THE PHASE TRANSITION PHENOMENON OF SHUFFLED REGRESSION

Anonymous authors
Paper under double-blind review

ABSTRACT

We study the phase transition phenomenon inherent in the shuffled (permuted) regression problem, which has found numerous applications in databases, privacy, data analysis, etc. For the permuted regression task: $\mathbf{Y} = \mathbf{\Pi}^\natural \mathbf{X}\mathbf{B}^\natural$, the goal is to recover the permutation matrix $\mathbf{\Pi}^\natural$ as well as the coefficient matrix $\mathbf{B}^\natural$. It has been empirically observed in prior studies that when recovering $\mathbf{\Pi}^\natural$, there exists a phase transition phenomenon: the error rate drops to zero rapidly once the parameters reach certain thresholds. In this study, we aim to precisely identify the locations of the phase transition points by leveraging techniques from *message passing* (MP).

In our analysis, we first transform the permutation recovery problem into a probabilistic graphical model. Then, we leverage the analytical tools rooted in the message passing (MP) algorithm and derive an equation to track the convergence of the MP algorithm. By linking this equation to the branching random walk process, we are able to characterize the impact of the *signal-to-noise-ratio* (snr) on the permutation recovery. Depending on whether the signal is given or not, we separately investigate the oracle case and the non-oracle case. The bottleneck in identifying the phase transition regimes lies in deriving closed-form formulas for the corresponding critical points, but only in rare scenarios can one obtain such precise expressions. To tackle this challenge, we propose the Gaussian approximation method, which allows us to obtain the closed-form formulas in almost all scenarios. In the oracle case, our method can fairly accurately predict the phase transition snr. In the non-oracle case, our proposed algorithm can predict the maximum allowed number of permuted rows and uncover its dependency on the sample number.

Numerical experiments reveal that the observed phase transition points are well aligned with our theoretical predictions. Our study will motivate exploiting MP algorithms (and related techniques) as an effective tool for permuted regression problems, which have found applications in machine learning, privacy, and databases.

1 INTRODUCTION

In this paper, we consider the following permuted (shuffled) linear regression problem:

$$\mathbf{Y} = \mathbf{\Pi}^\natural \mathbf{X}\mathbf{B}^\natural + \sigma \mathbf{W}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{n\times m}$ denotes the matrix of observations, $\mathbf{\Pi}^\natural \in \{0,1\}^{n\times n}$ is the permutation matrix, $\mathbf{X} \in \mathbb{R}^{n\times p}$ is the design matrix, $\mathbf{B}^\natural \in \mathbb{R}^{p\times m}$ is the matrix of signals (regressors), $\mathbf{W} \in \mathbb{R}^{n\times m}$ denotes the additive noise matrix (with unit variance), and σ^2 is the noise variance. The task is to recover both the signal matrix $\mathbf{B}^\natural$ and the permutation matrix $\mathbf{\Pi}^\natural$. The research on this challenging permuted regression problem dates back at least to 1970s under the name "broken sample problem" (DeGroot et al., 1971; Goel, 1975; DeGroot & Goel, 1976; 1980; Bai & Hsing, 2005). Recent years have witnessed a revival of this problem due to its broad spectrum of applications in (e.g.,) privacy protection, data integration, etc. (Unnikrishnan et al., 2015; Pananjady et al., 2018; Slawski & Ben-David, 2019; Pananjady et al., 2017; Slawski et al., 2020; Zhang & Li, 2020).

Specifically, this paper will focus on studying the "phase transition" phenomenon in recovering the whole permutation matrix $\mathbf{\Pi}^\natural$: the error rate for the permutation recovery sharply drops to zero once the parameters reach certain thresholds. In particular, we leverage techniques in the *message passing* (MP) algorithm literature to identify the precise positions of the phase transition thresholds. The

bottleneck in identifying the phase transition regimes lies in deriving closed-form formulas for the corresponding critical points. This is a highly challenging task because only in rare scenarios can one obtain such precise expressions. To tackle the difficulty, we propose the Gaussian approximation method which allows us to obtain the closed-form formula in almost all scenarios. We should mention that, in previous studies (Slawski et al., 2020; Slawski & Ben-David, 2019; Pananjady et al., 2017; Zhang et al., 2022; Zhang & Li, 2020), this phase transition phenomenon was empirically observed.

Related work. The problem we study in this paper simultaneously touches two distinct areas of research: (A) permutation recovery, and (B) message passing (MP). In the literature of permuted linear regression, essentially all existing works used the same setting (1). Pananjady et al. (2018); Slawski & Ben-David (2019) consider the single observation model (i.e., $m = 1$) and prove that the *signal-to-noise-ratio* (snr) for the correct permutation recovery is $\mathbb{O}_{\mathrm{P}}\left(n^c\right)$, where $c > 0$ is some positive constant. Slawski et al. (2020); Zhang & Li (2020); Zhang et al. (2022) investigate the multiple observations model (i.e., $m > 1$) and suggest that the snr requirement can be significantly decreased, from $\mathbb{O}_{\mathrm{P}}\left(n^c\right)$ to $\mathbb{O}_{\mathrm{P}}\left(n^{c/m}\right)$. In particular, Zhang & Li (2020) develop an estimator which we will leverage and analyze for studying the phase transition phenomenon. Our analysis leads to the precise identification of the locations of the phase transition thresholds.

Another line of related research comes from the field of statistical physics. For example, using the replica method, Mézard & Parisi (1985; 1986) study the *linear assignment problem* (LAP), i.e., $\min_{\mathbf{\Pi}} \sum_{i,j} \mathbf{\Pi}_{ij}\mathbf{E}_{ij}$ where $\mathbf{\Pi}$ denotes a permutation matrix and $\mathbf{E}_{ij}$ is i.i.d random variable uniformly distributed in $[0, 1]$. Martin et al. (2005) then generalize LAP to multi-index matching and presented an investigation based on MP algorithm. Recently, Caracciolo et al. (2017); Malatesta et al. (2019) extend the distribution of $\mathbf{E}_{ij}$ to a broader class. However, all the above works exhibit no phase transition. In Chertkov et al. (2010), this method is extended to the particle tracking problem, where a phase transition phenomenon is observed. Later, Semerjian et al. (2020) modify it to fit the graph matching problem, which paves way for our work in studying the permuted linear regression problem.

Our contributions. We propose the first framework to identify the precise locations of phase transition thresholds associated with permuted linear regression. In the oracle case where $\mathbf{B}^\natural$ is known, our scheme is able to determine the phase transition snr. In the non-oracle case where $\mathbf{B}^\natural$ is not given, our method will also predict the maximum allowed number of permuted rows and uncover its dependence on the ratio p/n. In our analysis, we identify the precise positions of the phase transition points in the large-system limit, e.g., n, m, p all approach to infinity with $m/n \to \tau_m$, $p/n \to \tau_p$. Interestingly, numerical results well match predictions even when n, m, p are not large. There is one additional contribution. In the graphical model based on the linear assignment problem, we can modify the graph and design a scheme for partial recovery, which is a separate contribution and may be further analyzed for future study.

Here, we would also like to briefly mention the technical challenges. Compared with the previous works (Mezard & Montanari, 2009; Talagrand, 2010; Linusson & Wästlund, 2004; Mézard & Parisi, 1987; 1986; Parisi & Ratiéville, 2002; Semerjian et al., 2020), where the edge weights are relatively simple, our edge weights usually involve high-order interactions across Gaussian random variables and are densely correlated. To tackle this issue, our proposed approximation method to compute the phase transition thresholds consists of three parts: 1) performing Gaussian approximation; 2) modifying the leave-one-out technique; and 3) performing size correction. A detailed explanation can be found in Section D. Hopefully, our approximation method will serve independent technical interests for researchers in the machine learning community.

Notations. In this paper, $a \xrightarrow{\text{a.s.}} b$ denotes a converges almost surely to b. We denote $f(n) \simeq g(n)$ when $\lim_{n\to\infty} f(n)/g(n) = 1$, and $f(n) = \mathbb{O}_{\mathrm{P}}\left(g(n)\right)$ if the sequence $f(n)/g(n)$ is bounded in probability, and $f(n) = o_{\mathrm{P}}\left(g(n)\right)$ if $f(n)/g(n)$ converges to zero in probability. The inner product between two vectors (resp. matrices) are denoted as $\langle \cdot, \cdot \rangle$. For two distributions d_1 and d_2, we write $d_1 \cong d_2$ if they are equal up to normalization. Moreover, $\mathcal{P}_n$ denotes the set of all possible permutation matrices: $\mathcal{P}_n \triangleq \{\mathbf{\Pi} \in \{0,1\}^{n\times n}, \sum_i \mathbf{\Pi}_{ij} = 1, \sum_j \mathbf{\Pi}_{ij} = 1\}$. The *signal-to-noise-ratio* is $\mathsf{snr} = \frac{\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}{m\cdot\sigma^2}$, where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm and σ^2 is the variance of the sensing noise.

2 PERMUTATION RECOVERY USING THE MESSAGE PASSING ALGORITHM

Inspired by Mezard & Montanari (2009); Chertkov et al. (2010); Semerjian et al. (2020), we leverage tools from the statistical physics to identify the locations of the phase transition threshold. We start this section with a brief review of the *linear assignment problem* (LAP), which reads as

$$\widehat{\mathbf{\Pi}} = \ \text{argmin}_{\mathbf{\Pi}\in\mathcal{P}_n} \langle \mathbf{\Pi}, \mathbf{E} \rangle \,, \tag{2}$$

where $\mathbf{E} \in \mathbb{R}^{n\times n}$ is a fixed matrix and $\mathcal{P}_n$ denotes the set of all possible permutation matrices. We follow the approach in Mezard & Montanari (2009); Semerjian et al. (2020) and introduce a probability measure over the permutation matrix $\mathbf{\Pi}$, which is written as

$$\mu(\mathbf{\Pi}) = \ (1/Z) \prod_i \mathbb{1}\big(1 - \sum_j \mathbf{\Pi}_{ij}\big) \prod_j \mathbb{1}\big(1 - \sum_i \mathbf{\Pi}_{ij}\big) \times \exp\Big(-\beta \sum_{i,j} \mathbf{\Pi}_{ij}\mathbf{E}_{ij}\Big), \tag{3}$$

where $\mathbb{1}(\cdot)$ is the indicator function, Z is the normalization constant of the probability measure $\mu(\mathbf{\Pi})$, and $\beta > 0$ is an auxiliary parameter. It is easy to verify the following two properties:

- the ML estimator in (2) can be rewritten as $\widehat{\mathbf{\Pi}} = \text{argmax}_{\mathbf{\Pi}}\mu(\mathbf{\Pi})$;
- the probability measure $\mu(\mathbf{\Pi})$ concentrates on $\widehat{\mathbf{\Pi}}$ when letting $\beta \to \infty$.

In the next three subsections, we study the impact of $\{\mathbf{E}_{ij}\}$ on the reconstructed permutation $\widehat{\mathbf{\Pi}}$ with the *message passing* (MP) algorithm. First, we associate a probabilistic graphical model with the probability measure defined in (3). Then, we rewrite the solution in (2) in the language of the MP algorithm. Finally, we derive an equation (7) to track the convergence of the MP algorithm. By exploiting relation of (7) to the *branching random walk* (BRW) process, we identify the phase transition points corresponding to the LAP in (2).

2.1 CONSTRUCTION OF THE GRAPHICAL MODEL

Firstly, we construct the factor graph associated with the probability measure in (3). Adopting the same strategy as in Chapter 16 of Mezard & Montanari (2009), we conduct the following operations:

- associating each variable $\mathbf{\Pi}_{ij}$ a variable node v_{ij};
- connecting the variable node v_{ij} a function node representing the term $e^{-\beta\mathbf{\Pi}_{ij}\mathbf{E}_{ij}}$;
- linking each constraint $\sum_i \mathbf{\Pi}_{ij} = 1$ to a function node and similarly for the constraint $\sum_j \mathbf{\Pi}_{ij} = 1$.

A graphical representation is available in Figure 1.

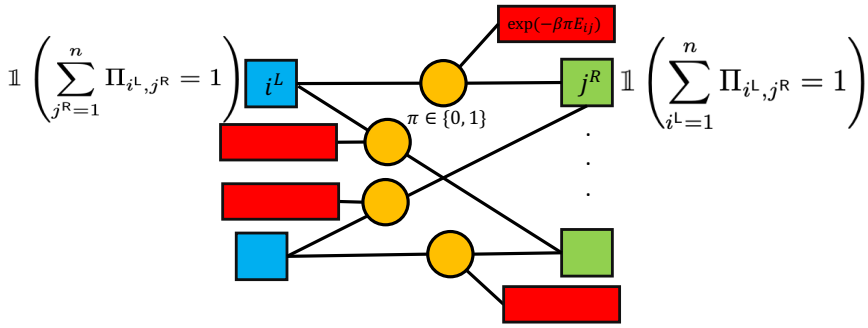

Figure 1: The constructed graphical model. The circle icons represent the variable nodes and the square icons represent the function nodes: a blue square for the constraint on the rows of $\mathbf{\Pi}$, a green square for the constraint on the columns of $\mathbf{\Pi}$, and a red square for the function $e^{-\beta\pi\mathbf{E}_{ij}}$.

Now we briefly review the MP algorithm. Informally speaking, MP is a local algorithm to compute the marginal probabilities over the graphical model. In each iteration, the variable node v transmits

the message to its incident function node f by multiplying all incoming messages except the message along the edge (v, f). The function node f transmits the message to its incident variable node v by computing the weighted summary of all incoming messages except the message along the edge (f, v). For a detailed introduction to MP, we refer readers to Kschischang et al. (2001), Chapter 16 in MacKay et al. (2003), and Chapter 14 in Mezard & Montanari (2009).

It is known that MP can obtain the exact marginals (Mezard & Montanari, 2009) for singly connected graphical models. For other types of graphs, however, whether MP can obtain the exact solution still remains an open problem (Cantwell & Newman, 2019; Kirkley et al., 2021). At the same time, numerical evidences have been witnessed to show that MP can yield meaningful results for graphs with loops; particular examples include applications in the coding theory (Chung, 2000; Richardson & Urbanke, 2001; 2008) and the LAP (which happens to be our case) (Mezard & Montanari, 2009; Chertkov et al., 2010; Caracciolo et al., 2017; Malatesta et al., 2019; Semerjian et al., 2020).

2.2 The message passing (MP) algorithm

Next, we perform permutation recovery via MP. The following derivation follows the standard procedure, which can be found in the previous works (Mezard & Montanari, 2009; Semerjian et al., 2020). We denote the message flow from the node i^{L} to the variable node $(i^{\mathsf{L}}, j^{\mathsf{R}})$ as $\widehat{m}_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(\cdot)$ and that from the edge $(i^{\mathsf{L}}, j^{\mathsf{R}})$ to node i^{L} as $m_{(i^{\mathsf{L}},j^{\mathsf{R}})\to i^{\mathsf{L}}}(\cdot)$. Similarly, we define $\widehat{m}_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(\cdot)$ and $m_{(i^{\mathsf{L}},j^{\mathsf{R}})\to j^{\mathsf{R}}}(\cdot)$ as the message flow transmitted between the functional node j^{R} and the variable node $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$. Here the superscripts L and R are used to indicate the positions of the node (left and right, respectively). Roughly speaking, these transmitted messages can be viewed as (unnormalized) conditional probability $\mathbb{P}(\Pi_{i,j} = \{0, 1\}|(\cdot))$ with the joint PDF being defined in (3). The message transmission process is to iteratively compute these conditional probabilities.

First, we consider the message flows transmitted between the functional node i^{L} and the variable node $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$, which are written as

$$m_{(i^{\mathsf{L}},j^{\mathsf{R}})\to i^{\mathsf{L}}}(\pi) \cong \widehat{m}_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(\pi) e^{-\beta\pi\mathbf{E}_{i^{\mathsf{L}},j^{\mathsf{R}}}},$$

$$\widehat{m}_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(\pi) \cong \sum_{\pi_{i^{\mathsf{L}},k^{\mathsf{R}}}} \prod_{k^{\mathsf{R}}\neq j^{\mathsf{R}}} \widehat{m}_{k^{\mathsf{R}}\to(i^{\mathsf{L}},k^{\mathsf{R}})}(\pi_{i^{\mathsf{L}},k^{\mathsf{R}}}) \cdot e^{-\beta\pi_{i^{\mathsf{L}},k^{\mathsf{R}}}\mathbf{E}_{i^{\mathsf{L}},k^{\mathsf{R}}}} \mathbb{1}(\pi + \sum_{k} \pi_{i^{\mathsf{L}},k^{\mathsf{R}}} = 1), \tag{4}$$

where $\pi \in \{0, 1\}$ is a binary value. Similarly, we can write the message flows between the functional node j^{R} and the variable node $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$, which are denoted as $m_{(i^{\mathsf{L}},j^{\mathsf{R}})\to j^{\mathsf{R}}}(\pi)$ and $\widehat{m}_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(\pi)$, respectively. With the parametrization approach, we define

$$h_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} \triangleq \frac{1}{\beta}\log\frac{\widehat{m}_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(1)}{\widehat{m}_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(0)}, \qquad h_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} \triangleq \frac{1}{\beta}\log\frac{\widehat{m}_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(1)}{\widehat{m}_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}(0)}.$$

Denote $\zeta_{i^{\mathsf{L}},j^{\mathsf{R}}}$ as $h_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} + h_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} - \mathbf{E}_{i^{\mathsf{L}},j^{\mathsf{R}}}$. We select the edge $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$ according to the probability $m_{(i^{\mathsf{L}},j^{\mathsf{R}})}(\pi) \triangleq \frac{\exp(\pi\cdot\beta\zeta_{i^{\mathsf{L}},j^{\mathsf{R}}})}{1+\exp(\beta\zeta_{i^{\mathsf{L}},j^{\mathsf{R}}})}$, $\pi \in \{0, 1\}$. Provided $m_{(i^{\mathsf{L}},j^{\mathsf{R}})}(1) > m_{(i^{\mathsf{L}},j^{\mathsf{R}})}(0)$, or equivalently,

$$\zeta_{i^{\mathsf{L}},j^{\mathsf{R}}} > 0, \tag{5}$$

we pick $\widehat{\pi}(i^{\mathsf{L}}) = j^{\mathsf{R}}$; otherwise, we have $\widehat{\pi}(i^{\mathsf{L}}) \neq j^{\mathsf{R}}$. Due to the fact that $\mu(\mathbf{\Pi})$ concentrates on $\widehat{\mathbf{\Pi}}$ when β is sufficiently large, we can thus rewrite the MP update equation as

$$h_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} = \min_{k^{\mathsf{R}}\neq j^{\mathsf{R}}} \mathbf{E}_{i^{\mathsf{L}},k^{\mathsf{R}}} - h_{k^{\mathsf{R}}\to(i^{\mathsf{L}},k^{\mathsf{R}})}, \qquad h_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})} = \min_{k^{\mathsf{L}}\neq i^{\mathsf{L}}} \mathbf{E}_{k^{\mathsf{L}},j^{\mathsf{R}}} - h_{k^{\mathsf{L}}\to(k^{\mathsf{L}},j^{\mathsf{R}})}, \tag{6}$$

which is attained by letting $\beta \to \infty$.

2.3 Identification of the phase transition threshold

To identify the phase transition phenomenon inherent in the MP update equation (6), we follow the strategy in Semerjian et al. (2020) and divide all edges $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$ into two categories according to

whether the edge $\left(i^{\mathsf{L}}, j^{\mathsf{R}}\right)$ corresponds to the ground-truth permutation matrix $\mathbf{\Pi}^{\natural}$ or not. Within each category, we assume the edges' weights and the message flows along them can be represented by independently identically distributed random variables.

For the edge $(i^{\mathsf{L}}, \pi^{\natural}(i^{\mathsf{L}}))$ for the ground-truth correspondence, we represent the random variable associated with the weight $\mathbf{E}_{ij}$ as Ω. The random variable for the message flow along this edge is denoted H (for both $h_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}$ and $h_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}$). For the rest of edges $(i^{\mathsf{L}}, j^{\mathsf{R}})$ $(j^{\mathsf{R}} \neq \pi^{\natural}(i^{\mathsf{L}}))$, we define the corresponding random variables for the edge weight and message flow as $\widehat{\Omega}$ and $\widehat{H}$, respectively. Then, we can rewrite (6) as

$$\widehat{H}^{(t+1)} = \min\left(\Omega - H^{(t)}, H^{'(t)}\right), \quad H^{(t+1)} = \min_{1\leq i\leq n-1} \widehat{\Omega}_i - \widehat{H}_i^{(t)}, \tag{7}$$

where $(\cdot)^{(t)}$ denotes the update in the t-th iteration, $H^{'}$ is an independent copy of H, $\{H_i^{(t)}\}_{1\leq i\leq n-1}$ and $\{\widehat{\Omega}_i\}_{1\leq i\leq n-1}$ denote the i.i.d. copies of random variables $H_{(\cdot)}^{(t)}$ and $\widehat{\Omega}_{(\cdot)}$. This equation (7) can be viewed as the analogous version of the *density evolution* and *state evolution*, which are used to analyze the convergence of the message passing and approximate message passing algorithm, respectively (Chung, 2000; Richardson & Urbanke, 2001; 2008; Donoho et al., 2009; Maleki, 2010; Bayati & Montanari, 2011; Rangan, 2011).

Remark 1. *We conjecture that the distribution difference in the edges' weights is a necessary component in capturing the phase transition. On one hand, according to Mézard & Parisi (1986; 1987); Parisi & Ratiéville (2002); Linusson & Wästlund (2004); Mezard & Montanari (2009); Talagrand (2010), there is no phase transition phenomenon in LAP if the edges' weights, i.e., $\mathbf{E}_{ij}$, are assumed to be i.i.d uniformly distributed in $[0, 1]$. On the other hand, Semerjian et al. (2020) show a phase transition phenomenon when assuming the weights $\mathbf{E}_{ij}$ follow different distributions among the edges associated with the ground-truth correspondence $\left(i^{\mathsf{L}}, \pi^{\natural}(i^{\mathsf{L}})\right)$ and the rest edges.*

Relation to *branching random walk* (BRW) process. Conditional on the event that the permutation can be perfectly reconstructed, i.e., $H + H^{'} > \Omega$ as in (5), we can simplify (7) as

$$H^{(t+1)} = \min_{1\leq i\leq n-1} H_i^{(t)} + \Xi_i, \tag{8}$$

where Ξ is defined as the difference between $\widehat{\Omega}$ and Ω, which is written as $\Xi \triangleq \widehat{\Omega} - \Omega$, and $\{H_i^{(t)}\}_{1\leq i\leq n-1}$ and $\{\Xi_i\}_{1\leq i\leq n-1}$ denote the i.i.d. copies of random variables $H_{(\cdot)}^{(t)}$ and $\Xi_{(\cdot)}$.

