clustering Lesieur et al. (2016); Even et al. (2024, 2025b) or Gaussian graph alignment which is the focus of this paper, state-of-the-art polynomial time algorithms fail to reach the statistical performance provably achievable by algorithms free of computational constraints. These observations have led the community to conjecture the existence of *statistical-computational gaps*, that is gaps between the best polynomial-time performance and the best performance with no computational constraints, for given problems. In light of this conjecture, a cornerstone of research in high-dimensional statistics is to determine rigorous lower bounds for specific classes of algorithms.

In order to address average-case complexity, lower bounds are obtained for different models of computation, through the S-o-S hierarchy (Hopkins et al. (2017)), the overlap gap property Gamarnik (2021), statistical query (Brennan and Bresler (2020)) and low-degree polynomials (Schramm and Wein (2022); Hopkins (2018); Wein (2025)).

In this paper, we will be interested in the low-degree model of computation: we consider the class of algorithms which can be written as multivariate polynomials in the observation with degree D. The low-degree model of computation has recently attracted considerable attention due to its ability to provide lower bounds matching state-of-the-art upper-bounds for polynomial time algorithms in different models: community detection (Hopkins et al. (2017); Sohn and Wein (2025); Carpentier et al. (2025a,b)), clustering (Even et al. (2024, 2025b)), and among others, Erdős-Rényi graph matching (Ding et al. (2023); Li (2025)). We refer to Wein (2025) for a recent survey. In this framework, the failure for a given task of polynomials of degree $D = O\left(\log(T)^{1+\eta}\right)$, where T is the dimension of the observation, is taken as an evidence for the failure of all polynomial-time algorithms.

For estimation problems, the seminal work by Schramm and Wein (2022) offers a starting point for analyzing the performance of the best low-degree polynomial of degree D. Our analysis heavily builds upon their Theorem 2.2, which provides a formula for the case of Gaussian additive models.

Our contribution. In this paper, we provide answers to questions (Q1) and (Q2) previously defined.

For (Q1), we generalize the IT threshold $\rho = \sqrt{\frac{8\log(n)}{np}}$ Vassaux and Massoulié (2025) to regimes where p is not necessarily constant. We characterize the informational thresholds for partial and perfect recovery, which are respectively

$$\rho \gtrsim \frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}} \ ,$$

and

$$\rho \gtrsim \frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}} \ ,$$

where $\gtrsim$ hides multiplicative constants. Moreover, between these two thresholds, we obtain an exponential decay of the error in terms of n, p and ρ.

In particular, when $\log(p) = \omega(\log(n))$, these two thresholds differ significantly (by more than a multiplicative constant), and the all-or-nothing phenomenon established in Vassaux and Massoulié (2025) vanishes; there is now a different IT threshold for partial and exact recovery.

For (Q2), we provide a computational threshold in the low-degree (LD) framework. We prove that low-degree polynomials fail at non-trivial recovery whenever

$$\rho \leq_{\log} 1,$$

where $\leq_{\log}$ hides logarithmic factors.

Main message. These results suggest that, while observing a large number $p \geq 2$ of correlated Gaussian graphs helps information-theoretically in estimating $\pi^\star$, the computational difficulty of the problem remains equivalent up to logarithmic factors to the case $p = 2$, that is to that of aligning two correlated Gaussian graphs.

Moreover, if the low degree conjecture is true, performing polynomial time alignment of p Gaussian graphs requires the correlation ρ to be at least of order $\frac{1}{\mathrm{polylog}(n,p)}$.

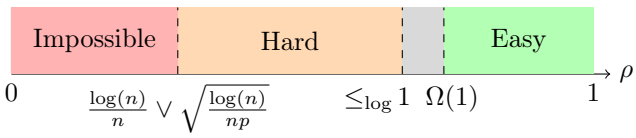

Figure 1: Statistical-computational landscape of partial recovery in multiple Gaussian graph alignment. The grey region is not fully covered by our results, since our lower bound involve logarithmic factors in n and p.

Therefore, an optimal method up to some logarithmic factors is to perform pairwise alignment of the p graphs using the algorithm of Ding and Li (2022). In other words, there is no significant gain at a computational level in having access to multiple graphs. We summarize our results in the phase diagram on Fig. 1.

Outline of the paper. We derive the maximum likelihood estimator and provide upper and lower information-theoretic bounds in Section 2, present the computational lower bound in Section 3, and discuss these results along with their connections to the literature in Section 4. All proofs are deferred to the appendices, while the main text provides an interpretation of the results.

Notations. Throughout this paper, we adopt the following notations. We write $f(n,p) \lesssim g(n,p)$ if there exists a positive constant $c > 0$ such that $f(n,p) \leq cg(n,p)$ for all $n, p \geq n_0$ for some constant $n_0 \geq 1$. We write $f(n,p) \asymp g(n,p)$ if both $f(n,p) \lesssim g(n,p)$ and $f(n,p) \gtrsim g(n,p)$. We write $f(n,p) \leq_{\log} g(n,p)$ if there exist polynomials P, Q independent of everything else such that $f(n,p) \leq \frac{P(\log(n),\log(p))}{Q(\log(n),\log(p))} g(n,p)$ for all $n, p \geq n_0$ for some constant $n_0 \geq 1$. For $n \geq 1$, We write $[n]$ for the set $\{1, \dots, n\}$, $\mathcal{S}_n$ for the set of permutations of $[n]$, and $N = \binom{n}{2}$ be the number of edges per graph. E denotes the edge set $\binom{[n]}{2}$.

We denote by $u, v, w, u', v', w' \in [n]$ the nodes of a graph, and by $j, k, \ell \in [p]$ the indices of graph copies. The undirected edges of a graph G_j are denoted by $e, e', f, f' \in E$, and we write G_{je} to refer to a particular edge e of graph G_j.

For each graph G_j, we consider node permutations $\pi_j \in \mathcal{S}_n$, and their corresponding edge permutations $\Pi_j \in \mathcal{S}_N$, defined by

$$\Pi_j(e) = \{\pi_j(u), \pi_j(v)\} \in E \quad \text{for } e = \{u, v\} \in E.$$

It is important to note that Π_j and π_j are distinct objects: the former acts on edges in E, while the latter acts on nodes in $[n]$.

We denote by $\mathbb{P}$ the joint law of $(\pi^\star, G_1, \dots, G_p)$ and, for $\pi \in (\mathcal{S}_n)^p$, we denote by $\mathbb{P}_\pi$ the law of $(\pi^\star, G_1, \dots, G_p)$ conditionally on $\pi^\star = \pi$. We write $\mathbb{E}$ (resp. $\mathbb{E}_\pi$) for expectation over $\mathbb{P}$ (resp. $\mathbb{P}_\pi$).

2 INFORMATION BARRIERS FOR EXACT AND PARTIAL RECOVERY

In this section, we provide the IT thresholds for partial and perfect recovery. In Section 2.1, we derive the maximum likelihood estimator and introduce useful objects for our analysis. We analyze its performance in Section 2.2, in regimes where p is moderately small (namely $p \lesssim \rho^{-1}$). In Section 2.3, we propose a different estimator for larger p along with different upper bounds. In Section 2.4, we provide lower bounds which match the previously obtained upper bounds, for partial and perfect recovery. The combination of those upper bounds and lower bounds fully characterize the information barriers for exact and partial recovery.

2.1 Derivation of the maximum likelihood estimator

In line with model (1), define

- $G = (G_{je})_{j \in [p], e \in E} \in \mathbb{R}^{pN \times 1}$, viewed as a column vector.
- $H \in \mathbb{R}^{pN \times 1}$ the vector with p copies of H_0, that is, for all $j \in [p]$, $e \in E$, $H_{je} = H_e$.
- $\Pi^\star \in \mathbb{R}^{pN \times pN}$ the aggregated block matrix of the edge permutations, that is, for all $j, j' \in [p]$, $e, e' \in E$, $\Pi^\star_{je, j'e'} = \mathbf{1}\{j = j'\} \mathbf{1}\{\Pi^\star_j(e) = e'\}$.
- $G^\star \in \mathbb{R}^{pN \times 1}$ defined as $G^\star = \Pi^\star H$. For all $j \in [p]$, $e \in E$, $G^\star_{je} = H_{\Pi^\star_j(e)}$.
- $Z \in \mathbb{R}^{pN \times 1}$ the vector of aggregated noise variables, $Z = (Z_{je})_{j \in [p], e \in E} \sim \mathcal{N}(0, I_{pN})$.

A compact formulation of model (1) is given by

$$G = \sqrt{\rho}\, \Pi^\star H + \sqrt{1-\rho}\, Z = \sqrt{\rho}\, G^\star + \sqrt{1-\rho}\, Z\,. \quad (5)$$

Under $\mathbb{P}_\pi$, G is a centered Gaussian vector with covariance matrix $\Sigma(\pi) \in \mathbb{R}^{pN \times pN}$ given by

$$\begin{aligned}\Sigma(\pi) &= \mathbb{E}[GG^T] = \rho \Pi \mathbb{E}[HH^T](\Pi)^T + (1-\rho)\mathbb{E}[ZZ^T] \\ &= \Pi[(\rho J_p + (1-\rho) I_p) \otimes I_N](\Pi)^T,\end{aligned}$$

where $\otimes$ denotes the Kronecker product, using that $\Pi\Pi^T = I_{pN}$. Unsurprisingly, $\Sigma(\pi)$ is invariant under

global permutation, just as the law of G. Note that for $0 \le \rho < 1$, $\Sigma(\pi)$ is invertible with inverse

$$\begin{aligned}\Sigma^{-1}(\pi) &= \Pi[(\rho J_p + (1-\rho)I_p)^{-1} \otimes I_N]\Pi^T \\ &= \Pi\left[\left(-\frac{\rho}{(1-\rho)(1+(p-1)\rho)}J_p + \frac{1}{1-\rho}I_p\right)\otimes I_N\right]\Pi^T \\ &= -\frac{\rho}{(1-\rho)(1+(p-1)\rho)}\Pi(J_p\otimes I_N)\Pi^T + \frac{1}{1-\rho}I_{pN}\,. \end{aligned} \tag{6}$$

Given the shape of the Gaussian likelihood and since $\det\Sigma(\pi)$ does not depend on π, the maximum likelihood estimator (MLE) for this problem is given by

$$\begin{aligned}\widehat{\pi} &\in \arg\min_{\pi\in(\mathcal{S}_n)^p} G^T\Sigma^{-1}(\pi)G \\ &= \arg\max_{\pi\in(\mathcal{S}_n)^p} G^T\Pi(J_p\otimes I_N)\Pi^T G\,. \end{aligned} \tag{7}$$

For any $\pi \in (\mathcal{S}_n)^p$, define its *normalized edge alignment matrix* $B(\pi) \in \mathbb{R}^{pN\times pN}$ as

$$[B(\pi)]_{je,j'e'} = \frac{1}{p}\mathbf{1}\left\{\Pi_j(e) = \Pi_{j'}(e')\right\}\,. \tag{8}$$

Define $\mathcal{B} := B((\mathcal{S}_n)^p)$. Note that $\pi \mapsto B(\pi)$ is a bijection from $(\mathcal{S}_n)^p/\equiv$ onto $\mathcal{B}$, where $\pi \equiv \pi'$ in $(\mathcal{S}_n)^p$ if and only there exists $\phi \in \mathcal{S}_n$ such that $\pi'_j = \phi\circ\pi_j$ for all $j \in [p]$.

Note also that any $B \in \mathcal{B}$ satisfies

$$B^T = B, B^2 = B, B\mathbf{1} = \mathbf{1}, \mathrm{Tr}(B) = N\,.$$

and that in (6),

$$\Pi^\star(J_p\otimes I_N)(\Pi^\star)^T = pB^\star, \tag{9}$$

where $B^\star := B(\pi^\star)$. Finding the MLE in (7) is thus equivalent to solving

$$\widehat{B} = \arg\max_{B\in\mathcal{B}}\langle GG^T, B\rangle, \tag{10}$$

and we find back the MLE $\widehat{\pi}$ taking any $\widehat{\pi} \in B^{-1}(\widehat{B})$.

2.2 Moderately small p: exact recovery via the MLE

In order to prove an IT upper bound, we show that the maximum likelihood estimator in its matrix version (10) enables to recover $B^\star$, and thereby $\pi^\star$, with the desired error guarantees. We are able to analyze directly the MLE in the case where $p \lesssim \rho^{-1}$, and will address the case $p \gtrsim \rho^{-1}$ slightly differently in the next section.

Theorem 2.1. *There exists numerical constants C, c_1, c_2 and n_0 such that the following holds. Whenever $\rho \ge c_1\sqrt{\frac{\log(n)}{np}}$, $p \le C\rho^{-1}$ and $n \ge n_0$, we have*

$$\mathbb{P}\left[\mathrm{err}(\widehat{\pi},\pi^\star) \neq 0\right] \le \exp\left(-c_2 np\rho^2\right),$$

where $\widehat{\pi}$ is the MLE defined in (7).

The above Theorem is proved in Appendix A. We remark that the probability in the right-hand side of the inequality of Theorem 2.1 is upper-bounded by $n^{-c_1c_2} = o(1)$, hence the feasibility of exact (and partial) recovery under these assumptions. In this regime, we recover the threshold from Vassaux and Massoulié (2025), up to some multiplicative numerical constant, which we generalize to higher values of p.

Note also that under the assumptions of Theorem 2.1, $p \lesssim \rho^{-1} \lesssim \sqrt{\frac{np}{\log(n)}}$, which gives $p \lesssim \frac{n}{\log(n)}$. We will refer to this as the *moderately small-p* regime.

2.3 For larger p: partial recovery upper bound

When $p \gtrsim \rho^{-1}$, the approach relies on computing the maximum likelihood estimator for the first p' graphs, with $p' \asymp \rho^{-1}$, and aligning the rest of the graphs pairwise with the aggregated signal on the first p' graphs.

Namely, we build an estimator $\hat{\pi} \in (\mathcal{S}_n)^p$ with the following procedure.

1. Let $p' = C\rho^{-1}$, where C is determined by Theorem 2.1, set $(\hat{\pi}_1,\dots,\hat{\pi}_{p'})$ to be the MLE of $(G_1,\dots,G_{p'})$ as defined in (7). Let $\widehat{G}'$ be the graph defined by
$$\widehat{G}'_e = \frac{1}{p'}\sum_{j\le p'} G_{j(\widehat{\Pi}_j)^{-1}(e)}\ ;$$
2. For all $j > p'$, compute the remaining $\widehat{\pi}_j$ aligning G_j with $\widehat{G}'$ with the MLE estimator (7) on $(\widehat{G}', G_j)$.
3. Return the estimator
$$\widehat{\pi} = (\widehat{\pi}_1,\dots,\widehat{\pi}_p)\,. \tag{11}$$

Theorem 2.1 suggests that, for ρ large enough, $(\hat{\pi}_1,\dots,\hat{\pi}_{p'})$ will recover exactly $(\pi^\star_1,\dots,\pi^\star_{p'})$ with high probability and thus that $\widehat{G}'$ will be isomorphic to the graph G' defined by

$$G'_e = \frac{1}{p'}\sum_{j\le p'} G_{j(\Pi^\star_j)^{-1}(e)}\ .$$

Then, appealing again to Theorem 2.1, one can control the probability of recovering $\pi^\star_j$, for $j > p'$, by aligning G_j with $\widetilde{G}'$. We deduce the next theorem, proved in Appendix B.

Theorem 2.2. *Let C be the constant defined in Theorem 2.1. There exists numerical constants c_1, c_2 and n_0 such that the following holds. Whenever $\rho \geq c_1 \left(\frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}} \right)$, $p > C\rho^{-1}$ and $n \geq n_0$, we have*

$$\mathbb{E}\left[\mathrm{err}(\hat{\pi}, \pi^\star)\right] \leq \exp\left(-c_2 n\rho\right) \quad ,$$

where $\widehat{\pi}$ is defined in (11).

The exponential term in the right-hand side of the inequality of Theorem 2.2 is upper-bounded by $n^{-c_1 c_2} = o(1)$, hence the feasibility of partial recovery under these assumptions.

Combining Theorem 2.1 and Theorem 2.2, we deduce that partial recovery is possible with high probability whenever

$$\rho \gtrsim \left(\frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}} \right) \quad .$$

When the correlation is even higher, namely when

$$\rho \gtrsim \frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}} \tag{12}$$

by a large enough constant, observe that

- the probability of error $\mathbb{P}\left[\mathrm{err}(\widehat{\pi}, \pi^\star) \neq 0\right]$ in Theorem 2.1 drops below $1/(np)^2$,
- the expectation $\mathbb{E}\left[\mathrm{err}(\hat{\pi}, \pi^\star)\right]$ in Theorem 2.2 drops below $1/(np)^3$, and by Markov's inequality,

$$\begin{aligned}\mathbb{P}\left[\mathrm{err}(\widehat{\pi}, \pi^\star) \neq 0\right] &= \mathbb{P}\left[\mathrm{err}(\widehat{\pi}, \pi^\star) \geq 1/(np)\right] \\ &\leq 1/(np)^2 \,.\end{aligned}$$

Since the MLE is always Bayes-optimal for perfect recovery, we deduce from the above that the MLE achieves perfect recovery with high probability when ρ is large enough[2].

Corollary 2.3. *There exists numerical constants c_1, c_2 and n_0 such that the following holds. Whenever $\rho \geq c_1 \left(\frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}} \right)$ and $n \geq n_0$, we have, with probability at least $1 - 1/(np)^2$,*

$$\mathrm{err}(\widehat{\pi}, \pi^\star) = 0 \,,$$

where $\widehat{\pi}$ is the MLE defined in (7).

