Gaussian Approximation for High-Dimensional Second-Order U and V-statistics with Size-Dependent Kernels under i.n.i.d. Sampling

Shunsuke Imai *[1]

[1]Graduate School of Economics, Kyoto University

November 13, 2025

Abstract

We develop Gaussian approximations for high-dimensional vectors formed by second-order U and V-statistics whose kernels depend on sample size under independent but not identically distributed (i.n.i.d.) sampling. Our results hold irrespective of which component of the Hoeffding decomposition is dominant, thereby covering both non-degenerate and degenerate regimes as special cases. By allowing i.n.i.d. sampling, the class of statistics we analyze includes weighted U- and V-statistics and two-sample U- and V-statistics as special cases, which cover estimators of parameters in regression models with many covariates, many-weak instruments as well as a broad class of smoothed two-sample tests, to name but a few. In addition, we extend sharp maximal inequalities for U-statistics with size-dependent kernels from the i.i.d. to the i.n.i.d. setting, which may be of independent interest.

Keywords: Adaptive test; high-dimensional central limit theorem; many covariates; many-weak instrumental variables; maximal inequalities; Stein's method; U-statistics.

*This research was supported by JSPS KAKENHI Grant Number 24KJ1472. I am grateful to Yuta Koike for numerous invaluable comments on many aspects of this paper. All errors are entirely my own.

1 Introduction

In modern applications, the number of target parameters of statistical inference can be large and one might wish to conduct simultaneous inference on such high-dimensional parameters, as noted by many authors (e.g. Belloni et al., 2018; Chernozhukov et al., 2022, 2023). For example, such situations arise when there are many outcomes, groups or time points (or combination thereof), and parameters are estimated for each outcome, group, and time point (Belloni et al., 2018, Section 1.1); when economic theory implies a large number of testable conditions (e.g. Chernozhukov et al., 2019); or when one seeks to conduct uniform inference over some parameters for the purpose of adaptive inference (Horowitz and Spokoiny, 2001), sensitivity analysis/robustness check Armstrong and Kolesár (2018) or post-selection inference Kuchibhotla et al. (2022).

As a useful tool for such problems, Chernozhukov et al. (2013) established Gaussian approximation for the sum of high-dimensional independent random vectors. Notably, their approximation is valid even when the dimension p is much larger than the sample size n and does not restrict correlations of the coordinates of the random vectors. Since then, many authors have extended the remarkable result to several directions including U-statistics type statistics (Chen, 2018; Chen and Kato, 2019, 2020; Song et al., 2019, 2023; Cheng et al., 2022; Chiang et al., 2023; Koike, 2023; Imai and Koike, 2025). Among them, the most closely related work to ours is Imai and Koike (2025). In Imai and Koike (2025), they provide Gaussian approximations for high-dimensional with size-dependent kernels under i.i.d. sampling. The point of their results is that the approximations remain valid regardless of which term in the Hoeffding decomposition is dominant and even in the case such dominant component is absent.

In this study, we establish Gaussian approximations for high-dimensional second-order U and V-statistics with kernel functions which depend on sample size under independent but not identically distributed (i.n.i.d.) sampling. Our bounds on Gaussian approximation errors are analytical and explicitly express the dependence on the dimension p and remain valid regardless of whether the first- or second-order Hoeffding component dominates or such dominant component does not exist. As special cases, we allow non-degenerate cases and degenerate cases with known degrees.

At first glance, the results in this paper appear to be straightforward extensions of the Gaussian approximation results for high-dimensional U-statistics under the i.i.d. setting by Imai and Koike (2025) to the i.n.i.d. setting. However, allowing i.n.i.d. observations substantially enlarges the class of applications. First, the extension enables us to directly handles settings where each observation can have its own distribution. Second and more importantly, the resulting U-statistics have important subclasses such as weighted U-statistics and two-sample U-statistics, etc (See Section 4).

Two-sample U-statistics typically appears as test statistics for the equivalence of parameters or underlying distributions across two samples, and for related problems; for example, our considered class of U-statistics covers the smoothed version of classical statistics for two sample problem such as Mann–Whitney and Cramér-von-Mises type statistics and their extensions. Among them, in Section 4.2, we consider a maximum mean discrepancy (MMD) based nonparametric adaptive homogeneity test (Gretton et al., 2012) which concerns whether two independent samples come from a common population as an application of our Gaussian approximation for high-dimensional U-statistics with size-dependent kernels. The high dimensionality in this application is induced by adaptation; In the framework of adaptive tests, we construct test statistic uniformly in smoothness over some function space, because smoothness of some functions associated with data generating processes are typically unknown (cf. Chetverikov et al., 2021; Imai and Koike, 2025). In our application in Section 4.2, the resulting test statistic is defined as maximum of single test statistic over the space of bandwidths whose cardinal diverges to infinity as the sample size grows.

In the recent econometrics and statistics, the high-dimensional covariates and many-weak instrumental variables (IVs) settings are commonplace and second-order weighted V and U-statistics play an important role in such settings. We first explain the motivation for using many covariates. As noted by (Cattaneo et al., 2018a, Section 1), one of the motivations using many covariates is to make unconfundedness assumption more credible by conditioning on a rich set of controls (potentially including interactions or/and non-linear functional of covariates and fixed effects) in observational studies. Another motivation is efficiency gain in randomized controlled trials (Imbens and Rubin, 2015, Section 7) and the theoretical analysis have been conducted under high-dimensional covariates settings Lei and Ding (2021); Chiang et al. (2025). We next document that many-weak IVs situations frequently occur in practice. In terms of the strength of IVs, Andrews et al. (2019) and Lee et al. (2022) document that a nontrivial share of papers in American Economic Review report F-statistics below the conventional threshold of 10 and conclude that weak instruments are frequently encountered in practice. Also, many-IV situations frequently arise in application, for example, Quarter-of-Birth in the famous paper Angrist and Krueger (1991), Hausman IVs (Berry et al., 1995; Nevo, 2001), judge IVs (Kling, 2006; Dahl et al., 2014; Dobbie et al., 2018; Mikusheva and Sun, 2022; Frandsen et al., 2023), Bartik IVs (Adao et al., 2019; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022, 2025), wind direction IVs (Deryugina et al., 2019; Bondy et al., 2020), granular IVs (Gabaix and Koijen, 2024) and Mendelian randomization (Davies et al., 2015) etc. Motivated by such surroundings, as concrete examples of application of weighted V and U-statistics, we consider the many-weak IVs asymptotics by Chao et al. (2012) and the many covariate asymptotics by Cattaneo et al. (2018a) (also closely related to Cattaneo et al., 2018b; Jochmans, 2022), respectively. In this framework, the estimators of parameters of interest have second-order V and U-statistic forms and whether the linear term or the quadratic term is dominant depends on the assumption. In particular, in the many-weak IV asymptotics, it depends on the ratio of IVs to the dimension of IVs. Also, in the many-covariate asymptotics by Cattaneo et al. (2018a), the ratio of the number of covariates to the sample size. One advantage of these asymptotic frameworks is that the required restrictions on these ratios are relaxed relative to classical asymptotic regime. In addition, Cattaneo et al. (2024) show that Gaussian approximations is more accurate in the case of density weighted average derivatives estimators (Powell et al., 1989) under an asymptotic framework called small bandwidth asymptotics (Cattaneo et al., 2014), whose mathematical structure is shared by many-covariate and many-weak IV asymptotics (Cattaneo et al., 2018a, Section 2.1-2.2 and Cattaneo et al., 2024, Section 5), we expect comparable improvements. In Section 4.1, we suppose that the high-dimensionality comes from a large number of outcomes groups and/or moment conditions, etc following (Belloni et al., 2018, Section 1).

On the technical side, we strongly rely on technical lemmas and the proof of Theorem 2 in Imai and Koike (2025). In particular, we directly utilize their high-dimensional central limit theorem via exchangeable pairs (Imai and Koike, 2025, Theorem 6) as the starting point for our analysis of Gaussian approximation error. However, in evaluating the Gaussian approximation error, Imai and Koike (2025) make substantial use of auxiliary results by Döbler and Peccati (2019) and several lemmas they develop, some of which can be applied only in the i.i.d. setting. Accordingly, we reconstruct parts of the argument and several auxiliary lemmas tailored to the i.n.i.d. setting. In the i.n.i.d. case, the underlying probability measures are not common across observations and this loss of common measure symmetry makes the multi-index summations and σ-field conditioning substantially more delicate for the analysis of U-statistics. Among these, the most important technical contribution is an extension of the maximal inequalities for U-statistics with size dependent kernel (Imai and Koike, 2025, Theorem 7 and 8) to the i.n.i.d. setting. We state these results as Theorem 2 and Theorem 3 and they may be of independent interest beyond our application. Nevertheless, as noted above, several technical lemmas due to Imai and Koike (2025) still play an important role. In particular, we can directly use their maximal inequalities for high-dimensional nonnegative adapted sequences and for high-dimensional martingales (Imai and Koike, 2025, Lemma 1 and 2), with due care in the construction of the filtration, to retain the Gaussian-approximation error bounds as sharp as in the i.i.d. case.

Organization: The rest of the paper is organized as follows. Sections 2 and 3 introduce the formal setup and state the main theoretical results, respectively. Section 4 discusses several concrete examples of potential applications. Section A presents the two key building blocks of the proofs of main results, and the proofs of main results are in Section B. In Section C, we prove the auxiliary results.

Notations: For a positive integer m, we write $[m] := \{1, \dots, m\}$. We also set $[0] := \emptyset$ by convention. Given a vector $x \in \mathbb{R}^p$, its j-th component is denoted by x_j. Also, we set $|x| := \sqrt{\sum_{j=1}^p x_j^2}$ and $\|x\|_\infty := \max_{j\in[p]} |x_j|$. Given a $p \times q$ matrix A, its (j, k)-th entry is denoted by A_{jk}. Also, we set $\|A\|_\infty := \max_{j\in[p],k\in[q]} |A_{jk}|$. $\mathcal{R}_p$ denotes the set of all rectangles in $\mathbb{R}^p$. For a normed space $\mathfrak{X}$, its norm is denoted by $\|\cdot\|_{\mathfrak{X}}$. We interpret $\max\emptyset$ as 0 unless otherwise stated. For two random variables ξ and η, we write $\xi \lesssim \eta$ or $\eta \gtrsim \xi$ if there exists a *universal* constant $C > 0$ such that $\xi \leq C\eta$. Given parameters $\theta_1, \dots, \theta_m$, we use $C_{\theta_1,\dots,\theta_m}$ to denote positive constants, which depend only on $\theta_1, \dots, \theta_m$ and may be different in different expressions.

2 Notation and Setting

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $X_1, \dots, X_n$ be independent random variables taking values in respective measurable spaces $(S_1, \mathcal{S}_1), \dots, (S_n, \mathcal{S}_n)$. Write P_i for the marginal distribution of X_i. For $r \geq 1$ and each $(i_1, \dots, i_r)$, we say that a function $\psi_{(i_1,\dots,i_r)} : \prod_{s=1}^r S_{i_s} \to \mathbb{R}$ is *symmetric* if $\psi_{(i_1,\dots,i_r)}(x_{i_1}, \dots, x_{i_r}) = \psi_{(i_{\sigma(1)},\dots,i_{\sigma(r)})}(x_{\sigma(1)}, \dots, x_{\sigma(r)})$ for all $(x_1, \dots, x_r) \in \prod_{s=1}^r S_{i_s}$ and $\sigma \in \mathbb{S}_r$, where $\mathbb{S}_r$ is the symmetric group of degree r. To lighten notation we write $\psi(X_{i_1}, \dots, X_{i_r})$ or $\psi_{(i_1,\dots,i_r)}$ for $\psi_{(i_1,\dots,i_r)}(X_{i_1}, \dots, X_{i_r})$ whenever no confusion can arise. Let $\boldsymbol{\psi} := \{\psi_{(i_1,\dots,i_r)}\}_{(i_1,\dots,i_r)\in I_{n,r}}$ be a family of $\otimes_{s=1}^r \mathcal{S}_{i_s}$-measurable symmetric functions $\psi_{(i_1,\dots,i_r)} : \prod_{s=1}^r S_{i_s} \to \mathbb{R}$, we define r-th order U-statistics is defined as follows

$$J_r(\boldsymbol{\psi}) := \sum_{1\leq i_1<\cdots<i_r\leq n} \psi(X_{i_1}, \dots, X_{i_r}) = \frac{1}{r!} \sum_{(i_1,\dots,i_r)\in I_{n,r}} \psi(X_{i_1}, \dots, X_{i_r}),$$

where $I_{n,r} := \{(i_1, \dots, i_r) : 1 \leq i_1, \dots, i_r \leq n,\ i_s \neq i_t \text{ for all } s \neq t\}$. We define a function $\prod_{l=1}^{r-s} P_{k_l}\psi : \prod_{t=1}^s S_{i_t} \to \mathbb{R}$ as

$$\prod_{l=1}^{r-s} P_{k_l}\psi(X_{i_1}, \dots, X_{i_s}) := \int \psi(X_{i_1}, \dots, X_{i_s}, x_{k_1}, \dots, x_{k_{r-s}}) \prod_{l=1}^{r-s} dP_{k_l}(x_{k_l}).$$