Adopting the same viewpoint of Semerjian et al. (2020), we treat (8) as a *branching random walk* (BRW) process, which enjoys the following property.

Theorem 1 (Hammersley (1974); Kingman (1975); Semerjian et al. (2020))**.** *Consider the* recursive distributional equation $K^{(t+1)} = \min_{1\leq i\leq n} K_i^{(t)} + \Xi_i$, *where $K_i^{(t)}$ and Ξ_i are i.i.d copies of random variables $K_{(\cdot)}^{(t)}$ and $\Xi_{(\cdot)}$, we have $\frac{K^{(t+1)}}{t} \xrightarrow{\text{a.s.}} -\inf_{\theta>0} \frac{1}{\theta} \log\left[\sum_{i=1}^{n} \mathbb{E}e^{-\theta\Xi_i}\right]$, conditional on the event that $\lim_{t\to\infty} K^{(t)} \neq \infty$.*

With Theorem 1, we can compute phase transition point for the correct (full) permutation recovery, i.e., $H + H^{'} > \Omega$, by letting $\inf_{\theta>0} \frac{1}{\theta} \log\left[\sum_{i=1}^{n} \mathbb{E}e^{-\theta\Xi_i}\right] = 0$, since otherwise the condition in (5) will be violated (see a detailed explanation in Appendix). In practice, directly computing the infimum of $\inf_{\theta>0} \frac{1}{\theta} \log\left[\sum_{i=1}^{n} \mathbb{E}e^{-\theta\Xi_i}\right]$ is only possible for limited scenarios. In the next section, we propose an approximate computation method for the phase transition points, which is capable of covering a broader class of scenarios.

3 Analysis of the Phase Transition Points

Recall that, in this paper, we consider the following linear regression problem with permuted labels

$$\mathbf{Y} = \mathbf{\Pi}^{\natural}\mathbf{X}\mathbf{B}^{\natural} + \sigma\mathbf{W},$$

where $\mathbf{Y} \in \mathbb{R}^{n\times m}$ represents the matrix of observations, $\mathbf{\Pi}^{\natural} \in \mathcal{P}_n$ denotes the permutation matrix to be reconstructed, $\mathbf{X} \in \mathbb{R}^{n\times p}$ is the sensing matrix with each entry $\mathbf{X}_{ij}$ following the i.i.d standard

normal distribution, $\mathbf{B}^\natural \in \mathbb{R}^{p\times m}$ is the matrix of signals, and $\mathbf{W} \in \mathbb{R}^{n\times m}$ represents the additive noise matrix and its entries $\mathbf{W}_{ij}$ are i.i.d standard normal random variables. In addition, we denote h as the number of permuted rows corresponding to the permutation matrix $\mathbf{\Pi}^\natural$.

In this work, we focus on studying the "phase transition" phenomenon in recovering $\mathbf{\Pi}^\natural$ from the pair $(\mathbf{Y}, \mathbf{X})$. That is, the error rate for the permutation recovery sharply drops to zero once certain parameters reach the thresholds. In particular, our analysis will identify the precise positions of the phase transition points in the large-system limit, i.e., n, m, p, and h all approach to infinity with $m/n \to \tau_m$, $p/n \to \tau_p$, $h/n \to \tau_h$. We will separately study the phase transition phenomenon in 1) the oracle case where $\mathbf{B}^\natural$ is given as a prior, and 2) the non-oracle case where $\mathbf{B}^\natural$ is unknown.

In this section, we consider the oracle scenario, as a warm-up example. To reconstruct the permutation matrix $\mathbf{\Pi}^\natural$, we adopt the following *maximum-likelihood* (ML) estimator:

$$\widehat{\mathbf{\Pi}}^{\mathsf{oracle}} = \operatorname{argmin}_{\mathbf{\Pi}} \left\langle \mathbf{\Pi}, -\mathbf{Y}\mathbf{B}^{\natural\top}\mathbf{X}^\top \right\rangle, \quad \text{s.t.} \quad \sum_i \mathbf{\Pi}_{ij} = 1, \sum_j \mathbf{\Pi}_{ij} = 1, \mathbf{\Pi} \in \{0,1\}^{n\times n}. \tag{9}$$

Denoting the variable $\mathbf{E}_{ij}^{\mathsf{oracle}}$ as $-\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathbf{B}^{\natural\top}\mathbf{X}_j - \sigma \mathbf{W}_i^\top \mathbf{B}^{\natural\top}\mathbf{X}_j$, $(1 \le i, j \le n)$, we can transform the objective function in (9) as the canonical form of LAP, i.e., $\sum_{i,j} \mathbf{\Pi}_{ij}\mathbf{E}_{ij}^{\mathsf{oracle}}$.

3.1 The phase transition threshold for the oracle case

In the oracle case where $\mathbf{B}^\natural$ is known, we define the following random variable Ξ:

$$\Xi = \boldsymbol{x}^\top \mathbf{B}^\natural \mathbf{B}^{\natural\top} (\boldsymbol{x} - \boldsymbol{y}) + \sigma \boldsymbol{w} \mathbf{B}^{\natural\top} (\boldsymbol{x} - \boldsymbol{y}), \tag{10}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ follow the distribution $\mathsf{N}(\mathbf{0}, \mathbf{I}_{p\times p})$, and $\boldsymbol{w}$ follows the distribution $\mathsf{N}(\mathbf{0}, \mathbf{I}_{m\times m})$.

Recalling Theorem 1, we predict the phase transition point by letting

$$\inf_{\theta>0} {}^{1}/_{\theta} \cdot \log \left(\sum_{i=1}^{n} \mathbb{E} e^{-\theta \Xi_i} \right) = \inf_{\theta>0} {}^{1}/_{\theta} \cdot \left(\log n + \log \mathbb{E} e^{-\theta\Xi} \right) = 0. \tag{11}$$

The computation procedure consists of two stages:

- **Stage I.** We compute the optimal θ_*, which is written as $\theta_* = \operatorname{argmin}_{\theta>0} {}^{1}/_{\theta} \cdot \left(\log n + \log \mathbb{E} e^{-\theta\Xi_i} \right)$.
- **Stage II.** We plug the optimal θ^* into (11) and obtain the phase transition snr accordingly.

The following context illustrates the computation details.

Stage I: Determine θ_*. The key in determining θ_* lies in the computation of $\mathbb{E}e^{-\theta\Xi}$, which is summarized in the following proposition.

Proposition 1. *For the random variable Ξ defined in (10), we can write its expectation as*

$$\mathbb{E}e^{-\theta\Xi} = \prod_{i=1}^{\operatorname{rank}(\mathbf{B}^\natural)} \left[1 + 2\theta\lambda_i^2 - \theta^2\lambda_i^2 \left(\lambda_i^2 + 2\sigma^2 \right) \right]^{-\frac{1}{2}}, \tag{12}$$

provided that

$$\theta^2\sigma^2\lambda_i^2 < 1 \text{ and } \theta^2\lambda_i^2 \left(\lambda_i^2 + 2\sigma^2 \right) \le 1 + 2\theta\lambda_i^2 \tag{13}$$

hold for all singular values λ_i of $\mathbf{B}^\natural$, $1 \le i \le \operatorname{rank}(\mathbf{B}^\natural)$.

Remark 2. *When the conditions in (13) is violated, we have the expectation $\mathbb{E}e^{-\theta\Xi}$ diverge to infinity, which suggests the optimal θ_* for $\inf_{\theta>0} {}^{\log\left(n\cdot\mathbb{E}e^{-\theta\Xi}\right)}/_{\theta}$ cannot be achieved.*

With (12), we can compute the optimal θ_* by setting the gradient $\frac{\partial\left[\log(n\cdot\mathbb{E}e^{-\theta\Xi})/\theta\right]}{\partial\theta} = 0$. However, a closed-form of the exact solution for θ^* is out of reach. As a mitigation, we resort to approximating $\log \mathbb{E}e^{-\theta\Xi}$ by its lower-bound, which reads as

$$\log \mathbb{E}e^{-\theta\Xi} \ge \frac{\theta^2}{2} \left[\left\| \mathbf{B}^{\natural\top}\mathbf{B}^\natural \right\|_{\mathrm{F}}^2 + 2\sigma^2 \left\| \mathbf{B}^\natural \right\|_{\mathrm{F}}^2 \right] - \theta \left\| \mathbf{B}^\natural \right\|_{\mathrm{F}}^2.$$

The corresponding minimum value $\widetilde{\theta}_*$ is thus obtained by minimizing the lower-bound, which is written as $\widetilde{\theta}_* = 2\log n / \left(\left\| \mathbf{B}^{\natural\top}\mathbf{B}^\natural \right\|_{\mathrm{F}}^2 + 2\sigma^2 \left\| \mathbf{B}^\natural \right\|_{\mathrm{F}}^2 \right)$.

Stage II: Compute the phase transition snr**.** We predict the phase transition point $\mathsf{snr}_{\text{oracle}}$ by letting the lower bound being zero, which can be written as

$$\frac{\log n}{\theta^*} - \left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2 + \frac{\theta^*}{2}\left(\left\|\mathbf{B}^{\natural\top}\mathbf{B}^\natural\right\|_{\mathrm{F}}^2 + 2\sigma^2\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2\right) = 0.$$

With standard algebraic manipulations, we obtain the equation

$$2(\log n)\mathsf{snr}_{\text{oracle}} \cdot \left\|\mathbf{B}^{\natural\top}/\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}} \cdot \mathbf{B}^\natural/\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}\right\|_{\mathrm{F}}^2 + {}^{4\log n}/_{m} = \mathsf{snr}_{\text{oracle}}. \tag{14}$$

To evaluate the accuracy of our predicted phase transition threshold, we compare the predicted values with the numerical values. The results are shown in Table 1, from which we can conclude the phase transition threshold snr can be predicted to a good extent. In addition, we observe that the gap between the theoretical values and the numerical values keeps shrinking as m increases.

Table 1: Comparison between the predicted value of the phase transition threshold $\mathsf{snr}_{\text{oracle}}$ and its simulated value when $n = 500$. **P** denotes the predicted value while **S** denotes the simulated value (i.e., mean $\pm$ std). **S** corresponds to the snr when the error rate drops below 0.05. A detailed description of the numerical method can be found in the appendix (code also included).

m	20	30	40	50	60	70
P	3.283	1.415	0.902	0.662	0.523	0.432
S	2.529 ± 0.079	1.290 ± 0.054	0.872 ± 0.034	0.649 ± 0.012	0.515 ± 0.016	0.429 ± 0.015
m	100	110	120	130	140	150
P	0.284	0.255	0.231	0.211	0.195	0.181
S	0.282 ± 0.008	0.256 ± 0.006	0.232 ± 0.006	0.212 ± 0.004	0.196 ± 0.006	0.183 ± 0.005

3.2 Gaussian approximation of the phase transition threshold

From the above analysis, we can see that deriving a closed-form expression of the infimum value θ of ${}^{\log(n\mathbb{E}e^{-\theta\Xi})}/_{\theta}$ can be difficult. In fact, in certain scenarios, even obtaining a closed-form expression of $\mathbb{E}e^{-\theta\Xi}$ is difficult. To handle such challenge, we propose to approximate random variable Ξ by a Gaussian $\mathsf{N}(\mathbb{E}\Xi, \mathrm{Var}\Xi)$, namely,

$$\mathbb{E}e^{-\theta\Xi} \approx \exp\left(-\theta\mathbb{E}\Xi + \frac{\theta^2}{2}\mathrm{Var}\Xi\right). \tag{15}$$

With this approximation, we can express $\theta_* \triangleq \inf {}^{\log\left(n\cdot\mathbb{E}e^{-\theta\Xi}\right)}/_{\theta}$ in a closed form, which is $\sqrt{{}^{2\log n}/_{\mathrm{Var}\Xi}}$. Thus, the critical point corresponding to the phase transition in (11) is written as

$$2(\log n)\cdot\mathrm{Var}\Xi = \left(\mathbb{E}\Xi\right)^2. \tag{16}$$

Comparison with (14). To verify that this approximation can yield meaningful results, we revisit the oracle case and have

$$\mathbb{E}\Xi = \left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2, \quad \mathrm{Var}\Xi = 3\left\|\mathbf{B}^\natural\mathbf{B}^{\natural\top}\right\|_{\mathrm{F}}^2 + 2\sigma^2\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2. \tag{17}$$

Plugging (17) into (16) then yields the relation

$$6(\log n)\mathsf{snr}_{\text{oracle}} \cdot \left\|\mathbf{B}^{\natural\top}/\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}} \cdot \mathbf{B}^\natural/\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}\right\|_{\mathrm{F}}^2 + {}^{4\log n}/_{m} = \mathsf{snr}_{\text{oracle}}, \tag{18}$$

from which we can determine the critical point of snr.

Example 1 (Scaled identity matrix)**.** *We consider the scenario where* $\mathbf{B}^\natural = \lambda\mathbf{I}_{m\times m}$*. Then, we have* ${}^{\mathbf{B}^\natural}/\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}} = m^{-1/2}\mathbf{I}$*. The phase transition threshold* $\mathsf{snr}_{\text{oracle}}$ *in* (14) *is then* ${}^{4\log n}/_{(m-2\log n)}$*, and the phase transition threshold* $\widetilde{\mathsf{snr}}_{\text{oracle}}$ *in* (18) *as* ${}^{4\log n}/_{(m-6\log n)}$*. This solution is almost identical to* (14) *in the limit as* $\mathsf{snr}_{\text{oracle}} \approx \widetilde{\mathsf{snr}}_{\text{oracle}} \approx {}^{4\log n}/_{m} \simeq n^{\frac{4}{m}} - 1$.

Moreover, we should mention that 1) our approximation method applies to a general matrix $\mathbf{B}^\natural$, not limited to a scaled identity matrix; and 2) our approximation method can also predict the phase transition thresholds to a good extent when the entries $\mathbf{X}_{ij}$ are sub-Gaussian. The numerical experiments are given in Table 2, from which we can conclude that the predicted values are well aligned with the simulation results. The details of the numerical method can be found in the appendix.

Table 2: Comparison between the predicted value of the phase transition threshold $\widetilde{\mathsf{snr}}_{\mathsf{oracle}}$ and its simulated value when $n = 600$. In (**Case 1**), half of singular values are with λ and the other half are with $\lambda/2$; while in (**Case 2**), half of the singular values are with λ and the other half are with $(3\cdot\lambda)/4$. **Gauss** refers to $\mathbf{X}_{ij} \overset{\text{i.i.d}}{\sim} \mathsf{N}(0,1)$ and **Unif** refers to $\mathbf{X}_{ij} \overset{\text{i.i.d}}{\sim} \text{Unif}[-1,1]$. **P** denotes the predicted value and **S** denotes the simulated value (i.e., mean $\pm$ std). **S** corresponds to the snr when the error rate drops below 0.05. Averaged over 20 repetitions.

m		100	110	120	130	140	150
(**Case 1**)	**P**	0.297	0.266	0.241	0.220	0.203	0.188
(**Gauss**)	**S**	0.307 ± 0.009	0.275 ± 0.005	0.246 ± 0.006	0.227 ± 0.007	0.210 ± 0.005	0.194 ± 0.004
(**Unif**)	**S**	0.294 ± 0.008	0.266 ± 0.005	0.239 ± 0.008	0.216 ± 0.004	0.201 ± 0.005	0.189 ± 0.006
(**Case 2**)	**P**	0.310	0.276	0.249	0.227	0.209	0.193
(**Gauss**)	**S**	0.294 ± 0.008	0.266 ± 0.006	0.241 ± 0.005	0.220 ± 0.004	0.204 ± 0.006	0.190 ± 0.003
(**Unif**)	**S**	0.287 ± 0.007	$0.255 \pm .0043$	0.234 ± 0.007	0.213 ± 0.005	0.197 ± 0.003	0.185 ± 0.005

4 Extension to Non-Oracle Case

Having analyzed the oracle case in the previous section, we now extend the analysis to the non-oracle case, where the value of $\mathbf{B}^\natural$ is not given. Different from the oracle case, the ML estimator reduces to a *quadratic assignment problem* (QAP) as opposed to LAP. As a mitigation, we adopt the estimator in Zhang & Li (2020), which reconstructs the permutation matrix within the LAP framework, i.e.,

$$\widehat{\mathbf{\Pi}}^{\mathsf{oracle}} = \text{argmin}_{\mathbf{\Pi}} \left\langle \mathbf{\Pi}, -\mathbf{Y}\mathbf{Y}^\top \mathbf{X}\mathbf{X}^\top \right\rangle, \quad \text{s.t.} \quad \sum_i \mathbf{\Pi}_{ij} = 1, \sum_j \mathbf{\Pi}_{ij} = 1, \mathbf{\Pi} \in \{0,1\}^{n\times n}. \tag{19}$$

We expect this estimator can yield good insights of the permuted linear regression since

- this estimator can reach the statistical optimality in a broad range of parameters;
- this estimator exhibits a phase transition phenomenon, which follows a similar pattern to that in the oracle case.

Following the same procedure as in Section 3, we identify the phase transition threshold snr with Theorem 1. To begin with, we write the random variable Ξ as

$$\Xi \cong \mathbf{Y}_i \mathbf{Y}^\top \mathbf{X} \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j \right)^\top,$$

where i and j are uniformly distributed among the set $\{1, 2, \cdots, n\}$. Afterwards, we adopt the Gaussian approximation scheme illustrated in Subsection 3.2 and determine the phase transition points by first computing $\mathbb{E}\Xi$ and $\text{Var}\Xi$, respectively.

Theorem 2. *For the random variable Ξ defined in (21), its mean $\mathbb{E}\Xi$ and variance $\text{Var}\Xi$ are*

$$\begin{aligned}
\mathbb{E}\Xi \simeq\ & n(1-\tau_h)\left[(1+\tau_p)\left|\!\left|\!\left|\mathbf{B}^\natural\right|\!\right|\!\right|_{\mathrm{F}}^2 + m\tau_p\sigma^2\right], \\
\text{Var}\Xi \simeq\ & n^2\tau_h(1-\tau_h)\tau_p^2\left[\left|\!\left|\!\left|\mathbf{B}^\natural\right|\!\right|\!\right|_{\mathrm{F}}^2 + m\sigma^2\right]^2 + n^2\left[2\tau_p + 3(1-\tau_h)^2\right]\left|\!\left|\!\left|\mathbf{B}^{\natural\top}\mathbf{B}^\natural\right|\!\right|\!\right|_{\mathrm{F}}^2 \\
& + n^2\left[6\tau_p(1-\tau_h)^2 + (3-\tau_h)\tau_p^2\right]\left|\!\left|\!\left|\mathbf{B}^{\natural\top}\mathbf{B}^\natural\right|\!\right|\!\right|_{\mathrm{F}}^2,
\end{aligned}$$

respectively, where the definitions of τ_p and τ_h can be found in Section 3.

The proof of Theorem 2 is quite complicated, involving Wick's theorem, Stein's lemma, the conditional technique, and the leave-one-out technique, etc. We defer the technical details to Appendix.

4.1 An Illustrating Example

Afterwards, we predict the phase transition points. Unlike the oracle case, we notice the edge weights $\mathbf{E}_{ij}$ are strongly correlated, especially when $j = \pi^\natural(j)$, which corresponds to the non-permuted rows. To factor out these dependencies, we only take the permuted rows into account and correct the sample

size from n to $\tau_h n$. The prediction $\mathsf{snr}_{\text{non-oracle}}$ is then computed by solving $2\log(n\tau_h)\text{Var}\Xi = (\mathbb{E}\Xi)^2$, where $\mathbb{E}\Xi$ and $\text{Var}\Xi$ are in Theorem 2.

To illustrate the prediction accuracy, we consider the case where $\mathbf{B}^\natural$'s singular values are of the same order, i.e., $\frac{\lambda_i(\mathbf{B}^\natural)}{\lambda_j(\mathbf{B}^\natural)} = O(1),\ 1 \le i, j \le m$, where $\lambda_i(\cdot)$ denotes the i-th singular value. Then, we obtain the $\mathsf{snr}_{\text{non-oracle}}$, which is written as

$$\mathsf{snr}_{\text{non-oracle}} \approx \eta_1/\eta_2. \tag{20}$$

Here, η_1 and η_2 are defined as

$$\begin{aligned}\eta_1 &\triangleq 2\tau_h\tau_p^2\log(n\tau_h) - \tau_p(\tau_p+1)(1-\tau_h) + \tau_p\sqrt{2(1-\tau_h)\tau_h\cdot\log(n\tau_h)},\\ \eta_2 &\triangleq (1-\tau_h)(\tau_p+1)^2 - 2\tau_h\tau_p^2\log(n\tau_h).\end{aligned}$$

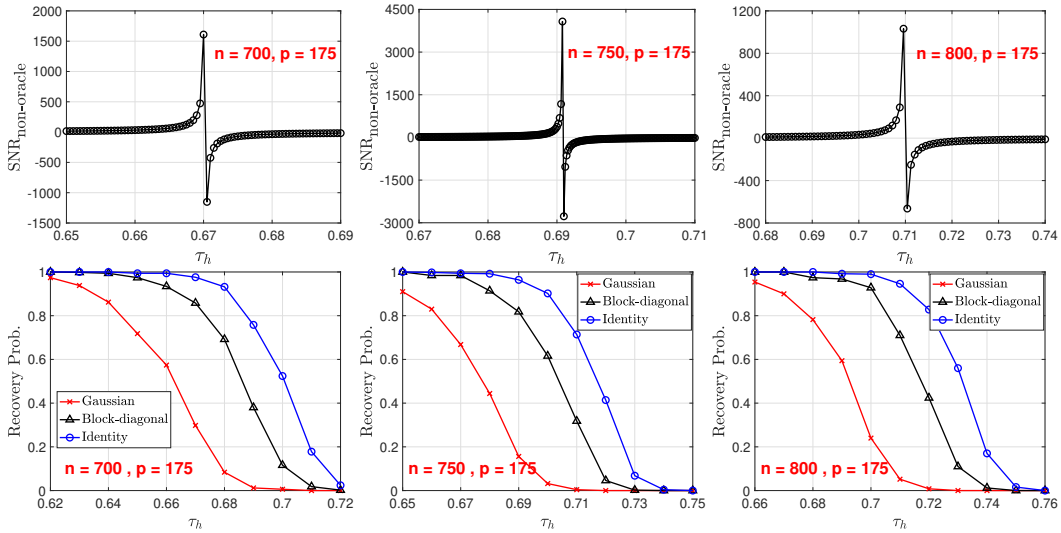

Figure 2: **Upper panel**: Predicted phase transition points $\mathsf{snr}_{\text{non-oralce}}$. **Lower panel**: Plot of the recovery rate under the noiseless setting, i.e., $\mathsf{snr} = \infty$. **Gaussian**: $\mathbf{B}^\natural_{ij} \overset{\text{i.i.d}}{\sim} \mathsf{N}(0,1)$; **Identity**: $\mathbf{B}^\natural = \mathbf{I}_{p\times p}$; **Block-diagonal**: $\mathbf{B}^\natural = \text{diag}\{1,\cdots,1,0.5,\cdots,0.5\}$. We observe that the correct recovery rates drop sharply within the regions of our predicted value.