[2]Theorems 2.1 and 2.2 show that only $\rho \gtrsim \frac{\log(np)}{n} \vee \sqrt{\frac{\log(n)}{np}}$ is needed for exact recovery. Our statement is equivalent since $\frac{\log(np)}{n} \vee \sqrt{\frac{\log(n)}{np}} \asymp \frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}}$.

Remark 2.1 (On the exponential decay)**.** *Theorems 2.1 and 2.2 show in particular that the error $\mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right]$ exhibits an exponential decay with respect to $n\left(\rho \wedge p\rho^2\right)$. Such exponential behavior is reminiscent of what is observed in several other high-dimensional inference problems, including clustering (Ndaoud (2022); Giraud and Verzelen (2019); Even et al. (2024)), multiple feature matching (Even et al. (2025a)), and supervised classification (Giraud and Verzelen (2019)).*

2.4 Informational lower bounds

The next result, proved in Appendix C, provides an information-theoretic lower bound for partial recovery.

Theorem 2.4. *There exists numerical constants $n_0, c, c' > 0$ such that the following holds. If $n \geq n_0$ and $\rho \leq c \left(\frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}} \right)$, then*

$$\inf_{\widehat{\pi}} \mathbb{E}[\mathrm{err}(\widehat{\pi}(G), \pi^\star)] \geq c',$$

where the infimum is taken over all measurable functions $\widehat{\pi} : \mathbb{R}^{E \times p} \to (\mathcal{S}_n)^p$. Thus, partial recovery is impossible.

Theorems 2.1, 2.2 and 2.4 show that the estimator defined in (11) is optimal up to some numerical constant and that the IT threshold for partial recovery is

$$\rho \gtrsim \frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}} \quad . \tag{13}$$

Let us now give some intuition for why the IT lower bound (13) takes this particular form. Consider a genie-aided version of the problem in which we are given $\pi_2^\star = \cdots = \pi_p^\star = \mathrm{id}$ and we seek to recover $\pi_1^\star$. A sufficient statistic for estimating $\pi_1^\star$ is $(G_1, \overline{G})$, where

$$\overline{G}_e = \frac{1}{p-1} \sum_{j=2}^{p} G_{je} \,.$$

In these two graphs, the edge-weight variances are 1 and $\frac{1+(p-2)\rho}{p-1}$, respectively, while each correlated edge pair has covariance ρ. Thus, the problem reduces to aligning the two correlated Gaussian graphs G_1 and $\overline{G}$, with effective edge correlation

$$\rho' = \frac{\sqrt{p-1}\,\rho}{\sqrt{1+(p-2)\rho}} \lesssim \sqrt{\rho} \vee \sqrt{p}\,\rho \,. \tag{14}$$

For $p = 2$, the previously known information-theoretic threshold (4) implies that a necessary condition for partial recovery of $\pi_1^\star$ is

$$\rho' \gtrsim \sqrt{\frac{\log n}{n}},$$

and combining this with (14) yields exactly (13).

In the regime where $p \lesssim n/\log(n)$, the threshold for partial recovery reduces to $\rho \gtrsim \sqrt{\frac{\log(n)}{np}}$, which matches, up to constants, the condition obtained by Vassaux and Massoulié (2025) for fixed p. While our result is not constant sharp, it naturally generalizes the previous bounds to settings where p may also be moderately large, that is for $p \lesssim n/\log(n)$.

When the number of graphs becomes large ($p \gtrsim n/\log(n)$), however, the threshold saturates at $\rho \gtrsim \frac{\log(n)}{n}$ and no longer decreases with p: the abundance of observed graphs stops to help the statistician.

We also provide an information-theoretic lower bound for perfect recovery in the next result.

Theorem 2.5. *For all $\varepsilon > 0$, there exists constants c, n_0 depending only on ε, such that, if $\rho \leq c\left(\frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}}\right)$ and $n \geq n_0$, then*

$$\inf_{\widehat{\pi}} \mathbb{P}[\mathrm{err}(\widehat{\pi}, \pi^\star) \neq 0] \geq 1 - \varepsilon \ ,$$

where the infimum is taken over all measurable functions $\widehat{\pi} : \mathbb{R}^{E \times p} \to (\mathcal{S}_n)^p$. Thus, perfect recovery is impossible.

Taken together, Corollary 2.3 and Theorem 2.1 characterize the IT threshold for perfect recovery, which is

$$\rho \gtrsim \frac{\log(np)}{n} \vee \sqrt{\frac{\log(np)}{np}} \ . \tag{15}$$

In regimes where $\log(p)$ is much larger than $\log(n)$, we observe a gap between the IT threshold for partial recovery (13) and perfect recovery (15). As mentioned in the introduction, this contrasts with the regime studied by Vassaux and Massoulié (2025) (constant p), where an all-or-nothing phenomenon arises: either non-trivial recovery is impossible, or perfect recovery occurs with high probability.

3 LOW DEGREE LOWER BOUND

As discussed in the introduction, even when $p = 2$, the performance of the best known polynomial-time algorithm Ding and Li (2022) does not match the optimal information-theoretic threshold (4). This suggests the existence of a statistical-computational gap for (multiple) Gaussian graphs alignment. In this section, we provide evidence of this statistical-computational gap by establishing a low degree polynomial lower bound, and we further characterize–up to logarithmic factors–the computational barrier for general p.

As already pointed out by Even et al. (2025b), low-degree polynomials are ill-suited to capture combinatorial constraints and therefore cannot be directly employed to construct estimators of the permutations $\pi_1^\star, \dots, \pi_p^\star$. Instead, we concentrate on the task of estimating the *node alignment matrix* $M^\star \in \mathbb{R}^{(n\times p)\times(n\times p)}$, defined as

$$M^\star_{(u,j),(u',j')} = \mathbf{1}\left\{\pi_j^\star(u) = \pi_{j'}^\star(u')\right\} \ .$$

The next proposition, slightly adapted from Proposition 2.1 of Even et al. (2025a), shows that computational hardness for estimating $M^\star$ for the L^2 loss implies computational hardness for estimating $\pi_1^\star, \dots, \pi_p^\star$ for the err loss. We refer to Appendix F for a proof.

Proposition 3.1. *Suppose that*

$$\begin{aligned}\mathrm{MMSE}_{poly} := &\inf_{\hat{M}\ \mathrm{poly-time}} \frac{1}{p(p-1)n^2}\,\mathbb{E}\left[\|\hat{M} - M^\star\|_F^2\right] \\ =&\frac{1}{n}(1-\varepsilon),\end{aligned}$$

with $0 \leq \varepsilon \leq 1$. Then, for all polynomial-time estimator $\hat{\pi}$ of $\pi^\star$, one has

$$\mathbb{E}\left[(1 - \mathrm{err}(\hat{\pi}, \pi^\star))^2\right] \leq \sqrt{\varepsilon} \ .$$

Therefore, in order to provide evidence of a computational barrier for the problem of estimating $M^\star$ (and hence $\pi^\star$), we provide a low-degree polynomial lower bound for estimating $M^\star$. We consider the *degree-D minimum mean squared error* (Schramm and Wein (2022)) defined by

$$\mathrm{MMSE}_{\leq D} := \inf_{\widehat{M} \in \mathbb{R}_D[G_{1:p}]} \frac{1}{p(p-1)n^2}\,\mathbb{E}\left[\|\widehat{M} - M^\star\|_F^2\right], \tag{16}$$

where $\mathbb{R}_D[G_{1:p}]$ is the space of polynomials in the entries of $(G_1, \dots, G_p)$ of degree at most D.

We observe that the trivial estimator, $\mathbb{E}[M^\star]$, has loss

$$\frac{1}{p(p-1)n^2}\,\mathbb{E}\left[\|\,\mathbb{E}[M^\star] - M^\star\|^2\right] = \frac{1}{n} - \frac{1}{n^2} \,.$$

Our next result characterizes a regime on which no low-degree polynomial performs significantly better than this trivial estimator. We refer to Section E for a proof of this theorem.

Theorem 3.2. *Let $D \leq n-2$ and suppose*

$$\zeta := \frac{D^3\sqrt{\rho}}{1-\sqrt{\rho}}\sqrt{1+D/2}\,\frac{2}{\left(1-\frac{D+1}{n}\right)^2} < 1 \,.$$

Then,

$$\mathrm{MMSE}_{\leq D} \geq \frac{1}{n} - \frac{1}{n^2} - \frac{2}{(n-1-D)^2}\zeta\frac{1+\zeta}{1-\zeta} \ .$$

In particular, taking $D = O\left(\log(np)^{1+\eta}\right)$ for a small $\eta > 0$, we have that, whenever $\rho = o\left(\frac{1}{\log(np)^{7(1+\eta)}}\right)$,

$$\mathrm{MMSE}_{\leq D} = \frac{1}{n} - \frac{1}{n^2}(1 + o(1)) .$$

Since the failure of $O\left(\log(np)^{1+\eta}\right)$-degree polynomial estimators is taken as evidence of hardness for polynomial-time algorithms, Theorem 3.2 suggests that estimating $M^\star$ (and hence $\pi^\star$) is computationally hard whenever

$$\rho \leq_{\log} 1 .$$

On the contrary, it is well known that as soon as ρ is non-vanishing, alignment of two correlated Gaussian graphs is possible in polynomial-time (Ding and Li (2022)). Thus, when ρ is not vanishing, a naïve strategy for aligning p correlated random graphs in polynomial time would be two fix $\pi_1^\star = \mathrm{id}$ and to successively align every other graph with G_1 pairwise with the iterative algorithm developed by Ding and Li (2022).

Theorem 3.2 suggests that this naïve strategy is optimal up to logarithmic factors. In particular, Theorem 3.2 highlights that the presence of multiple graphs does not simplify the problem computationally, and, in fact, enlarges the statistical–computational gap.

Theorem 1.1. from Ding and Li (2022) directly implies the following.

Corollary 3.3. *If $\rho > \varepsilon$ with $\varepsilon > 0$ an independent constant, then there exists a constant $C = C(\varepsilon) > 0$ and an algorithm with $O\left(p \times n^C\right)$-running time such that $\mathbb{E}\left[\mathrm{err}\left(\widehat{\pi}, \pi^\star\right)\right] \to 0$ when $n \to \infty$, that is, achieving partial recovery.*

High-level proof strategy of Theorem 3.2 In Gaussian additive models, Theorem 2.2 of Schramm and Wein (2022) provides a lower bound of the $\mathrm{MMSE}_{\leq D}$ with respect to a sum squared joint cumulants of the signal (see Proposition E.1 for a statement of this formula with the notation of our model). We refer to Novak (2014) for a backround on cumulants. However, for general p, relying only on this formula is not sufficient. Thus, we proceed in two steps:

1. We consider the case $p = 2$. In this case, it is sufficient to appeal to Theorem 2.2 of Schramm and Wein (2022) and to compute the cumulants involved. We point out that for analyzing those cumulants, we build upon the work of Even et al. (2025b) which allows us to consider cumulants of Bernoulli variables depending only on the random permutations $\pi_1^\star, \pi_2^\star$.
2. We observe that the problem of aligning p correlated graphs with correlation ρ is harder than the problem of aligning two graphs with correlation $\sqrt{\rho}$. Plugging the first step is then sufficient for concluding the proof of the theorem for general p.

4 DISCUSSION

Comparison with n-dimensional feature matching. The problem of multiple feature matching (Even et al. (2025a)) is another statistical problem where permutations are hidden in the data. Here, we observe p correlated matrices $Y_1, \ldots, Y_p \in \mathbb{R}^{n \times n}$ where the rows of each matrix Y_j is relabeled according to some $\pi_j^\star$. This model comes with a Gaussian mixture flavor : we suppose that the rows of each matrix are drawn independently, where row u of matrix j is a Gaussian vector $\mathcal{N}\left(\mu_{\pi_j^\star(u)}, I_n\right)$ of dimension n. Assume that the row relabelings $\pi_1^\star, \ldots, \pi_p^\star$ and the centers $\mu_1, \ldots, \mu_n$ are hidden. Even et al. (2025a) characterizes the statistical-computational landscape of this problem, generalizing the results from Collier and Dalayan (2016) which studied only the case $p = 2$.

To clarify the comparison between the two problems, we assume that the centers $\mu_1, \ldots, \mu_n$ are taken with a gaussian prior. Then, $Y_1, \ldots, Y_p$ are correlated (non-symmetric) Gaussian matrices with $\mathrm{Cov}\left(Y_{juv}, Y_{j'u'v'}\right) = \rho \mathbf{1}\left\{\pi_j^\star(u) = \pi_j^\star(u')\right\} \mathbf{1}\left\{v = v'\right\}$, for some parameter $\rho > 0$ depending on the variance of the centers $\mu_1, \ldots, \mu_n$. Heuristically, the difference between multiple feature matching and multiple gaussian graph alignment is that in the first problem, only one dimension (the rows of the matrices) is permuted, whereas in the second problem, both coordinates are permuted. One would therefore expect the second problem to be more difficult.

Yet, at the information level, the thresholds for perfect recovery and partial recovery coincide, up to some numerical constant (see Even et al. (2025b)). Moreover, we observe the same exponential decay of the error above the threshold for partial recovery. With no computational constraints, both problems are equivalently difficult, but the optimal estimator is the solution to a combinatorial optimization problem over all p-tuple of permutations (or p'-tuple of permutation with $p' \asymp \rho^{-1}$ in the case of Gaussian matrix alignment), and is not computable in polynomial time for $p \geq 3$, and even for $p = 2$ in Gaussian graphs alignment.

However, Corollary 2.3 of Even et al. (2025a) shows that there exists polynomial-time algorithms for multiple feature matching which exactly recover $\pi^\star$ as soon as

$$\rho \geq_{\log} \frac{1}{\sqrt{n}} \wedge \frac{1}{\sqrt{p}} .$$

This suggests that the problem of multiple feature matching is much easier that the problem of multiple gaussian graph alignment where our computational barrier gives the evidence that recovering $\pi^\star$ is computationally tractable only when $\rho \geq_{\log} 1$.

Hence, the presence of a bi-dimensional structure hardens the problem at a computational level. In our problem, the permutation of both coordinates breaks the low-rank structure of the signal and prevents any spectral method from being informative. On the contrary, for large p, in the case of multiple feature matching, the computational barrier coincides with a spectral barrier known as the *BBP transition* Baik et al. (2005), below which the spectrum of the Gram matrix YY^T no longer carries information about the signal.

Hardness for complex bi-dimensional structures. The computational intractability of estimating bi-dimensional structures in matrix problems has also been observed in Even et al. (2025b), where the authors compare the statistical–computational landscapes of high-dimensional clustering with related problems exhibiting hidden bi-dimensional structure, such as sparse clustering and bi-clustering. Without computational constraints, the full structure can be exploited, whereas polynomial-time algorithms fail to leverage the dependencies across both rows and columns simultaneously.

This phenomenon also underlies the statistical–computational gaps in planted submatrix estimation (Schramm and Wein (2022)) and graphon estimation (Luo and Gao (2023)).

Multiple Gaussian graph alignment thus provides another instance of a problem in which a complex hidden bi-dimensional structure gives rise to a wide statistical–computational gap.

Limitation: non-sharpness of the computational lower bound. Our low-degree lower bound matches the state-of-the art polynomial-time algorithm only up to some logarithmic factors. A recent line of work succeeds in some problems to remove logarithmic factors in low-degree lower-bounds Sohn and Wein (2025); Even et al. (2025b), but at the price of a more subtle analysis. In Even et al. (2025b), the analysis highly rely on the independencies of the latent variables, which should not be expected to work for our problem, due to the fact that the values of the hidden permutations $\pi_1^\star, \ldots, \pi_p^\star$ are weakly dependent. We did not investigate the approach of Sohn and Wein (2025).

However, to the best of our knowledge, in estimation problems where the low-degree lower bound exactly matches the performance of the best known polynomial-time algorithm, the computational barrier is characterized by a constant (e.g., the Kesten–Stigum threshold for the SBM Sohn and Wein (2025), the BBP transition for clustering Even et al. (2025b), etc.). This leads to a sharp phase transition: the problem is efficiently solvable above the threshold, and low-degree hard below it.

In contrast, for the problem of multiple Gaussian graph alignment, the state-of-the-art algorithm Ding and Li (2022) can recover $\pi^\star$ in polynomial time as soon as the correlation ρ is non-vanishing. Rather than exhibiting a sharp phase transition, the upper bound of Ding and Li (2022) reveals a continuum of hardness: the number of iterations required, and hence the time complexity, grows as ρ decreases. This suggests that closing the remaining (polylogarithmic) gap between the state-of-the-art polynomial-time algorithm and the low-degree lower bound is likely to be very challenging.

Open questions. An intriguing direction is the alignment of correlated Erdős-Rényi graphs. For $p = 2$, it is believed that the Otter constant characterizes the computational threshold in terms of the correlation parameter (Ding et al. (2023); Ganassali et al. (2024b); Li (2025)). However, the situation for $p \geq 3$ remains unclear. Does the Otter constant continue to govern the computational barrier in this multi-graph setting, or does a different phenomenon arise?

References

Ameen, T. and Hajek, B. (2025). Detecting correlation between multiple unlabeled gaussian networks.

Baik, J., Arous, G. B., and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643 – 1697.

Barak, B., Hopkins, S., Kelner, J., Kothari, P. K., Moitra, A., and Potechin, A. (2019). A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735.

Brennan, M. and Bresler, G. (2020). Reducibility and statistical-computational gaps from secret leakage. In Abernethy, J. and Agarwal, S., editors, *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 648–847. PMLR.

Carpentier, A., Giancola, S. M., Giraud, C., and Verzelen, N. (2025a). Low-degree lower bounds via almost orthonormal bases.