For a kernel $\psi_{(i_1,\dots,i_r)}$ with $\prod_{l=1}^r P_{i_l}\psi_{(i_1,\dots,i_r)} = 0$, we say $\psi_{(i_1,\dots,i_r)}$ is degenerate if $P_{i_l}\psi_{(i_1,\dots,i_r)} = 0$ for every $1 \leq l \leq r$. For each $(i_1, \dots, i_s)$, define $K_{n,r-s}(i_1, \dots, i_s) := \{1 \leq k_1, \dots, k_{r-s} \leq n, k_l \neq k_m \text{ for all } l \neq m \text{ and } k_l \notin \{i_1, \dots, i_s\} \text{ for all } 1 \leq l \leq r-s\}$. Then, for $1 \leq s \leq r$, Hoeffding projections are defined as follows

$$\begin{aligned}
&\pi_{s,(i_1,\dots,i_s)}\psi(X_{i_1}, \dots, X_{i_s}) \\
&:= \sum_{(k_1,\dots,k_{r-s})\in K_{n,r-s}(i_1,\dots,i_s)} \left(\prod_{l=1}^{s} (\delta_{X_{i_l}} - P_{i_l}) \prod_{m=1}^{r-s} P_{k_m} \right) \psi(X_{i_1}, \dots, X_{i_s}, X_{k_1}, \dots, X_{k_{r-s}}).
\end{aligned}$$

The Hoeffding decomposition of $J_r(\boldsymbol{\psi})$ is given by

$$J_r(\boldsymbol{\psi}) - \mathbb{E}[J_r(\boldsymbol{\psi})] = \sum_{s=1}^{r} \frac{1}{s!} \sum_{(i_1,\dots,i_s)\in I_{n,s}} \pi_{s,(i_1,\dots,i_s)}\psi(X_{i_1}, \dots, X_{i_s}).$$

Notations for Second-Order U-statistics: In this paper, we focus on Gaussian approximation in the case of $r = 2$, though some technical results for higher-order U-statistics are required. For reader's convenience, we restate, in the case of $r = 2$, some definitions introduced above. First, the second order U-statistics is given by

$$J_2(\boldsymbol{\psi}) := \sum_{1\leq i_1<i_2\leq n} \psi(X_{i_1}, X_{i_2}) = \frac{1}{2} \sum_{(i_1,i_2)\in I_{n,2}} \psi(X_{i_1}, X_{i_2}).$$

Then, for $s = \{1, 2\}$, Hoeffding projections are defined as follows

$$\begin{aligned}\pi_{1,i}\psi(X_i) &= \sum_{m\neq i}(\delta_{X_i} - P_i)P_m\psi(X_i, X_m),\\ \pi_{2,im}\psi(X_i, X_m) &= (\delta_{X_i} - P_i)(\delta_{X_m} - P_m)\psi(X_i, X_m).\end{aligned} \tag{1}$$

Finally, Hoeffding decomposition of $J_2(\boldsymbol{\psi})$ is

$$J_2(\boldsymbol{\psi}) - \mathbb{E}[J_2(\boldsymbol{\psi})] = \sum_{i=1}^{n}\pi_{1,i}\psi(X_i) + \frac{1}{2}\sum_{(i,m)\in I_{n,2}}\pi_{2,im}\psi(X_i, X_m).$$

3 Main Results

In this section, we provide main results, that is, Gaussian Approximation for high-dimensional second-order U-statistics and V-statistics with size dependent kernels. We prove theorems in this section in Section B.

3.1 Gaussian Approximation for High-dimensional U-statistics

Let p be a positive integer. We assume $p \geq 3$ so that $\log p > 1$. For every $j \in [p]$, we consider a family of kernels $\boldsymbol{\psi}_j := \{\psi_{j,(i,m)} : S_i \times S_m \to \mathbb{R}\}_{(i,m)\in I_{n,2}}$, which are $\mathcal{S}_i \otimes \mathcal{S}_m$-measurable, symmetric, and satisfy $\psi_{j,(i,m)} \in L^4(P_i \otimes P_m)$ for all $(i, m) \in I_{n,2}$. We suppress (i, m) in $\psi_{j,(i,m)}(x_i, x_m)$ when no confusion can arise and simply write $\psi_j(x_i, x_m)$. Also, $\sigma_j := \sqrt{\mathrm{Var}[J_2(\psi_j)]} > 0$. Define

$$W := (J_2(\boldsymbol{\psi}_1) - \mathbb{E}[J_2(\boldsymbol{\psi}_1)], \ldots, J_2(\boldsymbol{\psi}_p) - \mathbb{E}[J_2(\boldsymbol{\psi}_p)])^\top.$$

Our main result is an explicit error bound on the normal approximation of $\mathbb{P}(W \in A)$ uniformly over $A \in \mathcal{R}_p$. To state the result concisely, we introduce some notation. Set

$$\begin{aligned}\Delta_1^{(0)} &:= n^2 \max_{(j,k)\in[p]^2}\max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}}\frac{\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}}{\sigma_j\sigma_k},\\ \Delta_1^{(1)} &:= n^{3/2}\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\frac{\|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}}{\sigma_j\sigma_k},\end{aligned}$$

where we define $\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k(X_m, X_l) := \int \pi_{2,im}\psi_j(x_i, X_m)\pi_{2,il}\psi_k(x_i, X_l)dP_i(x_i)$ and $\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k(X_m) := \int \pi_{1,i}\psi_j(x_i)\pi_{2,im}\psi_k(x_i, X_m)dP_i(x_i)$, and

$$\Delta_{2,*}^{(1)}(1) := n\max_{j\in[p]}\max_{i\in[n]}\frac{\|\pi_{1,i}\psi_j\|_{L^4(P_i)}^4}{\sigma_j^4},\quad \Delta_{2,q}^{(2)}(1) := n^{4/q}\left\|\max_{j\in[p]}\max_{i\in[n]}\frac{|\pi_{1,i}\psi_j|}{\sigma_j}\right\|_{L^q(\mathbb{P})}^4\log p$$

and

$$\Delta_{2,*}^{(1)}(2) := n^2\max_{j\in[p]}\max_{i\in[n]}\frac{\|\pi_{2,im}\psi_j\|_{L^4(P_i\otimes P_m)}^4}{\sigma_j^4}\log^3 p,\quad \Delta_{2,*}^{(2)}(2) := n^3\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{\|P_m(|\pi_{2,im}\psi_j|^2)\|_{L^2(P_i)}^2}{\sigma_j^4}\log^2 p,$$

$$\Delta_{2,*}^{(3)}(2) := n\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{P_m(\pi_{2,im}\psi_j)^4(X_i)}{\sigma_j^4}\right]\log^4 p,$$

$$\Delta_{2,q}^{(4)}(2) = n^{4/q}\left\|\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{|\pi_{2,im}\psi_j|}{\sigma_j}\right\|_{L^q(\mathbb{P})}^4\log^5(np),$$

$$\Delta_{2,q}^{(5)}(2) = n^{2+4/q}\left\|\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{P_m(|\pi_{2,im}\psi_j|^2)}{\sigma_j^2}\right\|_{L^{q/2}(\mathbb{P})}^2\log^3(np).$$

Then, the error bound on the normal approximation is explicitly given as follows. The following Gaussian approximation is an extension of Theorem 2 in Imai and Koike (2025) and is valid regardless of whether the first- or second-order Hoeffding component is dominant or neither is as long as the approximation error bound converges to 0 as $n \to \infty$.

Theorem 1. *There exists a universal constant C such that*

$$\sup_{A\in\mathcal{R}_p}|\mathbb{P}(W\in A)-\mathbb{P}(Z\in A)| \le C\left(\sqrt{\Delta_1'}+\left\{(\Delta_{2,q}(1)+\Delta_{2,q}(2))\log^5 p\right\}^{1/4}\right), \tag{2}$$

where $Z \sim N(0, \mathrm{Cov}(W))$ and

$$\Delta_{2,q}(1) := \Delta_{2,*}^{(1)}(1)+\Delta_{2,q}^{(2)}(1), \quad \Delta_{2,q}(2) := \sum_{\ell=1}^{3}\Delta_{2,*}^{(\ell)}(2)+\sum_{\ell=4}^{5}\Delta_{2,q}^{(\ell)}(2),$$

$$\Delta_1' := \Delta_1^{(0)}\log^3 p+\Delta_1^{(1)}\log^{5/2}p+n^{1/2}\max_{j\in[p]}\max_{m\in[n]}\frac{\|\pi_{1,m}\psi_j\|_{L^2(P_m)}}{\sigma_j}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4}.$$

As noted by Imai and Koike (2025), it is often easy to evaluate the quantities in terms of kernel functions rather than those of Hoeffding projections, in application. The following corollary is an extension of Corollary 2 in Imai and Koike (2025) and is useful for the purpose.

Corollary 1. *Under the assumption of Theorem 1, there exists a universal constant C such that*

$$\sup_{A\in\mathcal{R}_p}|\mathbb{P}(W\in A)-\mathbb{P}(Z\in A)| \le C\left(\sqrt{\tilde\Delta_1'}+\left\{(\tilde\Delta_{2,q}(1)+\tilde\Delta_{2,q}(2))\log^5 p\right\}^{1/4}\right),$$

where

$$\tilde\Delta_{2,q}(1) := \tilde\Delta_{2,*}^{(1)}(1)+\tilde\Delta_{2,q}^{(2)}(1), \quad \tilde\Delta_{2,q}(2) := \sum_{\ell=1}^{3}\tilde\Delta_{2,*}^{(\ell)}(2)+\sum_{\ell=4}^{5}\tilde\Delta_{2,q}^{(\ell)}(2),$$

$$\tilde\Delta_1' := \tilde\Delta_1^{(0)}\log^3 p+\tilde\Delta_1^{(1)}\log^{5/2}p+n^{3/2}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{\sqrt{\mathrm{Var}[P_m\psi_j(X_i,X_m)]}}{\sigma_j}\left(\tilde\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4}.$$

with

$$\tilde\Delta_1^{(0)} := n^2\max_{(j,k)\in[p]^2}\max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}}\frac{\|\psi_{j,(i,m)}\star_i^1\psi_{k,(i,l)}\|_{L^2(P_m\otimes P_l)}}{\sigma_j\sigma_k},$$

$$\tilde{\Delta}_1^{(1)} \coloneqq n^{3/2} \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{\sqrt{\mathrm{Var}[P_m \psi_j(X_i, X_m)]}}{\sigma_j} \sqrt{\tilde{\Delta}_1^{(0)}},$$

where we define $\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}(X_m, X_l) \coloneqq \int \psi_j(x_i, X_m)\psi_k(x_i, X_l) dP_i(x_i)$, *and*

$$\tilde{\Delta}_{2,*}^{(1)}(1) \coloneqq n^5 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{\|P_m \psi_{j,(i,m)}\|_{L^4(P_i)}^4}{\sigma_j^4}, \ \ \tilde{\Delta}_{2,q}^{(2)}(1) \coloneqq n^{4+4/q} \left\| \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{|P_m \psi_{j,(i,m)}|}{\sigma_j} \right\|_{L^q(\mathbb{P})}^4 \log p,$$

and

$$\tilde{\Delta}_{2,*}^{(1)}(2) \coloneqq n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{\|\psi_{j,(i,m)}\|_{L^4(P_i\otimes P_m)}^4}{\sigma_j^4} \log^3 p, \ \ \tilde{\Delta}_{2,*}^{(2)}(2) \coloneqq n^3 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{\|P_m(|\psi_{j,(i,m)}|^2)\|_{L^2(P_i)}^2}{\sigma_j^4} \log^2 p,$$

$$\tilde{\Delta}_{2,*}^{(3)}(2) \coloneqq n\, \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{P_m(\psi_{j,(i,m)})^4(X_i)}{\sigma_j^4 \psi_j} \right] \log^4 p, \ \ \tilde{\Delta}_{2,q}^{(4)}(2) = n^{4/q} \left\| \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{|\psi_{j,(i,m)}|}{\sigma_j} \right\|_{L^q(\mathbb{P})}^4 \log^5(np),$$

$$\tilde{\Delta}_{2,q}^{(5)}(2) = n^{2+4/q} \left\| \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \frac{P_m(|\psi_{j,(i,m)}|^2)}{\sigma_j^2} \right\|_{L^{q/2}(\mathbb{P})}^2 \log^3(np).$$

3.2 Gaussian Approximation for High-dimensional V-statistics

In applications, we often need to handle not only U-statistics but also V-statistics. Once Theorem 1 is established, the Gaussian approximation for the second-order high-dimensional V-statistic with size-dependent kernels follows immediately. To facilitate applications, we state it explicitly as a corollary.