Notice that the predicted $\mathsf{snr}_{\text{non-oracle}}$ varies for different τ_h and τ_p. Viewing $\mathsf{snr}_{\text{non-oracle}}$ as a function of τ_h, we observe a singularity point of τ_h, which corresponds to the case when $\eta_2 = 0$. This suggests a potential phase transition phenomenon w.r.t. τ_h. To validate the predicted phenomenon, we consider the noiseless case, i.e., $\mathsf{snr} = \infty$, and reconstruct the permutation matrix $\mathbf{\Pi}^\natural$ with (2). Numerical experiments in Figure 2 confirm our prediction.

Due to the space limit, this section only presents a glimpse of our results in the non-oracle case. The technical details and additional numerical experiments can be found in the supplementary material.

5 Conclusion

This is the first work that can identify the precise location of phase transition thresholds of permuted linear regressions. For the oracle case where the signal $\mathbf{B}^\natural$ is given as a prior, our analysis can predict the phase transition threshold $\mathsf{snr}_{\text{oracle}}$ to a good extent. For the non-oracle case where $\mathbf{B}^\natural$ is not given, we modified the leave-one-out technique to approximately compute the phase critical $\mathsf{snr}_{\text{non-oracle}}$ value for the phase transition, as the precise computation becomes significantly complicated as the high-order interaction between Gaussian random variables is involved. Moreover, we associated the singularity point in $\mathsf{snr}_{\text{non-oracle}}$ with a phase transition point w.r.t the maximum allowed number of permuted rows. Moreover, we present numerous numerical experiments to corroborate the accuracy of our theoretical predictions.

REFERENCES

Zhidong Bai and Tailen Hsing. The broken sample problem. *Probability Theory and Related Fields*, 131(4):528–552, 2005.

Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

Mohsen Bayati and Andrea Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. Inf. Theory*, 57(2):764–785, 2011.

George T. Cantwell and M. E. J. Newman. Message passing on networks with loops. *Proceedings of the National Academy of Sciences*, 116(47):23398–23403, 2019.

Sergio Caracciolo, Matteo P D'Achille, Enrico M Malatesta, and Gabriele Sicuro. Finite-size corrections in the random assignment problem. *Physical Review E*, 95(5):052129, 2017.

Michael Chertkov, Lukas Kroc, F Krzakala, M Vergassola, and L Zdeborová. Inference in particle tracking experiments by passing messages between images. *Proceedings of the National Academy of Sciences*, 107(17):7663–7668, 2010.

Sae-Young Chung. *On the construction of some capacity-approaching coding schemes*. PhD thesis, Massachusetts Institute of Technology, 2000.

Morris H. DeGroot and Prem K. Goel. The matching problem for multivariate normal data. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 38(1):14–29, 1976.

Morris H. DeGroot and Prem K. Goel. Estimation of the correlation coefficient from a broken random sample. *The Annals of Statistics*, 8(2):264–278, 03 1980.

Morris H. DeGroot, Paul I. Feder, and Prem K. Goel. Matchmaking. *The Annals of Mathematical Statistics*, 42(2):578–593, 04 1971.

David L Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

Prem K. Goel. On re-pairing observations in a broken random sample. *The Annals of Statistics*, 3(6): 1364–1369, 11 1975.

John M Hammersley. Postulates for subadditive processes. *The Annals of Probability*, 2(4):652–680, 1974.

Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

Svante Janson. *Gaussian Hilbert Spaces*. Cambridge Tracts in Mathematics. Cambridge University Press, 1997. doi: 10.1017/CBO9780511526169.

Noureddine El Karoui. Asymptotic behavior of unregularized and ridge-regularized high-dimensional robust regression estimators: rigorous results. *arXiv preprint arXiv:1311.2445*, 2013.

Noureddine EL Karoui. On the impact of predictor geometry on the performance on high-dimensional ridge-regularized generalized robust regression estimators. *Probability Theory and Related Fields*, 170(1-2):95–175, 2018.

John Frank Charles Kingman. The first birth problem for an age-dependent branching process. *The Annals of Probability*, pp. 790–801, 1975.

Alec Kirkley, George T. Cantwell, and M. E. J. Newman. Belief propagation for networks with loops. *Science Advances*, 7(17):eabf1211, 2021. doi: 10.1126/sciadv.abf1211.

Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.

Svante Linusson and Johan Wästlund. A proof of parisi's conjecture on the random assignment problem. *Probability Theory and Related Fields*, 128(3):419–440, 2004.

David JC MacKay, David JC Mac Kay, et al. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

Enrico M Malatesta, Giorgio Parisi, and Gabriele Sicuro. Fluctuations in the random-link matching problem. *Physical Review E*, 100(3):032102, 2019.

Arian Maleki. *Approximate message passing algorithms for compressed sensing*. PhD thesis, Stanford University, 2010.

OC Martin, M Mézard, and Olivier Rivoire. Random multi-index matching problems. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(09):P09006, 2005.

Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

Marc Mézard and Giorgio Parisi. Replicas and optimization. *Journal de Physique Lettres*, 46(17): 771–778, 1985.

Marc Mézard and Giorgio Parisi. Mean-field equations for the matching and the travelling salesman problems. *EPL (Europhysics Letters)*, 2(12):913, 1986.

Marc Mézard and Giorgio Parisi. On the solution of the random link matching problems. *Journal de Physique*, 48(9):1451–1459, 1987.

Heinz Neudecker and Tom Wansbeek. Fourth-order properties of normally distributed random matrices. *Linear Algebra and Its Applications*, 97:13–21, 1987.

Ashwin Pananjady, Martin J Wainwright, and Thomas A Courtade. Denoising linear models with permuted data. In *Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT)*, pp. 446–450, Aachen, Germany, 2017.

Ashwin Pananjady, Martin J. Wainwright, and Thomas A. Courtade. Linear regression with shuffled data: Statistical and computational limits of permutation recovery. *IEEE Trans. Inf. Theory*, 64(5): 3286–3300, 2018.

Giorgio Parisi and Matthieu Ratiéville. On the finite size corrections to some random matching problems. *The European Physical Journal B-Condensed Matter and Complex Systems*, 29(3): 457–468, 2002.

Sundeep Rangan. Generalized approximate message passing for estimation with random linear mixing. In *Proceedings of the 2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 2168–2172, St. Petersburg, Russia, 2011.

Thomas J. Richardson and Rüdiger L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Inf. Theory*, 47(2):599–618, 2001.

Tom Richardson and Ruediger Urbanke. *Modern coding theory*. Cambridge university press, 2008.

Guilhem Semerjian, Gabriele Sicuro, and Lenka Zdeborová. Recovery thresholds in the sparse planted matching problem. *Physical Review E*, 102(2):022304, 2020.

Martin Slawski and Emanuel Ben-David. Linear regression with sparsely permuted data. *Electronic Journal of Statistics*, 1:1–36, 2019.

Martin Slawski, Emanuel Ben-David, and Ping Li. Two-stage approach to multivariate linear regression with sparsely mismatched data. *J. Mach. Learn. Res.*, 21(204):1–42, 2020.

Pragya Sur, Yuxin Chen, and Emmanuel J Candès. The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square. *Probability Theory and Related Fields*, 175(1):487–558, 2019.

Michel Talagrand. *Mean field models for spin glasses: Volume I: Basic examples*, volume 54. Springer Science & Business Media, 2010.

Jayakrishnan Unnikrishnan, Saeid Haghighatshoar, and Martin Vetterli. Unlabeled sensing: Solving a linear system with unordered measurements. In *Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 786–793, Monticello, IL, 2015.

Hang Zhang and Ping Li. Optimal estimator for unlabeled linear regression. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 11153–11162, Virtual Event, 2020.

Hang Zhang, Martin Slawski, and Ping Li. The benefits of diversity: Permutation recovery in unlabeled sensing from multiple measurement vectors. *IEEE Trans. Inf. Theory*, 68(4):2509–2529, 2022.

A COMMENTS ON (7) AND (8)

This paper considers the phase transition phenomenon w.r.t. the full permutation recovery. Informally speaking, this can be partly deduced from (7) and (8).

Here, we regard message flows $h_{i^{\mathsf{L}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}$ and $h_{j^{\mathsf{R}}\to(i^{\mathsf{L}},j^{\mathsf{R}})}$ i.i.d. samples from certain distributions (represented by the random variable H). When studying the evolution behavior of the random variable H, we track the behaviors of all message flows. Hence, if we find an arbitrary sample H that will yield the correct recovery, we can say that the correspondence between all pairs is correct. On the other hand, we can say that there exist some pairs with wrong correspondence if H leads to incorrect recovery. This can explain why the phase transition phenomenon exists.

B ANALYSIS OF ORACLE CASE: PROOF OF PROPOSITION 1

Proof. Denote the singular values of $\mathbf{B}^{\natural}$ as $\{\lambda_i\}_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})}$. We exploit the rotation invariance property of Gaussian random variables; and have $\boldsymbol{\Xi}$ be identically distributed as

$$\boldsymbol{\Xi} = \sum_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \lambda_i^2 x_i \left(x_i - y_i\right) + \sigma \sum_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \lambda_i w_i \left(x_i - y_i\right).$$

Due to the independence across $\boldsymbol{w}$, $\boldsymbol{x}$, and $\boldsymbol{y}$, we have

$$\begin{aligned}
\mathbb{E}e^{-\theta\Xi} &= \prod_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \mathbb{E}_{x,y,w} \exp\left[-\theta\lambda_i^2 x\left(x-y\right) - \theta\sigma\lambda_i w\left(x-y\right)\right] \\
&= \prod_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \mathbb{E}_{x,y} \exp\left(\frac{\theta\lambda_i^2(x-y)\left(\theta\sigma^2(x-y)-2x\right)}{2}\right) \\
&\overset{①}{=} \prod_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \mathbb{E}_{x} \frac{\exp\left(\frac{\theta\lambda_i^2 x^2\left(\theta\left(\lambda_i^2+\sigma^2\right)-2\right)}{2-2\theta^2\lambda_i^2\sigma^2}\right)}{\sqrt{1-\theta^2\lambda_i^2\sigma^2}} \\
&\overset{②}{=} \prod_{i=1}^{\mathrm{rank}(\mathbf{B}^{\natural})} \left(1+2\theta\lambda_i^2 - \theta^2\lambda_i^2\left(\lambda_i^2+2\sigma^2\right)\right)^{-\frac{1}{2}},
\end{aligned}$$

where in ① we use the fact $\theta^2\sigma^2\lambda_i^2 < 1$ and in ② we use the fact $\theta^2\lambda_i^2\left(\lambda_i^2+2\sigma^2\right) \le 1+2\theta\lambda_i^2$. □

C NUMERICAL METHODS FOR THE PHASE TRANSITION POINTS

In this section, we present the numerical method to compute the phase transition points. Notice that the correct recovery rate is in monotonic non-decreasing relation with the snr, we adopt a binary-search-based method.

First, we fix the snr and run the experiments for 100 times. Then, we calculate the error rate of permutation recovery (full permutation recovery). If the error rate is below 0.05, we regard the corresponding snr as above the phase transition point and try a smaller value. Otherwise, we regard the snr as below the phase transition point and try a larger value. The detailed description is given in Algorithm 1.

In our numerical experiments, we run 20 times of Algorithm 1 for each parameter setting. Then, we estimate its mean and the standard deviation from these estimated phase transition points.

Algorithm 1: Numerical method to compute the phase transition points.

1: **Initialization.** Set the initial search range for snr as $[l, r]$. Define the precision threshold ε.
2:
3: **while** $|l - r| > \varepsilon$ **do**
4: Set $\mathsf{snr}_{\text{middle}} = \frac{l+r}{2}$.
5: Run experiments 100 times for this $\mathsf{snr}_{\text{middle}}$.
6: Compute the error rate of full permutation recovery.
7:
8: **if** the error rate is below 0.05 **then**
9: $\mathsf{snr}_{\text{middle}} \to r - \varepsilon$, # we have $\mathsf{snr}_{\text{middle}}$ be greater than the phase transition point
10: **else**
11: $\mathsf{snr}_{\text{middle}} \to l + \varepsilon$. # we have $\mathsf{snr}_{\text{middle}}$ be no greater than the phase transition point
12: **end if**
13:
14: **end while**
15:
16: **Output.** Return the phase transition point $\mathsf{snr}_{\text{middle}}$.

Complexity analysis. For a given precision threshold ε, each iteration (Line 3 to Line 14 in Algorithm 1) takes $O(\log \frac{1}{\varepsilon})$ rounds to converge and it runs the permutation recovery algorithm 100 times in each round. Hence, we run $20 \times 100 \times \log_2(10^4)$ permutation recovery experiments for each parameter.

D ANALYSIS OF THE NON-ORACLE CASE

This section presents the technical details in analyzing the non-oracle case.

D.1 ADDITIONAL NUMERICAL RESULTS

We consider the same settings as in Subsection 4.1. Here, we present additional numerical results to evaluate the prediction accuracy of our method.

D.1.1 VERIFICATION OF PHASE TRANSITION POINTS

For the predicted phase transition $\mathsf{snr}_{\text{non-oracle}}$, we notice an increasing gap between the predicted value and the simulated value, unlike in the oracle case. This might be caused by the strong correlation across the edge weights $\{\mathbf{E}_{ij}\}_{1\le i,j\le n}$, or due to the error with the approximation relation $\mathbb{E}e^{-\theta\Xi} \approx \mathbb{E}\exp\left(\theta\mathbb{E}\Xi - {}^{\theta^2\mathrm{Var}\Xi}/_{2}\right)$.

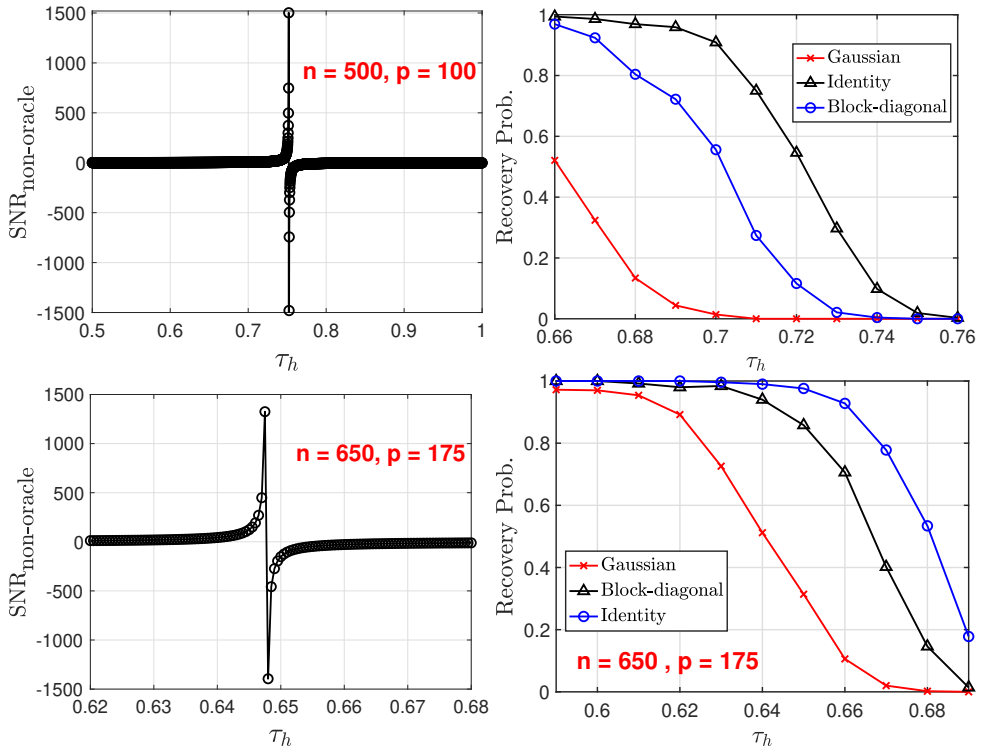

Figure 3: In the top row, **Left panel**: Predicted $\mathsf{snr}_{\text{non-oralce}}$. **Right panel**: Plot of recovery rate under the noiseless setting, i.e., $\mathsf{snr} = \infty$. **Gaussian**: $\mathbf{B}^{\natural}_{ij} \overset{\text{i.i.d}}{\sim} \mathsf{N}(0,1)$; **Identity**: $\mathbf{B}^{\natural} = \mathbf{I}_{p\times p}$; **Block-diagonal**: $\mathbf{B}^{\natural} = \mathrm{diag}\,\{1,\cdots,1,0.5,\cdots,0.5\}$.

Table 3: Comparison between the predicted value of the phase transition threshold τ_h and its simulated value when $n = 500$. **P** denotes the **predicted value** while **S** denotes the **simulated value**. **S** corresponds to the τ_h when the error rate drops below 0.05.

p	75	100	125	150	175	200
P	0.82	0.73	0.68	0.62	0.56	0.52
S	0.77	0.75	0.7	0.66	0.61	0.57

Additional experiments are availabe in Table 3, from which we conclude the solution (20) can predict the phase transition point w.r.t. τ_h to a good extent.

D.1.2 Impact of n on the phase transition point

We study the impact of n on τ_h. The numerical experiment is shown in the top row of Figure 3, from which we can see the predicted phase transition τ_h matches to a good extent to the numerical experiments. Then, we fix the p and study the impact of n on τ_h. We observe that the phase transition τ_h increases together with the sample number n, which is also captured by our formula in (20).

D.1.3 Limits of τ_h

We consider the limiting behavior of τ_h when τ_p approaches 0, or equivalently, $p = o_{\mathrm{P}}(n)$. We can simplify $\mathbb{E}\Xi$ and VarΞ in Theorem 2 as

$$\begin{aligned}\mathbb{E}\Xi &\simeq n(1-\tau_h)\left\|\mathbf{B}^{\natural}\right\|_{\mathrm{F}}^2,\\ \mathrm{Var}\Xi &\simeq 3n^2(1-\tau_h)^2\left\|\mathbf{B}^{\natural\top}\mathbf{B}^{\natural}\right\|_{\mathrm{F}}^2.\end{aligned}$$

We notice that the singularity point in (20) disappears. In other words, we can have the correct permutation matrix $\mathbf{\Pi}^{\natural}$ even when $h \approx n$. This is (partly) verified by Figure 4, from which we observe that the phase transition point w.r.t. τ_h approaches to one, or equivalently, h approaches n, as τ_p decreases to zero.

D.2 Analysis of Non-Oracle Case: Proof of Theorem 2

This subsection presents the computational details of Theorem 2. To begin with, we decompose the random variable Ξ as

$$\Xi = \Xi_1 + \sigma(\Xi_2 + \Xi_3) + \sigma^2\Xi_4, \tag{21}$$

where Ξ_i $(1 \le i \le 4)$ are respectively defined as

$$\begin{aligned}\Xi_1 &\triangleq \mathbf{X}_{\pi^{\natural}(i)}^{\top}\mathbf{B}^{\natural}\mathbf{B}^{\natural\top}\mathbf{X}^{\top}\mathbf{\Pi}^{\natural\top}\mathbf{X}(\mathbf{X}_{\pi^{\natural}(i)} - \mathbf{X}_j),\\ \Xi_2 &\triangleq \mathbf{X}_{\pi^{\natural}(i)}^{\top}\mathbf{B}^{\natural}\mathbf{W}^{\top}\mathbf{X}(\mathbf{X}_{\pi^{\natural}(i)} - \mathbf{X}_j),\\ \Xi_3 &\triangleq \mathbf{W}_i^{\top}\mathbf{B}^{\natural\top}\mathbf{X}^{\top}\mathbf{\Pi}^{\natural\top}\mathbf{X}(\mathbf{X}_{\pi^{\natural}(i)} - \mathbf{X}_j),\\ \Xi_4 &\triangleq \mathbf{W}_i^{\top}\mathbf{W}^{\top}\mathbf{X}(\mathbf{X}_{\pi^{\natural}(i)} - \mathbf{X}_j).\end{aligned}$$

Unlike the oracle case, obtaining a closed-form expression of $\mathbb{E}e^{-\theta\Xi}$ would be too difficult. Hence, we adopt the Gaussian approximation method as presented in Section 3.2. The task then transforms to computing the expectation and variance of Ξ. Before delving into the technical details, we give a glimpse of our proof strategy.