Carpentier, A., Giraud, C., and Verzelen, N. (2025b).

Phase transition for stochastic block model with more than $\sqrt{n}$ communities.

Collier, O. and Dalayan, A. S. (2016). Minimax rates in permutation estimation for feature matching. *Journal of machine learning reasearch 17*.

Conte, D., Foggia, P., Vento, M., and Sansone, C. (2004). Thirty Years Of Graph Matching In Pattern Recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 18(3):265–298.

Ding, J., Du, H., and Li, Z. (2023). Low-degree hardness of detection for correlated erdos-rényi graphs.

Ding, J. and Li, Z. (2022). A polynomial time iterative algorithm for matching gaussian matrices with non-vanishing correlation.

Even, B., Giraud, C., and Verzelen, N. (2024). Computation-information gap in high-dimensional clustering. In Agrawal, S. and Roth, A., editors, *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 1646–1712. PMLR.

Even, B., Giraud, C., and Verzelen, N. (2025a). Computational barriers for permutation-based problems, and cumulants of weakly dependent random variables.

Even, B., Giraud, C., and Verzelen, N. (2025b). Computational lower bounds in latent models: clustering, sparse-clustering, biclustering.

Fan, Z., Mao, C., Wu, Y., and Xu, J. (2020). Spectral graph matching and regularized quadratic relaxations: Algorithm and theory. In III, H. D. and Singh, A., editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2985–2995. PMLR.

Gamarnik, D. (2021). The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences*, 118(41):e2108492118.

Ganassali, L. (2022). Sharp threshold for alignment of graph databases with gaussian weights. In Bruna, J., Hesthaven, J., and Zdeborova, L., editors, *Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference*, volume 145 of *Proceedings of Machine Learning Research*, pages 314–335. PMLR.

Ganassali, L., Lelarge, M., and Massoulié, L. (2024a). Correlation detection in trees for planted graph alignment. *The Annals of Applied Probability*, 34(3):2799 – 2843.

Ganassali, L., Lelarge, M., and Massoulié, L. (2022). Spectral alignment of correlated gaussian matrices. *Advances in Applied Probability*, 54(1):279–310.

Ganassali, L., Massoulié, L., and Semerjian, G. (2024b). Statistical limits of correlation detection in trees. *The Annals of Applied Probability*, 34(4):3701–3734.

Giraud, C. (2008). Estimation of Gaussian graphs by model selection. *Electron. J. Stat.*, 2:542–563.

Giraud, C. (2021). *Introduction to high-dimensional statistics*, volume 168 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL.

Giraud, C. and Verzelen, N. (2019). Partial recovery bounds for clustering with the relaxed k-means. *Mathematical Statistics and Learning*, 1(3):317–374.

Haghighi, A., Ng, A., and Manning, C. (2005). Robust textual inference via graph matching. In Mooney, R., Brew, C., Chien, L.-F., and Kirchhoff, K., editors, *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 387–394, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

Hopkins, S. (2018). *Statistical inference and the sum of squares method*. PhD thesis, Cornell University.

Hopkins, S. B., K. Kothari, P., A. Potechin, A., Raghavendra, P., chramm, T., and Steurer, D. (2017). The power of sum-of-squares for detecting hidden structures. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 720–731, Los Alamitos, CA, USA. IEEE Computer Society.

Lesieur, T., De Bacco, C., Banks, J., Krzakala, F., Moore, C., and Zdeborová, L. (2016). Phase transitions and optimal algorithms in high-dimensional gaussian mixture clustering. In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 601–608. IEEE.

Li, Z. (2025). Algorithmic contiguity from low-degree conjecture and applications in correlated random graphs.

Luo, Y. and Gao, C. (2023). Computational lower bounds for graphon estimation via low-degree polynomials.

Makarychev, K., Manokaran, R., and Sviridenko, M. (2014). Maximum quadratic assignment problem: Reduction from maximum label cover and lp-based approximation algorithm. *CoRR*, abs/1403.7721.

Mao, C., Rudelson, M., and Tikhomirov, K. (2021). Random graph matching with improved noise robustness. In Belkin, M. and Kpotufe, S., editors, *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pages 3296–3329. PMLR.

Mao, C., Wu, Y., Xu, J., and Yu, S. H. (2023). Random graph matching at otter's threshold via counting chandeliers. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1345–1356, New York, NY, USA. Association for Computing Machinery.

Narayanan, A. and Shmatikov, V. (2009). De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187.

Ndaoud, M. (2022). Sharp optimal recovery in the two component Gaussian mixture model. *The Annals of Statistics*, 50(4):2096 – 2126.

Novak, J. (2014). Three lectures on free probability. *Random matrix theory, interacting particle systems, and integrable systems*, 65(309-383):13.

Pardalos, P. M., Rendl, F., and Wolkowicz, H. (1993). *The Quadratic Assignment Problem: A Survey and Recent Developments*, volume 16 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 1–42. DIMACS/AMS.

Schramm, T. and Wein, A. S. (2022). Computational barriers to estimation from low-degree polynomials. *The Annals of Statistics*, 50(3):1833–1858.

Singh, R., Xu, J., and Berger, B. (2008). Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences*, 105(35):12763–12768.

Sohn, Y. and Wein, A. S. (2025). Sharp phase transitions in estimation with low-degree polynomials. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 891–902, New York, NY, USA. Association for Computing Machinery.

Vassaux, L. and Massoulié, L. (2025). The feasibility of multi-graph alignment: a bayesian approach.

Wein, A. S. (2025). Computational complexity of statistics: New insights from low-degree polynomials.

Wu, Y., Xu, J., and Yu, S. H. (2022). Settling the sharp reconstruction thresholds of random graph matching. *IEEE Trans. Inf. Theor.*, 68(8):5391–5417.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. **Yes.**
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. **Not Applicable.**
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. **Not Applicable.**
2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. **Yes.**
   (b) Complete proofs of all theoretical results. **Yes.**
   (c) Clear explanations of any assumptions. **Yes.**
3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). **Not Applicable.**
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). **Not Applicable.**
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). **Not Applicable.**
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). **Not Applicable.**
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator if your work uses existing assets. **Not Applicable.**
   (b) The license information of the assets, if applicable. **Not Applicable.**
   (c) New assets either in the supplemental material or as a URL, if applicable. **Not Applicable.**
   (d) Information about consent from data providers/curators. **Not Applicable.**
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. **Not Applicable.**
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. **Not Applicable.**
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. **Not Applicable.**

(c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. **Not Applicable.**

Statistical-computational gap in multiple Gaussian graph alignment – Supplementary Material

A Proof of Theorem 2.1

We will state and prove the following theorem, which will imply Theorem 2.1.

Theorem A.1. *There exists numerical constants $C, c_1, c_2, c_3 > 0$ and n_0 such that the following holds for all $\eta \geq 1$. Whenever $\rho \geq c_1\sqrt{\frac{\log(n)+\log(\eta)}{np}}$, $p \leq C\frac{n}{\log(n)+\log(\eta)}$ and $n \geq n_0$, we have, with probability at least $1 - \frac{c_2}{n^2\eta}$,*

$$\mathrm{err}(\widehat{\pi}, \pi^\star) = 0.$$

Let us first explain why Theorem A.1 implies Theorem 2.1. Suppose that Theorem A.1 is true, take $\rho > c_1\sqrt{\frac{\log(n)}{np}}$ and assume $p \leq C'\rho^{-1}$ for some C' to be fixed later. Let $\eta = \exp\left(\frac{np}{c_1^2}\rho^2 - \log(n)\right) \geq 1$, so that $\rho = c_1\sqrt{\frac{\log(n)+\log(\eta)}{np}}$. Since $p \leq C'\rho^{-1} = \frac{C'}{c_1}\sqrt{\frac{np}{\log(n)+\log(\eta)}}$, we have $p \leq (C'/c_1)^2\frac{n}{\log(n)+\log(\eta)}$ and thus, taking $C' = c_1\sqrt{C}$, Theorem A.1 implies

$$\mathbb{P}\left[\mathrm{err}(\widehat{\pi}, \pi^\star) \neq 0\right] \leq \frac{1}{n^2\eta} \leq \exp\left(-cnp\rho^2\right) \quad ,$$

for $c = 1/c_1^2$.

Let us now move on to the proof of Theorem A.1. In this section, for $u, u' \in [n], e, e' \in E, j, j' \in [p]$ we will use the notations $(j,e) \sim (j',e')$ (resp. $(j,e) \not\sim (j',e')$) when $\Pi_j^\star(e) = \Pi_{j'}^\star(e')$ (resp. $\Pi_j^\star(e) \neq \Pi_{j'}^\star(e')$) and $(j,u) \sim (j',u')$ (resp. $(j,u) \not\sim (j',u')$) when $\pi_j^\star(u) = \pi_{j'}^\star(u')$ (resp. $\pi_j^\star(u) \neq \pi_{j'}^\star(u')$).

We recall that deriving the MLE is equivalent to solving

$$\widehat{B} = \arg\max_{B \in \mathcal{B}} \langle GG^T, B\rangle, \tag{17}$$

with $\mathcal{B}$ the set of all matrices B of the form $B(\pi)$ as defined in (8), namely matrices B which can be written $B_{je,j'e'} = \frac{1}{p}\mathbf{1}\left\{\Pi_j(e) = \Pi_{j'}(e)\right\}$ for some $\pi \in (\mathcal{S}_n)^p$. After solving (17), the MLE can be found taking any $\widehat{\pi}$ such that $\widehat{B} = B(\widehat{\pi})$.

A.1 First definitions and useful lemmas

The main ideas for the proof of this informational upper bound are derived from Even et al. (2024). The key distinction in our framework lies in the fact that, unlike in Even et al. (2024), we study a quadratic form involving a vector $G \in \mathbb{R}^{N\times p}$, rather than a quadratic form of a matrix. Consequently, the technical arguments required to derive concentration inequalities for the various terms differ from those used in Even et al. (2024).

A first step is to control the error in π in terms of the error on the normalized edge alignment matrix B, which looks at the permutation at the edge level. Another useful error is measured looking at the permutation at the node level. For any $\pi \in (\mathcal{S}_n)^p$ we introduce its *normalized vertex alignment matrix* $A(\pi) \in \mathbb{R}^{pn\times pn}$ as

$$[A(\pi)]_{ju,j'u'} = \frac{1}{p}\mathbf{1}_{\pi_j(u)=\pi_{j'}(u')}\,. \tag{18}$$

Define $\mathcal{A} := A((\mathcal{S}_n)^p)$. Note that, similarly to B, $\pi \mapsto A(\pi)$ is a bijection from $(\mathcal{S}_n)^p/\equiv$ onto $\mathcal{A}$, where $\equiv$ is defined in Section 2.1 after (8). Denote $A^\star := A(\pi^\star)$.

We start by stating a lemma that will be instrumental in the proofs to come.

Lemma A.1. *For all $B \in \mathcal{B}$,*

$$p\mathrm{Tr}(B^\star - B^\star B) = \sum_{(j,e)\not\sim(j',e')} B_{je,j'e'} = \frac{1}{2}\|B^\star - B^\star B\|_1 \,.$$

$$\|B^\star - B^\star B\|_F \leq \sqrt{\|B^\star - B^\star B\|_\infty}\sqrt{\|B^\star - B^\star B\|_1} \leq \sqrt{2/p}\sqrt{\|B^\star - B^\star B\|_1} \,.$$

For all $A \in \mathcal{A}$,

$$p\mathrm{Tr}(A^\star - A^\star A) = \sum_{(j,u)\not\sim(j',u')} A_{ju,j'u'} = \frac{1}{2}\|A^\star - A^\star A\|_1 \,.$$

Another useful fact is that errors in B are equivalent to errors in A for a common $\pi \in (\mathcal{S}_n)^p$.

Lemma A.2 (Equivalence of errors in A and B)**.** *We have for all $\pi \in (\mathcal{S}_n)^p$, denoting $A = A(\pi)$ and $B = B(\pi)$,*

$$\frac{1}{4n}\|B^\star - B^\star B\|_1 \leq \|A^\star - A^\star A\|_1 \leq \frac{1}{(n-1)}\|B^\star - B^\star B\|_1 \,.$$

Lemma A.3 (Errors in π, A and B)**.** *We have for all $\pi \in (\mathcal{S}_n)^p$, denoting $A = A(\pi)$ and $B = B(\pi)$,*

$$\mathrm{err}(\pi, \pi^\star) \leq \frac{2}{np} \times \|A^\star - A^\star A\|_1 \leq \frac{2}{n(n-1)p} \times \|B^\star - B^\star B\|_1 \,.$$

Thanks to these results, all we have to do in order to prove Theorem 2.1 is to control $\|A^\star - A^\star \widehat{A}\|_1$ with high probability, where $\widehat{A} = A(\widehat{\pi})$ with $\widehat{\pi}$ the MLE defined in (10).

A.2 Control of $\|A^\star - A^\star \widehat{A}\|_1$

To do so, recall that by definition,

$$\langle GG^T, \widehat{B} - B^\star \rangle \geq 0,$$

which gives after plugging (5):

$$\rho\langle G^\star(G^\star)^T, B^\star - \widehat{B}\rangle \leq \sqrt{\rho(1-\rho)}\langle G^\star Z^T + Z(G^\star)^T, \widehat{B} - B^\star\rangle + (1-\rho)\langle ZZ^T, \widehat{B} - B^\star\rangle \,. \tag{19}$$

We will next refer to $\langle G^\star(G^\star)^T, B^\star - \widehat{B}\rangle$ as the *signal term*, $\langle G^\star Z^T + Z(G^\star)^T, \widehat{B} - B^\star\rangle$ as the *cross term*, and $\langle ZZ^T, \widehat{B} - B^\star\rangle$ as the *quadratic term*. We will control each of these terms separately, uniformly when $\widehat{B}$ runs in $\mathcal{B}$.

For concentration bounds involving quadratic forms of Gaussian vectors, an instrumental result is the Hanson-Wright inequality. We refer to e.g. Giraud (2008) for a proof.

Lemma A.4 (Hanson-Wright inequality)**.** *Let $d \geq 1$, $X \sim \mathcal{N}(0, I_d)$ and M a deterministic $d \times d$ matrix. Then, for all $x > 0$,*

$$\mathbb{P}\left(|X^T M X - \mathrm{Tr}(M)| > \sqrt{8\|M\|_F^2 x} \vee (8\|M\|_{op} x)\right) \leq 2e^{-x} \,.$$

Using Hanson-Wright inequality, we are able to derive the following uniform controls on the three terms in (19):

Proposition A.1. *There exists numerical constants $C, c_1, c_2, c_3 > 0$ and $n_0 \geq 1$ such that the following holds for all $\eta \geq 1$. Assume $\rho \geq c_3\sqrt{\frac{\log(n)+\log(\eta)}{np}}$, $p \leq C\frac{n}{\log(n)+\log(\eta)}$ and $n \geq n_0$. Then, with probability at least $1 - \frac{c_1}{n^2\eta}$, for all $B \in \mathcal{B}$,*

- *(signal term) For some numerical constant $c > 0$,*

$$\langle G^\star(G^\star)^T, B^\star - B\rangle \geq \|B^\star - B^\star\widehat{B}\|_1/4 \ ,$$

- *(cross term)*

$$|\langle G^\star Z^T + Z(G^\star)^T, \widehat{B} - B^\star\rangle| \leq c_2\sqrt{\langle G^\star(G^\star)^T, B^\star - B\rangle}\sqrt{\frac{\|B^\star - B^\star B\|_1}{n}\left(\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)\right)},$$

- *(quadratic term) for $A = A(\pi)$ where π is such that $B = B(\pi)$,*

$$|\langle ZZ^T, B - B^\star\rangle| \leq c_2\|B^\star - B^\star B\|_1\left(\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)}{n} + \sqrt{\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)}{np}}\right).$$

A.3 Conclusion

In the following, we restrict ourselves to the event of probability at least $1 - c_1/(n^2\eta)$ on which bounds of Proposition A.1 hold. First, plugging the upper bounds on the cross and quadratic terms in Equation (19) to $A = \widehat{A}$ defined previously and $B = \widehat{B}$ yields

$$\rho\langle G^\star(G^\star)^T, B^\star - \widehat{B}\rangle \lesssim \sqrt{\rho}\sqrt{\langle G^\star(G^\star)^T, \widehat{B} - B^\star\rangle}\sqrt{\frac{\|B^\star - B^\star\widehat{B}\|_1}{n}\left(\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)\right)}$$
$$+ \|B^\star - B^\star\widehat{B}\|_1\left(\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{n} + \sqrt{\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{np}}\right).$$

Discussing on which of the two terms on the right-hand side is greater, this reduces to

$$\rho\langle G^\star(G^\star)^T, \widehat{B} - B^\star\rangle \lesssim \|B^\star - B^\star\widehat{B}\|_1\left(\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{n} + \sqrt{\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{np}}\right).$$

If $\widehat{B}$ is such that $\|B^\star - B^\star\widehat{B}\|_1 = 0$, then $\|A^\star - A^\star\widehat{A}\|_1 = 0$ and $\mathrm{err}(\widehat{\pi}, \pi^\star) = 0$ by Lemma A.3, and there is nothing to prove. If $\|B^\star - B^\star\widehat{B}\|_1 > 0$, the lower bound on the signal term in Proposition A.1 entails

$$\rho \lesssim \frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{n} + \sqrt{\frac{\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{np}}.$$

Under our assumptions, $\rho \geq c_3\sqrt{\frac{\log(n)+\log(\eta)}{np}}$ and $p \leq C\frac{n}{\log(n)+\log(\eta)}$, so $\frac{\log(n)+\log(\eta)}{n} \leq \frac{\sqrt{C}}{c_3}\rho$. Thus, if c_3 is large enough, the previous inequality yields

$$\rho \lesssim \frac{\log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{n} + \sqrt{\frac{\log\left(\frac{np}{\|A^\star - A^\star\widehat{A}\|_1}\right)}{np}},$$

which leads to the existence of some numerical constants c', c satisfying

$$\|A^\star - A^\star\widehat{A}\| \leq np\exp\left(-c'n\left(\rho \wedge p\rho^2\right)\right) \leq np\exp\left(-cnp\rho^2\right).$$

Using Lemma A.3, we deduce that

$$\mathrm{err}(\widehat{\pi}, \pi^\star) \leq 2\exp\left(-cnp\rho^2\right).$$

When $\rho \geq c_3\sqrt{\frac{\log(n)+\log(\eta)}{np}}$ with c_3 large enough, the error drops strictly below $\frac{1}{n^2} \leq \frac{1}{np}$, and thus equals 0. This concludes the proof of Theorem A.1.