Before stating the corollary, we introduce notations used below. We write the second order V-statistics as $J_2^V(\boldsymbol{\psi})$.

$$J_2^V(\boldsymbol{\psi}) \coloneqq \sum_{i_1=1}^n \sum_{i_2=1}^n \psi(X_{i_1}, X_{i_2}).$$

For $s = \{1, 2\}$, Hoeffding projections are defined as follows

$$\begin{aligned} \pi_{1,i}^V \psi(X_i) &= \frac{1}{2}(\delta_{X_i} - P_i)\psi(X_i, X_i) + \sum_{m\neq i}(\delta_{X_i} - P_i) P_m \psi(X_i, X_m), \\ \pi_{2,im}\psi(X_i, X_m) &= (\delta_{X_i} - P_i)(\delta_{X_m} - P_m)\psi(X_i, X_m). \end{aligned} \tag{3}$$

Then, Hoeffding decomposition of $J_2^V(\boldsymbol{\psi})$ is

$$J_2^V(\boldsymbol{\psi}) - \mathbb{E}[J_2^V(\boldsymbol{\psi})] = \sum_{i=1}^n \pi_{1,i}^V \psi(X_i) + \frac{1}{2} \sum_{(i,m)\in I_{n,2}} \pi_{2,im}\psi(X_i, X_m).$$

Then, the error bound on the normal approximation is explicitly given as follows. As in the case of U-statistics, the following Gaussian approximations are valid regardless of whether the first- or second-order Hoeffding component is dominant or neither is as long as the approximation error bound converges to 0 as $n \to \infty$.

Corollary 2. *There exists a universal constant C such that*

$$\sup_{A\in\mathcal{R}_p} |\mathbb{P}(W^V \in A) - \mathbb{P}(Z \in A)| \leq C\left(\sqrt{(\Delta_1')^V} + \left\{(\Delta_{2,q}^V(1) + \Delta_{2,q}(2))\log^5 p\right\}^{1/4}\right), \tag{4}$$

where

$$\Delta_{2,q}^V(1) \coloneqq (\Delta_{2,*}^{(1)})^V(1) + (\Delta_{2,q}^{(2)})^V(1),$$
$$(\Delta_1')^V \coloneqq \Delta_1^{(0)}\log^3 p + (\Delta_1^{(1)})^V \log^{5/2} p + n^{1/2}\max_{j\in[p]}\max_{m\in[n]}\frac{\|\pi_{1,m}^V\psi_j\|_{L^2(P_m)}}{\sigma_j}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4},$$

with

$$(\Delta_1^{(1)})^V \coloneqq n^{3/2}\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\frac{\|\pi_{1,i}^V\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}}{\sigma_j\sigma_k},$$

and

$$(\Delta_{2,*}^{(1)})^V(1) \coloneqq n^5\max_{j\in[p]}\max_{i\in[n]}\frac{\|\pi_{1,i}^V\psi_j\|_{L^4(P_i)}^4}{\sigma_j^4}, \quad (\Delta_{2,q}^{(2)})^V(1) \coloneqq n^{4/q}\left\|\max_{j\in[p]}\max_{i\in[n]}\frac{|\pi_{1,i}^V\psi_j|}{\sigma_j}\right\|_{L^q(\mathbb{P})}^4\log p.$$

Proof. Replacing $\pi_1\psi_j$ in quantities associated with Theorem 1 with $\pi_1^V\psi_j$ completes the proof. □

To facilitate applications, we formulate an analogue of our result expressed solely through bounds on quantities in terms of kernel functions, for the V-statistic case as well.

Corollary 3. *There exists a universal constant C such that*

$$\sup_{A\in\mathcal{R}_p} |\mathbb{P}(W^V \in A) - \mathbb{P}(Z \in A)| \leq C\left(\sqrt{(\tilde{\Delta}_1')^V} + \left\{(\tilde{\Delta}_{2,q}^V(1) + \tilde{\Delta}_{2,q}(2))\log^5 p\right\}^{1/4}\right),$$

where

$$\tilde{\Delta}_{2,q}^V(1) \coloneqq (\tilde{\Delta}_{2,*}^{(1)})^V(1) + (\tilde{\Delta}_{2,q}^{(2)})^V(1),$$
$$(\tilde{\Delta}_1')^V \coloneqq \Delta_1^{(0)}\log^3 p + (\tilde{\Delta}_1^{(1)})^V \log^{5/2} p + n^{1/2}\max_{j\in[p]}\max_{m\in[n]}\frac{\|\pi_{1,m}^V\psi_j\|_{L^2(P_m)}}{\sigma_j}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4},$$

with

$$(\tilde{\Delta}_1^{(1)})^V \coloneqq n^{3/2}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\sqrt{\mathrm{Var}[P_m\psi_{j,(i,m)}(X_i)]}\sqrt{\tilde{\Delta}_1^{(0)}}$$
$$+ n^{1/2}\max_{j\in[p]}\max_{i\in[n]}\sqrt{\mathrm{Var}[\psi_{j,(i,i)}(X_i, X_i)]}\sqrt{\tilde{\Delta}_1^{(0)}},$$

and

$$(\tilde{\Delta}_{2,*}^{(1)})^V(1) \coloneqq n^5\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{\|\psi_{j,(i,m)}\|_{L^4(P_i)}^4}{\sigma_j^4} + n\max_{j\in[p]}\max_{i\in[n]}\frac{\|\psi_{j,(i,i)}\|_{L^4(P_i\otimes P_i)}^4}{\sigma_j^4},$$
$$(\tilde{\Delta}_{2,q}^{(2)})^V(1) \coloneqq n^{4+4/q}\left\|\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{|\psi_{j,(i,m)}|}{\sigma_j}\right\|_{L^q(\mathbb{P})}^4\log p + n^{4/q}\left\|\max_{j\in[p]}\max_{i\in[n]}\frac{|\psi_{j,(i,i)}|}{\sigma_j}\right\|_{L^q(\mathbb{P})}^4\log p.$$

4 Special Cases and Potential Applications

U-statistics constructed from i.n.i.d. observations admit several subclasses of practical importance. In what follows, describe these subclasses in the second-order case. Alongside this discussion, we present concrete application examples rather than provide full technical proofs.

4.1 Weighted U-statistics

Suppose we have an i.i.d. sample $X \coloneqq (X_i)_{i=1}^n$ and a non-random or i.i.d. sample (independent of X) $(w_{(i,m)})_{(i,m)\in I_{n,2}}$. Then, the second-order weighted U-statistics $\sum_{(i,m)\in I_{n,2}} w_{(i,m)}\varphi(X_i, X_m)$ is exactly our defined U-statistics associated with $\psi_{(i,m)}(X_i, X_m) = w_{(i,m)}\varphi(X_i, X_m)$, where φ is an index-independent kernel in the sense that φ does not depend on (i, m). See also (Döbler and Peccati, 2017, pp. 7) for this point.

As an application of high-dimensional weighted U-statistics, we consider designs where the target parameter becomes high-dimensional due to a large number of outcomes, groups and moment conditions, etc. In the case of a single sum of independent variables, Belloni et al. (2018) consider these problems as application of high-dimensional central limit theorems and name the framework "many approximating mean (MAM)" framework. In the U-statistics counterparts of this framework, we consider (i) many-weak instrumental variables asymptotics by Chao et al. (2012) and (ii) many covariates asymptotics by Cattaneo et al. (2018a) for the partially linear models (also closely related to Cattaneo et al., 2018b and Jochmans, 2022).

Example 1 (Many-Weak Instrumental Variables Asymptotics in MAM framework)**.** Chao et al. (2012) proposed an asymptotic framework for instrumental variable models which is robustly valid under a wide range of identification strength of instrumental variables. Our Gaussian approximation can be used to extend the asymptotic theory of Chao et al. (2012) within MAM framework.

Suppose the following instrumental variable model for each $j \in [p]$.

$$\begin{aligned}
&Y_{ij} = X_{ij}^\top \theta_j + u_{ij},\\
&X_{ij} = \gamma_{n,j}^\top Z_{ij} + \varepsilon_{ij}, \quad \mathbb{E}[u_{ij} \mid Z_{ij}] = 0, \quad \mathbb{E}[\varepsilon_{ij} \mid Z_{ij}] = 0.
\end{aligned}$$

where Y_{ij} is scalar dependent variable, u_{ij} and ε_{ij} are scalar and d_j-dimensional error terms respectively, X_{ij} is a d_j-dimensional vector of explanatory variable, Z_{ij} is a $K_{n,j}$-dimensional vector of instrumental variables and $\theta_j \in \mathbb{R}^{d_j}$ and $\gamma_{n,j} \in \mathbb{R}^{K_{n,j}}$ are non-random coefficients. For the coefficient $\gamma_{n,j}$, we assume that

$$\gamma_{n,j} \coloneqq \frac{1}{\sqrt{n}}\mu_{n,j}\pi_j \tag{5}$$

where $\mu_{n,j}$ is a scalar n-dependent sequence and π_j is a $(K_{n,j} \times d_j)$-dimensional matrix of constants.

In this paragraph, we consider the JIVE2 proposed by Angrist et al. (1999):

$$\hat{\theta}_{n,j} = \left(\sum_{i=1}^{n-1}\sum_{m\neq i}^{n} X_{ij}P_{im,j}X_{mj}^{\top}\right)^{-1}\sum_{i=1}^{n-1}\sum_{m\neq i}^{n} X_{ij}P_{im,j}Y_{mj},$$

where $P_{im,j}$ is (i,m)-th element of the projection matrix $P_j := Z_j(Z_j^{\top}Z_j)^{-1}Z_j^{\top}$. We show that $\hat{\theta}_{n,j} - \theta_j$ has a weighted U-statistic form which does not always have Hoeffding dominant component. We assume $\sqrt{K_{n,j}}/\mu_{n,j}^2 \to 0$ as $n \to \infty$ to ensure the consistency of an estimator defined latter. Since $Y_{ij} = X_{ij}^{\top}\theta_j + u_{ij}$, it can be seen that

$$\hat{\theta}_{n,j} := \theta_j + \hat{\Gamma}_{n,j}^{-1}S_{n,j},$$

where $\hat{\Gamma}_{n,j} := \mu_{n,j}^{-2}\sum_{i=1}^{n-1}\sum_{m\neq i}^{n} X_{ij}P_{im,j}X_{mj}^{\top}$ and $S_{n,j} := \mu_{n,j}^{-2}\sum_{i=1}^{n-1}\sum_{m\neq i}^{n} X_{ij}P_{im,j}u_{mj}$. Some evaluations using our developed maximal inequality introduced later (Theorem 2) give

$$\hat{\theta}_{n,j} - \theta_j \approx \Gamma_{n,j}^{-1}\left(\frac{1}{\mu_{n,j}\sqrt{n}}\sum_{i=1}^{n}(I - P_{ii,j})(Z_{i,j}^{\top}\pi_j)u_{ij} + \frac{1}{\mu_{n,j}^2}\sum_{i=1}^{n}\sum_{i<m}^{n} P_{im,j}\Big(\varepsilon_{ij}u_{mj} + \varepsilon_{mj}u_{ij}\Big)\right),$$

with $\Gamma_{n,j} := \frac{1}{n}\sum_{i=1}^{n}(1 - P_{ii,j})(\pi_j^{\top}Z_{ij})(\pi_j^{\top}Z_{ij})^{\top} = O_p(1)$. Note that the convergence rate of the first term is $O_p(1/\mu_{n,j})$ and of the second term is $O_p(\sqrt{K_{n,j}}/\mu_{n,j}^2)$. Then, it can be seen that

$$\begin{aligned}
&\text{Case (i)} : K_{n,j} \text{ is fixed and } \mu_{n,j} = O(n^{1/2}) \iff \text{the linear term is dominant,}\\
&\text{Case (ii)} : K_{n,j} \to \infty \text{ and } K_{n,j}/\mu_{n,j}^2 \to \kappa \in (0,\infty) \iff \text{the linear term} \asymp \text{the quadratic term,}\\
&\text{Case (iii)} : K_{n,j} \to \infty \text{ and } K_{n,j}/\mu_{n,j}^2 \to \infty \iff \text{the quadratic term is dominant.}
\end{aligned}$$

These three cases are corresponding to (i) the ordinary instrumental variable regression model, (ii) linear regression model with many instrumental variables, and (iii) linear regression model with many-weak instrumental variables. In U-statistic terms, Case (i) corresponds to dominance of the linear term, Case (ii) to absence of the dominant term and Case (iii) to dominance of the quadratic term. Therefore, $\hat{\theta}_{n,j} - \theta_j$ has a weighted U-statistic form which does not always have Hoeffding dominant component.