Computation of the mean $\mathbb{E}\Xi$. For the computation of the mean $\mathbb{E}\Xi$, we can verify that $\mathbb{E}\Xi_2$ and $\mathbb{E}\Xi_3$ are both zero, due to the independence between $\mathbf{X}$ and $\mathbf{W}$. For $\mathbb{E}\Xi_1$ and $\mathbb{E}\Xi_4$, we adopt Wick's theorem to obtain

$$\begin{aligned}\mathbb{E}\Xi_1 &= n(1-\tau_h)(1+\tau_p)\left[1+o_{\mathrm{P}}(1)\right]\left\|\mathbf{B}^{\natural}\right\|_{\mathrm{F}}^2,\\ \mathbb{E}\Xi_4 &= nm\tau_p(1-\tau_h)\left[1+o_{\mathrm{P}}(1)\right].\end{aligned}$$

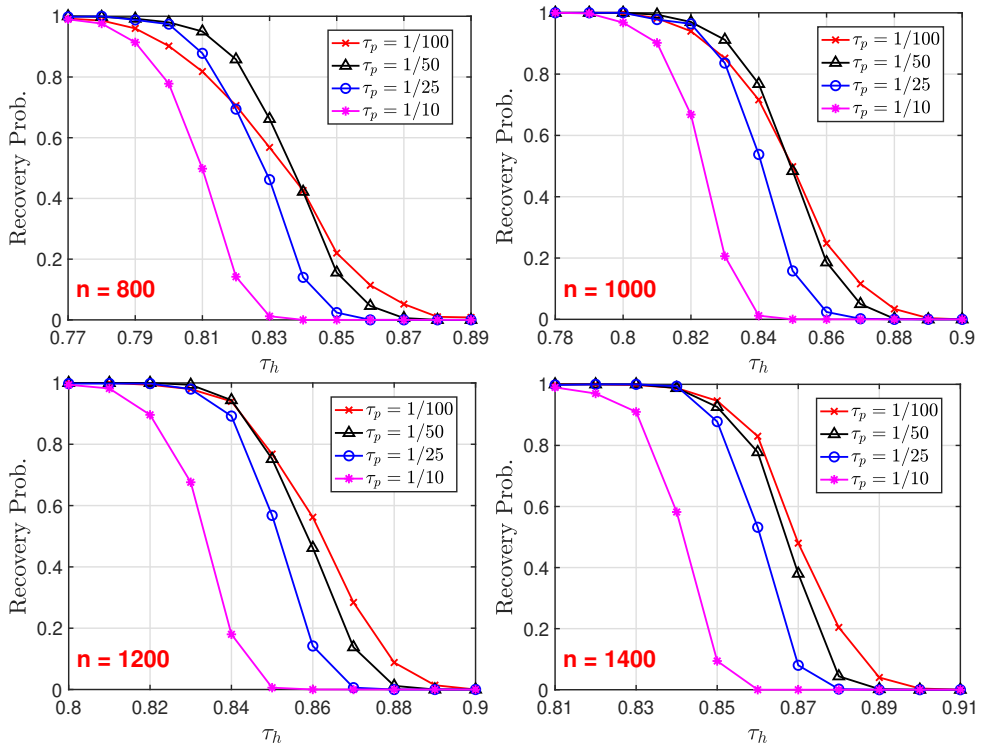

Figure 4: Plot of correct recovery rate w.r.t. τ_h. We consider the noiseless scenario (i.e., $\mathsf{snr} = \infty$) and pick $n = \{800, 1000, 1200, 1400\}$.

Computation of the variance VarΞ. Since $\mathrm{Var}\Xi = \mathbb{E}\Xi^2 - (\mathbb{E}\Xi)^2$, we just need to compute $\mathbb{E}\Xi^2$, which can be expanded into the following six terms

$$\mathbb{E}\Xi^2 = \mathbb{E}\Xi_1^2 + \sigma^2\mathbb{E}\Xi_2^2 + \sigma^2\mathbb{E}\Xi_3^2 + \sigma^4\mathbb{E}\Xi_4^2 + 2\sigma^2\mathbb{E}\Xi_1\Xi_4 + 2\sigma^2\mathbb{E}\Xi_2\Xi_3.$$

The computation of above terms turns out to be quite complex due to the high order Gaussian random variables. For example, the term $\mathbb{E}\Xi_1^2$ involves the eighth-order Gaussian moments, the terms $\mathbb{E}\Xi_2^2, \mathbb{E}\Xi_3^2, \mathbb{E}\Xi_1\Xi_4$ and $\mathbb{E}\Xi_2\Xi_3$ all involve the sixth-order Gaussian variables, etc. To handle the difficulties in computing $\mathbb{E}\Xi^2$, we propose the following computation procedure, which can be roughly divided into 3 phases.

- **Phase I: Leave-one-out decomposition.** The major technical difficulty comes from the correlation between the product $\mathbf{X}^\top\mathbf{\Pi}^\natural\mathbf{X}$ and the difference $\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j$. We decouple this correlation by first rewriting the matrix $\mathbf{X}^\top\mathbf{\Pi}^\natural\mathbf{X}$ as the sum $\sum_\ell \mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top$. Then we collect all terms $\mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top$ independent of $\mathbf{X}_{\pi^\natural(i)}$ and $\mathbf{X}_j$ in the matrix $\mathbf{\Sigma}$ and leave the remaining terms to the matrix $\mathbf{\Delta}$, i.e., $\mathbf{\Delta} \triangleq \mathbf{X}^\top\mathbf{\Pi}^\natural\mathbf{X} - \mathbf{\Sigma}$. This decomposition shares the same spirit as the leave-one-out technique (Karoui, 2013; Bai & Silverstein, 2010; Karoui, 2018; Sur et al., 2019). Then, we divide all terms in $\mathbb{E}\Xi^2$ into 3 categories: 1) terms only containing matrix $\mathbf{\Sigma}$; 2) terms containing both $\mathbf{\Sigma}$ and $\mathbf{\Delta}$; and 3) terms only containing $\mathbf{\Delta}$.
- **Phase II: Conditional technique.** Concerning the terms in the first two categories, which covers the majority of terms, we can exploit the independence among the rows in the sensing matrix $\mathbf{X}$. With the conditional technique, we can reduce the order of Gaussian moments by separately taking the expectation w.r.t $\mathbf{\Sigma}$ and w.r.t vectors $\mathbf{X}_{\pi^\natural(i)}$ and $\mathbf{X}_j$.
- **Phase III: Direct computation.** For the few terms in the third category (i.e., terms only containing $\mathbf{\Delta}$), we compute the high-order Gaussian moments by exhausting all terms and iterative applying of Wick's Theorem and Stein's Lemma, which can reduce the higher-order Gaussian moments to lower-orders.

The computation details are attached as follows.

D.2.1 NOTATIONS

Note that our analysis can involve the terms containing $(\mathbf{X}_{\pi^\natural}(i) - \mathbf{X}_j)$ and $\mathbf{X}^\top \mathbf{\Pi}^\natural \mathbf{X}$ simultaneously. To decouple the dependence between $(\mathbf{X}_{\pi^\natural}(i) - \mathbf{X}_j)$ and $\mathbf{X}^\top \mathbf{\Pi}^\natural \mathbf{X}$, we first rewrite the matrix $\mathbf{X}^\top \mathbf{\Pi}^\natural \mathbf{X}$ as the sum $\sum_\ell \mathbf{X}_\ell \mathbf{X}_{\pi^\natural(\ell)}^\top$ and then collect all terms $\mathbf{X}_\ell \mathbf{X}_{\pi^\natural(\ell)}^\top$ independent of $\mathbf{X}_{\pi^\natural(i)}$ and $\mathbf{X}_j$ in the matrix $\mathbf{\Sigma}$, which is written as

$$\mathbf{\Sigma} \triangleq \sum_{\ell, \pi^\natural(\ell) \neq \pi^\natural(i), j} \mathbf{X}_\ell \mathbf{X}_{\pi^\natural(\ell)}^\top. \tag{22}$$

The rest terms are then put in the matrix $\mathbf{\Delta}$ such that $\mathbf{X}^\top \mathbf{\Pi}^\natural \mathbf{X} = \mathbf{\Sigma} + \mathbf{\Delta}$. Note that the expression of $\mathbf{\Delta}$ varies under different cases such that

- **Case** (s, s)**:** $i = \pi^\natural(i)$ **and** $j = \pi^\natural(j)$. We have

$$\mathbf{\Delta} = \mathbf{\Delta}^{(s,s)} = \mathbf{X}_i \mathbf{X}_i^\top + \mathbf{X}_j \mathbf{X}_j^\top. \tag{23}$$

- **Case** (s, d)**:** $i = \pi^\natural(i)$ **and** $j \neq \pi^\natural(j)$. We have

$$\mathbf{\Delta} = \mathbf{\Delta}^{(s,d)} = \mathbf{X}_i \mathbf{X}_i^\top + \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top. \tag{24}$$

- **Case** (d, s)**:** $i \neq \pi^\natural(i)$ **and** $j = \pi^\natural(j)$. We have

$$\mathbf{\Delta} = \mathbf{\Delta}^{(d,s)} = \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top + \mathbf{X}_j \mathbf{X}_j^\top. \tag{25}$$

- **Case** (d, d)**:** $i \neq \pi^\natural(i)$ **and** $j \neq \pi^\natural(j)$. We have

$$\mathbf{\Delta} = \mathbf{\Delta}^{(d,d)} = \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top + \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top. \tag{26}$$

In addition, we define the matrix $\mathbf{M}$ as $\mathbf{B}^\natural \mathbf{B}^{\natural\top}$, and define the index sets $\mathcal{S}, \mathcal{D}$, and $\mathcal{D}_{\text{pair}}$ as

$$\mathcal{S} \triangleq \left\{ \ell \mid \ell \neq i \text{ or } j,\ \ell = \pi^\natural(\ell) \right\}, \tag{27}$$

$$\mathcal{D} \triangleq \left\{ \ell \mid \ell, \pi^\natural(\ell) \neq i \text{ or } j,\ \ell \neq \pi^\natural(\ell) \right\}, \tag{28}$$

$$\mathcal{D}_{\text{pair}} \triangleq \left\{ (\ell_1, \ell_2) : \ell_1 = \pi^\natural(\ell_2), \ell_2 = \pi^\natural(\ell 1), \ell_1, \ell_2 \in \mathcal{D} \right\}, \tag{29}$$

respectively.

D.2.2 MAIN COMPUTATION

In this case, we can write Ξ as

$$\begin{aligned}\Xi = \ & \underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathbf{B}^{\natural\top} \mathbf{X}^\top \mathbf{\Pi}^{\natural\top} \mathbf{X} \left[\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j \right]}_{\triangleq \Xi_1} + \sigma \underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathbf{W}^\top \mathbf{X} \left[\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j \right]}_{\triangleq \Xi_2} \\ & + \sigma \underbrace{\mathbf{W}_i^\top \mathbf{B}^{\natural\top} \mathbf{X}^\top \mathbf{\Pi}^{\natural\top} \mathbf{X} \left[\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j \right]}_{\triangleq \Xi_3} + \sigma^2 \underbrace{\mathbf{W}_i^\top \mathbf{W}^\top \mathbf{X} \left[\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j \right]}_{\triangleq \Xi_4}.\end{aligned}$$

The following context separately computes its expectation $\mathbb{E}\Xi$ and its variance $\text{Var}\Xi$.

Expectation. We can easily verify that both $\mathbb{E}\Xi_2$ and $\mathbb{E}\Xi_3$ are zero. Then our goal turns to calculating the expectation of $\mathbb{E}\Xi_1$ and $\mathbb{E}\Xi_4$. First, we have

$$\mathbb{E}\Xi_1 = \mathbb{E} \sum_{\ell = \pi^\natural(\ell)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_\ell \mathbf{X}_\ell^\top \mathbf{X}_{\pi^\natural(i)} - \mathbb{E} \sum_\ell \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(\ell)} \mathbf{X}_\ell^\top \mathbf{X}_j.$$

With Lemma 15 and Lemma 16, we conclude

$$\begin{aligned}\mathbb{E}\Xi_1 &= (n-h)\operatorname{Tr}(\mathbf{M}) + (p+1)\mathbb{E}\mathbb{1}_{i=\pi^\natural(i)} \operatorname{Tr}(\mathbf{M}) - \left(p\mathbb{E}\mathbb{1}_{i=j} + \mathbb{E}\mathbb{1}_{j=\pi^{\natural 2}(i)} \right) \operatorname{Tr}(\mathbf{M}) \\ &= (n + p - h - hp/n)\left[1 + o(1)\right] \operatorname{Tr}(\mathbf{M}).\end{aligned} \tag{30}$$

Meanwhile, we have

$$\begin{aligned}\mathbb{E}\Xi_4 = &\ \mathbb{E}\left[\mathbf{W}_i^\top\mathbf{W}_1 \ \cdots\ \mathbf{W}_i^\top\mathbf{W}_i \ \cdots \mathbf{W}_i^\top\mathbf{W}_n\right]\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\\ = &\ m\mathbb{E}\mathbf{X}_i^\top\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right) = mp\left(\mathbb{E}\mathbb{1}_{i=\pi^\natural(i)} - \mathbb{E}\mathbb{1}_{i=j}\right) = \frac{mp(n-h)\sigma^2}{n}\left[1+o(1)\right]. \quad (31)\end{aligned}$$

Combining (30) and (31) and neglecting the $o(1)$ terms yields

$$\mathbb{E}\Xi \approx (n+p)\left(1-h/n\right)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2 + \frac{mp(n-h)\sigma^2}{n}.$$

Variance. Then we study the variance of Ξ. With the relation $\mathrm{Var}(\Xi) = \mathbb{E}\Xi^2 - (\mathbb{E}\Xi)^2$, our goal reduces to computing $\mathbb{E}\Xi^2$, which can be written as

$$\mathbb{E}\Xi^2 = \mathbb{E}\Xi_1^2 + \sigma^2\mathbb{E}\Xi_2^2 + \sigma^2\mathbb{E}\Xi_3^2 + \sigma^4\mathbb{E}\Xi_4^2 + 2\sigma^2\mathbb{E}\Xi_1\Xi_4 + 2\sigma^2\mathbb{E}\Xi_2\Xi_3.$$

The following context separately computes each terms

$$\begin{aligned}\mathbb{E}\Xi_1^2 \approx &\ (n-h)^2\left(1+\frac{2p}{n}+\frac{p^2}{n(n-h)}\right)\left[\mathrm{Tr}(\mathbf{M})\right]^2\\ &+ n^2\left[\frac{2p}{n}+3\left(1-\frac{h}{n}\right)^2+\frac{6(n-h)^2p}{n^3}+\frac{(3n-h)p^2}{n^3}\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}),\\ \mathbb{E}\Xi_2^2 \approx &\ 2np\left(1+p/n\right)\mathrm{Tr}(\mathbf{M}),\\ \mathbb{E}\Xi_3^2 \approx &\ 2n^2\left(\frac{p}{n}+\left(1-\frac{h}{n}\right)^2+\frac{p^2}{n^2}+\frac{4p(n-h)^2}{n^3}\right)\mathrm{Tr}(\mathbf{M}),\\ \mathbb{E}\Xi_4^2 \approx &\ \frac{(n-h)m^2p^2}{n},\\ \mathbb{E}\Xi_1\Xi_4 \approx &\ \frac{mp(n-h)(n+p-h)}{n}\mathrm{Tr}(\mathbf{M}),\\ \mathbb{E}\Xi_2\Xi_3 \approx &\ \frac{p(n-h)(n+p-h)}{n}\mathrm{Tr}(\mathbf{M}).\end{aligned}$$

The detailed computation is attached as follows.

Lemma 1. *We have*

$$\begin{aligned}\mathbb{E}\Xi_1^2 = &\ (n-h)^2\left(1+\frac{2p}{n}+\frac{p^2}{n(n-h)}+o(1)\right)\left[\mathrm{Tr}(\mathbf{M})\right]^2\\ &+ n^2\left[\frac{2p}{n}+3\left(1-\frac{h}{n}\right)^2+\frac{6(n-h)^2p}{n^3}+\frac{(3n-h)p^2}{n^3}+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}),\end{aligned}$$

where Ξ_1 is defined in (21).

Proof. We begin the proof by decomposing Ξ_1^2 as

$$\begin{aligned}\mathbb{E}\Xi_1^2 = &\ \mathbb{E}\underbrace{\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top\boldsymbol{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\boldsymbol{\Sigma}^\top\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)}_{\Lambda_1}\\ &+ 2\mathbb{E}\underbrace{\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top\boldsymbol{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\boldsymbol{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)}_{\Lambda_2}\\ &+ \mathbb{E}\underbrace{\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top\boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\boldsymbol{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)}_{\Lambda_3},\end{aligned}$$

and separately bound each term as in Lemma 2, Lemma 3, and Lemma 4.

□

Lemma 2. *We have*

$$\begin{aligned}&\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top \mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Sigma}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\\ =\ &(n-h)^2\left(1+o(1)\right)\left[\mathrm{Tr}(\mathbf{M})\right]^2+\ n^2\left[\frac{2p}{n}+3\left(1-\frac{h}{n}\right)^2+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}).\end{aligned}$$

Proof. Due to the independence among different rows of the sensing matrix $\mathbf{X}$, we condition on $\mathbf{\Sigma}$ and take expectation w.r.t. $\mathbf{X}_{\pi^\natural(i)}$ and $\mathbf{X}_j$, which leads to

$$\mathbb{E}\Lambda_1=\ \mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Sigma}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{1,1}}+\mathbb{E}\underbrace{\mathbf{X}_j^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Sigma}^\top\mathbf{X}_j}_{\Lambda_{1,2}}.$$

For $\mathbb{E}\Lambda_{1,1}$, we obtain

$$\begin{aligned}\mathbb{E}\Lambda_{1,1}\ &\overset{①}{=}\ \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{M}\right)\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{M}\right)\right]+\mathbb{E}\,\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{M}\mathbf{M}^\top\mathbf{\Sigma}^\top\right)+\mathbb{E}\,\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma}\mathbf{M}\right)\\ &\overset{②}{=}\ (n-h)^2\left[1+o(1)\right]\left[\mathrm{Tr}(\mathbf{M})\right]^2+\ n^2\left[\frac{p}{n}+2\left(1-\frac{h}{n}\right)^2+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}),\end{aligned}$$

where ① is due to (65), and ② is due to Lemma 12, Lemma 13, and Lemma 14. As for $\mathbb{E}\Lambda_{1,2}$, we have

$$\mathbb{E}\Lambda_{1,2}=\mathbb{E}\,\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{M}\mathbf{M}^\top\mathbf{\Sigma}^\top\right)=n^2\left[\frac{p}{n}+\left(1-\frac{h}{n}\right)^2+o(1)\right]\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right),$$

and hence complete the proof. □

Lemma 3. *We have*

$$\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\approx\frac{(n-h)^2p}{n}\left[\left(\mathrm{Tr}(\mathbf{M})\right)^2+3\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right].$$

Proof. Similar as above, we first expand Λ_2 as

$$\begin{aligned}\mathbb{E}\Lambda_2=\ &(n-h)\mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,1}}+(n-h)\mathbb{E}\underbrace{\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Delta}^\top\mathbf{X}_j}_{\Lambda_{2,2}}\\ &-(n-h)\mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Delta}^\top\mathbf{X}_j}_{\Lambda_{2,3}}-(n-h)\mathbb{E}\underbrace{\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,4}}.\end{aligned}$$

Case (s,s): $i=\pi^\natural(i)$ and $j=\pi^\natural(j)$. We first compute $\Lambda_{2,1}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,1}=\ &\underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2}+\underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_i}_{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2}\\ =\ &(p+5)\left[\left(\mathrm{Tr}(\mathbf{M})\right)^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right].\end{aligned}$$

We consider $\Lambda_{2,2}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,2}=\ &\underbrace{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)}_{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2}+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j\right)}_{\mathbb{E}\|\mathbf{X}_j\|_2^2\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j}\\ =\ &\left(\mathrm{Tr}(\mathbf{M})\right)^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)+(p+2)\,\mathrm{Tr}\left(\mathbf{M}\mathbf{M}\right).\end{aligned}$$

As for $\Lambda_{2,3}$ and $\Lambda_{2,4}$, we can verify that they are both zero, which gives

$$\mathbb{E}\Lambda_2=\ (n-h)p\left[1+o(1)\right]\left[\mathrm{Tr}(\mathbf{M})\right]^2+3(n-h)p\left[1+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}).\tag{32}$$

Case (s,d)**:** $i=\pi^\natural(i)$ **and** $j\neq\pi^\natural(j)$**.** We can compute $\Lambda_{2,1}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,1} =\ & \mathbb{E}\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}}_{0}\\ &+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^\natural(i)}}_{0}\\ =\ & (p+4)\left[(\mathrm{Tr}(\mathbf{M}))^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right].\end{aligned}$$

We consider $\Lambda_{2,2}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,2} =\ & \mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{0}\\ &+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{0}\\ =\ & \mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2=(\mathrm{Tr}(\mathbf{M}))^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right).\end{aligned}$$

Similarly, we can verify that both $\mathbb{E}\Lambda_{2,3}$ and $\mathbb{E}\Lambda_{2,4}$ are zero and hence have

$$\mathbb{E}\Lambda_2 = (n-h)p\left[1+o(1)\right]\left[\mathrm{Tr}(\mathbf{M})\right]^2+2(n-h)p\left[1+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}). \tag{33}$$

Case (d,s)**:** $i\neq\pi^\natural(i)$ **and** $j=\pi^\natural(j)$**.** We compute $\Lambda_{2,1}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,1} =\ & \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}}_{0}+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_{\pi^\natural(i)}}_{0}\\ &+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}=(\mathrm{Tr}(\mathbf{M}))^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right).\end{aligned}$$

We consider $\Lambda_{2,2}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{2,2} =\ & \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{0}+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_j}_{0}\\ &+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_j\|_2^2\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j}=(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}).\end{aligned}$$

As for $\mathbb{E}\Lambda_{2,3}$ and $\Lambda_{2,4}$, we can follow the same strategy and prove they are both zero, which yields

$$\mathbb{E}\Lambda_2 = (n-h)\left[\mathrm{Tr}(\mathbf{M})\right]^2+(n-h)p\left[1+o(1)\right]\mathrm{Tr}(\mathbf{M}\mathbf{M}). \tag{34}$$

Case (d,d)**:** $i\neq\pi^\natural(i)$ **and** $j\neq\pi^\natural(j)$**.** Contrary to the previous cases, we have $\mathbb{E}\Lambda_{2,1}$ and $\mathbb{E}\Lambda_{2,2}$ to be zero in this case rather than $\mathbb{E}\Lambda_{2,3}$ and $\mathbb{E}\Lambda_{2,4}$.