A.4 Proof of Proposition A.1

We now give the proof of Proposition A.1. We fix $\eta \geq 1$, which may depend on n. Throughout this proof, for brevity, we will denote

$$\delta_B := \|B^\star - B^\star B\|_1 .$$

Step 1: the signal term. We have

$$\langle G^\star (G^\star)^T, B^\star - B\rangle = (G^\star)^T (B^\star - B) G^\star,$$

and $\mathrm{Var}(G^\star) = pB^\star$. By Hanson-Wright inequality (Lemma A.4), with probability $1 - 2e^{-x}$,

$$\begin{aligned}\langle G^\star (G^\star)^T, B^\star - B\rangle &\geq p\mathrm{Tr}(B^\star(B^\star - B)) - \sqrt{8p^2\|B^\star(B^\star - B)B^\star\|_F^2 x} \vee (8p\|B^\star(B^\star - B)B^\star\|_{op} x) \\ &= \frac{1}{2}\|B^\star - B^\star B\|_1 - \sqrt{8p^2\|B^\star(B^\star - B)B^\star\|_F^2 x} \vee (8p\|B^\star(B^\star - B)B^\star\|_{op} x),\end{aligned}$$

where Lemma A.1 justifies the equality. The symmetry of B and $B^\star$ together with Lemma A.1 yields $\|B^\star(B^\star - B)B^\star\|_F = \|B^\star(B^\star - B)\|_F \leq \sqrt{2\delta_B/p}$. Moreover, B and $B^\star$ are projections, thus $\|B^\star(B^\star - B)B^\star\|_{op} \leq 2$. Besides, one always have $\|B^\star(B^\star - B)B^\star\|_{op} \leq \|B^\star(B^\star - B)B^\star\|_F$. Thus,

$$\langle G^\star (G^\star)^T, B^\star - B\rangle \geq \delta_B/2 - \left((4\sqrt{p\delta_B x}) \vee (16p(1 \wedge \sqrt{\delta_B/p})x)\right) .$$

Let, for all $1 \leq t \leq 2np$,

$$\mathcal{B}_t := \left\{ B \in \mathcal{B}, \quad \delta_B \in \left](t-1)\frac{n-1}{2}, t\frac{n-1}{2}\right]\right\} . \tag{20}$$

Now, our strategy to bound $\langle G^\star (G^\star)^T, B - B^\star\rangle$ uniformly on $\mathcal{B}_t$ and then do a union bound on $1 \leq t \leq 2np$. For $B \in \mathcal{B}_t$ for a fixed t, we will apply the previous inequality for $x = x_t \gtrsim \log 2 + \log|\mathcal{B}_t| + \log(n^3 p) + \log(\eta)$ so that $2|\mathcal{B}_t|e^{-x} \leq \frac{1}{n^3 p\eta}$. Elementary combinatorial arguments enable us to bound $|\mathcal{B}_t|$.

Lemma A.5 (Control of the size of $\mathcal{A}_t$ and $\mathcal{B}_t$)**.** *Assume $n \geq 2$. Let, for all $1 \leq t \leq 2np$,*

$$\mathcal{A}_t := \left\{ A \in \mathcal{A}, \quad \|A^\star - A^\star A\|_1 \in \left[\frac{t-1}{2}, \frac{t}{2}\right]\right\} .$$

For all $1 \leq t \leq 2np$,

$$|\mathcal{A}_t| \leq \binom{np}{t \wedge np} n^{t \wedge np} \quad \textit{and} \quad |\mathcal{B}_t| \leq t\left(\frac{16n^2 pe}{t \wedge np}\right)^{t \wedge np} .$$

In turn,

$$\log|\mathcal{B}_t| \lesssim t\log n + t\log\left(\frac{16npe}{t}\right) .$$

According to Lemma A.5, for $1 \leq t \leq 2np$, it is enough to take $x = x_t \asymp t\log(n) + t\log(2np/t) + \log(n^3 p) + \log(\eta)$, where $\asymp$ denotes equality up to multiplicative constant, to get that with probability at least $1 - \frac{1}{n^3 p\eta}$, for all $B \in \mathcal{B}_t$,

$$\begin{aligned}\langle G^\star (G^\star)^T, B^\star - B\rangle \geq& \frac{\delta_B}{2} - (4\sqrt{p\delta_B x_t}) \vee (16p\left(1 \wedge \sqrt{\delta_B/p}\right) x_t)) \\ \geq& \frac{\delta_B}{2} - c\left(\sqrt{p\delta_B\left(t\log(n) + t\log(2np/t) + \log(n^3 p) + \log(\eta)\right)}\right) \\ &- cp\left(t\log(n) + t\log(2np/t) + \log(n^3 p) + \log(\eta)\right),\end{aligned}$$

with c some numerical constant, where we used $a \vee (b \wedge c) \leq a + c$ in the second line. We have, keeping in mind that $t \asymp \delta_B/n \asymp \|A^\star - A^\star A\|_1$ (by definition (20) and Lemma A.2) and that when $B \neq B^\star$ (resp. $A \neq A^\star$), $\delta_B \gtrsim n$ (resp. $\|A^\star - A^\star A\|_1 \gtrsim 1$);

- $pt\log(n) \lesssim \frac{\delta_B}{n} p\log(n) \leq \delta_B/256$, since we assumed $p \leq C\frac{n}{\log(n)}$ taking C small enough;

- $pt\log(2np/t) \lesssim \frac{\delta_B}{n} p\log(2np) \le \delta_B/256$, for the same reasons;
- $p\log(n^3p) \lesssim \frac{\delta_B}{n\|A^\star - A^\star A\|_1} p\log(n^3p) \le \delta_B/256$, for the same reasons;
- $p\log(\eta) \lesssim \frac{\delta_B}{n} p\log(\eta) \le \delta_B/256$, since we assumed $p \le C\frac{n}{\log(\eta)}$, taking C small enough;

We deduce that, for some fixed $1 \le t \le 2np$, under our assumptions on p and ρ, with probability at least $1 - \frac{1}{n^3p\eta}$,

$$\forall B \in \mathcal{B}_t, \quad \langle G^\star (G^\star)^T, B^\star - B\rangle \ge \delta_B/4 \ .$$

We conclude the proof of the control of the signal term with an union bound on $1 \le t \le 2np$.

Step 2: the cross term For $B \in \mathcal{B}$, the cross term writes

$$\begin{aligned}
\langle G^\star Z^T + Z(G^\star)^T, \widehat{B} - B^\star\rangle &= \sum_{(j,e),(j',e')} (G^\star_{je} Z_{j'e'} + G^\star_{j'e'} Z_{je})(B_{je,j'e'} - B^\star_{je,j'e'}) \\
&= \sum_{(j,e),(j',e')} (G^\star_{je} - G^\star_{j'e'})(Z_{j'e'} - Z_{je})(B_{je,j'e'} - B^\star_{je,j'e'}) \\
&= \sum_{(j,e),\not\sim(j',e')} (G^\star_{je} - G^\star_{j'e'})(Z_{j'e'} - Z_{je}) B_{je,j'e'},
\end{aligned}$$

where the second equality comes from the identity $B\mathbf{1} = B^\star\mathbf{1} = \mathbf{1}$, and the last equality comes from the fact that $G^\star_{je} = G^\star_{j'e'}$ when $(j,e) \sim (j',e')$. Since the previous quantity can be split into two terms which are equal in distribution, it is enough to control, for all $B \in \mathcal{B}$,

$$\sum_{(j,e)\not\sim(j',e')} (G^\star_{je} - G^\star_{j'e'}) Z_{je} B_{je,j'e'} \ .$$

Again, we proceed by bounding this linear form uniformly for all $B \in \mathcal{B}_t$, where $\mathcal{B}_t$ is defined in (20), and we then apply a union bound over $t \in [1, 2np]$. Let us first fix $t \in [1, 2np]$ and $B \in \mathcal{B}_t$. Then, conditioning on the signal $G^\star$, $\sum_{(j,e)\not\sim(j',e')} (G^\star_{je} - G^\star_{j'e'}) Z_{je} B_{je,j'e'}$ is a centered Gaussian random variable with variance given by

$$\begin{aligned}
\sum_{(j,e)} \left(\sum_{(j',e')\not\sim(j,e)} \left(G^\star_{je} - G^\star_{j'e'}\right) B_{je,j'e'} \right)^2 &= \sum_{(j,e)} \left(\sum_{(j',e')} \left(G^\star_{je} - G^\star_{j'e'}\right) B_{je,j'e'} \right)^2 \\
&\le \sum_{(j,e)} \sum_{(j',e')} \left(G^\star_{je} - G^\star_{j'e'}\right)^2 B_{je,j'e'} \\
&= \sum_{(j,e)} \sum_{(j',e')} \left(G^\star_{je} - G^\star_{j'e'}\right)^2 (B_{je,j'e'} - B^\star_{je,j'e'}) \\
&= -2 \sum_{(j,e)} \sum_{(j',e')} G^\star_{je} G^\star_{j'e'} (B_{je,j'e'} - B^\star_{je,j'e'}) = 2\langle G^\star (G^\star)^T, B^\star - B\rangle,
\end{aligned}$$

where the first inequality comes from Jensen's inequality, since for all (j,e), $\sum_{(j',e')} B_{je,j'e'} = 1$, third line comes from the fact that $G^\star_{je} = G^\star_{j'e'}$ whenever $B^\star_{je,j'e'} = 1$, and the last line is a consequence of the identity $B\mathbf{1} = B^\star\mathbf{1} = \mathbf{1}$.

Standard Gaussian concentration gives that, with probability at least $1 - 2e^{-x}$,

$$\left| \sum_{(j,e)\not\sim(j',e')} (G^\star_{je} - G^\star_{j'e'}) Z_{je} B_{je,j'e'} \right| \lesssim \sqrt{\langle G^\star (G^\star)^T, B^\star - B\rangle} \sqrt{x}$$

Taking again $x = x_t \asymp t\log(n) + t\log(2np/t) + \log(n^3p) + \log(\eta)$, we have, with probability at least $1 - \frac{1}{n^3p\eta}$, for all $B \in \mathcal{B}_t$,

$$|\sum_{(j,e)\not\sim(j',e')}(G^\star_{je}-G^\star_{j'e'})Z_{je}B_{je,j'e'}| \lesssim \sqrt{\langle G^\star(G^\star)^T, B^\star - B\rangle}\sqrt{t\log(n)+t\log(2np/t)+\log(n^3p)+\log(\eta)}$$
$$\leq \sqrt{\langle G^\star(G^\star)^T, B^\star - B\rangle}\sqrt{\frac{\delta_B}{n}\left(\log(n)+\log(\eta)+\log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)\right)},$$

where the last inequality comes from the fact $p \leq C\frac{n}{\log(n)}$ and that $t \asymp \delta_B/n \asymp \|A^\star - A^\star A\|_1$ (by definition (20) and Lemma A.2). We conclude the proof of the control of the cross term with an union bound on $1 \leq t \leq 2np$.

Step 3: the quadratic noise term.

We seek to control $\langle ZZ^T, B^\star - B\rangle$ for $B \in \mathcal{B}$. We shall suppose without loss of generality that $\pi^\star = \mathrm{id}$. To any $B \in \mathcal{B}$, we identify $B = B(\pi)$ with $\pi \in (\mathcal{S}_n)^p$ minimizing $\mathrm{der}(\pi) := \sum_{j\in[p]}\sum_{u\in[n]}\mathbf{1}\{\pi_j(u) \neq u\}$. From Lemma A.3, we know that $\mathrm{der}(\pi) = \mathrm{err}(\pi, \pi^\star) \lesssim \frac{1}{n}\|B^\star - B^\star B\|_1$.

As previously, our strategy is to upper-bound the quadratic term uniformly on all $\mathcal{B}_t$ and then use the union bound on $1 \leq t \leq 2np$. First, we shall fix such a $1 \leq t \leq 2np$ and $B \in \mathcal{B}_t$. Using Hanson-Wright Lemma (Lemma A.4), gives that, with probability at least $1 - 2e^{-x}$,

$$|\langle ZZ^T, B - B^\star\rangle| \lesssim \|B - B^\star\|_{op}x \vee \|B - B^\star\|_F\sqrt{x}\,.$$

Bounds on $\|B - B^\star\|_{op}$ and $\|B - B^\star\|_F$ are given in the following Lemma.

Lemma A.6.

$$\|B - B^*\|_F \leq \sqrt{\delta_B/p} \quad \textit{and} \quad \|B - B^\star\|_{op} \leq 1\,.$$

Thus, with probability at least $1 - 2e^{-x}$, we have

$$|\langle ZZ^T, B - B^\star\rangle| \leq x \vee \sqrt{x\delta_B/p}\,.$$

Doing an union bound, we get that, with probability at least $1 - 2|\mathcal{B}_t|e^{-x}$, uniformly over all $B \in \mathcal{B}_t$,

$$|\langle ZZ^T, B - B^\star\rangle| \leq x \vee \sqrt{x\delta_B/p}\,.$$

As before, we choose $x = x_t \asymp t\log(n) + t\log(2np/t) + \log(n^3p) + \log(\eta)$ so that, with probability at least $1 - \frac{1}{n^3p\eta}$, for all $B \in \mathcal{B}_t$, uniformly over all $B \in \mathcal{B}_t$, we have

$$|\langle ZZ^T, B - B^\star\rangle| \lesssim \big(t\log(n) + t\log(2np/t) + \log(n^3p) + \log(\eta)\big)$$
$$+ \sqrt{\frac{\delta_B}{p}\left(t\log(n) + t\log(2np/t) + \log(n^3p) + \log(\eta)\right)}\,.$$

We remark that, as in step 1, keeping in mind that $t \asymp \delta_B/n \asymp \|A^\star - A^\star A\|_1$ (by definition (20) and Lemma A.2) and that when $B \neq B^\star$ (resp. $A \neq A^\star$), $\delta_B \gtrsim n$ (resp. $\|A^\star - A^\star A\|_1 \gtrsim 1$);

- $t\log(n) \lesssim \frac{\delta_B}{n}\log(n)$;
- $t\log(2np/t) \lesssim \frac{\delta_B}{n}\log(\frac{np}{\|A^\star - A^\star A\|_1})$;
- $\log(n^3p) \lesssim t\log(n) \lesssim \frac{\delta_B}{n}\log(n)$, since we assumed $p \leq C\frac{n}{\log(n)}$;
- $\log(\eta) \lesssim \frac{\delta_B}{n}\log(\eta)$.

Finally, we end up with

$$|\langle ZZ^T, B - B^\star\rangle| \lesssim \frac{\delta_B}{n}\left(\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)\right) + \delta_B\sqrt{\frac{1}{np}\left(\log(n) + \log(\eta) + \log\left(\frac{np}{\|A^\star - A^\star A\|_1}\right)\right)}\,.$$

We conclude the proof of the control of the quadratic noise term with an union bound on $1 \leq t \leq 2np$.

A.5 Further proofs of Appendix A

A.5.1 Proof of Lemma A.1

For all $B \in \mathcal{B}$, note that

$$(B^\star - B^\star B)_{je,j'e'} = \begin{cases} \frac{1}{p} - \frac{1}{p}\sum_{(k,f):(k,f)\sim(j,e)} B_{kf,j'e'} & \text{if } (j,e) \sim (j',e') \\ -\frac{1}{p}\sum_{(k,f):(k,f)\sim(j,e)} B_{kf,j'e'} & \text{if } (j,e) \not\sim (j',e') \end{cases}$$

Thus,

$$\begin{aligned} p\mathrm{Tr}(B^\star - B^\star B) &= p\sum_{(j,e)}\left(\frac{1}{p} - \frac{1}{p}\sum_{(k,f):(k,f)\sim(j,e)} B_{kf,je}\right) \\ &= pN - \left(pN - \sum_{(k,f):(k,f)\not\sim(j,e)} B_{kf,j'e'}\right) \\ &= \sum_{(j,e)\not\sim(j',e')} B_{je,j'e'}\,. \end{aligned}$$

On the other hand,

$$\begin{aligned} \|B^\star - B^\star B\|_1 &= \sum_{(j,e)}\left[\sum_{(j',e'):(j',e')\sim(j,e)}\left(\frac{1}{p} - \frac{1}{p}\sum_{(k,f):(k,f)\sim(j,e)} B_{kf,j'e'}\right) + \sum_{(j',e'):(j',e')\not\sim(j,e)} \frac{1}{p}\sum_{(k,f):(k,f)\sim(j,e)} B_{kf,j'e'}\right] \\ &= \frac{1}{p}\sum_{(j,e)}\sum_{(k,f):(k,f)\sim(j,e)}\left[\sum_{(j',e'):(j',e')\sim(j,e)}\left(\frac{1}{p} - B_{kf,j'e'}\right) + \sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'}\right] \\ &= \frac{1}{p}\sum_{(j,e)}\sum_{(k,f):(k,f)\sim(j,e)}\left[1 - \left(1 - \sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'}\right) + \sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'}\right] \\ &= \frac{2}{p}\sum_{(j,e)}\sum_{(k,f):(k,f)\sim(j,e)}\sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'} \\ &= \frac{2}{p}\sum_{(k,f)}\sum_{(j,e):(j,e)\sim(k,f)}\sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'} \\ &= 2\sum_{(k,f)}\sum_{(j',e'):(j',e')\not\sim(j,e)} B_{kf,j'e'} = 2\sum_{(j,e)\not\sim(j',e')} B_{je,j'e'}\,. \end{aligned}$$

Then, the last inequality comes from the fact that since $B^\star$ is stochastic,

$$\|B^\star - B^\star B\|_\infty = \|B^\star(B^\star - B)\|_\infty \le \|B^\star - B\|_\infty \le 2/p\,.$$

The proof works similarly for all $A \in \mathcal{A}$.