Applying their Lemma A.2, Chao et al. (2012) established normal approximation of $\hat{\theta}_{n,j}$ for a fixed j. Their approximation is valid for any of the three regimes above. We can use our high-dimensional CLT (Theorem 1) to extend this result to the $(d_j \times p)$-dimensional joint vector $\hat{\theta}_n := (\hat{\theta}_{n,1}, \ldots, \hat{\theta}_{n,p})$ with possibly diverging p. Notably, it accommodates heterogeneous regimes across coordinates, for example, Case (i) holds for coordinate j whereas Case (iii) holds for coordinate k.

Example 2 (Many Covariates Asymptotics in MAM framework)**.** Cattaneo et al. (2018a) proposed an asymptotic framework for the partially linear models with high-dimensional covariates. This framework enable us to conduct inference based on asymptotic normality of estimators for parameters of interest even when the dimension of covariates to grow as fast as sample size. Our Gaussian approximation can be used to extend the asymptotic theory by Cattaneo et al. (2018a) within MAM framework.

Suppose we have an i.i.d. sample $(Y_j, X_j, Z_j) \coloneqq (Y_{ij}, X_{ij}^\top, Z_{ij}^\top)_{i=1}^n$ ($j \in [p]$) and consider the following partially linear model

$$Y_{ij} = X_{ij}^\top \beta_j + g_j(Z_{ij}) + \varepsilon_{ij}, \quad \mathbb{E}[\varepsilon_{ij} \mid X_{ij}, Z_{ij}] = 0,$$

where Y_{ij} is a scalar dependent variable, X_{ij} and Z_{ij} are d_{X_j}-dimensional and d_{Z_j}-dimensional random vector. Let $\mathbf{p}_{K_{n,j}}(Z_j)$ be a $K_{n,j}$-dimensional vector of approximating functions, such as power series and splines, and its definition is given by as follows

$$\mathbf{p}_{K_{n,j}}(Z_j) \coloneqq [p_1(Z_j), \dots, p_{K_{n,j}}(Z_j)]^\top.$$

Also, let $P_{K_{n,j}}$ be $(n \times K_{n,j})$ matrix whose i-th row of $P_{K_{n,j}}$ is $\mathbf{p}_{n,j}(Z_{ij})$, that is

$$P_{K_{n,j}} \coloneqq [\mathbf{p}_{K_{n,j}}(Z_{1j}), \dots, \mathbf{p}_{K_{n,j}}(Z_{nj})]^\top.$$

To describe estimator, define $M_j \coloneqq I_n - P_{K_{n,j}}(P_{K_{n,j}}^\top P_{K_{n,j}})^{-1} P_{K_{n,j}}^\top$. Then, the series based estimator for β is given by

$$\hat{\beta}_j = \left(\sum_{i=1}^n \sum_{m=1}^n M_{im,j} X_{ij} X_{mj}^\top \right)^{-1} \sum_{i=1}^n \sum_{m=1}^n M_{im,j} X_{ij} Y_{mj},$$

where $M_{ij,j}$ is the (i,m)-th element of M_j.

Next, we show $\hat{\beta}_j - \beta_j$ has a weighted U-statistic form whose Hoeffding dominant component is absent. Since $Y_{ij} = X_{ij}^\top \beta_j + g_j(Z_{ij}) + \varepsilon_{ij}$, it holds that $\sqrt{n}(\hat{\beta}_j - \beta_j) \coloneqq \hat{\Gamma}_{n,j}^{-1} S_{n,j}$, where $\hat{\Gamma}_{n,j} \coloneqq n^{-1} \sum_{i=1}^n \sum_{m=1}^n M_{im,j} X_{ij} X_{mj}^\top$ and $S_{n,j} \coloneqq n^{-1/2} \sum_{i=1}^n \sum_{m=1}^n M_{im,j} X_{ij} (g_j(Z_{mj}) + \varepsilon_{mj})$. To describe the asymptotic representation of $\hat{\Gamma}_{n,j}$ and $S_{n,j}$, we define $h_j(Z_{ij}) \coloneqq \mathbb{E}[X_{ij} \mid Z_{ij}]$ and $v_{ij} \coloneqq X_{ij} - h_j(Z_{ij})$. Assume, for all $j \in [p]$, $rank(P_{K_{n,j}}) = K_{n,j}$. In addition, assume that there is a positive constant $C > 0$ such that $C \leq M_{ii,j}$ and for some $\alpha_{g,j}, \alpha_{h,j} > 0$ and there is a $C < \infty$ such that

$$\min_{\gamma_{g,j} \in \mathbb{R}^{K_{n,j}}} \mathbb{E}\left[|g_j(Z_{ij}) - \gamma_{g,j}^\top p_{K_{n,j}}(Z_{ij})|^2 \right] \leq C K_{n,j}^{-2\alpha_{g,j}},$$
$$\min_{\gamma_{h,j} \in \mathbb{R}^{K_{n,j} \times d_{X_j}}} \mathbb{E}\left[\|h_j(Z_{ij}) - \gamma_{h,j}^\top \mathbf{p}_{K_{n,j}}(Z_{ij})\|^2 \right] \leq C K_{n,j}^{-2\alpha_{h,j}},$$

for all $j \in [p]$. Then, an evaluation using our developed maximal inequality (Theorem 2) gives

$$\hat{\Gamma}_{n,j} \approx \frac{1}{n} \sum_{i=1}^n M_{ii,j} \mathbb{E}[v_{ij} v_{ij}^\top \mid Z_{ij}].$$

Also, Hoeffding decomposition of $S_{n,j}$ is given by as follows.

$$S_{n,j} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{n} M_{im,j} X_{ij}(g_j(Z_{mj}) + \varepsilon_{mj}) = B_{n,j} + \Psi_{n,j} + R_{n,j} + U_{n,j},$$

with

$$B_{n,j} := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{m=1}^{n} M_{im,j} h_j(Z_{ij}) g_j(Z_{mj}), \quad \Psi_{n,j} := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} M_{ii,j} v_{ij} \varepsilon_{ij},$$
$$R_{n,j} := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{m=1}^{n} M_{im,j}[v_{ij} g_j(Z_{mj}) + h_j(Z_{mj})\varepsilon_{ij}], \quad U_{n,j} := \frac{1}{2\sqrt{n}} \sum_{i=1}^{n} \sum_{m \neq i}^{n} M_{im,j}[v_{ij}\varepsilon_{mj} + v_{mj}\varepsilon_{ij}],$$

where $Q_{im,j}$ is (i,m)-th element of the matrix $P_{K_{n,j}}(P_{K_{n,j}}^\top P_{K_{n,j}})^{-1} P_{K_{n,j}}^\top$. From some evaluations using our developed maximal inequality (Theorem 2), we can see that $B_{n,j}$ and $R_{n,j}$ are negligible. Finally, we have the influential function:

$$S_{n,j} \approx \Psi_{n,j} + U_{n,j} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{m=1}^{n} M_{im,j} v_{ij} \varepsilon_{mj}.$$

In the fixed dimensional setting, Cattaneo et al. (2018a) apply Lemma A.2 of Chao et al. (2012) and show the asymptotic normality which captures the uncertainty of both the linear term Ψ_n and the quadratic term U_n. Similarly to the case of many-weak IV asymptotics, we can use our high-dimensional CLT (Corollary 2) to extend this result to the $(d_{X_j} \times p)$-dimensional joint vector $\hat{\beta}_n := (\hat{\beta}_{n,1}, \ldots, \hat{\beta}_{n,p})$ with possibly diverging p.

4.2 Two-sample U-statistics

Let $\mathcal{I} = \{1, \ldots, n\}$ and $\mathcal{J} = \{n+11, \ldots, n+m\}$ be index sets for two independent samples $(X_i)_{i \in \mathcal{I}} \overset{\text{i.i.d.}}{\sim} P$ and $(Y_j)_{j \in \mathcal{J}} \overset{\text{i.i.d.}}{\sim} Q$. We write $N := n + m$. Define $Z := (Z_a)_{a=1}^N$ by $Z_a = X_a$ for $1 \leq a \leq n$ and $Z_a = Y_{a-n}$ for $n < a \leq N$, and set the marginals $P_a = P$ for $a \leq n$ and $P_a = Q$ for $a > n$, then Z is i.n.i.d.observation.

Let $\varphi_1, \varphi_2, \varphi_3 : S \times S \to \mathbb{R}$ be index-independent symmetric, possibly different kernels and for $(a,b) \in I_{N,2}$ and some constants $c_1, c_2, c_3 \in \mathbb{R}$, define

$$\begin{aligned}\psi_{(a,b)}(z_a, z_b) &:= c_1 \varphi_1(z_a, z_b) \mathbf{1}\{(a,b) \in \mathcal{I}\} \\ &\quad + c_2 \varphi_2(z_a, z_b) \mathbf{1}\{(a,b) \in \mathcal{J}\} + c_3 \varphi_3(z_a, z_b) \mathbf{1}\{(a \in \mathcal{I}, b \in \mathcal{J}) \cup (a \in \mathcal{J}, b \in \mathcal{I})\}.\end{aligned}$$

Then, we can see that the two-sample U-statistic is exactly our defined second-order U-statistic associated with $\psi_{(a,b)}$.

As an application of high-dimensional two-sample U-statistics, we revisit theoretical analysis of Li and Yuan (2024) on an adaptive kernel based homogeneity test, following Imai and Koike (2025)'s theoretical analysis on an adaptive goodness-of-fit test.

Example 3 (Adaptive Kernel Based Homogeneity Test)**.** $\mathcal{I} = \{1, \ldots, n\}$ and $\mathcal{J} = \{n+1, \ldots, n+m\}$ be index sets for two independent samples $(X_i)_{i \in \mathcal{I}} \overset{\text{i.i.d.}}{\sim} P$ and $(Y_j)_{j \in \mathcal{J}} \overset{\text{i.i.d.}}{\sim} Q$ and define the densities of P and Q as f and g, respectively. We define $N := n + m$ and $Z := (Z_a)_{a=1}^N$ by $Z_a = X_a$ for $1 \leq a \leq n$ and $Z_a = Y_{a-n}$ for $n < a \leq N$, and set the marginals $P_a = P$ for $a \leq n$ and $P_a = Q$ for $a > n$, then Z is i.n.i.d.observation. We aim to test whether two independent sample X and Y come from a common population or not. Namely, we consider the following hypothesis testing problem:

$$H_0 : P = Q, \quad \text{vs} \quad H_1 : P \neq Q.$$

Let $K : \mathbb{R}^d \to \mathbb{R}$ be a bounded positive definite function. For every positive number $h > 0$, write

$$\varphi_h(x, y) := K_h(x - y), \quad \text{with} \quad K_h(u) := \frac{1}{h^d} K\left(\frac{u}{h}\right).$$

Lemma 6 in Gretton et al. (2012) give the following distance between P and Q;

$$\mathrm{MMD}^2(P||Q) := \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi_h(x, y)\{f(x) - g(x)\}\{f(y) - g(y)\} dx dy,$$

which can be seen as the squared maximum mean discrepancy (MMD) between P and Q, based on the kernel φ_{h_n} (cf. Eq(10) in Sriperumbudur et al. (2010)). In particular, $\mathrm{MMD}^2(P||Q) = 0$ if and only if $f = g$ a.e, provided that φ_{h_n} is a characteristic kernel in teh sense of (Sriperumbudur et al., 2010, Definition 6). This suggests rejecting the null hypothesis when an estimator for $\mathrm{MMD}^2(P||Q) = 0$ is large. As a sample analogue of $\mathrm{MMD}^2(P||Q)$, Gretton et al. (2012) propose;

$$\sum_{(i,j) \in I_{N,2}} w_{ij} \varphi_h(Z_i, Z_j) \ \text{ with } \ w_{ij} := \frac{1\{i, j \in \mathcal{I}\}}{n(n-1)} + \frac{1\{i, j \in \mathcal{J}\}}{m(m-1)} - \frac{1\{(i \in \mathcal{I}, j \in \mathcal{J}) \cup (i \in \mathcal{J}, j \in \mathcal{I}\}}{mn}.$$