Hence our focus turns to the calculation of $\mathbb{E}\Lambda_{2,3}$ and that of $\mathbb{E}\Lambda_{2,4}$. For $\Lambda_{2,3}$, we have

$$\begin{aligned}\mathbb{E}\Lambda_{2,3} =\ & \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{p\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}\\ &+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{p\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}+\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}\\ =\ & 2\left(p\mathbb{1}_{i=j}+\mathbb{1}_{j=\pi^{\natural 2}(i)}\right)\left[(\mathrm{Tr}(\mathbf{M}))^2+\mathrm{Tr}(\mathbf{M}\mathbf{M})+\mathrm{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right].\end{aligned}$$

Then we turn to the calculation of $\mathbb{E}\Lambda_{2,4}$, which proceeds as

$$\begin{aligned}\mathbb{E}\Lambda_{2,4} = & \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}(i=j)\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_2^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}} \\ & + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}(i=j)\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_2^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}} \\ = & \ 2\mathbb{1}_{j=\pi^{\natural 2}(i)}(p+2)\operatorname{Tr}(\mathbf{M}\mathbf{M}) + 2\mathbb{1}(i=j)\left[(\operatorname{Tr}(\mathbf{M}))^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right].\end{aligned}$$

Then we conclude

$$\begin{aligned}\mathbb{E}\Lambda_2 = & -2(n-h)\left[(p+1)\mathbb{1}_{i=j} + \mathbb{1}_{j=\pi^{\natural 2}(i)}\right]\left[(\operatorname{Tr}(\mathbf{M}))^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right] \\ & -2(n-h)(p+2)\mathbb{1}_{j=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M}\mathbf{M}). \qquad (35)\end{aligned}$$

The proof is thus completed by combining (32), (33), (34), and (35).

□

Lemma 4. *We have*

$$\begin{aligned}& \mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\boldsymbol{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right) \\ = & \left(1 - \frac{h}{n} + o(1)\right)p^2\left[\operatorname{Tr}(\mathbf{M})\right]^2 + \left(3 - \frac{h}{n} + o(1)\right)p^2\operatorname{Tr}(\mathbf{M}\mathbf{M}).\end{aligned}$$

Proof. We begin the proof by expanding Λ_3 as

$$\begin{aligned}\mathbb{E}\Lambda_3 = & \ \mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{3,1}} + \mathbb{E}\underbrace{\mathbf{X}_j^\top \boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j}_{\Lambda_{3,2}} \\ & - \mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j}_{\Lambda_{3,3}} - \mathbb{E}\underbrace{\mathbf{X}_j^\top \boldsymbol{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{3,4}}.\end{aligned}$$

Case (s,s): $i = \pi^\natural(i)$ and $j = \pi^\natural(j)$. First we compute $\Lambda_{3,1}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{3,1} = & \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^4\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top \mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} \\ & + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top \mathbf{X}_i}_{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i\right)^2\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j} \\ = & \ (p+4)(p+8)\left[(\operatorname{Tr}(\mathbf{M}))^2 + 2\operatorname{Tr}(\mathbf{M}\mathbf{M})\right] + 2(\operatorname{Tr}(\mathbf{M}))^2 + (p+6)\operatorname{Tr}(\mathbf{M}\mathbf{M}).\end{aligned}$$

Then, we consider $\Lambda_{3,2}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{3,2} = & \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top \mathbf{X}_j}_{(p+2)\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} \\ & + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{X}_j}_{(p+2)\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_j\|_2^4\mathbf{X}_j^\top \mathbf{M}\mathbf{M}\mathbf{X}_j} \\ = & \ (3p+8)\left[(\operatorname{Tr}(\mathbf{M}))^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top\mathbf{M}\right)\right] + (p+2)(p+4)\operatorname{Tr}(\mathbf{M}\mathbf{M}).\end{aligned}$$

In addition, we can verify that $\mathbb{E}\Lambda_{3,3}$ and $\mathbb{E}\Lambda_{3,4}$ are both zero. Hence we conclude

$$\mathbb{E}\Lambda_3 = \left(p^2 + 15p + 42\right)\left[\operatorname{Tr}(\mathbf{M})\right]^2 + \left(3p^2 + 37p + 94\right)\operatorname{Tr}(\mathbf{M}\mathbf{M}). \qquad (36)$$

Case (s,d): $i=\pi^\natural(i)$ and $j\neq\pi^\natural(j)$. We can compute $\Lambda_{3,1}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,1} =\ & \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^4\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i}_{0} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\mathbf{X}_i^\top\mathbf{M}\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i}_{0} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\mathbf{X}_i^\top\mathbf{M}\mathbf{M}\mathbf{X}_i} \\
=\ & \left(p^2+10p+24+2\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\left[\left(\operatorname{Tr}(\mathbf{M})\right)^2+2\operatorname{Tr}(\mathbf{M}\mathbf{M})\right]+2(p+2)\operatorname{Tr}(\mathbf{M}\mathbf{M}).
\end{aligned}$$

We consider $\Lambda_{3,2}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,2} =\ & \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\right)^2} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_2^4\operatorname{Tr}(\mathbf{M}\mathbf{M})} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\|\mathbf{X}_j\|_2^2\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\|\mathbf{X}_j\|_2^2\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_j\|_2^2\mathbf{X}_j^\top\mathbf{M}\mathbf{M}\mathbf{X}_j} \\
=\ & (p+4)\left[\left(\operatorname{Tr}(\mathbf{M})\right)^2+2\operatorname{Tr}(\mathbf{M}\mathbf{M})\right]+(p+2)\left(p+1+2\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\operatorname{Tr}(\mathbf{M}\mathbf{M}).
\end{aligned}$$

As for $\Lambda_{3,3}$ and $\Lambda_{3,4}$, we can prove that they are both zero in this case. Hence we conclude

$$\begin{aligned}
\mathbb{E}\Lambda_3 =\ & \left(p^2+11p+28+2\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\left(\operatorname{Tr}(\mathbf{M})\right)^2 \\
& + \left[3p^2+27p+62+2\left(p+4\right)\mathbb{1}_{j=\pi^{\natural 2}(j)}\right]\operatorname{Tr}(\mathbf{M}\mathbf{M}).
\end{aligned} \tag{37}$$

Case (d,s): $i\neq\pi^\natural(i)$ and $j=\pi^\natural(j)$. We consider the term $\Lambda_{3,1}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,1} =\ & \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2\left(\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2\left(\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}}_{0}
\end{aligned}$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi\natural(i)}}_{\mathbb{1}_{i=\pi\natural 2(i)} \mathbb{E}\left\|\mathbf{X}_{\pi\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)}\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural 2(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_{\pi\natural(i)}}_{\mathbb{E}\|\mathbf{X}_i\|_2^4 \mathbf{X}_i^\top \mathbf{M}\mathbf{M}\mathbf{X}_i} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_{\pi\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural 2(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_{\pi\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_{\pi\natural(i)}}_{\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_j\right)^2 \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j} \\
&= (p+6)\left(\mathrm{Tr}(\mathbf{M})\right)^2 + \left(p^2+9p+22\right)\mathrm{Tr}(\mathbf{M}\mathbf{M}) + 2\mathbb{1}_{i=\pi\natural 2(i)}(p+4)\left[\left(\mathrm{Tr}(\mathbf{M})\right)^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right].
\end{aligned}
$$

We consider the term $\Lambda_{3,2}$ as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,2} = &\ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{p\times\mathbb{E}\left(\mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)}\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural 2(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi\natural 2(i)}\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi\natural 2(i)}\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural 2(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{M}\mathbf{X}_i} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural 2(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi\natural 2(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi\natural(i)} \mathbf{X}_{\pi\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^4 \times \mathbf{X}_j^\top \mathbf{M}\mathbf{M}\mathbf{X}_j} \\
= &\ p\left(\mathrm{Tr}(\mathbf{M})\right)^2 + \left(p^2+9p+10\right)\mathrm{Tr}(\mathbf{M}\mathbf{M}) + 2\mathbb{1}_{i=\pi\natural 2(i)}\left[\left(\mathrm{Tr}(\mathbf{M})\right)^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right].
\end{aligned}
$$

As for $\Lambda_{3,3}$ and $\Lambda_{3,4}$, easily we can verify they are both zero and hence

$$
\begin{aligned}
\mathbb{E}\Lambda_3 = &\ 2(p+3)\left(\mathrm{Tr}(\mathbf{M})\right)^2 + 2\left(p^2+9p+16\right)\mathrm{Tr}(\mathbf{M}\mathbf{M}) \\
&+ 2\mathbb{1}_{i=\pi\natural 2(i)}(p+5)\left[\left(\mathrm{Tr}(\mathbf{M})\right)^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right]. \qquad (38)
\end{aligned}
$$

Case (d,d): $i \neq \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. First, We compute $\mathbb{E}\Lambda_{3,1}$ as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,1} = \; & \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=i} \mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=i} \mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=i} \mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2 + \mathbb{1}_{i\neq j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(j)} \left[\mathbb{1}_{i=j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2 + \mathbb{1}_{i\neq j} \mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2\right]} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)} \\
& + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(j)} \left[\mathbb{1}_{i=j} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2 + \mathbb{1}_{i\neq j} \mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right)^2\right]}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E} \left\| \mathbf{X}_{\pi^\natural(i)} \right\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \right) + \mathbb{1}_{j \neq \pi^{\natural 2}(i)} \mathbb{E} \left\| \mathbf{X}_{\pi^\natural(i)} \right\|_{\mathrm{F}}^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \right)} \\
&= \left(1 + 2\mathbb{1}_{i=\pi^{\natural 2}(i)} + 3\mathbb{1}_{i=j} + 6\mathbb{1}_{i=j}\mathbb{1}_{j=\pi^{\natural 2}(i)}\right)(p+4)\left[(\mathrm{Tr}(\mathbf{M}))^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right] \\
&+ \left(1 + 3\mathbb{1}_{j=\pi^{\natural 2}(i)}\right)(p+2)(p+4)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}) \\
&+ 2\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{1}_{i \neq j}\left[(\mathrm{Tr}(\mathbf{M}))^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right] \\
&+ \mathbb{1}_{j \neq \pi^{\natural 2}(i)}(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}) + \mathbb{1}_{i \neq j}(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}). \qquad (39)
\end{aligned}
$$

Then we calculate $\mathbb{E}\Lambda_{3,2}$ as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,2} = &\underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2 + p\mathbb{1}_{i \neq j}\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} + \mathbb{1}_{i \neq j}\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2\right]} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} + \mathbb{1}_{i \neq j}\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2\right]} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{j \neq \pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_{\pi^\natural(i)}\|_{\mathrm{F}}^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{M} \mathbf{X}_{\pi^\natural(i)} + \mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)}\right)^2 + \mathbb{1}_{i \neq j}\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^4 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{M} \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\left[\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} + \mathbb{1}_{i \neq j}\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^2 \mathbf{X}_j^\top \mathbf{M} \mathbf{M} \mathbf{X}_j\right]}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j\right)^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\left[\mathbb{1}_{i=j}\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} + \mathbb{1}_{i\neq j}\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^2 \mathbf{X}_j^\top \mathbf{M}\mathbf{M}\mathbf{X}_j\right]} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \left(\mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top\right) \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\left(\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top\right) \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)^2 \left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j\right)^2 + \mathbb{1}_{j\neq\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^2 \mathbf{X}_j^\top \mathbf{M}\mathbf{M}\mathbf{X}_j} \\
&= 4\mathbb{1}_{i=j} p(p+2)\left[[\mathrm{Tr}(\mathbf{M})]^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right] + \mathbb{1}_{i\neq j} p(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}) \\
&+ 8\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}(p+2)\left[(\mathrm{Tr}(\mathbf{M}))^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right] \\
&+ 4\mathbb{1}_{j=\pi^{\natural 2}(i)}\left[2\,[\mathrm{Tr}(\mathbf{M})]^2 + (p+6)\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right] \\
&+ 2\mathbb{1}_{j\neq\pi^{\natural 2}(i)}(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}) + 2\mathbb{1}_{i\neq j}\mathbb{1}_{j=\pi^{\natural 2}(j)}(p+2)\,\mathrm{Tr}(\mathbf{M}\mathbf{M}) \\
&+ \mathbb{1}_{i\neq j}\left(p + 2\mathbb{1}_{i=\pi^{\natural 2}(i)}\right)\left[(\mathrm{Tr}(\mathbf{M}))^2 + 2\,\mathrm{Tr}(\mathbf{M}\mathbf{M})\right]. \qquad (40)
\end{aligned}
$$

The term $\Lambda_{3,3}$ is computed as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,3} = &\ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^\natural(j)} p\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 3}(i)}\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^\natural(j)}\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 3}(i)} \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j}_{0} \\
&= \left[(p+1)\mathbb{1}_{i=\pi^\natural(j)} + \mathbb{1}_{j=\pi^{\natural 3}(i)}\right] [\mathrm{Tr}(\mathbf{M})]^2 \\
&+ \left[2(p+1)\mathbb{1}_{i=\pi^\natural(j)} + (p+4)\mathbb{1}_{j=\pi^{\natural 3}(i)}\right] \mathrm{Tr}(\mathbf{M}\mathbf{M}). \qquad (41)
\end{aligned}
$$

Then, we consider the term $\mathbb{E}\Lambda_{3,4}$, which can be written as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,4} = &\ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^\natural(j)} \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 3}(i)} \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^\natural(j)} p \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\right]^2} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural2(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^\natural3(i)} \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_{\mathrm{F}}^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{M} \mathbf{X}_{\pi^\natural(i)}} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&= \left[(p+1)\mathbb{1}_{i=\pi^\natural(j)} + \mathbb{1}_{j=\pi^\natural3(i)}\right] [\mathrm{Tr}(\mathbf{M})]^2 \\
&+ \left[2(p+1)\mathbb{1}_{i=\pi^\natural(j)} + (p+4)\mathbb{1}_{j=\pi^\natural3(i)}\right] \mathrm{Tr}(\mathbf{M}\mathbf{M}). \qquad (42)
\end{aligned}
$$

Combing (39), (40), (41), and (42) together then yields

$$
\mathbb{E}\Lambda_3 = \frac{4hp(p+2)}{n^2} [1+o(1)] [\mathrm{Tr}(\mathbf{M})]^2 + 2p^2 [1+o(1)] \mathrm{Tr}(\mathbf{M}\mathbf{M}). \qquad (43)
$$

The proof is thus completed by summarizing the computations thereof. □

.

Lemma 5. *We have*

$$
\mathbb{E}\Xi_2^2 = 2\left[(p+2)(p+3) + (n-2)(p+1)\right] \left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2 = 2np\left(1 + p/n + o(1)\right) \left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2,
$$

where Ξ_2 is defined in (21).

Proof. We have

$$
\begin{aligned}
\mathbb{E}\Xi_2^2 &= \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathbf{W}^\top \mathbf{X} \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right) \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}^\top \mathbf{W} \mathbf{B}^{\natural\top} \mathbf{X}_{\pi^\natural(i)}\right] \\
&= \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathrm{Tr}\left[\mathbf{X} \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right) \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}^\top\right] \mathbf{B}^{\natural\top} \mathbf{X}_{\pi^\natural(i)}\right] \\
&= \mathbb{E}\left[\left\|\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\right\|_2^2 \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural \mathbf{B}^{\natural\top} \mathbf{X}_{\pi^\natural(i)}\right] \\
&= \mathbb{E}\left[\left\|\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\right\|_2^2 \times \left\|\mathbf{B}^{\natural\top} \mathbf{X}_{\pi^\natural(i)}\right\|_2^2\right].
\end{aligned}
$$

For the conciseness of notation, we assume $\pi^\natural(i) = 1$ and $j = 2$ w.l.o.g. Decomposing the term $\left\|\mathbf{X}\left(\mathbf{X}_1 - \mathbf{X}_2\right)\right\|_2^2$ as

$$
\left\|\mathbf{X}\left(\mathbf{X}_1 - \mathbf{X}_2\right)\right\|_{\mathrm{F}}^2 = \underbrace{\left[\mathbf{X}_1^\top \left(\mathbf{X}_1 - \mathbf{X}_2\right)\right]^2}_{\mathcal{T}_1} + \underbrace{\left[\mathbf{X}_2^\top \left(\mathbf{X}_1 - \mathbf{X}_2\right)\right]^2}_{\mathcal{T}_2} + \underbrace{\sum_{i=3}^n \left[\mathbf{X}_i^\top \left(\mathbf{X}_1 - \mathbf{X}_2\right)\right]^2}_{\mathcal{T}_3},
$$

we then separately bound the above three terms. For the first term $\mathbb{E}\mathcal{T}_1 \left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2$, we have

$$
\begin{aligned}
\mathbb{E}\mathcal{T}_1 \left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2 &= \mathbb{E}\left[\left(\left\|\mathbf{X}_1\right\|_2^4 + \left(\mathbf{X}_1^\top \mathbf{X}_2\right)^2\right) \left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2\right] \\
&= \underbrace{\mathbb{E}\left\|\mathbf{X}_1\right\|_2^4 \left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2}_{(p+2)(p+4)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\left(\mathbf{X}_1^\top \mathbf{X}_2\right)^2 \left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2}_{(p+2)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2} \overset{①}{=} (p+2)(p+5)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2,
\end{aligned} \qquad (44)
$$

where ① is due to (66) and (67).

Similarly, for term $\mathbb{E}\mathcal{T}_2\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2$, we invoke (66) and (67), which gives

$$\mathbb{E}\mathcal{T}_2\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2 = \underbrace{\mathbb{E}\left[\left(\mathbf{X}_1^\top\mathbf{X}_2\right)^2\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2\right]}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\|\mathbf{X}_2\|_2^4 \times \mathbb{E}\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2}_{p(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} \tag{45}$$

$$\overset{②}{=} (p+1)(p+2)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2, \tag{46}$$

where ② is due to (66).

For the last term $\mathbb{E}\mathcal{T}_3\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2$, we exploit the independence among the rows of matrix $\mathbf{X}$ and have

$$\begin{aligned}
\mathbb{E}T_3\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2 &= \sum_{i\geq 3}\mathbb{E}\left[\left(\mathbf{X}_i^\top\left(\mathbf{X}_1-\mathbf{X}_2\right)\right)^2\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2\right] \\
&= \sum_{i\geq 3}\mathbb{E}\left[\|\mathbf{X}_1-\mathbf{X}_2\|_2^2\cdot\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2\right] \\
&= \sum_{i\geq 3}\mathbb{E}\left[\left(\|\mathbf{X}_1\|_2^2+\|\mathbf{X}_2\|_2^2\right)\cdot\left\|\mathbf{B}^{\natural\top}\mathbf{X}_1\right\|_{\mathrm{F}}^2\right] \\
&= 2\sum_{i\geq 3}(p+1)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2 = 2(n-2)(p+1)\left\|\mathbf{B}^\natural\right\|_{\mathrm{F}}^2.
\end{aligned} \tag{47}$$

The proof is then completed by combining (44), (46), and (47).

□

Lemma 6. *We have*

$$\mathbb{E}\Xi_3^2 = 2n^2\left[\frac{p}{n}+\left(1-\frac{h}{n}\right)^2+\frac{p^2}{n^2}+\frac{4p(n-h)^2}{n^3}+o(1)\right]\mathrm{Tr}(\mathbf{M}),$$

where Ξ_3 is defined in (21).

Proof. To begin with, we decompose the term $\mathbb{E}\Xi_3^2$ as

$$\begin{aligned}
\mathbb{E}\Xi_3^2 = {}& \mathbb{E}\underbrace{\left[\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\boldsymbol{\Sigma}\mathbf{M}\boldsymbol{\Sigma}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right]}_{\triangleq\Lambda_1} + 2\mathbb{E}\underbrace{\left[\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\boldsymbol{\Sigma}\mathbf{M}\boldsymbol{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right]}_{\triangleq\Lambda_2} \\
&+\mathbb{E}\underbrace{\left[\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\boldsymbol{\Delta}\mathbf{M}\boldsymbol{\Delta}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right]}_{\triangleq\Lambda_3}.
\end{aligned} \tag{48}$$

Step I. First we consider $\mathbb{E}\Lambda_1$, which can be written as

$$\mathbb{E}\Lambda_1 = 2\mathbb{E}\,\mathrm{Tr}\left(\boldsymbol{\Sigma}\mathbf{M}\boldsymbol{\Sigma}^\top\right) \overset{①}{=} 2n^2\left[\frac{p}{n}+\left(1-\frac{h}{n}\right)^2+o(1)\right]\mathrm{Tr}(\mathbf{M}), \tag{49}$$

where ① is due to Lemma 12.