A.5.2 Proof of Lemma A.2

Assume without loss of generality that $\pi_j^\star = \mathrm{id}$ for all $j \in [p]$. By Lemma A.1,

$$\|B^\star - B^\star B\|_1 = \frac{2}{p}\sum_{e\neq e'}\sum_{j\neq j'} \mathbf{1}\left\{\Pi_j(e) = \Pi_{j'}(e')\right\}$$

and

$$\|A^\star - A^\star A\|_1 = \frac{2}{p}\sum_{u\neq u'}\sum_{j\neq j'} \mathbf{1}\left\{\pi_j(u) = \pi_{j'}(u')\right\}\,.$$

Let us fix j, u and j', u' such that $u \neq u'$ and $\pi_j(u) = \pi_{j'}(u')$. For all $v \notin \{u, u'\}$, we have $\Pi_j(\{u, v\}) = \Pi_{j'}\left(\left\{u', \pi_{j'}^{-1} \circ \pi_j(v)\right\}\right)$, with $\{u, v\} \neq \left\{u', \pi_{j'}^{-1} \circ \pi_j(v)\right\}$. Thus, the number of edge pairs $((j, e), (j', e'))$ with $e \neq e'$, such that u is a node of e, u' is a node of e', and $\Pi_j(e) = \Pi_{j'}(e')$, is $|[n] \setminus \{u, u'\}|$ which is either $n-1$ or n. Moreover, each such pair $((j, e), (j', e'))$ can be obtained by at most four distinct choices of u, u'. We conclude that

$$\frac{1}{4n}\|B^\star - B^\star B\|_1 \leq \|A^\star - A^\star A\|_1 \leq \frac{1}{(n-1)}\|B^\star - B^\star B\|_1 .$$

A.5.3 Proof of Lemma A.3

In light of Lemma A.2, it is sufficient to prove the first inequality. This inequality is a direct consequence of Lemma 9 of Even et al. (2024). Since our framework and our notation are differ from the one of Even et al. (2024), let us nonetheless prove Lemma Lemma A.3. Let $\pi = \pi_1, \ldots, \pi_p$. Our goal is to bound the proportion of misclassified points

$$\mathrm{err}(\pi, \pi^\star) := \min_{\psi \in \mathcal{S}_n} \frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\left\{\psi(\pi_j(u)) \neq \pi_j^\star(u)\right\} \quad ,$$

with respect to $\|A^\star - A^\star A\|_1$. For that purpose, let us chose some relevant permutation $\psi \in \mathcal{S}_n$ such that

$$\frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\left\{\psi(\pi_j(u)) \neq \pi_j^\star(u)\right\} \leq \frac{2}{np} \times \|A^\star - A^\star A\|_1 \quad .$$

Let us suppose without loss of generality that $\pi^\star = id$.

Lemma A.7. *There exists a permutation $\psi \in \mathcal{S}_n$ satisfying; for all $u \neq v \in [n]$, $|\{j \in [p], \psi \circ \pi_j(u) = v\}| \leq p/2$.*

Let $\psi \in \mathcal{S}_n$ given by Lemma A.7. Without loss of generality, we suppose that $\psi = id$. By definition, we have

$$\mathrm{err}(\pi, \pi^\star) \leq \frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\{\pi_j(u) \neq u\} = \frac{1}{np} \sum_{u \neq v} |\{j \in [p], \pi_j(u) = v\}| \quad .$$

Let us fix $u \neq v$. By Lemma A.7, we have $|\{j \in [p], \pi_j(u) = v\}| \leq \frac{p}{2}$. Hence, all elements (j, u) with $\pi_j(u) = v$ satisfy $\sum_{j' \neq j} \sum_{v' \neq u} A_{ju, j'v'} \geq \frac{1}{2}$. We arrive at

$$\begin{aligned}
\frac{1}{np} \sum_{u \neq v} |\{j \in [p], \pi_j(u) = v\}| =& \frac{1}{np} \sum_{u \neq v} |\{j \in [p], \pi_j(u) = v\}| \\
\leq& \frac{1}{np} \sum_{u \neq v} 2 \sum_{j \in [p], \pi_j(u) = v} \sum_{v' \neq u} A_{ju, j'v'} \\
\leq& \frac{2}{np} \sum_{u \neq v'} \sum_{j \neq j'} A_{ju, j'v'} \\
\leq& \frac{2}{np} \|A^\star - A^\star A\|_1 \quad .
\end{aligned}$$

This concludes the proof of the lemma.

Proof of Lemma A.7. Let us construct $\phi = \psi^{-1}$. For u, v with $|\{j \in [p], \pi_j(u) = v\}| > p/2$, let us take $\phi(v) = \pi_j(u)$. This is an bijection from all the elements v such that there exists u with $|\{j \in [p], \pi_j(u) = v\}| > p/2$ to a subset of $[n]$. This bijection can be extanded to a bijection $\phi : [n] \to [n]$. The choice $\psi = \phi^{-1}$ satisfies the condition of the lemma. □

A.5.4 Proof of Lemma A.5

Fix $1 \leq t \leq 2np$, $A \in \mathcal{A}_t$ and $\pi \in (\mathcal{S}_n)^p$ such that $A = A(\pi)$. Without loss of generality, assume that $\psi = \mathrm{id}$ reaches the minimum in the definition of $\mathrm{err}(\pi, \pi^\star)$.

Remark that since $A \in \mathcal{A}_t$, thanks to Lemma A.3,

$$\frac{t}{2} \geq \|A^\star - A^\star A\|_1 \geq \frac{np}{2}\mathrm{err}(\pi, \pi^\star) = \frac{1}{2}|\{(j,u) \in [p] \times [n], \pi_j(u) \neq \pi_j^\star(u)\}|,$$

thus, we have at most t pairs (j,u) such that $\pi_j(u) \neq \pi_j^\star(u)$. Since this number is also upper bounded by np, we have at most $t \wedge np$ such pairs. Choosing these pairs among the np choices, and choosing values $\pi_j(u)$ for each of these (j,u) among n possibilities gives

$$|\mathcal{A}_t| \leq \binom{np}{t \wedge np} n^{t \wedge np}.$$

Then, by Lemma A.2, if $B(\pi)$ is an element of $\mathcal{B}_t$, then $A(\pi)$ lies in some A_s with $t/16 \leq s \leq t$ as soon as $n \geq 2$. This gives

$$|\mathcal{B}_t| \leq \sum_{s=t/16}^{t} \binom{np}{s \wedge np} n^{s \wedge np} \leq \sum_{s=t/16}^{t} \left(\frac{npe}{s \wedge np}\right)^{s \wedge np} n^{s \wedge np} \leq t \left(\frac{16n^2pe}{t \wedge np}\right)^{t \wedge np}.$$

This gives in turn that

$$\begin{aligned}
\log|\mathcal{B}_t| &\lesssim \log t + (t \wedge np)\log\left(\frac{16n^2pe}{t \wedge np}\right) \\
&\leq \log(2np) + t\log\left(\frac{16n^2pe}{t}\right) + t\log(16ne) \\
&\lesssim t\log n + t\log\left(\frac{16npe}{t}\right).
\end{aligned}$$

A.5.5 Proof of Lemma A.6

Since both B and $B^\star$ are projection matrices, we directly deduce $\|B - B^\star\|_{op} \leq 2$. For controlling $\|B - B^\star\|_F$, we proceed by counting the number of non-zero entries of $B - B^\star$, combined with the fact that $\|B - B^\star\|_\infty = \frac{1}{p}$. Let $e = \{u, v\}$, $e' = \{u', v'\}$ and j, j'. Denote

$$\mathcal{C}_j := \{u \in [n], \pi_j(u) \neq \pi_j^\star(u)\}.$$

In order to have $(B - B^\star)_{je,j'e'} \neq 0$, we need to have either:

1. $e = e'$ and either $u \in \mathcal{C}_j$, $v \in \mathcal{C}_j$, $u' \in \mathcal{C}_{j'}$ or $v' \in \mathcal{C}_{j'}$. The number of such pairs $(j,e),(j',e')$ is at most $2np\sum_{j\in[p]}|\mathcal{C}_j| = 2n^2p^2 \times \mathrm{err}(\pi,\pi^\star) \lesssim p\|B^\star - B^\star B\|_1$;
2. $e \neq e'$ but we must have $\{\Pi_j(e) = \Pi_{j'}(e')\}$ and, a fortiori, either $u \in \mathcal{C}_j$, $v \in \mathcal{C}_j$, $u' \in \mathcal{C}_{j'}$ or $v' \in \mathcal{C}_{j'}$. The number of such pairs is at most $2np \times \mathrm{err}(\pi,\pi^\star) \lesssim p\|B^\star - B^\star B\|_1$.

We can conclude with $\|B - B^\star\|_F \lesssim \sqrt{\delta_B/p}$.

B Proof of Theorem 2.2

Assume $p \geq C\rho^{-1}$ and $\rho \geq c_1\left(\frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}}\right)$ for some large enough numerical constant c_1. Take $p' = C\rho^{-1}$, where C is the constant given in Theorem 2.1. It is clear that, for this new p',

$$\rho \geq c_1\frac{\log n}{n} \implies \rho^2 \geq c_1\rho\frac{\log(n)}{n} = c_1C\frac{\log(n)}{np'} \implies \rho \geq \sqrt{c_1C}\sqrt{\frac{\log(n)}{p'n}}.$$

We denote in the following G' and $\widehat{G}'$ the graphs defined by

$$G'_e = \frac{1}{p'} \sum_{j \le p'} G_{j\left(\Pi_j^\star\right)^{-1}(e)} \quad \text{and} \quad \widehat{G}'_e = \frac{1}{p'} \sum_{j \le p'} G_{j\left(\widehat{\Pi}_j\right)^{-1}(e)} \ .$$

Using Theorem 2.1, we deduce that the MLE applied to $G_1, \dots, G_{p'}$ achieves perfect recovery with high probability.

Corollary B.1. *Suppose $\rho \ge c\left(\frac{\log(n)}{n} + \sqrt{\frac{\log(n)}{np}}\right)$ with c large enough and $n \ge n_0$ with n_0 some numerical constant. Then,*

$$\mathbb{P}\left(\mathrm{err}\left(\left(\widehat{\pi}_j\right)_{j \le p'}, \left(\pi_j^\star\right)_{j \le p'}\right) \neq 0\right) \le \exp\left(-c' n \rho\right) \ ,$$

with c' some numerical constant.

On the event of high probability of Corollary B.1, the graphs G' and $\widehat{G'}$ are isomorphic. Thus, it is sufficient to analyse the alignment of a given graph G_j, for $j > p'$, with the graph G'. We remark that the edge correlation between two such graphs is equal to

$$\rho' = \frac{\sqrt{p'-1}\,\rho}{\sqrt{1+(p'-2)\rho}} \gtrsim \sqrt{\rho} \ge c\sqrt{\frac{\log(n)}{n}} \ ,$$

with c some numerical constant that can be taken arbitrary large. Hence, we can apply Theorem 2.1 for the specific case $p = 2$.

Formally, we compute $\widehat{\pi}_j$ the unique element such that $(\widehat{\pi}_1, \widehat{\pi}_j)$ is the MLE of $(\widehat{G'}, G_j)$. Let $\widehat{\pi}_j^{\mathrm{oracle}}$ the unique element such that $(\pi_1^\star, \widehat{\pi}_j^{\mathrm{oracle}})$ is the MLE of (G', G_j). Next corollary controls the error of $\widehat{\pi}_j^{\mathrm{oracle}}$.

Corollary B.2. *There exists numerical constants c, c', n_0 such that the following holds for all $j > p'$ whenever $n \ge n_0$ and $\rho' \ge c\sqrt{\frac{\log(n)}{n}}$. The MLE applied to G' and G_j satisifies*

$$\mathbb{P}\left[\delta_H(\widehat{\pi}_j^{\mathrm{oracle}}, \pi_j^\star) \neq 0\right] \le \exp\left(-c' n \rho\right) \ ,$$

with δ_H the Hamming distance. We denote $\mathcal{E}_j$ the event in the LHS above.

If $\rho \ge c\left(\frac{\log(n)}{n} + \sqrt{\frac{\log(n)}{np}}\right)$ for some large enough numerical constant c, then Corollary B.2 and Corollary B.1 hold.

Then, let us restrict ourselves to the event $\mathcal{A}$ on which $\mathrm{err}\left(\left(\widehat{\pi}_j\right)_{j \le p'}, \left(\pi_j^\star\right)_{j \le p'}\right) = 0$. This means that there exists $\psi \in \mathcal{S}_n$ such that $\widehat{\pi}_j = \psi \circ \pi_j^\star$ for all $j \le p'$. Then, for all $j > p'$, we have $\widehat{\pi}_j = \psi \circ \widehat{\pi}_j^{\mathrm{oracle}}$. We deduce from this the upper-bound on event $\mathcal{A}$

$$\mathrm{err}(\widehat{\pi}, \pi^\star) \le \mathbf{1}\left\{\mathrm{err}\left(\left(\widehat{\pi}_j\right)_{j \in [p']}, \left(\pi_j^\star\right)_{j \in [p']}\right) \neq 0\right\} + \frac{1}{p} \sum_{j > p'} \mathbf{1}\left(\mathcal{E}_j\right) \ .$$

Taking to the expectation, we deduce

$$\mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right] \le \mathbb{P}(\mathcal{A}) + \exp\left(-cn\rho\right) \ \le \exp\left(-c_2 n \rho\right),$$

for some numerical constant $c_2 > 0$. This concludes the proof of the theorem.

C Proof of Theorem 2.4

For an estimator $\widehat{\pi} = \widehat{\pi}_1, \dots, \widehat{\pi}_p$, we recall the definition of err:

$$\mathrm{err}(\widehat{\pi}, \pi^\star) = \min_{\psi \in \mathcal{S}_n} \frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\left\{\psi(\widehat{\pi}_j(u)) \neq \pi_j^\star(u)\right\} \ .$$

In order to lower-bound the optimal error $\mathrm{err}_{\mathrm{opt}} := \inf_{\widehat{\pi}: \mathbb{R}^{E \times p} \to (\mathcal{S}_n)^p} \mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right]$, we shall proceed in two steps;

- Step 1: we lower-bound $\mathrm{err}_{\mathrm{opt}}$ with respect to the optimal error of estimating $\pi_p^\star$ given all the other permutations,
- Step 2: we lower-bound the optimal error of estimating $\pi_p^\star$ using Fano's lemma, which requires to upper-bound the Kullback–Leibler divergence between the joint law conditionally on $\pi_p^\star = \mathrm{id}$ and the law conditionally on $\pi_p^\star = \pi$ for any $\pi \in \mathcal{S}_n$.