Recently, Li and Yuan (2024) have shown that the test based on the normalized version of this estimator is minimax optimal against smooth alternatives if K is Gaussian kernel and h is chosen appropriately. To be precise, denote by $\mathcal{P}_d$ the set of probability density functions on $\mathbb{R}^d$. Fix a constant $R > 0$. Given a constant $\alpha > 0$ and a sequence ρ_n of positive numbers tending to 0 as $n \to \infty$, we associate the sequence of alternatives as

$$H_1(\rho_n; \alpha) := \{f, g \in \mathcal{P}_d, \|f\|_{H^\alpha} \vee \|g\|_{H^\alpha} \leq R, \|f - g\|_{L^2(\mathbb{R}^d)} \geq \rho_n\}.$$

In Theorem 5(i) in Li and Yuan (2024), if we choose $h \asymp n^{-2/(4\alpha+d)}$, the aforementioned test is consistent for the alternative $f, g \in H_1(\rho_n; \alpha)$ as long as $\rho_n / \rho_n^*(\alpha) \to \infty$, where $\rho_n^*(\alpha) := n^{-2\alpha/(4\alpha+d)}$. Moreover, Theorem 5(ii) in Li and Yuan (2024), $\liminf_{n\to\infty} \rho_n / \rho_n^*(\alpha) < \infty$, there is no consistent test against $f, g \in H_1(\rho_n; \alpha)$ for some significant level. To conduct the test without prior knowledge on α, letting the family of kernels ψ_h be $\psi_h := \{w_{ij}\varphi_h\}_{(i,j) \in I_{N,2}}$, Li and Yuan (2024) considered the maximum of $J_2(\psi_h)$ over a range of h and showed that this test is adaptive to $\alpha > d/4$ up to a logarithmic factor; see Theorem 10 in Li and Yuan (2024). In the same way as Theorem 4 in Imai and Koike (2025), we can refine the theory of Li and Yuan (2024) using our theory (cf. Remark 6 in Imai and Koike, 2025).

Appendix

A Auxiliary Results

A.1 High-dimensional CLT via generalized exchangeable pairs

Since Theorem 6 in Imai and Koike (2025) does not assume identical distributions of observations, we can use the theorem even under i.n.i.d. setting. For convenience, we restate the theorem below.

Lemma 1 (Theorem 6 in Imai and Koike (2025))**.** *Let (Y, Y') be an exchangeable pair of random variables taking values in a measurable space $(E, \mathcal{E})$. Let $\mathsf{W} : E \to \mathbb{R}^p$ be $\mathcal{E}$-measurable and set $W := \mathsf{W}(Y)$, $W' := \mathsf{W}(Y')$, and $D := W' - W$. Suppose there exists an antisymmetric $\mathcal{E}^{\otimes 2}$-measurable function $\mathsf{G} : E^2 \to \mathbb{R}^p$ (i.e. $\mathsf{G}(Y, Y') = -\mathsf{G}(Y', Y)$) such that $G := \mathsf{G}(Y, Y')$ satisfies*

$$\mathbb{E}[\, G \mid Y \,] = -(W + R) \tag{6}$$

for some random vector $R \in \mathbb{R}^p$. Let Σ be a $p \times p$ positive semidefinite symmetric matrix and write $\underline{\sigma} := \min_{j \in [p]} \sqrt{\Sigma_{jj}} > 0$. Then there exists a universal constant $C > 0$ such that for any $\varepsilon > 0$,

$$\begin{aligned} &\sup_{A \in \mathcal{R}_p} \big|\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)\big| \\ &\le \frac{C}{\underline{\sigma}} \Big\{ \mathbb{E}\left[\|R^\varepsilon\|_\infty\right] \sqrt{\log p} + \varepsilon^{-1}\, \mathbb{E}\left[\|V^\varepsilon\|_\infty\right] (\log p)^{3/2} + \varepsilon^{-3}\, \mathbb{E}\left[\Gamma^\varepsilon\right] (\log p)^{7/2} + \varepsilon \sqrt{\log p} \Big\}, \end{aligned} \tag{7}$$

where $Z \sim N(0, \Sigma)$ and $\beta := \varepsilon^{-1} \log p$, and we define

$$R^\varepsilon := R + \mathbb{E}\big[\, G\, \mathbf{1}_{\{\|D\|_\infty > \beta^{-1}\}} \mid Y \big], \qquad V^\varepsilon := \frac{1}{2}\, \mathbb{E}\big[\, G D^\top\, \mathbf{1}_{\{\|D\|_\infty \le \beta^{-1}\}} \mid Y \big] - \Sigma,$$

$$\Gamma^\varepsilon := \max_{j,k,l,m \in [p]} \mathbb{E}\big[\, |G_j D_k D_l D_m|\, \mathbf{1}_{\{\|D\|_\infty \le \beta^{-1}\}} \mid Y \big].$$

A.2 Maximal Inequality

The following theorems are maximal inequalities for high-dimensional U-statistics based on i.n.i.d. observations.

Theorem 2. *Let $q \ge 1$ and $\boldsymbol{\psi}_j := \{\psi_{j,(i_1,\dots,i_r)}\}_{(i_1,\dots,i_r) \in I_{n,r}}$. Assume $\psi_{j,(i_1,\dots,i_r)} \in L^{q \vee 2}(\otimes_{s=1}^r P_{i_s})$ be degenerate, symmetric kernels of order $r \ge 1$ for all $(i_1, \dots, i_r) \in I_{n,r}$. Then there exists a constant C_r depending only on r such that*

$$\begin{aligned} &\left\| \max_{j \in [p]} |J_r(\boldsymbol{\psi}_j)| \right\|_{L^q(\mathbb{P})} \\ &\le C_r \max_{0 \le s \le r} n^{\frac{r-s}{2}} (q + \log p)^{\frac{r+s}{2}} \\ &\quad \times \left\| \max_{j \in [p]} \max_{(i_1,\dots,i_s) \in I_{n,s}} \max_{(k_1,\dots k_{r-s}) \in K_{n,r-s}(i_1,\dots,i_s)} \int \psi_j^2(X_{i_1}, \dots, X_{i_s}, x_{k_1}, \dots x_{k_{r-s}}) \prod_{l=1}^{r-s} P_{k_l}(dx_{k_l}) \right\|_{L^{1 \vee q/2}(\mathbb{P})}^{1/2}. \end{aligned}$$

Theorem 3. *Let $q \geq 1$ and, and $\boldsymbol{\psi}_j := \{\psi_{j,(i_1,\dots,i_r)}\}_{(i_1,\dots,i_r)\in I_{n,r}}$. Assume $\psi_{j,(i_1,\dots,i_r)} \in L^{q\vee 2}(\otimes_{s=1}^r P_{i_s})$ be non-negative, symmetric kernels of order $r \geq 1$ for all $(i_1,\dots,i_r) \in I_{n,r}$. Then there exists a constant c_r depending only on r such that*

$$
\begin{aligned}
&\left\|\max_{j\in[p]} J_r(\boldsymbol{\psi}_j)\right\|_{L^q(\mathbb{P})} \\
&\leq c_r \max_{0\leq s\leq r} n^{r-s}(q+\log p)^s \\
&\quad\times\left\|\max_{j\in[p]} \max_{(i_1,\dots,i_s)\in I_{n,s}} \max_{(k_1,\dots k_{r-s})\in K_{n,r-s}(i_1,\dots,i_s)} \int \psi_j(X_{i_1},\dots,X_{i_s},x_{k_1},\dots x_{k_{r-s}}) \prod_{l=1}^{r-s} P_{k_l}(dx_{k_l})\right\|_{L^q(\mathbb{P})}.
\end{aligned}
$$

The following inequalities are the extension of Lemma 3 in Imai and Koike (2025) to i.n.i.d. setting.

Lemma 2. *For $j \in [p]$ and $(i,m) \in I_{n,2}$, let $\psi_{j,(i,m)} \in L^4(P_i \otimes P_m)$ be degenerate symmetric kernels of order 2. There exists a universal constant C such that*

$$
\begin{aligned}
&\max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \left|\sum_{m\in[n]:m>i} \psi_j(x_i,X_m)\right|^4 dP_i(x_i)\right] \\
&\leq C\Bigg(n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(\psi_j^2)\|_{L^2(P_i)}^2 \log^2 p \\
&\qquad + n \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\psi_j\|_{L^4(P_i\otimes P_m)}^4 \log^3 p + \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i\psi_j^4(X_m)\right]\log^4 p\Bigg),
\end{aligned} \tag{8}
$$

and

$$
\begin{aligned}
&\max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \left|\sum_{m\in[n]:m<i} \psi_j(x_i,X_m)\right|^4 dP_i(x_i)\right] \\
&\leq C\Bigg(n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(\psi_j^2)\|_{L^2(P_i)}^2 \log^2 p \\
&\qquad + n \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\psi_j\|_{L^4(P_i\otimes P_m)}^4 \log^3 p + \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i\psi_j^4(X_m)\right]\log^4 p\Bigg),
\end{aligned} \tag{9}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\max_{j\in[p]} \sum_{i=1}^n \left|\sum_{m\in[n]:m\neq i} \psi_j(X_i,X_m)\right|^4\right] \\
&\leq C\Bigg(n^3 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(\psi_j^2)\|_{L^2(P_i)}^2 \log^2 p
\end{aligned}
$$

$$+ n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\psi_j\|^4_{L^4(P_i\otimes P_m)} \log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i\psi_j^4(X_m)\right] \log^4 p$$

$$+ n^2\,\mathbb{E}\left[\max_{(i,m)\in I_{n,2}} \max_{j\in[p]} \{P_m\psi_j^2(X_i)\}^2\right] \log^3(np) + \mathbb{E}\left[\max_{(i,m)\in I_{n,2}} \max_{j\in[p]} \psi_j^4(X_i, X_m)\right] \log^5(np) \Bigg). \tag{10}$$

B Proof of Main Result

B.1 Proof of Theorem 1

We may assume $\sigma_j = 1$ for all $j \in [p]$ without loss of generality. For the rest of the proof, we proceed in 5 steps.

Step 1 Regarding $X = (X_i)_{i=1}^n$ as a random element taking values in the measurable space $\otimes_{i=1}^n (S_i, \mathcal{S}_i)$, we are going to apply Lemma 1 to

$$\mathsf{W}(X) := (J_{2,X}(\boldsymbol{\psi}_1) - \mathbb{E}[J_{2,X}(\boldsymbol{\psi}_1)], \dots, J_{2,X}(\boldsymbol{\psi}_p) - \mathbb{E}[J_{2,X}(\boldsymbol{\psi}_p)])^\top.$$

For this purpose, we need to construct an appropriate exchangeable pair (X, X') and an antisymmetric function G. Let $X^* = (X_i^*)_{i=1}^n$ be an independent copy of $X = (X_i)_{i=1}^n$. Also, let α be a random index uniformly distributed on $[n]$ and such that X, X^* and α are independent. Then, define $X' = (X_i')_{i=1}^n$ as $X_i' := X_i^*$ if $i = \alpha$ and $X_i' := X_i$ otherwise. It is well-known that (X, X') is an exchangeable pair.