Step II. Then we turn to $\mathbb{E}\Lambda_2$, which can be written as

$$\begin{aligned}
\mathbb{E}\Lambda_2 = {}& (n-h)\mathbb{E}\underbrace{\left[\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{\Lambda_{2,1}} + (n-h)\mathbb{E}\underbrace{\left[\mathbf{X}_j^\top\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j\right]}_{\Lambda_{2,2}} \\
&-(n-h)\mathbb{E}\underbrace{\left[\mathbf{X}_{\pi^\natural(i)}{}^\top\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j\right]}_{\Lambda_{2,3}} - (n-h)\mathbb{E}\underbrace{\left[\mathbf{X}_j{}^\top\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{\Lambda_{2,4}}.
\end{aligned}$$

Case (s,s): $i=\pi^\natural(i)$ and $j=\pi^\natural(j)$. We have

$$\mathbb{E}\Lambda_{2,1} = \mathbb{E}\mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_i = (p+3)\operatorname{Tr}(\mathbf{M}),$$
$$\mathbb{E}\Lambda_{2,2} = \mathbb{E}\mathbf{X}_j^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_j = (p+3)\operatorname{Tr}(\mathbf{M}).$$

In addition, we can verify that $\mathbb{E}\Lambda_{2,2}$ and $\Lambda_{2,3}$ are both zero, which suggests that

$$\mathbb{E}\Lambda_2 = 2(n-h)(p+3)\operatorname{Tr}(\mathbf{M}). \tag{50}$$

Case (s,d): $i = \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. We have

$$\mathbb{E}\Lambda_{2,1} = \mathbb{E}\mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{X}_i = (p+2)\operatorname{Tr}(\mathbf{M});$$
$$\mathbb{E}\Lambda_{2,2} = \mathbb{E}\mathbf{X}_j^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{X}_j = \operatorname{Tr}(\mathbf{M}).$$

Moreover, we have both $\mathbb{E}\Lambda_{2,3}$ and $\mathbb{E}\Lambda_{2,4}$ be zero, which suggests that

$$\mathbb{E}\Lambda_2 = (n-h)(p+3)\operatorname{Tr}(\mathbf{M}). \tag{51}$$

Case (d,s): $i \neq \pi^\natural(i)$ and $j = \pi^\natural(j)$. We have

$$\mathbb{E}\Lambda_{2,1} = \mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}\right) = \operatorname{Tr}(\mathbf{M}),$$
$$\mathbb{E}\Lambda_{2,2} = \mathbb{E}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j = (p+2)\operatorname{Tr}(\mathbf{M}).$$

Similar as above, we can verify both $\mathbb{E}\Lambda_{2,3}$ and $\mathbb{E}\Lambda_{2,4}$ are zero, which suggests that

$$\mathbb{E}\Lambda_2 = (n-h)(p+3)\operatorname{Tr}(\mathbf{M}). \tag{52}$$

Case (d,d): $i \neq \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. Different from the above three cases, we have $\mathbb{E}\Lambda_{2,1}$ and $\mathbb{E}\Lambda_{2,2}$ be zero and focus on the calculation of $\mathbb{E}\Lambda_{2,3}$ and $\mathbb{E}\Lambda_{2,4}$, which proceeds as

$$\begin{aligned}
\mathbb{E}\Lambda_{2,3} = &\underbrace{\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}{}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_j\right]}_{p\mathbb{1}_{i=j}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}{}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_j\right]}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} \\
&+ \underbrace{\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}{}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j\right]}_{p\mathbb{1}_{i=j}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}{}^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j\right]}_{\mathbb{1}_{\pi^\natural(i)=\pi^{\natural-1}(j)}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} \\
= &\ 2\left[p\mathbb{1}_{i=j} + \mathbb{1}_{j=\pi^{\natural 2}(i)}\right]\operatorname{Tr}(\mathbf{M}); \\
\mathbb{E}\Lambda_{2,4} = &\underbrace{\mathbb{E}\left[\mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{\mathbb{1}_{i=j}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\left[\mathbf{X}_j{}^\top \mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{p\mathbb{1}_{j=\pi^{\natural 2}(i)}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} \\
&+ \underbrace{\mathbb{E}\left[\mathbf{X}_j{}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{\mathbb{1}_{i=j}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} + \underbrace{\mathbb{E}\left[\mathbf{X}_j{}^\top \mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^\natural(i)}\right]}_{p\mathbb{1}_{j=\pi^{\natural 2}(i)}\|\mathbf{B}^\natural\|_{\mathrm{F}}^2} \\
= &\ 2\left[p\mathbb{1}_{j=\pi^{\natural 2}(i)} + \mathbb{1}_{i=j}\right]\operatorname{Tr}(\mathbf{M}),
\end{aligned}$$

which suggests that

$$\mathbb{E}\Lambda_2 = -2(n-h)(p+1)\left(\mathbb{1}_{j=\pi^{\natural 2}(i)} + \mathbb{1}_{i=j}\right)\operatorname{Tr}(\mathbf{M}). \tag{53}$$

Combing (50), (51), (52), and (53), we conclude

$$\mathbb{E}\Lambda_2 = \frac{2p(n-h)^2}{n}\operatorname{Tr}(\mathbf{M})\left[1 + o(1)\right]. \tag{54}$$

Step III. Then we turn to the calculation of $\mathbb{E}\Lambda_3$. First we perform the following decomposition

$$\Lambda_3 = \underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \boldsymbol{\Delta}\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{3,1}} + \underbrace{\mathbf{X}_j^\top \boldsymbol{\Delta}\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j}_{\Lambda_{3,2}} - \underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \boldsymbol{\Delta}\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_j}_{\Lambda_{3,3}} - \underbrace{\mathbf{X}_j^\top \boldsymbol{\Delta}\mathbf{M}\boldsymbol{\Delta}^\top\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{3,4}}.$$

Case (s,s): $i=\pi^\natural(i)$ and $j=\pi^\natural(j)$. We have

$$\begin{aligned}\mathbb{E}\Lambda_{3,1} =\ & \underbrace{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\ \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i\right)}_{\mathbb{E}\|\mathbf{X}_i\|_2^4\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}+\underbrace{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\ \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_i\right)}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\ \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i\right)}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}+\underbrace{\mathbb{E}\left(\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\ \mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_i\right)}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}=(p+2)(p+7)\operatorname{Tr}(\mathbf{M});\\ \mathbb{E}\Lambda_{3,2} =\ & \mathbb{E}\mathbf{X}_j^\top\left(\mathbf{X}_i\mathbf{X}_i^\top+\mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top+\mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_j=(p+2)(p+7)\operatorname{Tr}(\mathbf{M}).\end{aligned}$$

As for $\mathbb{E}\Lambda_{3,3}$ and $\mathbb{E}\Lambda_{3,4}$, easily we can verify that they are both zero and hence have

$$\mathbb{E}\Lambda_3=2(p+2)(p+7)\operatorname{Tr}(\mathbf{M})=2p^2\operatorname{Tr}(\mathbf{M})\left[1+o(1)\right]. \tag{55}$$

Case (s,d): $i=\pi^\natural(i)$ and $j\neq\pi^\natural(j)$. We can write $\Lambda_{3,1}$ as

$$\begin{aligned}\mathbb{E}\Lambda_{3,1} =\ & \underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i\right)}_{\mathbb{E}\|\mathbf{X}_i\|_2^4\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i\right)}_{0}\\ &+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i\right)}_{0}\\ &+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i\right)}_{0}+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i\right)}_{p\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\\ &+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i\right)}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\operatorname{Tr}(\mathbf{M})}\\ &+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i\right)}_{0}+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_i\right)}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\operatorname{Tr}(\mathbf{M})}\\ &+\underbrace{\mathbb{E}\ \left(\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_i\right)}_{p\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\\ =\ & \left(p^2+8p+8+2\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\operatorname{Tr}(\mathbf{M}).\end{aligned}$$

Mean $\Lambda_{3,2}$ can be written as

$$\begin{aligned}\mathbb{E}\Lambda_{3,2} =\ & \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j\right)}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j\right)}_{0}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j\right)}_{0}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j\right)}_{0}+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j\right)}_{\mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j\right)}_{\mathbb{1}_{\pi^{\natural-1}(j)=\pi^\natural(j)}(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_j\right)}_{0}+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top\mathbf{X}_j\right)}_{\mathbb{1}_{\pi^{\natural-1}(j)=\pi^\natural(j)}(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\\ &+\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j\right)}_{(p+2)\|\mathbf{B}^\natural\|_{\mathrm{F}}^2}\end{aligned}$$

$$= (p+2)\left(p+2+2\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\operatorname{Tr}(\mathbf{M}).$$

And for $\mathbb{E}\Lambda_{3,3}$ and $\mathbb{E}\Lambda_{3,4}$, easily we can verify that they are both zero. Then we conclude

$$\mathbb{E}\Lambda_3 = 2\left(p^2+6p+6+(p+3)\mathbb{1}_{j=\pi^{\natural 2}(j)}\right)\operatorname{Tr}(\mathbf{M}) = 2p^2\operatorname{Tr}(\mathbf{M})\left[1+o(1)\right]. \tag{56}$$

Case (d,s): $i \neq \pi^{\natural}(i)$ and $j = \pi^{\natural}(j)$. In this case, we can write $\Lambda_{3,1}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,1} = &\ \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{0}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{0}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{0}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{0}\\
&+\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
= &\ (p+2)\left(p+2+2\mathbb{1}_{i=\pi^{\natural 2}(i)}\right)\operatorname{Tr}(\mathbf{M}).
\end{aligned}$$

We consider $\Lambda_{3,2}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,2} = &\ \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{p\operatorname{Tr}(\mathbf{M})}+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M})}\\
&+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j}_{0}\\
&+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M})}+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{p\operatorname{Tr}(\mathbf{M})}\\
&+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j}_{0}\\
&+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{0}+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{0}+\underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
= &\ (p^2+8p+8+2\mathbb{1}_{i=\pi^{\natural 2}(i)})\operatorname{Tr}(\mathbf{M}).
\end{aligned}$$

Similarly, as above, we can verify that $\mathbb{E}\Lambda_{3,3} = 0$ and $\mathbb{E}\Lambda_{3,4} = 0$. Hence, we can conclude

$$\mathbb{E}\Lambda_3 = 2\left(p^2 + 6p + 6 + (p+3)\mathbb{1}_{i=\pi^{\natural 2}(i)}\right)\operatorname{Tr}(\mathbf{M}) = 2p^2\operatorname{Tr}(\mathbf{M})\left[1 + o(1)\right]. \tag{57}$$

Case (d,d): $i \neq \pi^{\natural}(i)$ and $j \neq \pi^{\natural}(j)$. We write $\Lambda_{3,1}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,1} = {} & \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\left[\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right]} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\left(\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right)} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} \\
& + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\left[\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right]} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_{\pi^{\natural}(i)}}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{j\neq\pi^{\natural 2}(i)}p\operatorname{Tr}(\mathbf{M})} \\
= {} & \left(p^2 + 5p + 2\right)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=\pi^{\natural 2}(i)}2(p+2)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{j=\pi^{\natural 2}(i)}(3p^2 + 5p)\operatorname{Tr}(\mathbf{M}) \\
& + 2\mathbb{1}_{j=\pi^{\natural 2}(j)}\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j}2(p+3)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}2(3p+5)\operatorname{Tr}(\mathbf{M}).
\end{aligned}$$

We consider $\Lambda_{3,2}$ as

$$\begin{aligned}
\mathbb{E}\Lambda_{3,2} = {} & \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i\neq j}p\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right]} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right]} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i + \mathbb{1}_{j\neq\pi^{\natural 2}(i)}p\operatorname{Tr}(\mathbf{M})} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(i)}\mathbf{X}_i^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural 2}(i)}\mathbf{X}_{\pi^{\natural}(i)}^\top\mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_{\pi^{\natural}(j)}\mathbf{X}_j^\top\mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_{\pi^{\natural}(j)}^\top\mathbf{M}\mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top\mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}
\end{aligned}$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 2}(j)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_j}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} \\
&= \left(p^2+5p+2\right)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{j=\pi^{\natural 2}(j)} 2(p+2)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j}\left(3p^2+5p\right)\operatorname{Tr}(\mathbf{M}) \\
&+ 2\mathbb{1}_{i=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{j=\pi^{\natural 2}(i)} 2(p+3)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)} 2\left(3p+5\right)\operatorname{Tr}(\mathbf{M}).
\end{aligned}
$$

We consider $\Lambda_{3,3}$ as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,3} = &\ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^\natural(j)} p\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 3}(i)} p\operatorname{Tr}(\mathbf{M})} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{\mathbb{1}_{i=\pi^\natural(j)}\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_j}_{0} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j}_{\mathbb{1}_{j=\pi^{\natural 3}(i)}\operatorname{Tr}(\mathbf{M})} \\
&+ \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j}_{0} + \underbrace{\mathbb{E}\,\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural-1(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_j}_{0} \\
= &\ (p+1)\left[\mathbb{1}_{i=\pi^\natural(j)} + \mathbb{1}_{j=\pi^{\natural 3}(i)}\right]\operatorname{Tr}(\mathbf{M}).
\end{aligned}
$$

Then we consider $\Lambda_{3,4}$ as

$$
\begin{aligned}
\mathbb{E}\Lambda_{3,4} = &\ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^\natural(j)}\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{j=\pi^{\natural 3}(i)}\operatorname{Tr}(\mathbf{M})} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{p\mathbb{1}_{i=\pi^\natural(j)}\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^\natural-1(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}}_{p \mathbb{1}_{j=\pi^{\natural 3}(i)} \operatorname{Tr}(\mathbf{M})} \\
&+ \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural -1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}}_{0} \\
&= (p+1) \left(\mathbb{1}_{i=\pi^\natural(j)} + \mathbb{1}_{j=\pi^{\natural 3}(i)} \right) \operatorname{Tr}(\mathbf{M}).
\end{aligned}
$$

In summary, we have

$$
\begin{aligned}
\mathbb{E}\Lambda_3 = \; & 2\left(p^2+5p+2\right) \operatorname{Tr}(\mathbf{M}) + 2(p+3) \left[\mathbb{1}_{i=\pi^{\natural 2}(i)} + \mathbb{1}_{j=\pi^{\natural 2}(j)} \right] \operatorname{Tr}(\mathbf{M}) \\
& + \left(3p^2+7p+6\right) \left(\mathbb{1}_{i=j} + \mathbb{1}_{j=\pi^{\natural 2}(i)} \right) \operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j} \mathbb{1}_{i=\pi^{\natural 2}(i)} 4\left(3p+5\right) \operatorname{Tr}(\mathbf{M}) \\
& - 2(p+1) \left[\mathbb{1}_{i=\pi^\natural(j)} + \mathbb{1}_{j=\pi^{\natural 3}(i)} \right] \operatorname{Tr}(\mathbf{M}) = 2p^2 \operatorname{Tr}(\mathbf{M}) \left[1+o(1)\right]. \qquad (58)
\end{aligned}
$$

Combining (55), (56), (57), and (58) then yields

$$
\mathbb{E}\Lambda_3 = 2p^2 \operatorname{Tr}(\mathbf{M}) \left[1+o(1)\right]. \qquad (59)
$$

The proof is then completed by (48), (49), (54), and (59).

□

Lemma 7. *We have*

$$
\begin{aligned}
\mathbb{E}\Xi_4^2 = \; & m(m+1) \left[p\left(p+2\right) \left(\mathbb{1}_{i=\pi^\natural(i)} + \mathbb{1}_{i=j} \right) + p \left(\mathbb{1}_{i\neq\pi^\natural(i)} + \mathbb{1}_{i\neq j} \right) \right] + 2mp(n+p+1) \\
= \; & \frac{(n-h)m^2p^2}{n} \left[1+o(1)\right],
\end{aligned}
$$

where Ξ_4 is defined in (21).

Proof. For the conciseness of notation, we define $\mathbf{\Gamma}$ as $\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}^\top$ and hence have

$$
\mathbb{E}\Xi_4^2 = \mathbb{E}\mathbf{W}_i^\top \mathbf{W}^\top \mathbf{\Gamma} \mathbf{W} \mathbf{W}_i.
$$

We begin the discussion by expanding $\mathbf{W}\mathbf{W}_i$ as

$$
\begin{bmatrix} \mathbf{W}_1^\top \\ \mathbf{W}_2^\top \\ \cdots \\ \mathbf{W}_n^\top \end{bmatrix} \mathbf{W}_i = \begin{bmatrix} \mathbf{W}_1^\top \mathbf{W}_i \\ \mathbf{W}_2^\top \mathbf{W}_i \\ \cdots \\ \mathbf{W}_n^\top \mathbf{W}_i \end{bmatrix}.
$$

Then we obtain

$$
\begin{aligned}
\mathbb{E}\Xi_4^2 = \; & \sum_{s=1}^n \sum_{t=1}^n \Gamma_{ij} \mathbb{E}\left[\left(\mathbf{W}_s^\top \mathbf{W}_i\right) \left(\mathbf{W}_t^\top \mathbf{W}_i\right) \right] = \Gamma_{ii} \mathbb{E}\left(\mathbf{W}_i^\top \mathbf{W}_i\right)^2 + \underbrace{\sum_{s\neq i} \sum_{t\neq i} \Gamma_{st} \mathbb{E}\left(\mathbf{W}_s^\top \mathbf{W}_i \mathbf{W}_t^\top \mathbf{W}_i\right)}_{\sum_{s\neq i} \Gamma_{ss} \mathbb{E}(\mathbf{W}_s^\top \mathbf{W}_i)^2} \\
= \; & \Gamma_{ii} \mathbb{E}\left(\sum_{j=1}^m W_{ij}^2 \right)^2 + \sum_{s\neq i} \Gamma_{ss} \cdot m = m(m+1)\mathbb{E}\Gamma_{ii} + m\mathbb{E}\operatorname{Tr}(\mathbf{\Gamma}). \qquad (60)
\end{aligned}
$$

We can thus complete the proof by separately computing $\mathbb{E}\operatorname{Tr}(\mathbf{M})$ and $\mathbb{E}\Gamma_{ii}$. First we compute $\mathbb{E}\Gamma_{ii}$, which proceeds as

$$
\begin{aligned}
\mathbb{E}\Gamma_{ii} = \; & \mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}\right)^2 + \mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_j\right)^2 \\
= \; & \mathbb{1}_{i=\pi^\natural(i)} p(p+2) + \mathbb{1}_{i\neq\pi^\natural(i)} p + \mathbb{1}_{i=j} p(p+2) + \mathbb{1}_{i\neq j} p \\
= \; & p\left(p+2\right) \left[\mathbb{1}_{i=\pi^\natural(i)} + \mathbb{1}_{i=j} \right] + p \left[\mathbb{1}_{i\neq\pi^\natural(i)} + \mathbb{1}_{i\neq j} \right]. \qquad (61)
\end{aligned}
$$

Then we turn to the computation of $\mathbb{E}\operatorname{Tr}(\mathbf{M})$, which proceeds as

$$
\mathbb{E}\operatorname{Tr}(\mathbf{\Gamma}) = \left\| \mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right) \right\|_{\mathrm{F}}^2
$$

$$
\begin{aligned}
&= \mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right\|_2^2+\mathbb{E}\left\|\mathbf{X}_j^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right\|_2^2+\sum_{s\neq\pi^\natural(i),j}\mathbb{E}\left\|\mathbf{X}_s^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\right\|_2^2\\
&= 2\mathbb{E}\left\|\mathbf{X}_{\pi^\natural(i)}\right\|_2^4+2\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_j\right)^2+2\sum_{s\neq\pi^\natural(i),j}\mathbb{E}\|\mathbf{X}_s\|_2^2\\
&= 2p(p+3)+2(n-2)p=2p(n+p+1). \qquad (62)
\end{aligned}
$$

The proof is thus completed by combing (60), (62), and (61). □

Lemma 8. *We have*

$$
\mathbb{E}\Xi_1\Xi_4=\frac{mp(n-h)(n+p-h)}{n}\left[1+o(1)\right]\operatorname{Tr}(\mathbf{M}),
$$

where Ξ_1 and Ξ_4 are defined in (21).

Proof. We have

$$
\mathbb{E}\Xi_1\Xi_4=\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}^\top\mathbf{\Pi}^{\natural\top}\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\mathbf{X}^\top}_{\triangleq\boldsymbol{v}^\top}\mathbf{W}\mathbf{W}_i.
$$

First we conditional on $\mathbf{X}$. Expanding the product $\mathbf{W}\mathbf{W}_i$ as

$$
\mathbb{E}\begin{bmatrix}\mathbf{W}_1^\top\mathbf{W}_i\\ \mathbf{W}_2^\top\mathbf{W}_i\\ \cdots\\ \mathbf{W}_i^\top\mathbf{W}_i\\ \cdots\\ \mathbf{W}_n^\top\mathbf{W}_i\end{bmatrix}=\begin{bmatrix}0\\0\\ \cdots\\ m\\ \cdots\\ 0\end{bmatrix},
$$

we can compute $\mathbb{E}\Xi_1\Xi_2$ w.r.t. $\mathbf{W}$ as

$$
\mathbb{E}\left(\Xi_1\Xi_2\right)=\mathbb{E}\boldsymbol{v}^\top\mathbf{W}\mathbf{W}_i=m\mathbb{E}\boldsymbol{v}_i,
$$

where $\boldsymbol{v}_i$ denotes the i-th entry of $\boldsymbol{v}$ and can be written as

$$
\boldsymbol{v}_i=\underbrace{\mathbf{X}_i^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_1}+\underbrace{\mathbf{X}_i^\top\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)\left(\mathbf{X}_{\pi^\natural(i)}-\mathbf{X}_j\right)^\top\mathbf{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_2}.
$$

For Λ_1, we conclude

$$
\begin{aligned}
\mathbb{E}\Lambda_1 &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}+\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\\
&\quad-\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_j^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}-\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{\Sigma}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}\\
&= \mathbb{1}_{i=\pi^\natural(i)}(p+3)\mathbb{E}\operatorname{Tr}(\mathbf{\Sigma}\mathbf{M})-\mathbb{1}_{i=j}(p+1)\mathbb{E}\operatorname{Tr}(\mathbf{\Sigma}\mathbf{M})\\
&= (n-h)\left(\mathbb{1}_{i=\pi^\natural(i)}(p+3)-\mathbb{1}_{i=j}(p+1)\right)\operatorname{Tr}(\mathbf{M})=\frac{p(n-h)^2}{n}\left[1+o(1)\right]\operatorname{Tr}(\mathbf{M}).
\end{aligned}
$$

Then we turn to $\mathbb{E}\Lambda_2$ and obtain

$$
\begin{aligned}
\mathbb{E}\Lambda_2 = {}& \mathbb{E}\underbrace{\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,1}}+\mathbb{E}\underbrace{\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,2}}\\
&-\mathbb{E}\underbrace{\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_j^\top\mathbf{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,3}}-\mathbb{E}\underbrace{\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{\Delta}\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\Lambda_{2,4}}.
\end{aligned}
$$

We compute the value of $\mathbb{E}\Lambda_2$ under the four different cases.