Step 1. Let us now lower-bound the $\mathrm{err}_{\mathrm{opt}}$ with respect to the optimal error of estimating $\pi_p^\star$. For $j \neq j' \in [p]$ and a global estimator $\widehat{\pi}$, we define the error of $\widehat{\pi}$ between G_j and $G_{j'}$ as

$$\mathrm{err}_{j,j'}(\widehat{\pi}, \pi^\star) := \min_{\psi \in \mathcal{S}_n} \frac{1}{2n} \sum_{u=1}^{n} \mathbf{1}\left\{\psi(\widehat{\pi}_j(u)) \neq \pi_j^\star(u)\right\} + \mathbf{1}\left\{\psi(\widehat{\pi}_{j'}(u)) \neq \pi_{j'}^\star(u)\right\} .$$

We have, for any $\widehat{\pi}$ and a minimizer ψ in the definition of $\mathrm{err}(\widehat{\pi}, \pi^\star)$,

$$\begin{aligned}
\mathrm{err}(\widehat{\pi}, \pi^\star) =& \frac{1}{np} \sum_{u=1}^{n} \sum_{j=1}^{p} \mathbf{1}\left\{\psi(\widehat{\pi}_j(u)) \neq \pi_j^\star(u)\right\} \\
=& \frac{1}{np(2p-2)} \sum_{1 \leq j \neq j' \leq p} \sum_{u=1}^{n} \mathbf{1}\left\{\psi(\widehat{\pi}_j(u)) \neq \pi_j^\star(u)\right\} + \mathbf{1}\left\{\psi(\widehat{\pi}_{j'}(u)) \neq \pi_{j'}^\star(u)\right\} \\
\geq& \frac{1}{p(p-1)} \sum_{1 \leq j \neq j' \leq p} \mathrm{err}_{j,j'}\left(\widehat{\pi}, \pi^\star\right),
\end{aligned}$$

which leads to, taking to the expectation,

$$\mathbb{E}\left[\mathrm{err}(\widehat{\pi}, \pi^\star)\right] \geq \frac{1}{p(p-1)} \sum_{1 \leq j \neq j' \leq p} \mathbb{E}\left[\mathrm{err}_{j,j'}\left(\widehat{\pi}, \pi^\star\right)\right] .$$

Step 2. Fix $(j, j') = (p-1, p)$ by symmetry and let us find a lower bound for $\inf_{\widehat{\pi}: \mathbb{R}^{E \times p} \to (\mathcal{S}_n)^p} \mathbb{E}\left[\mathrm{err}_{p-1,p}\left(\widehat{\pi}, \pi^\star\right)\right]$. Remark that $\mathrm{err}_{p-1,p}\left(\widehat{\pi}, \pi^\star\right)$ only depends on the two last permutations $\widehat{\pi}_{p-1}, \widehat{\pi}_p$ so we will lower-bound

$$\mathrm{err}_{p-1,p}^{\mathrm{opt}} := \inf_{\widehat{\pi}_{p-1}, \widehat{\pi}_p \in \mathcal{S}_n} \mathbb{E}\left[\mathrm{err}\left(\left(\widehat{\pi}_{p-1}, \widehat{\pi}_p\right), \left(\pi_{p-1}^\star, \pi_p^\star\right)\right)\right] .$$

Let $\left(\widehat{\pi}_{p-1}^{\mathrm{opt}}, \widehat{\pi}_p^{\mathrm{opt}}\right)$ minimizing the above quantity. For $\pi \in \mathcal{S}_n$, let $\mathbb{P}_\pi := \mathbb{P}_{\mathrm{id},\ldots,\mathrm{id},\pi}$ the law of $G_1, \ldots, G_p$ conditionally on $\pi_1^\star = \mathrm{id}, \ldots, \pi_{p-1}^\star = \mathrm{id}, \pi_p^\star = \pi$. We recall that we denote $\mathbb{P}$ the law of $G_1, \ldots, G_p$ with $\pi_1^\star, \ldots, \pi_p^\star$ taken independently and uniformly on $\mathcal{S}_n$. Consider the following procedure for recovering π under $\mathbb{P}_\pi$.

1. draw $\sigma_1, \ldots, \sigma_p$ uniformly and independently over $\mathcal{S}_n$,
2. relabel $G_1, \ldots, G_p$ accordingly to $\sigma_1, \ldots, \sigma_p$, and denote the relabeled graphs $\widetilde{G}_1, \ldots, \widetilde{G}_p$
3. compute $\left(\widehat{\pi}_{p-1}^{\mathrm{opt}}, \widehat{\pi}_p^{\mathrm{opt}}\right)(\widetilde{G}_1, \ldots, \widetilde{G}_p)$;
4. Take $\widehat{\pi} = \left(\widehat{\pi}_{p-1}^{\mathrm{opt}}\right)^{-1} \circ \sigma_{p-1} \circ (\sigma_p)^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}} \in \mathcal{S}_n$ as an estimator for $\hat{\pi}$.

We remark that, under the law $\mathbb{P}_\pi$ for $G_1, \ldots, G_p$, the permuted graphs $\widetilde{G}_1, \ldots, \widetilde{G}_p$ have joint law $\mathbb{P}$. Thus,

$$\mathbb{E}_\pi\left[\mathrm{err}\left(\left(\widehat{\pi}_{p-1}^{\mathrm{opt}}, \widehat{\pi}_p^{\mathrm{opt}}\right)\left(\widetilde{G}_1, \ldots, \widetilde{G}_p\right), (\sigma_{p-1}, \sigma_p \circ \pi)\right)\right] = \mathrm{err}_{p-1,p}^{\mathrm{opt}} .$$

In the following, we denote $\delta_H(\pi, \pi') := \frac{1}{n} \sum_{u=1}^{n} \mathbf{1}\left\{\pi(u) \neq \pi'(u)\right\}$ the Hamming distance between two permutations. We have

$$\begin{aligned}
\mathrm{err}\left(\left(\widehat{\pi}_{p-1}^{\mathrm{opt}}, \widehat{\pi}_p^{\mathrm{opt}}\right), (\sigma_{p-1}, \sigma_p \circ \pi)\right) =& \mathrm{err}\left(\left(\sigma_p^{-1} \circ \widehat{\pi}_{p-1}^{\mathrm{opt}}, \sigma_{p-1}^{\mathrm{opt}} \circ \widehat{\pi}_p^{\mathrm{opt}}\right), (\mathrm{id}, \pi)\right) \\
=& \frac{1}{2} \min_{\psi \in \mathcal{S}_n} \left[\delta_H\left(\psi \circ \sigma_{p-1}^{-1} \circ \widehat{\pi}_{p-1}^{\mathrm{opt}}, \mathrm{id}\right) + \delta_H\left(\psi \circ \sigma_p^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}, \pi\right)\right] .
\end{aligned}$$

Let us now fix $\psi \in \mathcal{S}_n$. Then, using the triangular equality,

$$\begin{aligned}\delta_H\left(\widehat{\pi}, \pi\right) =& \delta_H\left(\left(\widehat{\pi}_{p-1}^{\mathrm{opt}}\right)^{-1} \circ \sigma_{p-1} \circ \left(\sigma_p\right)^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}, \pi\right) \\ \leq& \delta_H\left(\psi \circ \sigma_p^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}, \pi\right) + \delta_H\left(\psi \circ \sigma_p^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}, \left(\widehat{\pi}_{p-1}^{\mathrm{opt}}\right)^{-1} \circ \sigma_{p-1} \circ \left(\sigma_p\right)^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}\right) \\ =& \delta_H\left(\psi \circ \sigma_p^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}, \pi\right) + \delta_H\left(\psi \circ \sigma_p^{-1} \circ \widehat{\pi}_p^{\mathrm{opt}}\right),\end{aligned}$$

and, taking the infimum over all $\psi \in \mathcal{S}_n$ and taking to the expectation, we get

$$\mathbb{E}_\pi\left[\delta_H\left(\widehat{\pi}, \pi\right)\right] \leq 2\mathrm{err}_{p-1,p}^{\mathrm{opt}} \leq 2\mathrm{err}_{\mathrm{opt}}\,.$$

We conclude the proof of the theorem with the next lemma, whose proof is postponed to Section C.1.

Lemma C.1. *Suppose $n \geq n_0$ with n_0 some numerical constant. There exists numerical constants c, c' such that, when $\rho \leq c\left(\frac{\log(n)}{n} \vee \sqrt{\frac{\log(n)}{np}}\right)$,*

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to\mathcal{S}_n} \sup_{\pi\in\mathcal{S}_n} \mathbb{E}_\pi\left[\delta_H(\widehat{\pi}, \pi)\right] \geq c\,.$$

C.1 Proof of Lemma C.1

We recall that, under $\mathbb{P}_\pi$, we observe $G_1, \ldots, G_p$ conditionally over $\pi_j^\star = \mathrm{id}$ for $j \in [p-1]$ and over $\pi_p^\star = \pi$. We shall lower-bound

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to\mathcal{S}_n} \sup_{\pi\in\mathcal{S}_n} \mathbb{E}_\pi\left[\delta_H(\widehat{\pi}, \pi)\right]\,.$$

To do so, we shall use an adaptation of Fano's lemma, stated as Lemma C.2 (see e.g. Section 3.2.2 of Giraud (2021) for a proof).

Lemma C.2. *For any finite set $A \subseteq \mathcal{S}_n$, we have the following inequality*

$$\inf_{\widehat{\pi}} \frac{2}{|A|} \sum_{\pi\in A} \mathbb{E}_\pi\left[\delta_H(\widehat{\pi}, \pi)\right] \geq \left(\min_{\pi\neq\pi'\in A} \delta_H(\pi, \pi')\right) \times \left(1 - \frac{1 + \frac{1}{|A|}\sum_{\pi\in A} \mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right)}{\log(|A|)}\right)\,.$$

In the light of this lemma, we shall find a set of permutations which are all at great Hamming distance one from the others, but such that $\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right)$ remains small. Such a set is given Lemma 16 of Collier and Dalayan (2016), which we state here.

Lemma C.3. *For any interger $n \geq 4$, there exists a set $A \subset \mathcal{S}_n$ such that*

1. $|A| \geq \left(\frac{n}{24}\right)^{n/6}$;
2. *For every $\pi \neq \pi' \in A$, $\delta_H(\pi, \pi') \geq 3/8$.*

It remains to compute the Kullback-Leibler divergence $\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right)$ for any $\pi \in A$. We refer to Section C.1.1 for a proof of the next lemma.

Lemma C.4. *Let A be a subset of $\mathcal{S}_n$ satisfying conditions of Lemma C.3. For any $\pi \in A$,*

$$\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right) \leq \frac{n(n-1)\rho^2(p-1)}{(1-\rho)(1+\rho(p-1))}\,.$$

Combining Lemmas C.4, C.3 and C.2, we get, whenever $n \geq 4$,

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to\mathcal{S}_n} \sup_{\pi\in A} \mathbb{E}_\pi\left[\delta_H\left(\widehat{\pi}, \pi\right)\right] \geq \frac{3}{8}\left(1 - \frac{1 + \frac{n(n-1)\rho^2(p-1)}{(1-\rho)(1+\rho(p-1))}}{\frac{n}{6}\log(\frac{n}{24})}\right)\,.$$

First, assume $\rho \leq c\sqrt{\frac{\log(n)}{np}}$, for a small enough numerical constant c. Then,

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to\mathcal{S}_n} \sup_{\pi\in A} \mathbb{E}_\pi\left[\delta_H\left(\widehat{\pi},\pi\right)\right] \geq \frac{3}{8}\left(1-\frac{6+12n^2\rho^2p}{n\log(n/24)}\right) \geq \frac{3}{16},$$

provided the constant c is small enough and n large enough.

Second, assume $\rho \leq c\frac{\log(n)}{n}$ with c a small enough numerical constant. This gives

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to\mathcal{S}_n} \sup_{\pi\in A} \mathbb{E}_\pi\left[\delta_H\left(\widehat{\pi},\pi\right)\right] \geq \frac{3}{8}\left(1-\frac{6+\frac{12n^2\rho^2(p-1)}{\rho(p-1)}}{n\log(n/24)}\right) \geq \frac{3}{8}\left(1-\frac{6+12n^2\rho}{n\log(n/24)}\right) \geq \frac{3}{16},$$

provided n is large enough and c small enough. This concludes the proof of the lemma.

C.1.1 Proof of Lemma C.4

Let us fix $\pi \in A$ and let us compute $\mathrm{KL}(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}) = \mathbb{E}_\pi\left[\log\left(\frac{d\,\mathbb{P}_\pi}{d\,\mathbb{P}_{\mathrm{id}}}\right)\right]$. We recall that the law $\mathbb{P}_\pi$ (resp. $\mathbb{P}_{\mathrm{id}}$) can be generated as follows.

1. draw $(G_e^\star)_{e\in E}$ independently and $\mathcal{N}(0,1)$;

2. Independently from $(G_e^\star)_{e\in E}$, draw $(Z_{je})_{e\in E, j\in[p]}$ i.i.d. $\mathcal{N}(0,1)$;

3. For $j \leq p-1$, take $G_{je} = \sqrt{\rho}G_e^\star + \sqrt{1-\rho}Z_{je}$;

4. For $j = p$, take $G_{pe} = \sqrt{\rho}G_{\Pi(e)}^\star + \sqrt{1-\rho}Z_{pe}$ (resp. $G_{pe} = \sqrt{\rho}G_e^\star + \sqrt{1-\rho}Z_{pe}$), where $\Pi(e) = \{\pi(u), \pi(u')\}$ for $e = \{u, u'\}$.

In order to compute $\mathrm{KL}(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}})$, we first compute the likelihood ratio $\frac{d\,\mathbb{P}_\pi}{d\,\mathbb{P}_{\mathrm{id}}}$. We have, since the $G_e^\star$'s are drawn independently,

$$\begin{aligned}\frac{d\,\mathbb{P}_\pi}{d\,\mathbb{P}_{\mathrm{id}}} &= \prod_{e\in E} \frac{\mathbb{E}_{G^\star}\left[\exp\left(\frac{-1}{2(1-\rho)}\left[\sum_{j=1}^{p-1}\left(G_{je}-\sqrt{\rho}G_e^\star\right)^2 + \left(G_{p\Pi^{-1}(e)}-\sqrt{\rho}G_e^\star\right)^2\right]\right)\right]}{\mathbb{E}_{G^\star}\left[\exp\left(\frac{-1}{2(1-\rho)}\sum_{j=1}^{p}\left(G_{je}-\sqrt{\rho}G_e^\star\right)^2\right)\right]} \\ &= \prod_{e\in E} \frac{\mathbb{E}_{G^\star}\left[\exp\left(\frac{-p}{2(1-\rho)}\left[(\sqrt{\rho}G_e^\star)^2 - 2\sqrt{\rho}G_e^\star\overline{G_{\Pi^{-1}(e)}}\right]\right)\right]}{\mathbb{E}_{G^\star}\left[\exp\left(\frac{-p}{2(1-\rho)}\left[(\sqrt{\rho}G_e^\star)^2 - 2\sqrt{\rho}G_e^\star\overline{G_e}\right]\right)\right]},\end{aligned}$$

where we denote $\overline{G_{\Pi^{-1}(e)}} = \frac{1}{p}\left(\sum_{j=1}^{p-1}G_{je} + G_{p\Pi^{-1}(e)}\right)$ and $\overline{G}_e = \frac{1}{p}\sum_{j=1}^{p}G_{je}$. Let us compute the numerator of

the above equality. Recall that $G_e^\star \sim \mathcal{N}(0,1)$, so that

$$\begin{aligned}
\mathbb{E}_{G^\star}\left[\exp\left(\frac{-p}{2(1-\rho)}\left[(\sqrt{\rho}G_e^\star)^2 - 2\sqrt{\rho}G_e^\star\overline{G_{\Pi^{-1}(e)}}\right]\right)\right] =& \frac{1}{\sqrt{2\pi}}\int \exp\left(\frac{-x^2}{2}\right)\exp\left(\frac{-p}{2(1-\rho)}\left[\rho x^2 - 2\sqrt{\rho}x\overline{G_{\Pi^{-1}(e)}}\right]\right)dx \\
=& \frac{1}{\sqrt{2\pi}}\int \exp\left(\frac{-1}{2}\left[\left(1+\frac{\rho p}{1-\rho}\right)x^2 - 2p\frac{\sqrt{\rho}\overline{G_{\Pi^{-1}(e)}}}{1-\rho}x\right]\right)dx \\
=& \frac{1}{\sqrt{2\pi}}\int \exp\left(\frac{-(1+\rho(p-1))}{2(1-\rho)}\left[x^2 - 2p\frac{\sqrt{\rho}\overline{G_{\Pi^{-1}(e)}}}{1+\rho(p-1)}x\right]\right)dx \\
=& \frac{1}{\sqrt{2\pi}}\int \exp\left(\frac{-(1+\rho(p-1))}{2(1-\rho)}\left[\left(x - p\frac{\sqrt{\rho}\overline{G_{\Pi^{-1}(e)}}}{1+\rho(p-1)}\right)^2\right]\right)dx \\
&\times \exp\left(\frac{(1+\rho(p-1))}{2(1-\rho)}\frac{p^2\rho\overline{G_{\Pi^{-1}(e)}}^2}{(1+\rho(p-1))^2}\right)dx \\
=& \sqrt{\frac{1+\rho(p-1)}{1-\rho}}\exp\left(\frac{p^2\rho\overline{G_{\Pi^{-1}(e)}}^2}{2(1-\rho)(1+\rho(p-1))}\right),
\end{aligned}$$

and, similarly, we have

$$\mathbb{E}_{G^\star}\left[\exp\left(\frac{-p}{2(1-\rho)}\left[(\sqrt{\rho}G_e^\star)^2 - 2\sqrt{\rho}G_e^\star\overline{G_e}\right]\right)\right] = \sqrt{\frac{1+\rho(p-1)}{1-\rho}}\exp\left(\frac{p^2\rho\overline{G_e}^2}{2(1-\rho)(1+\rho(p-1))}\right),$$

so, wrapping things up,

$$\frac{d\,\mathbb{P}_\pi}{d\,\mathbb{P}_{\mathrm{id}}} = \prod_{e\in E}\exp\left(\frac{p^2\rho\left(\overline{G_{\Pi^{-1}(e)}}^2 - \overline{G_e}^2\right)}{2(1-\rho)(1+\rho(p-1))}\right). \tag{21}$$

Plugging this equality in the definition of the KL divergence leads to

$$\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right) = \frac{p^2\rho}{2(1-\rho)(1+\rho(p-1))}\sum_{e\in E}\mathbb{E}_\pi\left[\overline{G_{\Pi^{-1}(e)}}^2 - \overline{G_e}^2\right]. \tag{22}$$

Fix $e \in E$. We have

$$\mathbb{E}_\pi\left[\overline{G_{\Pi^{-1}(e)}}^2\right] = \frac{1}{p^2}\mathbb{E}_\pi\left[\left(\sum_{j\in[p-1]} G_{je} + G_{p\Pi^{-1}(e)}\right)^2\right] = \frac{1}{p^2}\left(p + \rho p(p-1)\right),$$

and on the other hand,

$$\mathbb{E}_\pi\left[\overline{G_e}^2\right] \geq \frac{1}{p^2}\left(p + \rho(p-1)(p-2)\right).$$

Hence,

$$\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right) \leq \frac{n(n-1)\rho^2(p-1)}{(1-\rho)(1+\rho(p-1))},$$

which concludes the proof of the lemma.