In addition, define a random vector $G = \mathsf{G}(X, X')$ in $\mathbb{R}^p$ as $G_j := nD_{j,1} + \frac{n}{2} D_{j,2}$ for $j = 1, \dots, p$, where

$$\begin{aligned} D_{j,1} &:= \sum_{i=1}^n \{\pi_{1,i}\psi_j(X_i') - \pi_{1,i}\psi_j(X_i)\}, \\ D_{j,2} &:= \sum_{1\le i<m\le n} \{\pi_{2,im}\psi_j(X_i', X_m') - \pi_{2,im}\psi_j(X_i, X_m)\}. \end{aligned} \tag{11}$$

G is antisymmetric by construction. From Lemma 3.2 in Döbler and Peccati (2017), it holds that

$$\mathbb{E}[D_{j,1} \mid X] = -\frac{1}{n} \sum_{i=1}^n \pi_{1,i}\psi_j(X_i), \quad \mathbb{E}[D_{j,2} \mid X] = -\frac{2}{n} \sum_{1\le i<m\le n} \pi_{2,im}\psi_j(X_i, X_m),$$

so we have $\mathbb{E}[G \mid X] = -W$. Therefore, applying Lemma 1, we obtain for any $\varepsilon > 0$

$$\begin{aligned} &\sup_{A\in\mathcal{R}_p} |\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)| \\ &\lesssim \mathbb{E}\left[\|R^\varepsilon\|_\infty\right] \sqrt{\log p} + \varepsilon^{-1}\,\mathbb{E}\left[\|V^\varepsilon\|_\infty\right] (\log p)^{3/2} + \varepsilon^{-3}\,\mathbb{E}\left[\Gamma^\varepsilon\right] (\log p)^{7/2} + \varepsilon\sqrt{\log p}, \end{aligned}$$

where R^ε, V^ε and Γ^ε are defined in the same way as in Lemma 1 with $R = 0$ and (Y, Y') replaced by (X, X'). From the same argument as proof of Theorem 2 in Imai and Koike (2025),

$$\sup_{A\in\mathcal{R}_p} |\mathbb{P}(W \in A) - \mathbb{P}(Z \in A)|$$

$$\lesssim \frac{n\sqrt{\log p}}{\beta \wedge \beta^3}\mathbb{P}(\|D\|_\infty > \beta^{-1}) + \varepsilon^{-1}\,\mathbb{E}[\|V\|_\infty](\log p)^{3/2} + \varepsilon^{-3}\,\mathbb{E}[\Gamma_1+\Gamma_2](\log p)^{7/2} + \varepsilon\sqrt{\log p},$$

where $V := \frac{1}{2}\,\mathbb{E}[GD^\top \mid X] - \Sigma$ and $\Gamma_s := n\max_{j\in[p]}\mathbb{E}[|D_{j,s}|^4 \mid X]$ for $s = 1,2$. In the remaining proof, we will bound the quantities on the right-hand side and then choose ε appropriately.

Step 2 In this step, we bound $\mathbb{E}[\|V\|_\infty](\log p)^{3/2}$. First, we derive a U-statistic representation of $V_{jk} = \frac{1}{2}\,\mathbb{E}[G_jD_k \mid X] - \mathbb{E}[W_jW_k]$ for $j,k \in [p]$. For $j,k\in[p]$, observe that

$$G_jD_k = \left(nD_{j,1} + \frac{n}{2}D_{j,2}\right)(D_{k,1}+D_{k,2})\,.$$

From Eq. (11), we can see that

$$\begin{aligned}
&n\,\mathbb{E}[D_{j,1}D_{k,1} \mid X]\\
&= n\,\mathbb{E}\left[\sum_{i=1}^n\sum_{m=1}^n \{\pi_{1,i}\psi_j(X_i') - \pi_{1,i}\psi_j(X_i)\}\{\pi_{1,m}\psi_k(X_m') - \pi_{1,m}\psi_k(X_m)\} \mid X\right]\\
&= n\,\mathbb{E}\left[\sum_{i=1}^n\sum_{m=1}^n 1_{\{\alpha=i\}}1_{\{\alpha=m\}}\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}\{\pi_{1,m}\psi_k(X_m^*) - \pi_{1,m}\psi_k(X_m)\} \mid X\right]\\
&= n\sum_{i=1}^n \mathbb{E}\left[1_{\{\alpha=i\}}\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}\{\pi_{1,i}\psi_k(X_i^*) - \pi_{1,i}\psi_k(X_i)\} \mid X\right]\\
&= n\sum_{i=1}^n \mathbb{E}[1_{\{\alpha=i\}}]\Big(\mathbb{E}[\pi_{1,i}\psi_j(X_i^*)\pi_{1,i}\psi_k(X_i^*)] + \pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\Big)\\
&= \sum_{i=1}^n\Big(P_i\{\pi_{1,i}\psi_j(X_i^*)\pi_{1,i}\psi_k(X_i^*)\} + \pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\Big)\\
&= \sum_{i=1}^n\Big(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\} + \pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\Big),
\end{aligned}$$

where the second equality follows from the definition of $(X_i')_{i=1}^n$, the third equality holds because the event $\{\alpha=i\}$ and $\{\alpha=m\}$ are disjoint for $i\neq m$, the fourth equality follows from the degeneracy of Hoeffding projections ($\mathbb{E}[\pi_{1,i}\psi_j(X_i^*)]=0$), the fifth equality holds since α is uniformly distributed on $[n]$, and the final equality follows from the fact that $(X_i^*)_{i=1}^n$ is an independent copy of $(X_i)_{i=1}^n$.

Similarly, from Eq. (11), we can see that

$$\begin{aligned}
&n\,\mathbb{E}[D_{j,1}D_{k,2} \mid X]\\
&= n\,\mathbb{E}\left[\left(\sum_{i=1}^n\{\pi_{1,i}\psi_j(X_i') - \pi_{1,i}\psi_j(X_i)\}\right)\left(\sum_{1\le i<m\le n}\{\pi_{2,im}\psi_k(X_i',X_m') - \pi_{2,im}\psi_k(X_i,X_m)\}\right) \mid X\right]\\
&= n\,\mathbb{E}\left[\left(\sum_{i=1}^n 1_{\{\alpha=i\}}\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}\right)\left(\sum_{1\le i<m\le n} 1_{\{\alpha=i\}}\{\pi_{2,im}\psi_k(X_i^*,X_m) - \pi_{2,im}\psi_k(X_i,X_m)\}\right.\right.
\end{aligned}$$

$$
\left. \left. + \sum_{1\le i<m\le n} 1_{\{\alpha=m\}}\{\pi_{2,im}\psi_k(X_i,X_m^*) - \pi_{2,im}\psi_k(X_i,X_m)\}\right) \mid X\right]
$$

$$
\begin{aligned}
&= n \sum_{1\le i<m\le n} \mathbb{E}\left[1_{\{\alpha=i\}}\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}\{\pi_{2,im}\psi_k(X_i^*,X_m) - \pi_{2,im}\psi_k(X_i,X_m)\} \mid X\right] \\
&\quad + n \sum_{1\le i<m\le n} \mathbb{E}\left[1_{\{\alpha=m\}}\{\pi_{1,m}\psi_j(X_m^*) - \pi_{1,m}\psi_j(X_m)\}\{\pi_{2,im}\psi_k(X_i,X_m^*) - \pi_{2,im}\psi_k(X_i,X_m)\} \mid X\right] \\
&= \sum_{1\le i<m\le n} \Big(P_i\{\pi_{1,i}\psi_j(X_i^*)\pi_{2,im}\psi_k(X_i^*,X_m)\} + \pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\Big) \\
&\quad + \sum_{1\le i<m\le n} \Big(P_m\{\pi_{1,m}\psi_j(X_m^*)\pi_{2,im}\psi_k(X_i,X_m^*)\} + \pi_{1,m}\psi_j(X_m)\pi_{2,im}\psi_k(X_i,X_m)\Big) \\
&= \sum_{i=1}^{n}\sum_{m\neq i}^{n-1} \Big(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\} + \pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\Big),
\end{aligned}
$$

where the final equality follows from the symmetric property of the Hoeffding projections ($\pi_{2,im}\psi_k(X_i,X_m) = \pi_{2,mi}\psi_k(X_m,X_i)$). In the same way, it holds that

$$
\begin{aligned}
&\frac{n}{2}\,\mathbb{E}[D_{j,2}D_{k,1} \mid X] \\
&\quad = \frac{1}{2}\sum_{i=1}^{n}\sum_{m\neq i}^{n-1} \Big(P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\} + \pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\Big).
\end{aligned}
$$

Also, from Eq. (11), we can see that

$$
\begin{aligned}
&\frac{n}{2}\,\mathbb{E}[D_{j,2}D_{k,2} \mid X] \\
&= \frac{n}{2}\,\mathbb{E}\Bigg[\Bigg(\sum_{1\le i<m\le n}\{\pi_{2,im}\psi_j(X_i',X_m') - \pi_{2,im}\psi_j(X_i,X_m)\}\Bigg) \\
&\qquad\qquad \times \Bigg(\sum_{1\le i<m\le n}\{\pi_{2,im}\psi_k(X_i',X_m') - \pi_{2,im}\psi_k(X_i,X_m)\}\Bigg) \mid X\Bigg] \\
&= \frac{n}{2}\,\mathbb{E}\Bigg[\Bigg(\sum_{1\le i<m\le n} 1_{\{\alpha=i\}}\{\pi_{2,im}\psi_j(X_i^*,X_m) - \pi_{2,im}\psi_j(X_i,X_m)\} \\
&\qquad + 1_{\{\alpha=m\}}\{\pi_{2,im}\psi_j(X_i,X_m^*) - \pi_{2,im}\psi_j(X_i,X_m)\}\Bigg)\Bigg(\sum_{1\le i<m\le n} 1_{\{\alpha=i\}}\{\pi_{2,im}\psi_k(X_i^*,X_m) - \pi_{2,im}\psi_k(X_i,X_m)\} \\
&\qquad + 1_{\{\alpha=m\}}\{\pi_{2,im}\psi_k(X_i,X_m^*) - \pi_{2,im}\psi_k(X_i,X_m)\}\Bigg) \mid X\Bigg] \\
&= \frac{n}{2}\sum_{i}\sum_{m:i<m}\sum_{l:i<l}\mathbb{E}\left[1_{\{\alpha=i\}}\{\pi_{2,im}\psi_j(X_i^*,X_m) - \pi_{2,im}\psi_j(X_i,X_m)\}\{\pi_{2,il}\psi_k(X_i^*,X_l) - \pi_{2,il}\psi_k(X_i,X_l)\} \mid X\right] \\
&\ + \frac{n}{2}\sum_{i}\sum_{m:i<m}\sum_{l:l<i}\mathbb{E}\left[1_{\{\alpha=i\}}\{\pi_{2,im}\psi_j(X_i^*,X_m) - \pi_{2,im}\psi_j(X_i,X_m)\}\{\pi_{2,li}\psi_k(X_l,X_i^*) - \pi_{2,li}\psi_k(X_l,X_i)\} \mid X\right]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{n}{2}\sum_{i}\sum_{m:i<m}\sum_{l:m<l}\\
&\qquad\mathbb{E}\left[1_{\{\alpha=m\}}\{\pi_{2,im}\psi_j(X_i,X_m^*)-\pi_{2,im}\psi_j(X_i,X_m)\}\{\pi_{2,ml}\psi_k(X_m^*,X_l)-\pi_{2,ml}\psi_k(X_m,X_l)\}\mid X\right]\\
&+\frac{n}{2}\sum_{i}\sum_{m:i<m}\sum_{l:l<m}\\
&\qquad\mathbb{E}\left[1_{\{\alpha=m\}}\{\pi_{2,im}\psi_j(X_i,X_m^*)-\pi_{2,im}\psi_j(X_i,X_m)\}\{\pi_{2,lm}\psi_k(X_l,X_m^*)-\pi_{2,lm}\psi_k(X_l,X_m)\}\mid X\right]\\
&=\frac{1}{2}\sum_{i=1}^{n}\sum_{m\neq i}^{n-1}\sum_{l\neq i}^{n-1}\Big(P_i\{\pi_{2,im}\psi_j(X_i^*,X_m)\pi_{2,il}\psi_k(X_i^*,X_l)\}+\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\Big).
\end{aligned}
$$

Then, since $\mathbb{E}[W_jW_k]=\frac{1}{2}\,\mathbb{E}[G_jD_k]$, we have

$$
2V_{jk}=\sum_{i=1}^{n}\chi_{i,jk}^{(1,1)}(X_i)+\sum_{i=1}^{n}\sum_{m\neq i}^{n-1}\{\chi_{im,jk}^{(1,2)}(X_i,X_m)+\chi_{im,jk}^{(2,1)}(X_i,X_m)\}+\sum_{i=1}^{n}\sum_{m\neq i}^{n-1}\sum_{l\neq i}^{n-1}\chi_{iml,jk}^{(2,2)}(X_i,X_m,X_l),
$$

with

$$
\begin{aligned}
&\chi_{i,jk}^{(1,1)}(X_i)\coloneqq\left(\pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)-P_i\{\pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\}\right),\\
&\chi_{im,jk}^{(1,2)}(X_i,X_m)\coloneqq\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)+P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\},\\
&\chi_{im,jk}^{(2,1)}(X_i,X_m)\coloneqq\frac{1}{2}\left(\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)+P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\}\right),\\
&\chi_{iml,jk}^{(2,2)}(X_i,X_m,X_l)\coloneqq\frac{1}{2}\big(\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)-P_iP_mP_l\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\\
&\qquad+P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}-P_lP_m[P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}]\big).
\end{aligned}
$$