Case (s,s): $i=\pi^\natural(i)$ and $j=\pi^\natural(j)$. In this case, we have $\mathbf{\Delta}$ be

$$\boldsymbol{\Delta} = \boldsymbol{\Delta}^{(s,s)} = \mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top.$$

We have

$$\begin{aligned}
\mathbb{E}\Lambda_{2,1} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = (p+2)(p+5)\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,2} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = 2(p+2)\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,3} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_j^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = 0,\\
\mathbb{E}\Lambda_{2,4} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_i^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = 0,
\end{aligned}$$

which implies

$$\mathbb{E}\Lambda_2 = (p+2)\,(p+7)\operatorname{Tr}(\mathbf{M}).$$

Case (s,d): $i = \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. First we write $\boldsymbol{\Delta}$ as

$$\boldsymbol{\Delta}^{(s,d)} = \mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top.$$

Then we conclude

$$\begin{aligned}
\mathbb{E}\Lambda_{2,1} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_i^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = (p+2)(p+4)\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,2} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = (p+2)\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,3} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_i\mathbf{X}_j^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = 0,\\
\mathbb{E}\Lambda_{2,4} &= \mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_i^\top\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top\right)\mathbf{M}\mathbf{X}_i = 0,
\end{aligned}$$

which suggests that

$$\mathbb{E}\Lambda_2 = (p+2)(p+5)\operatorname{Tr}(\mathbf{M}).$$

Case (d,s): $i \neq \pi^\natural(i)$ and $j = \pi^\natural(j)$. In this case, $\boldsymbol{\Delta}$ reduces to

$$\boldsymbol{\Delta}^{(d,s)} = \mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top + \mathbf{X}_j\mathbf{X}_j^\top.$$

We have

$$\begin{aligned}
\mathbb{E}\Lambda_{2,1} =\ & \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\|\mathbf{X}_i\|_2^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_2^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i}\\
&+ \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} = (p+2)\left[1 + \mathbb{1}_{i=\pi^{\natural 2}(i)}\right]\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,2} =\ & \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{p\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M})}\\
&+ \underbrace{\mathbb{E}\mathbf{X}_i^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{X}_j\mathbf{X}_j^\top\mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} = \left(p + \mathbb{1}_{i=\pi^{\natural 2}(i)}\right)\operatorname{Tr}(\mathbf{M}),\\
\mathbb{E}\Lambda_{2,3} =\ & 0,\\
\mathbb{E}\Lambda_{2,4} =\ & 0,
\end{aligned}$$

which suggests

$$\mathbb{E}\Lambda_2 = 2(p+1)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=\pi^{\natural 2}(i)}\,(p+3)\operatorname{Tr}(\mathbf{M}).$$

Case (d,d): $i \neq \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. In this case, $\boldsymbol{\Delta}$ is written as

$$\boldsymbol{\Delta}^{(d,d)} = \mathbf{X}_i\mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_{\pi^\natural(i)}\mathbf{X}_{\pi^{\natural 2}(i)}^\top + \mathbf{X}_j\mathbf{X}_{\pi^\natural(j)}^\top + \mathbf{X}_{\pi^{\natural-1}(j)}\mathbf{X}_j^\top.$$

We have

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,1} = & \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i};
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,2} = & \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} p(p+2)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i\neq j} p \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \left[\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_2^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i + \mathbb{1}_{i\neq j} \operatorname{Tr}(\mathbf{M})\right]} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} p(p+2)\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=j} \mathbb{1}_{j=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}\|_2^2 \mathbf{X}^\top \mathbf{M}\mathbf{X}};
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,3} = & \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_j^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^\natural(j)} p \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_j^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0};
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,4} = & \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_{\pi^{\natural 2}(i)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0} \\
& + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(j)}^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{\mathbb{1}_{i=\pi^\natural(j)} \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\mathbf{X}_i^\top \mathbf{X}_j \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^{\natural-1}(j)} \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_{\pi^\natural(i)}}_{0}.
\end{aligned}
$$

Hence we conclude

$$
\begin{aligned}
\mathbb{E}\Lambda_2 = & \; 2(p+1)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j} 2\,(p+1)^2 \operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=\pi^\natural(j)}(p+1)\operatorname{Tr}(\mathbf{M}) \\
& + \mathbb{1}_{i=\pi^{\natural 2}(i)}(p+3)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}(3p+5)\operatorname{Tr}(\mathbf{M}).
\end{aligned}
$$

□

Lemma 9. *We have*

$$
\mathbb{E}\Xi_2\Xi_3 = \frac{p(n-h)(n+p-h)}{n}\operatorname{Tr}(\mathbf{M})\left[1+o(1)\right],
$$

where Ξ_2 and Ξ_3 are defined in (21).

Proof. To start with, we write the expectation as $\mathbb{E}\Xi_2\Xi_3$

$$
\begin{aligned}
\mathbb{E}\Xi_2\Xi_3 = & \; \mathbb{E}\underbrace{\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{B}^\natural}_{\boldsymbol{u}^\top} \mathbf{W}^\top \underbrace{\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)}_{\boldsymbol{p}\in\mathbb{R}^{n\times 1}} \mathbf{W}_i^\top \underbrace{\mathbf{B}^{\natural\top}\mathbf{X}^\top \mathbf{\Pi}^{\natural\top}\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)}_{\boldsymbol{v}} \\
= & \; \mathbb{E}\boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\mathbf{W}_i^\top \boldsymbol{v} = \mathbb{E}\left\langle \mathbf{W}_i, \boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}\right\rangle.
\end{aligned}
$$

Exploiting the independence among $\mathbf{X}$ and $\mathbf{W}$, we condition on $\mathbf{X}$ and have

$$
\mathbb{E}_{\mathbf{W}}\left\langle \mathbf{W}_i, \boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}\right\rangle = \mathbb{E}_{\mathbf{W}} \operatorname{Tr}\left(\nabla_{\mathbf{W}_i} \boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}\right).
$$

Note that only the diagonal entries of the Hessian matrix $\nabla_{\mathbf{W}_i} \boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}$ matters. For an arbitrary index s, we can compute the gradient of the s-th entry of $\boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}$ w.r.t. $\mathbf{W}_{i,s}$ as

$$
\frac{d}{dW_{i,s}}\left(\boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}_s\right) = \boldsymbol{v}_s \frac{d}{dW_{i,s}}\left(\boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\right) = \boldsymbol{v}_s \sum_{t=1}^n \frac{d}{dW_{i,s}}\left(\boldsymbol{p}_t \mathbf{W}_t^\top \boldsymbol{u}\right) = \boldsymbol{v}_s \frac{d}{dW_{i,s}} \boldsymbol{p}_i \mathbf{W}_i^\top \boldsymbol{u} = \boldsymbol{p}_i \boldsymbol{v}_s \boldsymbol{u}_s.
$$

Invoking the definitions of $\boldsymbol{p}$, $\boldsymbol{v}$ and $\boldsymbol{u}$, we have

$$\mathbb{E}_{\mathbf{W},\mathbf{X}}\left\langle \mathbf{W}_i, \boldsymbol{u}^\top \mathbf{W}^\top \boldsymbol{p}\boldsymbol{v}\right\rangle = \mathbb{E}_{\mathbf{X}}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}_i \sum_{s=1}^{m}\left[\mathbf{X}_{\pi^\natural(i)}^\top \left(\mathbf{B}^{\natural\top}\right)_s \left(\mathbf{B}^{\natural\top}\right)_s^\top \mathbf{X}^\top \mathbf{\Pi}^{\natural\top}\mathbf{X}\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\right]$$

$$\overset{①}{=} \underbrace{\mathbb{E}\left[\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\right]}_{\Lambda_1} + \underbrace{\mathbb{E}\left[\left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Delta}^\top \left(\mathbf{X}_{\pi^\natural(i)} - \mathbf{X}_j\right)\right]}_{\Lambda_2},$$

where in ① we use the relation $\sum_{s=1}^{m}\left(\mathbf{B}^{\natural\top}\right)_s\left(\mathbf{B}^{\natural\top}\right)_s^\top = \mathbf{B}^\natural\mathbf{B}^{\natural\top} = \mathbf{M}$.

For the first term Λ_1, we obtain

$$\begin{aligned}
\mathbb{E}\Lambda_1 =\ & \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=\pi^\natural(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=\pi^\natural(i)}\mathbb{E}\mathbf{X}_i^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_i} \\
& - \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=j}\mathbb{1}_{i\neq\pi^\natural(i)}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_{\pi^\natural(i)}} - \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{p\mathbb{1}_{i=j}\mathbb{1}_{i\neq\pi^\natural(i)}\mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Sigma}^\top \mathbf{X}_{\pi^\natural(i)}} \\
=\ & (n-h)\left[\mathbb{1}_{i=\pi^\natural(i)}(p+3) - (p+1)\mathbb{1}_{i=j}\mathbb{1}_{i\neq\pi^\natural(i)}\right]\mathrm{Tr}(\mathbf{M}).
\end{aligned}$$

Then we consider the second term Λ_2, which can be decomposed further into four sub-terms reading as

$$\begin{aligned}
\mathbb{E}\Lambda_2 =\ & \mathbb{E}\underbrace{\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Delta}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\Lambda_{2,1}} + \mathbb{E}\underbrace{\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Delta}^\top \mathbf{X}_j\right)}_{\Lambda_{2,2}} \\
& - \mathbb{E}\underbrace{\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Delta}^\top \mathbf{X}_j\right)}_{\Lambda_{2,3}} - \mathbb{E}\underbrace{\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\mathbf{\Delta}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\Lambda_{2,4}}.
\end{aligned}$$

Case (s,s): $i = \pi^\natural(i)$ and $j = \pi^\natural(j)$. In this case, we have $\mathbf{\Delta}$ be

$$\mathbf{\Delta}^{(s,s)} = \mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top.$$

Hence we conclude

$$\begin{aligned}
\mathbb{E}\Lambda_{2,1} =\ & \mathbb{E}\left[\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_i\right] = \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i + \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i \\
=\ & (p+2)(p+4)\,\mathrm{Tr}(\mathbf{M}) + (p+2)\,\mathrm{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,2} =\ & \mathbb{E}\left[\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_j\right] = \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i + \mathbb{E}\|\mathbf{X}_j\|_{\mathrm{F}}^2 \mathbf{X}_j^\top \mathbf{M}\mathbf{X}_j \\
=\ & 2\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i = 2(p+2)\,\mathrm{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,3} =\ & \mathbb{E}\left[\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_j\right] = 0, \\
\mathbb{E}\Lambda_{2,4} =\ & \mathbb{E}\left[\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_j\mathbf{X}_j^\top\right)\mathbf{X}_i\right] = 0,
\end{aligned}$$

which suggests $\mathbb{E}\Lambda_2 = (p+2)(p+7)\,\mathrm{Tr}(\mathbf{M})$.

Case (s,d): $i = \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. First we write $\mathbf{\Delta}$ as

$$\mathbf{\Delta}^{(s,d)\top} = \mathbf{X}_i\mathbf{X}_i^\top + \mathbf{X}_{\pi^\natural(j)}\mathbf{X}_j^\top + \mathbf{X}_j\mathbf{X}_{\pi^{\natural-1}(j)}^\top.$$

Then we conclude

$$\begin{aligned}
\mathbb{E}\Lambda_{2,1} =\ & \underbrace{\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_i\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_i\right)}_{0} \\
& + \underbrace{\mathbb{E}\left(\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_i\right)}_{0} = \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i = (p+2)(p+4)\,\mathrm{Tr}(\mathbf{M});
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,2} = & \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i \mathbf{X}_i^\top \mathbf{X}_j\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j\right)}_{0} \\
& + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j\right)}_{0} = \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i = (p+2)\operatorname{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,3} = & \ \mathbb{E}\left[\mathbf{X}_i^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i \mathbf{X}_i^\top + \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top + \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top\right) \mathbf{X}_j\right] = 0, \\
\mathbb{E}\Lambda_{2,4} = & \ \mathbb{E}\left[\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{M}\left(\mathbf{X}_i \mathbf{X}_i^\top + \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top + \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top\right) \mathbf{X}_i\right] = 0,
\end{aligned}
$$

which suggests $\mathbb{E}\Lambda_2 = (p+2)(p+5)\operatorname{Tr}(\mathbf{M})$.

Case (d,s): $i \neq \pi^\natural(i)$ and $j = \pi^\natural(j)$. In this case, $\mathbf{\Delta}$ reduces to

$$
\mathbf{\Delta}^{(d,s)\top} = \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top + \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_j \mathbf{X}_j^\top.
$$

Then we obtain

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,1} = & \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{0} = \left(1 + \mathbb{1}_{i=\pi^{\natural 2}(i)}\right)(p+2)\operatorname{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,2} = & \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i \neq j} p \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i + \mathbb{1}_{i \neq j} \operatorname{Tr}(\mathbf{M})\right]} \\
& + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_j^\top \mathbf{X}_j\right)}_{0} = \left(p + \mathbb{1}_{i=\pi^{\natural 2}(i)}\right)\operatorname{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,3} = & \ \mathbb{E}\left[\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\left(\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top + \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_j \mathbf{X}_j^\top\right) \mathbf{X}_j\right] = 0, \\
\mathbb{E}\Lambda_{2,4} = & \ \mathbb{E}\left[\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M}\left(\mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top + \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_j \mathbf{X}_j^\top\right) \mathbf{X}_{\pi^\natural(i)}\right] = 0,
\end{aligned}
$$

which suggests

$$
\mathbb{E}\Lambda_2 = 2(p+1)\operatorname{Tr}(\mathbf{M}) + (p+3)\mathbb{1}_{i=\pi^{\natural 2}(i)}\operatorname{Tr}(\mathbf{M}).
$$

Case (d,d): $i \neq \pi^\natural(i)$ and $j \neq \pi^\natural(j)$. In this case, $\mathbf{\Delta}$ is written as

$$
\mathbf{\Delta}^{(d,d)\top} = \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top + \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top + \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top + \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top.
$$

Then we have

$$
\begin{aligned}
\mathbb{E}\Lambda_{2,1} = & \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} \\
& + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=j} \mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} \\
= & \ (p+2)\left[1 + \mathbb{1}_{i=\pi^{\natural 2}(i)} + \mathbb{1}_{i=j} + \mathbb{1}_{i=\pi^{\natural 2}(i)} \mathbb{1}_{i=j}\right]\operatorname{Tr}(\mathbf{M}), \\
\mathbb{E}\Lambda_{2,2} = & \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=j} \mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \operatorname{Tr}(\mathbf{M}) + p \mathbb{1}_{i \neq j} \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=\pi^{\natural 2}(i)}[\mathbb{1}_{i=j}(p+2)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i \neq j} \operatorname{Tr}(\mathbf{M})]}
\end{aligned}
$$

$$
\begin{aligned}
&+ \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=j}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^4 \operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j\right)}_{\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)}\mathbb{E}\|\mathbf{X}_i\|_{\mathrm{F}}^2 \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i} \\
&= 2\mathbb{1}_{i=j} p(p+2)\operatorname{Tr}(\mathbf{M}) + p\mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=\pi^{\natural 2}(i)}\left[\mathbb{1}_{i=j} 2(p+2)\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i\neq j}\operatorname{Tr}(\mathbf{M})\right], \\
\mathbb{E}\Lambda_{2,3} = &\underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_j\right)}_{0} + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_j\right)}_{0} \\
&+ \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_j\right)}_{p\mathbb{1}_{i=\pi^\natural(j)}\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\left(\mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_j\right)}_{0}, \\
\mathbb{E}\Lambda_{2,4} = &\underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(i)} \mathbf{X}_i^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{0} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^{\natural 2}(i)} \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{0} \\
&+ \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_{\pi^\natural(j)} \mathbf{X}_j^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{\mathbb{1}_{i=\pi^\natural(j)}\operatorname{Tr}(\mathbf{M})} + \underbrace{\mathbb{E}\left(\mathbf{X}_j^\top \mathbf{X}_i \mathbf{X}_{\pi^\natural(i)}^\top \mathbf{M} \mathbf{X}_j \mathbf{X}_{\pi^{\natural-1}(j)}^\top \mathbf{X}_{\pi^\natural(i)}\right)}_{0},
\end{aligned}
$$

which gives

$$
\mathbb{E}\Lambda_2 = (p+1)\left[2 + 2(p+1)\mathbb{1}_{i=j} + \mathbb{1}_{i=\pi^\natural(j)}\right]\operatorname{Tr}(\mathbf{M}) + \mathbb{1}_{i=\pi^{\natural 2}(i)}\left[p + 3 + (3p+5)\mathbb{1}_{i=j}\right]\operatorname{Tr}(\mathbf{M}).
$$

□

D.2.3 Supporting Lemmas

First, we study the higher order expectations of Gaussian random vectors' inner product, which hopefully will serve independent interests.

Lemma 10. *Assume $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^p$ are Gaussian distributed random vectors whose entries follow the i.i.d. standard normal distribution, then we have*

$$
\begin{aligned}
\mathbb{E}\operatorname{Tr}\left(\boldsymbol{y}\boldsymbol{y}^\top \boldsymbol{x}\boldsymbol{x}^\top \mathbf{M}\right) &= \operatorname{Tr}(\mathbf{M}), && (63)\\
\mathbb{E}\operatorname{Tr}\left(\boldsymbol{y}\boldsymbol{y}^\top \boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right) &= \mathbb{E}\|\boldsymbol{y}\|_2^2 \operatorname{Tr}\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right) = p\operatorname{Tr}(\mathbf{M}), && (64)\\
\mathbb{E}\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 &= [\operatorname{Tr}(\mathbf{M})]^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top \mathbf{M}\right), && (65)\\
\mathbb{E}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) &= (p+2)\operatorname{Tr}(\mathbf{M}), && (66)\\
\mathbb{E}\|\boldsymbol{x}\|_2^4(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) &= (p+2)(p+4)\operatorname{Tr}(\mathbf{M}), && (67)\\
\mathbb{E}\|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 &= (p+4)\left[(\operatorname{Tr}(\mathbf{M}))^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top \mathbf{M}\right)\right], && (68)\\
\mathbb{E}\|\boldsymbol{x}\|_2^4\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 &= (p+4)(p+6)\left[(\operatorname{Tr}(\mathbf{M}))^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}) + \operatorname{Tr}\left(\mathbf{M}^\top \mathbf{M}\right)\right], && (69)\\
\mathbb{E}(\boldsymbol{x}^\top \boldsymbol{y})^2 \boldsymbol{y}^\top \mathbf{M}_1 \boldsymbol{x}\boldsymbol{x}^\top \mathbf{M}_2 \boldsymbol{y} &= 2\operatorname{Tr}(\mathbf{M}_1)\operatorname{Tr}(\mathbf{M}_2) + (p+4)\operatorname{Tr}(\mathbf{M}_1\mathbf{M}_2) + 2\operatorname{Tr}(\mathbf{M}_1\mathbf{M}_2^\top), && (70)
\end{aligned}
$$

where $\mathbf{M} \in \mathbb{R}^{p\times p}$ is a fixed matrix.

Remark 3. *If we assume $\mathbf{M} = \mathbf{I}_{p\times p}$, we can get $\mathbb{E}\|\boldsymbol{x}\|_2^4 = p(p+2)$, $\mathbb{E}\|\boldsymbol{x}\|_2^6 = p(p+2)(p+4)$, and $\mathbb{E}\|\boldsymbol{x}\|_2^8 = p(p+2)(p+4)(p+6)$.*

Proof. This lemma is proved by iteratively applying the Wick's theorem in Theorem 3, Stein's lemma in Lemma 19, and Lemma 18.

- **Proof of** (63) **and** (64). The proof can be conducted easily with the property such that $\operatorname{Tr}(\boldsymbol{u}\boldsymbol{v}^\top) = \boldsymbol{u}^\top \boldsymbol{v} = \operatorname{Tr}(\boldsymbol{v}\boldsymbol{u}^\top)$ holds for arbitrary vectors $\boldsymbol{u}$ and $\boldsymbol{v}$.
- **Proof of** (65). This property is a direct consequence of Neudecker & Wansbeek (1987) (Equation (3.2)), which is attached in Lemma 18 for the sake of self-containing.