D Proof of Theorem 2.5

In this section, we seek to lower-bound

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p}\mathbb{P}\left[\mathrm{err}(\widehat{\pi}, \pi^\star) \neq 0\right].$$

We will rely on Fano's lemma to do so. For that purpose, we need to consider the law of $G_1, \ldots, G_p$ conditionally on $\pi^\star$. For any $\pi \in (\mathcal{S}_n)^p$, we now write $\mathbb{P}_\pi$ the joint law of $G_1, \ldots, G_p$ conditionally over $\pi^\star = \pi$. We write $\mathbb{P}$ the joint law of $G_1, \ldots, G_p$ for $\pi^\star$ taken randomly.

Lemma D.1.

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \mathbb{P}\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] = \inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \sup_{\pi\in(\mathcal{S}_n)^p} \mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right]$$

Proof of Lemma D.1. Let $\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p$. Let $\pi\in(\mathcal{S}_n)^p$ and let $\sigma=(\sigma_1,\dots,\sigma_p)\in(\mathcal{S}_n)^p$ be i.i.d. uniformly drawn independently from everything else. Define $H_{je}:=G_{j\sigma_j(e)}$. Then, for $G=(G_1,\dots,G_p)\sim\mathbb{P}_\pi$, we have $H:=(H_1,\dots,H_p)\sim\mathbb{P}$. Thus,

$$\mathbb{P}\left[\mathrm{err}(\widehat{\pi}(G),\pi^\star)\neq 0\right] = \mathbb{P}_\pi\left[\mathrm{err}\left(\widehat{\pi}\left(H\right),\sigma\circ\pi\right)\neq 0\right] = \mathbb{P}_\pi\left[\mathrm{err}\left(\sigma^{-1}\circ\widehat{\pi}\left(H\right),\pi\right)\neq 0\right],$$

and this for all $\pi\in(\mathcal{S}_n)^p$ which in turn implies

$$\mathbb{P}\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] = \sup_{\pi\in(\mathcal{S}_n)^p} \mathbb{P}_\pi\left[\mathrm{err}\left(\sigma^{-1}\circ\widehat{\pi}\left(H\right),\pi\right)\neq 0\right].$$

We conclude with the sought equality

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \mathbb{P}\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] = \inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \sup_{\pi\in(\mathcal{S}_n)^p} \mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right]$$

□

In the light of Lemma D.1, it is sufficient to lower-bound $\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p}\sup_{\pi\in(\mathcal{S}_n)^p}\mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right]$. We shall appeal to Fano's Lemma that we recall here (see e.g p. 57 of Giraud (2021)).

Lemma D.2. *Let $(\mathbb{P}_s)_{s\in[\ell]}$ a set of probability distributions on the same space $\mathcal{Y}$. Let $\mathbb{Q}$ a probability distribution on $\mathcal{Y}$ such that $\mathbb{P}_s\ll\mathbb{Q}$ for all $s\in[\ell]$. Then,*

$$\min_{\widehat{J}:\mathcal{Y}\to[\ell]} \frac{1}{\ell}\sum_{s\in[\ell]}\mathbb{P}_s\left[\widehat{J}\neq s\right] \geq 1-\frac{1+\frac{1}{\ell}\sum_{s\in[\ell]}\mathrm{KL}\left(\mathbb{P}_s,\mathbb{Q}\right)}{\log(\ell)}.$$

Let us consider $\mathbb{Q}=\mathbb{P}_{\mathrm{id}}$ for applying Lemma D.2. Our goal is to find a large set of p-tuples of permutations such that $\mathrm{KL}\left(\mathbb{P}_\pi,\mathbb{P}_{\mathrm{id}}\right)$ remains small. To do so, for $j\in[p]$ and $e=\{u,u'\}\in E$, we define $\pi^{(j,e)}$ the p-tuple of permutations by $\pi^{(j,e)}_{j'}=\mathrm{id}$ whenever $j\neq j'$ and $\pi^{(j,e)}_j$ is the transposition swapping u and u'.

Lemma D.3. *For all $j\in[p]$ and $e\in E$, we have*

$$\mathrm{KL}\left(\mathbb{P}_{\pi^{(j,e)}},\mathbb{P}_{\mathrm{id}}\right) = \frac{(2n-3)\,\rho^2(p-1)}{(1-\rho)\left(1+\rho(p-1)\right)}.$$

Combining together Lemma D.3 and Lemma D.2, we deduce that

$$\min_{\widehat{J}:\mathbb{R}^{E\times p}\to[p]\times E} \sum_{j,e\in[p]\times E}\mathbb{P}_{\pi^{(j,e)}}\left[\widehat{J}\neq(j,e)\right] \geq 1-\frac{1+\frac{(2n-3)\rho^2(p-1)}{(1-\rho)(1+\rho(p-1))}}{\log(pn(n-1)/2)},$$

and thus,

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \sup_{\pi\in(\mathcal{S}_n)^p}\mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] \geq 1-\frac{1+\frac{(2n-3)\rho^2(p-1)}{(1-\rho)(1+\rho(p-1))}}{\log(pn(n-1)/2)}.$$

First, assume that $\rho\leq c\sqrt{\frac{\log(np)}{np}}$ for a small enough numerical constant $c>0$. Then,

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \sup_{\pi\in(\mathcal{S}_n)^p}\mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] \geq 1-\frac{1+8n\rho^2p}{\log(pn(n-1)/2)} \geq 1-\varepsilon,$$

provided c is small enough and n is large enough.

Second, assume that $\rho\leq c\frac{\log(np)}{n}$. Then,

$$\inf_{\widehat{\pi}:\mathbb{R}^{E\times p}\to(\mathcal{S}_n)^p} \sup_{\pi\in(\mathcal{S}_n)^p}\mathbb{P}_\pi\left[\mathrm{err}(\widehat{\pi},\pi^\star)\neq 0\right] \geq 1-\frac{1+8n\rho}{\log(np(n-1))} \geq 1-\varepsilon,$$

provided c is small enough and n is large enough. This concludes the proof of the theorem.

D.1 Proof of Lemma D.3

By symmetry, it is enough to prove the result for $j = p$ and $e = \{1, 2\}$. We denote τ the transposition swapping 1 and 2. Starting back from (22), we have

$$\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right) = \frac{p^2\rho}{2(1-\rho)(1+\rho(p-1))} \sum_{e\in E} \mathbb{E}_\pi \left[\overline{G_{\Pi^{-1}(e)}}^2 - \overline{G_e}^2\right],$$

where we recall that $\overline{G_e} = \frac{1}{p}\sum_{j\in[p]} G_{je}$ and $\overline{G_{\Pi^{-1}(e)}} = \frac{1}{p}\left(\sum_{j\in[p-1]} G_{je} + G_{p\tau(e)}\right)$. In the above sum, all the terms such that $e \cap \{u, u'\} = \varnothing$ make a null contribution. For the other terms – there are $\frac{n(n-1)}{2} - \frac{(n-2)(n-3)}{2} = 2n-3$ such terms – we use

$$\mathbb{E}_\pi \left[\overline{G_{\Pi^{-1}(e)}}^2\right] = \frac{1}{p^2}\left[p + \rho p(p-1)\right] \quad \text{and} \quad \mathbb{E}_\pi \left[\overline{G_e}^2\right] = \frac{1}{p^2}\left(p + \rho(p-1)(p-2)\right).$$

We can conclude

$$\mathrm{KL}\left(\mathbb{P}_\pi, \mathbb{P}_{\mathrm{id}}\right) = \frac{(2n-3)\,\rho^2(p-1)}{(1-\rho)\,(1+\rho(p-1))}.$$

E Proof of Theorem 3.2

The minimization problem (16) defining $\mathrm{MMSE}_{\leq D}$ is separable, and since the random variables $M^\star_{(u,j),(u',j')}$ are exchangeable, the $\mathrm{MMSE}_{\leq D}$ can be reduced to

$$\mathrm{MMSE}_{\leq D} := \inf_{f\in\mathbb{R}_D[G]} \mathbb{E}\left[(f(G) - x)^2\right],$$

where $x = \mathbf{1}\left\{\pi^\star_1(1) = \pi^\star_2(1)\right\}$. In order to lower-bound $\mathrm{MMSE}_{\leq D}$, we proceed in two steps;

1. We consider the case where $p = 2$ and we use the technics of Schramm and Wein (2022) and Even et al. (2025b);
2. We reduce the general problem of aligning p Gaussian graphs to the problem of aligning one graph with some signal graph $G^\star$ and we use the lower bound obtained for two graphs.

Interestingly, the method of Schramm and Wein (2022) is not sufficient for directly considering the general case with p Gaussian graphs. A technical analysis based directly on Schramm and Wein (2022), using the technics from Even et al. (2025a) for the weak dependencies, would only get the sought result for $p \leq n$. The reduction step to the problem of two graphs has therefore two benefits; (i) simplifying the computations (ii) allowing us to deal with the regime $p \geq n$.

E.1 Case $p = 2$.

We recall that the graphs G_1, G_2 can be sampled from a signal graph $G^\star$ with $(G^\star_e)_{e\in E}$ taken independently and $\mathcal{N}(0,1)$, and from noise graphs Z_1, Z_2 with also independent and $\mathcal{N}(0,1)$ edges. With those graphs, we construct $G_{je} = \sqrt{\rho}G^\star_{\Pi^\star_j(e)} + \sqrt{1-\rho}Z_{je}$.

For $j \in \{1, 2\}$ and $e \in E$, we denote $Y_{je} = \frac{\sqrt{\rho}}{\sqrt{1-\rho}}G^\star_{p\Pi^\star(e)} + Z_{je} = \frac{1}{\sqrt{1-\rho}}G_{je}$. The set of estimators which are polynomials of degree at most D of G_1, G_2 is equal to the set of polynomials of degree at most D of Y_1, Y_2. Hence,

$$\mathrm{MMSE}_{\leq D} = \inf_{g\in\mathbb{R}_D[Y_1,Y_2]} \mathbb{E}\left[(g(Y_1, Y_2) - x)^2\right].$$

The model considered is a particular instance of the Additive Gaussian Noise model considered in Schramm and Wein (2022). Therefore, we can use Theorem 2.2 from Schramm and Wein (2022) that we transcript here with our notation. Before that, let us introduce the notion of *joint cumulants* (see e.g Novak (2014) for more details).

Backround on cumulants. Given $Y_1, \dots, Y_l$ random variables on the same probability space $\mathcal{Y}$, we define their *joint cumulant* as the quantity

$$\text{cumul}\left(Y_1, \dots, Y_l\right) := \sum_{G \in \mathcal{P}([\ell])} m(G) \prod_{R \in G} \mathbb{E}\left[\prod_{s \in [\ell]} Y_s\right],$$

with $\mathcal{P}([\ell])$ the set of all partitions of $[\ell]$ and, for $G \in \mathcal{P}([\ell])$, $m(G) = (-1)^{|G|-1}\left(|G|-1\right)!$. The *joint cumulants* of $Y_1, \dots, Y_l$ can also be computed recursively with the formula

$$\text{cumul}\left(Y_1, \dots, Y_l\right) = \mathbb{E}\left[Y_1 \dots Y_l\right] - \sum_{A \subsetneq [2,l]} \text{cumul}\left(Y_1, (Y_s)_{s \in A}\right) \mathbb{E}\left[\prod_{s \in [2,l] \setminus A} Y_s\right]. \tag{23}$$

Proposition E.1. *Schramm and Wein (2022)*

$$\text{MMSE}_{\leq D} \geq \mathbb{E}\left[x^2\right] - \sum_{\substack{\alpha = (\alpha_1, \alpha_2) \in (\mathbb{N}^E)^2 \\ |\alpha| \leq D}} \left(\frac{\rho}{1-\rho}\right)^{|\alpha|} \frac{\kappa_{x,\alpha}^2}{\alpha!},$$

where, for $\alpha = (\alpha_1, \alpha_2) \in (\mathbb{N}^E)^2$, $\alpha! = \prod_{e \in E} \alpha_{1e}! \alpha_{2e}!$, $|\alpha| = \sum_{e \in E} \alpha_{1e} + \alpha_{2e}$, *and,*

$$\kappa_{x,\alpha} = \text{cumul}\left(x, G_{\alpha_1}^\star, G_{\alpha_2}^\star\right),$$

with G_{α_j} *the multiset containing* α_{je} *copies of* G_{je} *for all* $e \in E$.

In the light of Proposition E.1, it is sufficient to upper-bound the cumulant $\kappa_{x,\alpha}$ for all α with $|\alpha| \leq D$ in order to lower-bound $\text{MMSE}_{\leq D}$. Such an upper-bound is provided by the next lemma, whose proof is postponed to Section E.1. For $\alpha = (\alpha_1, \alpha_2) \in (\mathbb{N}^E)^2$ and $j \in \{1, 2\}$, we shall define

$$\text{supp}(\alpha_j) = \left\{u \in [n], (\alpha_j)_{u:} \neq 0 \quad or \quad (\alpha_j)_{:u} \neq 0\right\},$$

which is the number of nodes with positive degree when seeing α_j as a multigraph with vertex set $[n]$.

Lemma E.1. *For all* $\alpha = (\alpha_1, \alpha_2) \in (\mathbb{N}^E)^2$, *we have*

$$|\kappa_{x,\alpha}| \leq \mathbf{1}\left\{|\alpha_1| = |\alpha_2|\right\} \left(|\alpha|(1 + |\alpha|/2)\right)^{|\alpha|/2} \left(\frac{\sqrt{2}}{\sqrt{n-1-|\alpha|}}\right)^{|\{1\} \cup \text{supp}(\alpha_1)| + |\{1\} \cup \text{supp}(\alpha_2)|},$$

Note that when $(\alpha_1, \alpha_2) \neq (0, 0)$ with $|\alpha_1| = |\alpha_2|$, for $j = 1, 2$, we have, $2 \leq |\{1\} \cup \text{supp}(\alpha_j)| \leq |\alpha| + 1$.

Lemma E.2. *Let* $d \in [D]$ *and* $2 \leq m_1, m_2 \leq d+1$. *There are at most* $n^{m_1+m_2-2} d^{2d}$ *elements* α *satisfying* $|\alpha| = d$, $|\text{supp}(\alpha_1) \cup \{1\}| = m_1$ *and* $|\text{supp}(\alpha_2) \cup \{1\}| = m_2$.

Proof of Lemma E.2. There are at most $n^{m_1+m_2-2}$ possibilities for choosing two sets S_1 and S_2 both containing 1 and of size m_1 and m_2. Let us then count the number of multi-graph α_1 with $d/2$ edges and with $\text{supp}(\alpha_1) \cup \{1\} = S_1$;

- if $m_1 \leq d$, we choose for each edge two extremities in S_1. Since we have to choose $d/2$ edges, that makes at most d^d possibilities,
- if $m_1 = d+1$, we choose for each edge two extremities in $S_1 \setminus \{1\}$. Since we have to choose $d/2$ edges, that also makes at most d^d possibilities.

The same thing holds for the the number of multi-graph α_2 with $d/2$ edges and with $\text{supp}(\alpha_2) \cup \{1\} = S_2$. In total, we have at most d^{2d} possibilities for choosing a multigraph $\alpha = (\alpha_1, \alpha_2)$, given $\text{supp}(\alpha_1) \cup \{1\} = S_1$ and $\text{supp}(\alpha_2) \cup \{1\} = S_2$. □

Combining Lemma E.1, Lemma E.2 and Proposition E.1, we get

$$\begin{aligned}
\sum_{\substack{\alpha=(\alpha_1,\alpha_2)\in(\mathbb{N}^E)^2\\|\alpha|\leq D}} \frac{\rho}{1-\rho}\frac{\kappa_{x,\alpha}^2}{\alpha!} - \frac{1}{n^2} \leq & \sum_{d=1}^{D}\sum_{m_1,m_2=2}^{d+1} n^{m_1+m_2-2} d^{2d}\left(\frac{\rho}{1-\rho} d\left(1+d/2\right)\right)^d \left(\frac{\sqrt{2}}{n-1-d}\right)^{m_1+m_2} \\
\leq & \frac{1}{n^2}\sum_{d=1}^{D}\sum_{m_1,m_2=2}^{d+1}\left(\frac{D^3\rho}{1-\rho}\left(1+\frac{D}{2}\right)\right)^d \left(\frac{1}{1-\frac{D+1}{n}}\right)^{2d+2} \\
\leq & \frac{2}{(n-1-D)^2}\sum_{d=1}^{D} d^2\left(\frac{D^3\rho}{1-\rho}\left(1+D/2\right)\frac{2}{\left(1-\frac{D+1}{n}\right)^2}\right)^d \\
\leq & \frac{2}{(n-1-D)^2}\sum_{d=1}^{D} d^2\zeta^d \\
\leq & \frac{2}{(n-1-D)^2}\zeta\frac{1+\zeta}{1-\zeta},
\end{aligned}$$

whenever $\zeta := \frac{D^3\rho}{1-\rho}\left(1+D/2\right)\frac{2}{\left(1-\frac{D+1}{n}\right)^2} < 1$.

E.2 General case p

In order to lower-bound $\mathrm{MMSE}_{\leq D}$ in the general case with p Gaussian Correlated Graphs with correlation ρ, we shall reduce to the problem of aligning two correlateed Gaussian graphs with correlation $\sqrt{\rho}$. Given any polynomial $f \in \mathbb{R}_D[G_1,\ldots,G_p]$, we construct a random polynomial $g \in \mathbb{R}_D\left[G_1', G_2'\right]$, with G_1', G_2' having correlation $\sqrt{\rho}$ and being permuted with the same permutation $\pi_1^\star, \pi_2^\star$, and such that $E\left[(f-x)^2\right] = \mathbb{E}\left[(g-x)^2\right]$. Then, using the result for the case $p=2$, we can prove a lower bound on $E\left[(f-x)^2\right]$. This being valid for all polynomial of degree at most D, we deduce from this a lower bound on $\mathrm{MMSE}_{\leq D}$.