Observe that

$$
\begin{aligned}
&\chi_{im,jk}^{(1,2)}(X_i,X_m)\\
&=\Big(\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)-P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\}\Big)+2P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\}\\
&\eqqcolon\bar{\chi}_{im,jk}^{(1,2)}(X_i,X_m)+2P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\},
\end{aligned}
$$

and

$$
\begin{aligned}
&\chi_{im,jk}^{(2,1)}(X_i,X_m)\\
&=\frac{1}{2}\Big(\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)-P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\}\Big)+P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\}\\
&\eqqcolon\bar{\chi}_{im,jk}^{(2,1)}(X_i,X_m)+P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\},
\end{aligned}
$$

and

$$
\chi_{iml,jk}^{(2,2)}(X_i,X_m,X_l)
$$

$$
\begin{aligned}
&= \frac{1}{2}\Big(\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l) - P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\Big) \\
&\quad + \Big(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\} - P_iP_mP_l\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\Big) \\
&=: \bar{\chi}^{(2,2)}_{iml,jk}(X_i,X_m,X_l) + \Big(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\} - P_iP_mP_l\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\Big).
\end{aligned}
$$

Therefore, letting

$$
\begin{aligned}
\varphi^{(1,2)}_{m,jk}(X_m) &:= \sum_{i:m\neq i} 2P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\}, \\
\varphi^{(2,1)}_{m,jk}(X_m) &:= \sum_{i:m\neq i} P_i\{\pi_{1,i}\psi_k(X_i)\pi_{2,im}\psi_j(X_i,X_m)\}, \\
\varphi^{(2,2)}_{ml,jk}(X_m,X_l) &:= \sum_{i:m,l\neq i}\Big(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\} - P_iP_mP_l\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\Big),
\end{aligned}
$$

we have

$$
\begin{aligned}
2V_{jk} &= \sum_{i=1}^{n}\chi^{(1,1)}_{i,jk}(X_i) + \sum_{i=1}^{n}\sum_{m\neq i}^{n-1}\{\bar{\chi}^{(1,2)}_{im,jk}(X_i,X_m) + \bar{\chi}^{(2,1)}_{im,jk}(X_i,X_m)\} + \sum_{i=1}^{n}\sum_{m\neq i}^{n-1}\sum_{l\neq i}^{n-1}\bar{\chi}^{(2,2)}_{iml,jk}(X_i,X_m,X_l) \\
&\quad + \sum_{m=1}^{n}\{\varphi^{(1,2)}_{m,jk}(X_m) + \varphi^{(2,1)}_{m,jk}(X_m)\} + \sum_{m=1}^{n}\sum_{l=1}^{n}\varphi^{(2,2)}_{ml,jk}(X_m,X_l).
\end{aligned}
$$

Also, define

$$
\begin{aligned}
\tilde{\chi}^{(1,2)}_{im,jk}(X_i,X_m) &:= \frac{1}{2}\{\bar{\chi}^{(1,2)}_{im,jk}(X_i,X_m) + \bar{\chi}^{(1,2)}_{mi,jk}(X_m,X_i)\}, \\
\tilde{\chi}^{(2,1)}_{im,jk}(X_i,X_m) &:= \frac{1}{2}\{\bar{\chi}^{(2,1)}_{im,jk}(X_i,X_m) + \bar{\chi}^{(2,1)}_{mi,jk}(X_m,X_i)\}, \\
\tilde{\chi}^{(2,2),\mathtt{diag}}_{im,jk}(X_i,X_m) &:= \frac{1}{2}\{\bar{\chi}^{(2,2)}_{imm,jk}(X_i,X_m,X_m) + \bar{\chi}^{(2,2)}_{mii,jk}(X_m,X_i,X_i)\}, \\
\tilde{\chi}^{(2,2)}_{iml,jk}(X_i,X_m,X_l) &:= \frac{1}{6}\sum_{\sigma\in\mathbb{S}^3}\bar{\chi}^{(2,2)}_{\sigma(1)\sigma(2)\sigma(3),jk}(X_{\sigma(1)},X_{\sigma(2)},X_{\sigma(3)}), \\
\tilde{\varphi}^{(2,2)}_{ml,jk}(X_m,X_l) &:= \frac{1}{2}\{\varphi^{(2,2)}_{ml,jk}(X_m,X_l) + \varphi^{(2,2)}_{lm,jk}(X_l,X_m)\}, \quad \varphi^{(2,2),\mathtt{diag}}_{m,jk}(X_m) := \varphi^{(2,2)}_{mm,jk}(X_m,X_m),
\end{aligned}
$$

where $\mathbb{S}^3$ is the symmetric group of degree 3. We use boldface to indicate families of kernels over their sample coordinates; for example $\tilde{\boldsymbol{\chi}}^{(1,2)} := \{\tilde{\chi}^{(1,2)}_{im,jk}(X_i,X_m)\}_{(i,m)\in I_{n,2}}$. Noting that $\max_{(j,k)\in[p]^2} J_2(\tilde{\boldsymbol{\chi}}^{(2,1)}_{jk}) = \frac{1}{2}\max_{(j,k)\in[p]^2} J_2(\tilde{\boldsymbol{\chi}}^{\mathbf{(1,2)}}_{jk})$ and $\max_{(j,k)\in[p]^2} J_1(\boldsymbol{\varphi}^{(2,1)}_{\boldsymbol{jk}}) = \frac{1}{2}\max_{(j,k)\in[p]^2} J_1(\boldsymbol{\varphi}^{(1,2)}_{\boldsymbol{jk}})$, we can see that

$$
\begin{aligned}
&\mathbb{E}[\|2V\|_\infty] \\
&\le \mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_1(\boldsymbol{\chi}^{(1,1)}_{jk})|\right] + \frac{3}{2}\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\chi}}^{(1,2)}_{jk})|\right] + \mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\chi}}^{(2,2),\mathtt{diag}}_{jk})|\right] + \mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_3(\tilde{\boldsymbol{\chi}}^{(2,2)}_{jk})|\right] \\
&+ \frac{3}{2}\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_1(\boldsymbol{\varphi}^{(1,2)}_{jk})|\right] + \mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\varphi}}^{(2,2)}_{jk})|\right] + \mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_1(\boldsymbol{\varphi}^{(2,2),\mathtt{diag}}_{jk})|\right].
\end{aligned}
$$

In order to obtain a simple bound on $\mathbb{E}[\|V\|_\infty]$, we introduce the following lemma. The proof of this lemma is in Section C.3.

Lemma 3. *Under the assumptions for Theorem 1, there exists a universal constant C such that*

$$\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_1(\boldsymbol{\chi}_{jk}^{(1,1)})|\right](\log p)^{3/2} \le C\sqrt{\left(\Delta_{2,*}^{(1)}(1)+\Delta_{2,*}^{(2)}(1)\right)\log^5 p}, \tag{12}$$

$$\begin{aligned}\mathbb{E}&\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\chi}}_{jk}^{(1,2)})|\right](\log p)^{3/2} + \mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_1(\boldsymbol{\varphi}_{jk}^{(1,2)})|\right](\log p)^{3/2} \\ &\le C\Bigg(\Delta_1^{(1)}\log^{5/2} p + n^{1/2}\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|_{L^2(P_i)}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4} \\ &\qquad + \sqrt{(\Delta_{2,*}(1)+\Delta_{2,*}(2))\log^5 p}\Bigg),\end{aligned} \tag{13}$$

$$\begin{aligned}\mathbb{E}&\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\chi}}_{jk}^{(2,2),\texttt{diag}})|\right](\log p)^{3/2} + \mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_3(\tilde{\boldsymbol{\chi}}_{jk}^{(2,2)})|\right](\log p)^{3/2} \\ &\quad + \mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\varphi}}_{jk}^{(2,2)})|\right](\log p)^{3/2} + \mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_1(\boldsymbol{\varphi}_{jk}^{(2,2),\texttt{diag}})|\right](\log p)^{3/2} \\ &\le C\left(\Delta_1^{(0)}\log^3 p + \sqrt{\Delta_{2,*}(2)\log^5 p}\right).\end{aligned} \tag{14}$$

where $\Delta_{2,*}(1) \coloneqq \sum_{l=1}^2 \Delta_{2,*}^{(l)}(1)$ and $\Delta_{2,*}(2) \coloneqq \sum_{l=1}^5 \Delta_{2,*}^{(l)}(2)$ with

$$\begin{aligned}\Delta_{2,*}^{(2)}(1) &\coloneqq \mathbb{E}\left[\max_{j\in[p]}\max_{i\in[n]}\frac{(\pi_{1,i}\psi_j)^4(X_i)}{\sigma_j^4}\right]\log p, \\ \Delta_{2,*}^{(4)}(2) &\coloneqq \mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{(\pi_{2,im}\psi_j)^4(X_i,X_m)}{\sigma_j^4}\right]\log^5 p, \\ \Delta_{2,*}^{(5)}(2) &\coloneqq n^2\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\frac{\left(P_m(\pi_{2,im}\psi_j)^2(X_i)\right)^2}{\sigma_j^4}\right]\log^3 p.\end{aligned}$$

From these evaluations, it follows that

$$\mathbb{E}[\|V\|_\infty](\log p)^{3/2} \le C\left(\Delta_1' + \sqrt{(\Delta_{2,*}(1)+\Delta_{2,*}(2))\log^5 p}\right).$$

Step 3 In this step, we bound $\mathbb{E}[\Gamma_1]$ and $\mathbb{E}[\Gamma_2]$. Observe that

$$\begin{aligned}\Gamma_1 &= n\max_{j\in[p]}\mathbb{E}[|D_{j,1}|^4 \mid X] \\ &= n\max_{j\in[p]}\mathbb{E}\left[\left|\sum_{i=1}^n\{\pi_{1,i}\psi_j(X_i') - \pi_{1,i}\psi_j(X_i)\}\right|^4 \mid X\right]\end{aligned}$$

$$
\begin{aligned}
&= n \max_{j\in[p]} \mathbb{E}\left[\left| \sum_{i=1}^n 1_{\{\alpha=i\}} \{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\} \right|^4 \mid X \right] \\
&= \max_{j\in[p]} \sum_{i=1}^n \mathbb{E}\left[|\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}|^4 \mid X \right] \\
&\le 8 \max_{j\in[p]} \sum_{i=1}^n \left\{ P_i(\pi_{1,i}\psi_j)^4 + (\pi_{1,i}\psi_j)^4(X_i) \right\}.
\end{aligned}
$$

By Lemma 9 in Chernozhukov et al. (2015),

$$
\begin{aligned}
\mathbb{E}\left[\max_{j\in[p]} \sum_{i=1}^n (\pi_{1,i}\psi_j)^4(X_i) \right] &\le \max_{j\in[p]} \mathbb{E}\left[\sum_{i=1}^n (\pi_{1,i}\psi_j)^4(X_i) \right] + \mathbb{E}\left[\max_{i\in[n]} \max_{j\in[p]} (\pi_{1,i}\psi_j)^4(X_i) \right] \log p \\
&\le n \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^4(P_i)}^4 + \mathbb{E}\left[\max_{i\in[n]} \max_{j\in[p]} (\pi_{1,i}\psi_j)^4(X_i) \right] \log p \\
&\le n \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^4(P_i)}^4 + \left\| \max_{i\in[n]} \max_{j\in[p]} |\pi_{1,i}\psi_j| \right\|_{L^q(\mathbb{P})}^4 \log p \\
&\le \Delta_{2,*}^{(1)}(1) + \Delta_{2,q}^{(2)}(1) = \Delta_{2,q}(1).
\end{aligned}
$$