- **Proof of** (66). Invoking the Stein's lemma, we have

$$\mathbb{E}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) = \mathbb{E}\left[\nabla_{\boldsymbol{x}}(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\boldsymbol{x}\right].$$

Then our goal transforms to computing the trace of the Hessian matrix $\operatorname{Tr}\left[\nabla_{\boldsymbol{x}} \operatorname{Tr}(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\boldsymbol{x}\right]$. For the i-th entry of the gradient, we have

$$\frac{d}{dx_i}\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x} = \langle \mathbf{M}_i, \boldsymbol{x}\rangle + \langle (\mathbf{M}^\top)_i, \boldsymbol{x}\rangle,$$

where $\mathbf{M}_i$ is the i-th row (or column) of $\mathbf{M}$. Then we obtain

$$\frac{d}{dx_i}\left[x_i \operatorname{Tr}(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\right] = \boldsymbol{x}^\top \mathbf{M}\boldsymbol{x} + x_i\left[\langle \mathbf{M}_i, \boldsymbol{x}\rangle + \langle (\mathbf{M}^\top)_i, \boldsymbol{x}\rangle\right],$$

and hence

$$\begin{aligned}\mathbb{E}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) =& \sum_{i=1}^p \mathbb{E}(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) + \sum_{i=1}^p \mathbb{E}\left[x_i\left(\langle \mathbf{M}_i, \boldsymbol{x}\rangle + \langle (\mathbf{M}^\top)_i, \boldsymbol{x}\rangle\right)\right] \\ =& \; p\operatorname{Tr}(\mathbf{M}) + 2\sum_i \mathbf{M}_{ii} = (p+2)\operatorname{Tr}(\mathbf{M}).\end{aligned}$$

- **Proof of** (67). Following the same strategy as in proving (66), we have

$$\mathbb{E}\|\boldsymbol{x}\|_{\mathrm{F}}^4(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) = \mathbb{E}\left[\nabla_{\boldsymbol{x}}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\boldsymbol{x}\right].$$

Then our goal transforms to computing the trace of the Hessian matrix $\operatorname{Tr}\left[\nabla_{\boldsymbol{x}}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\boldsymbol{x}\right]$. For the i-th entry of the gradient, we obtain

$$\begin{aligned}\frac{d}{dx_i}\left[x_i\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\right] =& \|\boldsymbol{x}\|_2^2\cdot(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) + x_i\frac{d}{dx_i}\left[\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\right] \\ =& \; \|\boldsymbol{x}\|_2^2\cdot(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) + 2x_i^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) + x_i\|\boldsymbol{x}\|_2^2\left[\langle \mathbf{M}_i, \boldsymbol{x}\rangle + \langle (\mathbf{M}^\top)_i, \boldsymbol{x}\rangle\right],\end{aligned}$$

whose expectation reads as

$$\begin{aligned}&\mathbb{E}\|\boldsymbol{x}\|_2^2\cdot(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}) + 2\mathbb{E}\left[x_i^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\right] + \mathbb{E}x_i\|\boldsymbol{x}\|_2^2\left[\langle \mathbf{M}_i, \boldsymbol{x}\rangle + \langle (\mathbf{M}^\top)_i, \boldsymbol{x}\rangle\right] \\ =& \; (p+2)\operatorname{Tr}(\mathbf{M}) + 2\mathbb{E}\left[x_i^2 M_{ii}x_i^2 + x_i^2\left(\sum_{j\neq i} M_{jj}x_j^2\right)\right] + 2M_{ii}\mathbb{E}\left[x_i^2\|\boldsymbol{x}\|_2^2\right] \\ =& \; (p+2)\operatorname{Tr}(\mathbf{M}) + 2M_{ii}\left(\mathbb{E}x_i^4\right) + 2\sum_{j\neq i} M_{jj}(\mathbb{E}x_i^2)(\mathbb{E}x_j^2) + 2M_{ii}\left[\mathbb{E}(x_i^4) + \sum_{j\neq i}(\mathbb{E}x_i^2)(\mathbb{E}x_j^2)\right] \\ =& \; (p+2)\operatorname{Tr}(\mathbf{M}) + 6M_{ii} + 2\sum_{j\neq i} M_{jj} + 2M_{ii}(3+p-1) \\ =& \; (p+4)\operatorname{Tr}(\mathbf{M}) + 2(p+4)M_{ii}.\end{aligned}$$

Then we conclude

$$\begin{aligned}\mathbb{E}\operatorname{Tr}\left[\nabla_{\boldsymbol{x}}\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x})\boldsymbol{x}\right] =& \; p(p+2)\operatorname{Tr}(\mathbf{M}) + 2(p+4)\sum_i M_{ii} + 2p\operatorname{Tr}(\mathbf{M}) \\ =& \; (p+2)(p+4)\operatorname{Tr}(\mathbf{M}).\end{aligned}$$

- **Proof of** (68)**.** Invoking the Stein's lemma, we have

$$\mathbb{E}\|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 = \sum_i \frac{d}{dx_i}\left[x_i\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2\right] = p\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 + 4\sum_i x_i\left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle.$$

The proof is then completed by invoking Lemma 11.

- **Proof of** (69)**.** Following the same strategy as in proving (68), we consider the i-th gradient w.r.t x_i, which can be written as

$$\begin{aligned}
&\mathbb{E}\|\boldsymbol{x}\|_2^4 \left(\boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}\right)^2 = \sum_i \frac{d}{dx_i}\left[x_i\|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2\right] \\
= &\sum_i \|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2 + 2\sum_i x_i^2\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2 + 4\sum_i x_i\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x})\left\langle \mathbf{M}^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle \\
= &\sum_i \|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2 + 2\sum_i \mathbb{E}\frac{d}{dx_i}\left[x_i\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2\right] + 4\sum_i \mathbb{E}\frac{d}{dx_i}\left[\|\boldsymbol{x}\|_2^2(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x})\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle\right].
\end{aligned} \tag{71}$$

  Noticing the following relations

$$\frac{d}{dx_i}\left[x_i\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2\right] = \left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)^2 + 4x_i\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle, \tag{72}$$

$$\begin{aligned}
\frac{d}{dx_i}\left[\|\boldsymbol{x}\|_2^2\operatorname{Tr}(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x})\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle\right] = &\ 2x_i\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle \\
&+ 2\|\boldsymbol{x}\|_2^2\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle^2 + M_{ii}\|\boldsymbol{x}\|_2^2\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right),
\end{aligned} \tag{73}$$

  we can conclude the proof by combining (68), (71), (72), (73), and Lemma 11.

- **Proof of** (70)**.** Due to the independence between $\boldsymbol{x}$ and $\boldsymbol{y}$, we first condition on $\boldsymbol{x}$ and have

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{x}^\top\boldsymbol{y})^2\boldsymbol{y}^\top\mathbf{M}_1\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2\boldsymbol{y} = \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{y}}\boldsymbol{y}^\top\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{y}\boldsymbol{y}^\top\mathbf{M}_1\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2\boldsymbol{y} \\
\overset{①}{=}\ &\mathbb{E}_{\boldsymbol{x}}\operatorname{Tr}\left(\boldsymbol{x}\boldsymbol{x}^\top\right)\operatorname{Tr}\left(\mathbf{M}_1\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2\right) + \mathbb{E}_{\boldsymbol{x}}\operatorname{Tr}\left(\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_1\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2\right) + \mathbb{E}_{\boldsymbol{x}}\operatorname{Tr}\left(\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2^\top\boldsymbol{x}\boldsymbol{x}\mathbf{M}_1^\top\right) \\
=\ &\mathbb{E}_{\boldsymbol{x}}\|\boldsymbol{x}\|_2^2\boldsymbol{x}^\top\mathbf{M}_2\mathbf{M}_1\boldsymbol{x} + \mathbb{E}_{\boldsymbol{x}}\boldsymbol{x}^\top\mathbf{M}_1\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2\boldsymbol{x} + \mathbb{E}_{\boldsymbol{x}}\boldsymbol{x}^\top\mathbf{M}_1^\top\boldsymbol{x}\boldsymbol{x}^\top\mathbf{M}_2^\top\boldsymbol{x} \\
\overset{②}{=}\ &2\operatorname{Tr}(\mathbf{M}_1)\operatorname{Tr}(\mathbf{M}_2) + (p+4)\operatorname{Tr}(\mathbf{M}_1\mathbf{M}_2) + 2\operatorname{Tr}(\mathbf{M}_1\mathbf{M}_2^\top),
\end{aligned}$$

  where in ① and ② we both use Lemma 18.

□

Lemma 11. *For a fixed matrix $\mathbf{M} \in \mathbb{R}^{p\times p}$, we associate it with a symmetric matrix $\mathbf{M}^{(\mathsf{sym})}$ defined as $(\mathbf{M}+\mathbf{M}^\top)/2$. Consider the Gaussian distributed random vector $\boldsymbol{x} \sim \mathsf{N}(\mathbf{0}, \mathbf{I})$, we have*

$$\mathbb{E}\sum_i x_i(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x})\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle = (\operatorname{Tr}(\mathbf{M}))^2 + \|\mathbf{M}\|_{\mathrm{F}}^2 + \operatorname{Tr}(\mathbf{M}\mathbf{M}).$$

Proof. This lemma is a direct application of Wick's theorem, which is completed by showing

$$\begin{aligned}
&\mathbb{E}x_i\left(\boldsymbol{x}^\top\mathbf{M}\boldsymbol{x}\right)\left\langle \mathbf{M}_i^{(\mathsf{sym})}, \boldsymbol{x}\right\rangle = \mathbb{E}\sum_j\sum_{\ell_1,\ell_2} M_{i,j}^{(\mathsf{sym})}M_{\ell_1,\ell_2}x_ix_jx_{\ell_1}x_{\ell_2} \\
= &\underbrace{\mathbb{E}\sum_j\sum_{\ell_1,\ell_2}\mathbb{1}_{\ell_1=i}\mathbb{1}_{\ell_2=j}M_{i,j}^{(\mathsf{sym})}M_{\ell_1,\ell_2}x_ix_jx_{\ell_1}x_{\ell_2}}_{\sum_j M_{i,j}^{(\mathsf{sym})}M_{i,j}} + \underbrace{\mathbb{E}\sum_j\sum_{\ell_1,\ell_2}\mathbb{1}_{\ell_2=i}\mathbb{1}_{\ell_1=j}M_{i,j}^{(\mathsf{sym})}M_{\ell_1,\ell_2}x_ix_jx_{\ell_1}x_{\ell_2}}_{\sum_j M_{i,j}^{(\mathsf{sym})}M_{j,i}} \\
&+ \underbrace{\mathbb{E}\sum_j\sum_{\ell_1,\ell_2}\mathbb{1}_{i=j}\mathbb{1}_{\ell_1=\ell_2}M_{i,j}^{(\mathsf{sym})}M_{\ell_1,\ell_2}x_ix_jx_{\ell_1}x_{\ell_2}}_{\sum_\ell M_{i,i}^{(\mathsf{sym})}M_{\ell,\ell}} = \sum_j 2\left[M_{i,j}^{(\mathsf{sym})}\right]^2 + M_{i,i}\operatorname{Tr}(\mathbf{M}),
\end{aligned}$$

where $\mathbf{M}^{(\mathsf{sym})}$ is defined as $\left(\mathbf{M}+\mathbf{M}^\top\right)/2$. □

Then we study the properties of $\boldsymbol{\Sigma}$, which is defined as $\mathbf{X}^\top\boldsymbol{\Pi}^\natural\mathbf{X} - \boldsymbol{\Delta}$.

Lemma 12. *For a fixed matrix* $\mathbf{M}$*, we have*

$$\mathbb{E}\operatorname{Tr}\left(\mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma}^\top\right) = n^2\left[\frac{p}{n} + \left(1-\frac{h}{n}\right)^2 + o(1)\right]\operatorname{Tr}(\mathbf{M}),$$

where matrix $\mathbf{\Sigma}$ *is defined in* (22).

Proof. We conclude the proof by showing

$$\begin{aligned}
\mathbb{E}\operatorname{Tr}\left(\mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma}^\top\right) &\overset{①}{=} \sum_{\ell_1,\ell_2\in\mathcal{S}}\mathbb{E}\operatorname{Tr}\left[\mathbf{X}_{\ell_1}\mathbf{X}_{\ell_1}^\top\mathbf{M}\mathbf{X}_{\ell_2}\mathbf{X}_{\ell_2}^\top\right] + \sum_{\ell_1,\ell_2\in\mathcal{D}}\mathbb{E}\operatorname{Tr}\left[\mathbf{X}_{\ell_1}\mathbf{X}_{\pi^\natural(\ell_1)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(\ell_2)}\mathbf{X}_{\ell_2}^\top\right] \\
&= \sum_{\ell\in\mathcal{S}}\mathbb{E}\operatorname{Tr}\left(\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\right) + \sum_{\ell_1,\ell_2\in\mathcal{S},\ell_1\neq\ell_2}\mathbb{E}\operatorname{Tr}\left[\mathbf{X}_{\ell_1}\mathbf{X}_{\ell_1}^\top\mathbf{M}\mathbf{X}_{\ell_2}\mathbf{X}_{\ell_2}^\top\right] \\
&\quad + \sum_{\ell\in\mathcal{D}}\underbrace{\mathbb{E}\operatorname{Tr}\left[\mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(\ell)}\mathbf{X}_\ell^\top\right]}_{p\operatorname{Tr}(\mathbf{M})} + \sum_{(\ell_1,\ell_2)\in\mathcal{D}_{\text{pair}}}\underbrace{\mathbb{E}\operatorname{Tr}\left[\mathbf{X}_{\ell_1}\mathbf{X}_{\ell_2}^\top\mathbf{M}\mathbf{X}_{\ell_1}\mathbf{X}_{\ell_2}^\top\right]}_{\operatorname{Tr}(\mathbf{M})} \\
&= (n-h)(p+2)\operatorname{Tr}(\mathbf{M}) + (n-h)(n-h-1)\operatorname{Tr}(\mathbf{M}) + hp\operatorname{Tr}(\mathbf{M}) + |\mathcal{D}_{\text{pair}}|\operatorname{Tr}(\mathbf{M}) \\
&\overset{②}{=} n^2\left[\frac{p}{n} + \left(1-\frac{h}{n}\right)^2 + o(1)\right]\operatorname{Tr}(\mathbf{M}),
\end{aligned}$$

where ① is due to the definitions of index sets $\mathcal{S}$ and $\mathcal{D}$ (Equation (27) and in Equation (28)), and ② is because $|\mathcal{D}_{\text{pair}}| \leq h$. □

Lemma 13. *For a fixed matrix* $\mathbf{M}$*, we have*

$$\mathbb{E}\operatorname{Tr}(\mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma}\mathbf{M}) = (n-h+|\mathcal{D}_{\text{pair}}|)\left[\operatorname{Tr}(\mathbf{M})\right]^2 + (n-h)^2\operatorname{Tr}(\mathbf{M}\mathbf{M}) + n\operatorname{Tr}\left(\mathbf{M}\mathbf{M}^\top\right)$$

Proof. Following the same strategy as in proving Lemma 12, we complete the proof by showing

$$\begin{aligned}
\mathbb{E}\operatorname{Tr}(\mathbf{\Sigma}\mathbf{M}\mathbf{\Sigma}\mathbf{M}) &= \sum_{\ell=\pi^\natural(\ell)}\mathbb{E}\operatorname{Tr}\left(\mathbf{X}_\ell^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{M}\mathbf{X}_\ell\right) + \sum_{\ell_1,\ell_2\in\mathcal{S},\ell_1\neq\ell_2}\mathbb{E}\operatorname{Tr}\left(\mathbf{X}_{\ell_1}\mathbf{X}_{\ell_1}^\top\mathbf{M}\mathbf{X}_{\ell_2}\mathbf{X}_{\ell_2}^\top\mathbf{M}\right) \\
&\quad + \sum_{\ell\in\mathcal{D}}\mathbb{E}\operatorname{Tr}\left(\mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\right) + \sum_{\ell\in\mathcal{D}_{\text{pair}}}\mathbb{E}\operatorname{Tr}\left(\mathbf{X}_\ell\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(\ell)}\mathbf{X}_\ell^\top\mathbf{M}\right) \\
&= (n-h+|\mathcal{D}_{\text{pair}}|)\left[\operatorname{Tr}(\mathbf{M})\right]^2 + (n-h)^2\operatorname{Tr}(\mathbf{M}\mathbf{M}) + n\operatorname{Tr}\left(\mathbf{M}\mathbf{M}^\top\right).
\end{aligned}$$

□

Lemma 14. *For a fixed matrix* $\mathbf{M}$*, we have*

$$\mathbb{E}\left[\operatorname{Tr}(\Sigma\mathbf{M})\right]^2 = (n-h)^2\left[\operatorname{Tr}(\mathbf{M})\right]^2 + n\operatorname{Tr}\left(\mathbf{M}^\top\mathbf{M}\right) + (n-h+|\mathcal{D}_{\text{pair}}|)\operatorname{Tr}(\mathbf{M}\mathbf{M}).$$

Proof. We complete the proof by showing

$$\begin{aligned}
\mathbb{E}\left(\operatorname{Tr}(\Sigma\mathbf{M})\right)^2 &= \sum_{\ell\in\mathcal{S}}\mathbb{E}\left(\mathbf{X}_\ell^\top\mathbf{M}\mathbf{X}_\ell\right)^2 + \sum_{\ell_1,\ell_2\in\mathcal{S},\ell_1\neq\ell_2}\left(\operatorname{Tr}(\mathbf{M})\right)^2 + \sum_{\ell\in\mathcal{D}}\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{M}^\top\mathbf{X}_{\pi^\natural(\ell)}}_{\operatorname{Tr}(\mathbf{M}^\top\mathbf{M})} \\
&\quad + \sum_{\ell\in\mathcal{D}_{\text{pair}}}\underbrace{\mathbb{E}\mathbf{X}_{\pi^\natural(\ell)}^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{M}\mathbf{X}_{\pi^\natural(\ell)}}_{\operatorname{Tr}(\mathbf{M}\mathbf{M})} \\
&= (n-h)^2\left[\operatorname{Tr}(\mathbf{M})\right]^2 + n\operatorname{Tr}\left(\mathbf{M}^\top\mathbf{M}\right) + (n-h+|\mathcal{D}_{\text{pair}}|)\operatorname{Tr}(\mathbf{M}\mathbf{M}).
\end{aligned}$$

□

Lemma 15. *For a fixed matrix* $\mathbf{M}$*, we have*

$$\mathbb{E}\sum_{\ell=\pi^\natural(\ell)}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{X}_{\pi^\natural(i)} = (n-h)\operatorname{Tr}(\mathbf{M}) + (p+1)\mathbb{1}_{i=\pi^\natural(i)}\operatorname{Tr}(\mathbf{M}).$$

Proof. Provided that $i = \pi^\natural(i)$, we have

$$\begin{aligned}\mathbb{E}\Xi_{1,1} &= \mathbb{E}\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{X}_i + \sum_{\ell\neq i,\ \ell=\pi^\natural(\ell)} \mathbb{E}\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{X}_i \\ &= (p+2)\operatorname{Tr}(\mathbf{M}) + \sum_{\ell\neq i,\ \ell=\pi^\natural(\ell)} \operatorname{Tr}(\mathbf{M}) = (n-h+p+1)\operatorname{Tr}(\mathbf{M})\mathbb{1}_{i=\pi^\natural(i)}. \qquad (74)\end{aligned}$$

Provided that $i \neq \pi^\natural(i)$, we have

$$\mathbb{E}\Xi_{1,1} = \sum_{\ell=\pi^\natural(\ell)} \mathbb{E}\mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_\ell\mathbf{X}_\ell^\top\mathbf{X}_{\pi^\natural(i)} = (n-h)\operatorname{Tr}(\mathbf{M})\mathbb{1}_{i\neq\pi^\natural(i)}. \qquad (75)$$

Combining (74) and (75) then completes the proof. □

Lemma 16. *For a fixed* $\mathbf{M}$*, we have*

$$\mathbb{E}\sum_\ell \mathbf{X}_{\pi^\natural(i)}^\top\mathbf{M}\mathbf{X}_{\pi^\natural(\ell)}\mathbf{X}_\ell^\top\mathbf{X}_j = \left(p\mathbb{1}_{i=j}\mathbb{1}_{i\neq\pi^\natural(i)} + \mathbb{1}_{j=\pi^{\natural 2}(i)}\right)\operatorname{Tr}(\mathbf{M}).$$

We omit its proof as it is a direct application of Wick's theorem (Theorem 3).

Lemma 17. *We have*

$$\begin{aligned}\mathbb{E}\mathbb{1}(i=\pi^\natural(i)) &= \frac{n-h}{n}(1+o_{\mathrm{P}}(1)),\\ \mathbb{E}\mathbb{1}_{i=j} &= \frac{h}{n^2}(1+o_{\mathrm{P}}(1)),\\ \mathbb{E}\mathbb{1}_{j=\pi^{\natural 2}(i)} &= \frac{h}{n^2}(1+o_{\mathrm{P}}(1)),\\ \mathbb{E}\mathbb{1}_{i=j}\mathbb{1}_{i=\pi^{\natural 2}(i)} &= \frac{|\mathcal{D}_{\text{pair}}|}{n^2}(1+o_{\mathrm{P}}(1)).\end{aligned}$$

This lemma can be easily proved by assuming the indices i, j, $\pi^\natural(i)$, and $\pi^\natural(j)$ are uniformly sampled from the set $\{1, 2, \cdots, n\}$

E Useful Facts

This section collects some useful facts for the sake of self-containing.

Theorem 3 (Wick's theorem (Theorem 1.28 in Janson (1997))). *Considering the centered jointly normal variables* $g_1, g_2, \cdots, g_n$*, we conclude*

$$\mathbb{E}\left(g_1 g_2 \cdots g_n\right) = \sum_{\substack{\text{all possible disjoint}\\ \text{pairs } (i_k, j_k)\in\{1,2,\cdots,n\}}} \prod_k \mathbb{E}\left(g_{i_k} g_{j_k}\right).$$

With Wick's theorem, we can reduce the computation of high-order Gaussian moments to calculating the expectations of a series of low-order Gaussian moments.

Lemma 18 (Equation (3.2) in Neudecker & Wansbeek (1987)). *For a normally distributed random matrix* $\mathbf{G} \in \mathbb{R}^{n\times p}$ *which satisfies* $\mathbb{E}\mathbf{G} = \mathbf{0}$ *and* $\mathbb{E}\operatorname{vec}(\mathbf{G})\operatorname{vec}(\mathbf{G})^\top = \mathbf{U}\otimes\mathbf{V}$*, we have*

$$\begin{aligned}\mathbb{E}\left(\mathbf{G}^\top\mathbf{A}\mathbf{G}\mathbf{C}\mathbf{G}^\top\mathbf{B}\mathbf{G}\right) = &\operatorname{Tr}(\mathbf{A}\mathbf{U})\operatorname{Tr}(\mathbf{B}\mathbf{U})\mathbf{V}\mathbf{C}\mathbf{V} + \operatorname{Tr}\left(\mathbf{A}\mathbf{U}\mathbf{B}^\top\mathbf{U}\right)\mathbf{V}\mathbf{C}^\top\mathbf{V}\\ &+ \operatorname{Tr}(\mathbf{A}\mathbf{U}\mathbf{B}\mathbf{U})\operatorname{Tr}(\mathbf{C}\mathbf{V})\mathbf{V},\end{aligned}$$

where $\operatorname{vec}(\cdot)$ *is the vector operation;* $\otimes$ *is the Kronecker product (Horn & Johnson, 1990); and* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$ *are arbitrary fixed matrices.*

Lemma 19 (Stein's Lemma (cf. Section 1.3 in Talagrand (2010))). *Let* $g \sim \mathsf{N}(0,1)$*. Then for any differentiable function* $f : \mathbb{R} \mapsto \mathbb{R}$ *we have*

$$\mathbb{E}[gf(g)] = \mathbb{E}f^{'}(g),$$

where $\lim_{\|g\|\to\infty} f(g)e^{-a\|g\|_2^2} = 0$ *for any* $a > 0$.