Let $f \in \mathbb{R}_D[G_1,\ldots,G_p]$ any polynomial and let us lower-bound $\mathbb{E}\left[(f(G)-x)^2\right]$. We recall that we denote $\pi_1^\star,\ldots,\pi_p^\star$ the hidden partitions and that $x = \mathbf{1}\left\{\pi_1^\star(1) = \pi_2^\star(1)\right\}$. Let us generate G_1', G_2' two Gaussian graphs with, for all $e \in E$, $\mathbb{E}\left[G_{1e}' G_{2e'}'\right] = \sqrt{\rho}\mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(e')\right\}$. Then, sample $K_1,\ldots,K_p$ as follows;

1. Take $K_1 = G_1'$;
2. Independently from all the rest, sample $Z_2,\ldots,Z_p$ independent gaussian graphs;
3. Take $K_2 = \sqrt{\rho}G_2' + \sqrt{1-\rho}Z_2$;
4. For $j \in [3,p]$, take $K_{je} = \sqrt{\rho}G'_{2\Pi_j^\star(e)} + \sqrt{1-\rho}Z_{je}$.

We remark that $(x, K_1,\ldots,K_p)$ has the same joint law as $(x, G_1,\ldots,G_p)$. Let $g(G_1', G_2') = f(K_1,\ldots,K_p)$ a random polynomial of degree at most D. Then, plugging the result from the case $p=2$ with a correlation $\sqrt{\rho}$ instead of ρ, we get that, whenever

$$\zeta := \frac{D^3\sqrt{\rho}}{1-\sqrt{\rho}}\left(1+D/2\right)\frac{2}{\left(1-\frac{D+1}{n}\right)^2} < 1,$$

then,

$$\mathbb{E}\left[\left(g\left(G_1', G_2'\right)-x\right)^2\right] \geq \frac{1}{n} - \frac{1}{n^2} - \frac{2}{(n-1-D)^2}\zeta\frac{1+\zeta}{1-\zeta}.$$

Thus, under the same condition on ζ,

$$\mathbb{E}\left[\left(f\left(G_1,\ldots,G_p\right)-x\right)^2\right] = \mathbb{E}\left[\left(f\left(K_1,\ldots,K_p\right)-x\right)^2\right] = \mathbb{E}\left[\left(g\left(G_1', G_2'\right)-x\right)^2\right] \geq \frac{1}{n} - \frac{1}{n^2} - \frac{2}{(n-1-D)^2}\zeta\frac{1+\zeta}{1-\zeta}.$$

This being valid for all polynomial f of degree at most D, we conclude the proof of the theorem.

E.3 Proof of Lemma E.1

Let us fix $\alpha = (\alpha_1, \alpha_2)$ and let us upper bound $|\kappa_{x,\alpha}|$. The first ingredient is to apply Theorem 2.5 of Even et al. (2025b) which we adapt here with the specific notation of our problem. We recall that, for any edge $e = \{u, u'\} \in E$, we denote $\Pi_1^\star(e) = \{\pi_1^\star(u), \pi_1^\star(u')\}$ (resp. $\Pi_2^\star(e) = \{\pi_2^\star(u), \pi_2^\star(u')\}$).

Proposition E.2. *Even et al. (2025b) If $|\alpha_1| \neq |\alpha_2|$, then $\kappa_{x,\alpha} = 0$. If $|\alpha_1| = |\alpha_2|$, suppose that for all bijection $\psi : \alpha_1 \to \alpha_2$ (where we see α_1 and α_2 as multisets), we have*

$$|\mathrm{cumul}\left(x, (\Pi_1^\star(e) = \Pi_2^\star(\psi(e)))_{e\in\alpha_1}\right)| \leq C_\alpha,$$

then,

$$|\kappa_{x,\alpha}| \leq \left(\frac{|\alpha|}{2}\right)^{|\alpha|/2} C_\alpha\,.$$

In the light of Proposition E.2, we shall suppose that $|\alpha_1| = |\alpha_2|$ and we shall fix ψ a bijection from the multiset α_1 to the multiset α_2. Let us control

$$C_{\psi,\alpha} = \mathrm{cumul}\left(x, (\mathbf{1}\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\})_{e\in\alpha_1}\right)\,.$$

We refer to Section E.4 for a proof of the next lemma.

Lemma E.3. *For all $\alpha = (\alpha_1, \alpha_2)$ and all bijection $\psi : \alpha_1 \to \alpha_2$, we have*

$$|C_{\psi,\alpha}| \leq (1 + |\alpha|/2)^{|\alpha|/2} \left(\frac{\sqrt{2}}{\sqrt{n-1-|\alpha|}}\right)^{|\{1\}\cup\mathrm{supp}(\alpha_1)|+|\{1\}\cup\mathrm{supp}(\alpha_2)|}\,.$$

Combing Lemma E.3 and Proposition E.2 , we arrive at

$$|\kappa_{x,\alpha}| \leq (|\alpha|(1 + |\alpha|/2))^{|\alpha|/2} \left(\frac{\sqrt{2}}{\sqrt{n-1-|\alpha|}}\right)^{|\{1\}\cup\mathrm{supp}(\alpha_1)|+|\{1\}\cup\mathrm{supp}(\alpha_2)|}\,,$$

which concludes the proof of the lemma.

E.4 Proof of Lemma E.3

By symmetry of α_1 and α_2, it is sufficient to prove that, for all $\alpha = (\alpha_1, \alpha_2)$ and all bijection $\psi : \alpha_1 \to \alpha_2$, we have

$$|C_{\psi,\alpha}| \leq (1 + |\alpha_1|)^{|\alpha_1|} \left(\frac{\sqrt{2}}{n-1-|\alpha|}\right)^{|\{1\}\cup\mathrm{supp}(\alpha_1)|}\,.$$

For $\beta \subseteq \alpha_1$, we write

$$C_{\psi,(\beta,\psi(\beta))} := \mathrm{cumul}\left(x, (\mathbf{1}\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\})_{e\in\beta}\right)\,.$$

We shall proceed by induction over α_1 and prove that, for all $\beta \subseteq \alpha_1$, we have

$$|C_{\psi,(\beta,\psi(\beta))}| \leq (1 + |\beta|)^{|\beta|} \left(\frac{\sqrt{2}}{n-1-|\alpha|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}\,.$$

Initialization: $C_{x,0,0} = \mathrm{cumul}(x) = \mathbb{E}[x] = \frac{1}{n}$.

Induction step: Let $\varnothing \subsetneq \beta \subseteq \alpha_1$ and suppose that, for all $\gamma \subsetneq \beta$, we have

$$|C_{\psi,(\gamma,\psi(\gamma))}| \leq (1+|\gamma|)^{|\gamma|} \left(\frac{\sqrt{2}}{n-1-2|\gamma|}\right)^{|\{1\}\cup \mathrm{supp}(\gamma)|},$$

and let us prove that this inequality holds for β. Using Equation (23), we express the cumulant $C_{\psi,(\beta,\psi(\beta))}$ as a function of some mixed moments and of the cumulants $\left(C_{\psi,(\gamma,\psi(\gamma))}\right)_{\gamma\subsetneq\beta}$:

$$\begin{aligned} C_{\psi,(\beta,\psi(\beta))} =& \mathbb{E}\left[x \prod_{e\in\beta} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right] \\ &- \sum_{\gamma\subsetneq\beta} C_{\psi,(\gamma,\psi(\gamma))}\, \mathbb{E}\left[\prod_{e\in\beta\setminus\gamma} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right]. \end{aligned} \tag{24}$$

In the light of Equation (24), it is sufficient to upper-bound the mixed moments of the random variables $x, (\mathbf{1}\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\})_{e\in\beta}$.

Lemma E.4. *For all $\gamma \subsetneq \beta$;*

1. $|\mathbb{E}\left[x\prod_{e\in\beta}\mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right]| \leq \left(\frac{\sqrt{2}}{n-|\{1\}\cup\mathrm{supp}(\beta)|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}$
2. $|\mathbb{E}\left[\prod_{e\in\beta\setminus\gamma}\mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right]| \leq \left(\frac{\sqrt{2}}{n-|\mathrm{supp}(\beta\setminus\gamma)|}\right)^{|\mathrm{supp}(\beta\setminus\gamma)|}$.

Plugging Lemma E.4 in Equality (24), together with the induction hypothesis, leads to

$$\begin{aligned} |C_{\psi,(\beta,\psi(\beta)}| \leq& \left(\frac{\sqrt{2}}{n-|\{1\}\cup\mathrm{supp}(\beta)|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|} + \\ &+ \sum_{\gamma\subsetneq\beta}(1+|\gamma|)^{|\gamma|}\left(\frac{\sqrt{2}}{n-1-2|\gamma|}\right)^{|\{1\}\cup\mathrm{supp}(\gamma)|}\left(\frac{\sqrt{2}}{n-|\mathrm{supp}(\beta\setminus\gamma)|}\right)^{|\mathrm{supp}(\beta\setminus\gamma)|} \\ \leq& \left(\frac{\sqrt{2}}{n-1-2|\beta|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}\left(1+\sum_{\gamma\subsetneq\beta}(1+|\gamma|)^{|\gamma|}\right) \\ =& \left(\frac{\sqrt{2}}{n-1-2|\beta|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}\left(1+\sum_{l=0}^{|\beta|-1}\binom{|\beta|}{\ell}(1+l)^{\ell}\right) \\ =& \left(\frac{\sqrt{2}}{n-1-2|\beta|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}\left(1+\sum_{l=0}^{|\beta|-1}\binom{|\beta|}{\ell}|\beta|^{l}\right) \\ \leq& \left(\frac{\sqrt{2}}{n-1-2|\beta|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}\left(\sum_{l=0}^{|\beta|}\binom{|\beta|}{\ell}|\beta|^{l}\right) \\ =& (1+|\beta|)^{|\beta|}\left(\frac{\sqrt{2}}{n-1-2|\beta|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}, \end{aligned}$$

which concludes the induction and the proof of the lemma.

E.5 Proof of Lemma E.4

Let us prove the first point of the lemma. We recall $\mathrm{supp}(\beta) = \{u\in[n], (\beta)_{u:} \neq 0 \quad or \quad (\beta)_{:u} \neq 0\}$ is the set of nodes of positive degree when seeing β as a multigraph with vertex set $[n]$. We seek to control $\left|\mathbb{E}\left[x\prod_{e\in\beta}\mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right]\right|$.

Let us prove by induction that, for all $\beta \subseteq \alpha_1$,

$$\left| \mathbb{E}\left[x \prod_{e\in\beta} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right]\right| \leq \left(\frac{\sqrt{2}}{n - |\{1\}\cup \mathrm{supp}(\beta)|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|}.$$

Initialization: $\mathbb{E}[x] = \frac{1}{n}$.

Induction Step: Let $\beta \subseteq \alpha_1$. Let us suppose that the result holds for all $\gamma \subsetneq \beta$ and let us prove that it still holds for β. Let us fix $e_0 \in \beta$ and let us consider $\gamma_0 = \beta \setminus e_0$. Then,

$$\begin{aligned}\mathbb{E}\left[x \prod_{e\in\beta} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right] =& \mathbb{E}\left[x \prod_{e\in\gamma_0} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right] \\ &\times \mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\Big| \quad x\prod_{e\in\gamma_0} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\} = 1\right]. \end{aligned} \tag{25}$$

Let us work conditionally on the event $\mathcal{X} := \left\{x\prod_{e\in\gamma_0} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\} = 1\right\}$ (supposing it is of non-zero probability) and let us upper-bound $\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right]$. We distinguish three cases, according to the number of additional nodes brought by the edge $e_0 := \{u_0, u_0'\}$;

- If $\{u_0, u_0'\} \subset \{1\}\cup\mathrm{supp}(\gamma_0)$, then we have the trivial upper-bound
$$\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right] \leq 1,$$
- If $|\{u_0, u_0'\} \setminus \{\{1\}\cup\mathrm{supp}(\gamma_0)\}| = 1$, suppose by symmetry $u_0 \notin \{\{1\}\cup\mathrm{supp}(\gamma_0)\}$. Conditionally on $\mathcal{X}$, we know that, almost surely, $\pi_1^\star(u_0) \notin \{1\}\cup(\cup_{e\in\gamma_0}\psi(e))$ and $\pi_1^\star(u_0') \in \{1\}\cup(\cup_{e\in\gamma_0}\psi(e))$. Hence, for having a non-zero probability $\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right]$, we need to have $\psi(e_0) = (v_0, v_0')$ with $v_0 \notin \{1\}\cup(\cup_{e\in\gamma_0}\psi(e))$ and $v_0' \in \{1\}\cup(\cup_{e\in\gamma_0}\psi(e))$. In that case,,
$$\begin{aligned}\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right] =& \mathbb{P}\left[\pi_1^\star(u_0) = v_0|\mathcal{X}\right] \\ =& \frac{1}{n - |\{1\}\cup\mathrm{supp}(\gamma_0)|}.\end{aligned}$$
- If $|\{u_0, u_0'\} \setminus \{1\}\cup\mathrm{supp}(\gamma_0)| = 2$, then, for having $\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right] \neq 0$, we need to have $\psi(e_0) = \{v_0, v_0'\}$ with both $v_0, v_0' \notin \{1\}\cup(\cup_{e\in\gamma_0}\psi(e))$. In that case, conditionally on $\mathcal{X}$,
$$\begin{aligned}\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right] =& \mathbb{P}\left[\pi_1^\star(u_0) = v_0 \quad \textit{and} \quad \pi_1^\star(u_0') = v_0'\middle|\mathcal{X}\right] \\ &+ \mathbb{P}\left[\pi_1^\star(u_0) = v_0' \quad \textit{and} \quad \pi_1^\star(u_0') = v_0\middle|\mathcal{X}\right] \\ =& \frac{2}{(n - |\{1\}\cup\mathrm{supp}(\gamma_0)|)(n - 1 - |\{1\}\cup\mathrm{supp}(\gamma_0)|)}.\end{aligned}$$

For all those three cases, we have

$$\mathbb{P}\left[\Pi_1^\star(e_0) = \Pi_2^\star(\psi(e_0))\middle|\mathcal{X}\right] \leq \left(\frac{\sqrt{2}}{n - |\{1\}\cup\mathrm{supp}(\beta)|}\right)^{|\mathrm{supp}(e_0)\setminus(\{1\}\cup\mathrm{supp}(\gamma_0))|}.$$

Plugging this in Equation (25) together with the induction hypothesis, we end up with

$$\mathbb{E}\left[x \prod_{e\in\beta} \mathbf{1}\left\{\Pi_1^\star(e) = \Pi_2^\star(\psi(e))\right\}\right] \leq \left(\frac{\sqrt{2}}{n - |\{1\}\cup\mathrm{supp}(\beta)|}\right)^{|\{1\}\cup\mathrm{supp}(\beta)|},$$

which concludes the induction and the proof of the first point of the lemma. For the second point of the lemma, one could carry on the exact same proof, changing only the initialization step with $\mathbb{E}[1] = 1$ for β empty.

F Proof of Proposition 3.1

The proof of this proposition is adapted from the proof of Proposition 2.1 of Even et al. (2025a). Given a p-tuple of permutation $\pi = \pi_1, \dots, \pi_p$, we shall write M^π the matrix defined by $M^\pi_{(j,u),(j',u')} = \mathbf{1}\{\pi_j(u) = \pi_{j'}(u')\}$. We write $M^\star$ for $\pi^\star$. Then,

$$p(p-1)n^2 \mathrm{MMSE}_{poly} = \mathbb{E}\left[\|M^\star\|_F^2\right] - 1^2_{poly} = np^2 - 1^2_{poly},$$

where we define

$$1^2_{poly} = \sup_{\substack{\widehat{M} \quad poly-time \\ \mathbb{E}\left[\|\widehat{M}\|_F^2\right]=1}} \mathbb{E}\left[\langle M^\star, \widehat{M}\rangle_F\right]^2 .$$

Supposing that $\mathrm{MMSE}_{poly} = \frac{1}{n}(1-\varepsilon)$, with $0 \le \varepsilon \le 1$, we deduce that $1^2_{poly} = np + p(p-1)n\varepsilon$. In turn, we get, for all $\widehat{\pi}$ polynomial time estimator,

$$\begin{aligned}
\mathbb{E}\left[\|M^{\widehat{\pi}} - M^\star\|_F^2\right] =& \mathbb{E}\left[\|M^{\widehat{\pi}}\|_F^2\right] + \mathbb{E}\left[\|M^\star\|_F^2\right] - 2\,\mathbb{E}\left[\langle M^\star, M^{\widehat{\pi}}\rangle_F\right] \\
\ge& 2np^2 - 2\sqrt{np^2}\, corr_{poly} \\
\ge& 2np^2 - 2\sqrt{np^2}\sqrt{np + p(p-1)n\varepsilon} \\
\ge& 2np^2(1-\sqrt{\varepsilon})\,.
\end{aligned}$$

Using Lemma H.2 from Even et al. (2025a), we know that, for all estimator $\widehat{\pi}$,

$$[1 - \mathrm{err}(\widehat{\pi}, \pi^\star)]^2 \le 1 - \frac{\|M^{\widehat{\pi}} - M^\star\|_F^2}{2np^2},$$

which, in turn, implies

$$\mathbb{E}\left[(1 - \mathrm{err}(\widehat{\pi}, \pi^\star))^2\right] \le \sqrt{\varepsilon}\,.$$

This concludes the proof of the proposition.