Therefore

$$
\mathbb{E}[\Gamma_1] \lesssim \Delta_{2,q}(1).
$$

Next, observe that

$$
\begin{aligned}
\Gamma_2 &= n \max_{j\in[p]} \mathbb{E}[|D_{j,2}|^4 \mid X] \\
&= n \max_{j\in[p]} \mathbb{E}\left[\left| \sum_{1\le i<m\le n} \{\pi_{2,im}\psi_j(X_i', X_m') - \pi_{2,im}\psi_j(X_i, X_m)\} \right|^4 \mid X \right] \\
&= n \max_{j\in[p]} \mathbb{E}\Bigg[\Bigg| \sum_{i=1}^n 1_{\{\alpha=i\}} \Bigg(\sum_{m\in[n]:i<m} \{\pi_{2,im}\psi_j(X_i^*, X_m) - \pi_{2,im}\psi_j(X_i, X_m)\} \\
&\qquad + \sum_{m\in[n]:m<i} \{\pi_{2,mi}\psi_j(X_m, X_i^*) - \pi_{2,mi}\psi_j(X_m, X_i)\} \Bigg) \Bigg|^4 \mid X \Bigg] \\
&= \max_{j\in[p]} \sum_{i=1}^n \mathbb{E}\Bigg[\Bigg| \sum_{m\in[n]:i<m} \{\pi_{2,im}\psi_j(X_i^*, X_m) - \pi_{2,im}\psi_j(X_i, X_m)\} \\
&\qquad + \sum_{m\in[n]:m<i} \{\pi_{2,mi}\psi_j(X_m, X_i^*) - \pi_{2,mi}\psi_j(X_m, X_i)\} \Bigg|^4 \mid X \Bigg] \\
&\le 8 \left(\max_{j\in[p]} \sum_{i=1}^n \mathbb{E}\left[\left| \sum_{m\in[n]:m\ne i} \pi_{2,im}\psi_j(X_i^*, X_m) \right|^4 \mid X \right] + \max_{j\in[p]} \sum_{i=1}^n \left| \sum_{m\in[n]:m\ne i} \pi_{2,im}\psi_j(X_i, X_m) \right|^4 \right)
\end{aligned}
$$

$$\coloneqq 8(\Gamma_{2,1} + \Gamma_{2,2}).$$

Noting that $\{X_i^*\}_{i=1}^n$ is an i.n.i.d. sequence and independent of X, from Eq. (8) and Eq. (9), we have

$$\begin{aligned}
\mathbb{E}[\Gamma_{2,1}] &\lesssim \mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n \int \left|\sum_{m\in[n]:m\neq i} \pi_{2,im}\psi_j(x_i, X_m)\right|^4 dP_i(x_i)\right] \\
&\leq n \max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \left|\sum_{m\in[n]:m\neq i} \pi_{2,im}\psi_j(x_i, X_m)\right|^4 dP_i(x_i)\right] \\
&\lesssim n^3 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_i)} \log^2 p \\
&\quad + n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\pi_{2,im}\psi_j\|^4_{L^4(P_i\otimes P_m)} \log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right] \log^4 p.
\end{aligned}$$

Also, from Eq. (10), we have

$$\begin{aligned}
\mathbb{E}[\Gamma_{2,2}] &\lesssim n^3 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(\psi_j^2)\|^2_{L^2(P_i)} \log^2 p \\
&\quad + n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\psi_j\|^4_{L^4(P_i\otimes P_m)} \log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i\psi_j^4(X_m)\right] \log^4 p \\
&\quad + n^2\,\mathbb{E}\left[\max_{(i,m)\in I_{n,2}} \max_{j\in[p]} \{P_m\psi_j^2(X_i)\}^2\right] \log^3(np) + \mathbb{E}\left[\max_{(i,m)\in I_{n,2}} \max_{j\in[p]} \psi_j^4(X_i, X_m)\right] \log^5(np).
\end{aligned}$$

Therefore

$$\mathbb{E}[\Gamma_2] \lesssim \Delta_{2,q}(2).$$

Summing up

$$\mathbb{E}[\Gamma_1 + \Gamma_2] \lesssim \Delta_{2,q}(1) + \Delta_{2,q}(2)$$

Step 4 In this step, we bound $\mathbb{P}(\|D\|_\infty > \beta^{-1})$. By Markov's inequality

$$\Pr\left(\|D\|_\infty > \beta^{-1}\right) \leq \beta^q\,\mathbb{E}[\|D\|_\infty^q] \leq (2\beta)^q \left(\mathbb{E}\left[\max_{j\in[p]} |D_{j,1}|^q\right] + \mathbb{E}\left[\max_{j\in[p]} |D_{j,2}|^q\right]\right).$$

By definition, it holds that

$$\begin{aligned}
\mathbb{E}\left[\max_{j\in[p]} |D_{j,1}|^q\right] &= \mathbb{E}\left[\max_{j\in[p]} \left|\sum_{i=1}^n \{\pi_{1,i}\psi_j(X_i') - \pi_{1,i}\psi_j(X_i)\}\right|^q\right] \\
&= \mathbb{E}\left[\max_{j\in[p]} \left|\sum_{i=1}^n 1_{\{\alpha=i\}}\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}\right|^q\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\max_{j\in[p]} |\{\pi_{1,i}\psi_j(X_i^*) - \pi_{1,i}\psi_j(X_i)\}|^q\right]
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{2^q}{n}\sum_{i=1}^{n}\left\|\max_{j\in[p]}|\pi_{1,i}\psi_j|\right\|_{L^q(P_i)}^q \\
&\le 2^q\left\|\max_{i\in[n]}\max_{j\in[p]}|\pi_{1,i}\psi_j|\right\|_{L^q(\mathbb{P})}^q \le 2^q\left(\frac{\Delta_{2,q}^{(2)}(1)^{1/4}}{n^{1/q}(\log p)^{1/4}}\right)^q.
\end{aligned}
$$

Also,

$$
\begin{aligned}
\mathbb{E}\left[\max_{j\in[p]}|D_{j,2}|^q\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\max_{j\in[p]}\left|\sum_{m:m\neq i}\{\pi_{2,im}\psi_j(X_i,X_m)-\pi_{2,im}\psi_j(X_i,X_m^*)\}\right|^q\right] \\
&\le \max_{i\in[n]}\mathbb{E}\left[\max_{j\in[p]}\left|\sum_{m:m\neq i}\{\pi_{2,im}\psi_j(X_i,X_m)-\pi_{2,im}\psi_j(X_i,X_m^*)\}\right|^q\right] \\
&\le 2^q\max_{i\in[n]}\mathbb{E}\left[\max_{j\in[p]}\left|\sum_{m:m\neq i}\pi_{2,im}\psi_j(X_i,X_m)\right|^q\right].
\end{aligned}
$$

Then, Lemma 1 in Imai and Koike (2025) yields

$$
\begin{aligned}
&\max_{i\in[n]}\mathbb{E}\left[\max_{j\in[p]}\left|\sum_{m:m\neq i}\pi_{2,im}\psi_j(X_i,X_m)\right|^q\right] \\
&\le C\max_{i\in[n]}\left(\left\|\max_{j\in[p]}\sqrt{\sum_{m\neq i}P_m(|\pi_{2,im}\psi_j|^2)(X_i)}\right\|_{L^q(\mathbb{P})}^q\sqrt{\log^q(np)}+\left\|\max_{m\neq i}\max_{j\in[p]}|\pi_{2,im}\psi_j(X_i,X_m)|\right\|_{L^q(\mathbb{P})}^q\log^q(np)\right) \\
&\le C\Bigg(n^{q/2}\left\|\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\sqrt{P_m(|\pi_{2,im}\psi_j|^2)(X_i)}\right\|_{L^q(\mathbb{P})}^q\sqrt{\log^q(np)} \\
&\qquad\qquad +\left\|\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}|\pi_{2,im}\psi_j(X_i,X_m)|\right\|_{L^q(\mathbb{P})}^q\log^q(np)\Bigg) \\
&\le C\Bigg(\left(n\left\|\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_m(|\pi_{2,im}\psi_j|^2)(X_i)\right\|_{L^{q/2}(\mathbb{P})}\log(np)\right)^{q/2} \\
&\qquad\qquad +\left(\left\|\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}|\pi_{2,im}\psi_j(X_i,X_m)|\right\|_{L^q(\mathbb{P})}\log(np)\right)^q\Bigg) \\
&\le C\left(\frac{\{\Delta_{2,q}^{(5)}(2)+\Delta_{2,q}^{(4)}(2)\}^{1/4}}{n^{1/q}\log^{1/4}(np)}\right)^q.
\end{aligned}
$$

Consequently, in conjunction with $\beta=\varepsilon^{-1}\log p$, there exists a universal constant $C>0$ such that

$$
n\mathbb{P}\left(\|D\|_\infty>\beta^{-1}\right)\le\varepsilon^{-q}(\log p)^q\left(C\frac{\{\Delta_{2,q}(1)+\Delta_{2,q}(2)\}^{1/4}}{\log^{1/4}(np)}\right)^q.
$$

Step 5 In this step, we choose the value of ε appropriately and complete the proof. Let

$$
\varepsilon=\left(\Delta_1'+\sqrt{(\Delta_{2,*}(1)+\Delta_{2,*}(2))\log^5 p}\right)^{1/2}+C\left((\Delta_{2,q}(1)+\Delta_{2,q}(2))\log^3 p\right)^{1/4}.
$$

Then, from the same argument as Step 3 in the proof of Theorem 2 in Imai and Koike (2025), we have the desired result.

B.2 Proof of Corollary 1

Before the Proof of Corollary 1, we provide the extensions of Lemma 4 and Lemma 6 in Imai and Koike (2025) to i.n.i.d. setting. The first result is extension of Lemma 4 in Imai and Koike (2025).

Lemma 4. *Let $\psi_j := \{\psi_{j,(i_1,\dots,i_r)}\}_{(i_1,\dots,i_r)\in I_{n,r}}$. Assume $\psi_{j,(i_1,\dots,i_r)} \in L^1(\otimes_{s=1}^r P_{i_s})$ be symmetric kernels of order $r \geq 1$ for all $(i_1,\dots,i_r) \in I_{n,r}$.*

$$\mathbb{E}\left[\max_{j\in[p]} \max_{(i_1,\dots,i_{r-l})\in I_{n,r-l}} \max_{(k_1,\dots,k_l)\in K_{n,l}(i_1,\dots,i_{r-l})} \prod_{t=1}^{l} P_{k_t}\psi_{j,(i_1,\dots,i_{r-l},k_1,\dots,k_l)}(X_{i_1},\dots X_{i_{r-l}})\right] \\ \leq \frac{r!}{(r-l)!}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i_1,\dots,i_r)\in I_{n,r}} \psi_j(X_{i_1},\dots X_{i_r})\right]. \tag{15}$$

Proof of Lemma 4. In line with the proof of Lemma 4 in Imai and Koike (2025), the claim for general l follows from repeated applications of the claim for $l=1$. Hence, it suffices to consider the case of $l=1$. Also as in that proof, we can assume $p=1$ and $\psi_1 \geq 0$ without loss of generality.

Under this assumption, we prove the lemma by induction on r. When $r=1$,

$$\mathbb{E}\left[\max_{k\in[n]} P_k\psi_{(k)}\right] = \max_{k\in[n]} P_k\psi_{(k)} \leq \mathbb{E}\left[\max_{k\in[n]} \psi_{(k)}(X_k)\right],$$

so Eq. (15) holds. Next, suppose that $r>1$ and Eq. (15) holds for any symmetric kernel $\psi_{1,(i_1,\dots,i_s)}$ of order $s \leq r-1$. Classifying whether k is n or not,

$$\begin{aligned}
&\mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n,r-1}} \max_{k\in K_{n,1}(i_1,\dots,i_{r-1})} P_k\psi_{1,(i_1,\dots,i_{r-1},k)}(X_{i_1},\dots X_{i_{r-1}})\right] \\
&\leq \mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n-1,r-1}} P_n\psi_{1,(i_1,\dots,i_{r-1},n)}(X_{i_1},\dots X_{i_{r-1}})\right] \\
&\quad + \mathbb{E}\left[\max_{(i_1,\dots,i_{r-2})\in I_{n-1,r-2}} \max_{k\in K_{n-1,1}(i_1,\dots,i_{r-2})} P_k\psi_{1,(i_1,\dots,i_{r-2},k,n)}(X_{i_1},\dots X_{i_{r-2}},X_n)\right] \\
&=: I + II.
\end{aligned}$$

Since X_n is independent of $\mathcal{G} := \sigma(X_1,\dots,X_{n-1})$, we have

$$\begin{aligned}
I &= \mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n-1,r-1}} \mathbb{E}\left[\psi_{1,(i_1,\dots,i_{r-1},n)}(X_{i_1},\dots X_{i_{r-1}},X_n) \mid \mathcal{G}\right]\right] \\
&\leq \mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n-1,r-1}} \psi_{1,(i_1,\dots,i_{r-1},n)}(X_{i_1},\dots X_{i_{r-1}},X_n)\right] \\
&\leq \mathbb{E}\left[\max_{(i_1,\dots,i_r)\in I_{n,r}} \psi_{1,(i_1,\dots,i_r)}(X_{i_1},\dots X_{i_r})\right],
\end{aligned}$$

where the first inequality follows from Jensen's inequality.

$$II = \int \mathbb{E}\left[\max_{(i_1,\dots,i_{r-2})\in I_{n-1,r-2}} \max_{k\in K_{n-1,1}(i_1,\dots,i_{r-2})} P_k\psi_{1,(i_1,\dots,i_{r-2},k,n)}(X_{i_1},\dots X_{i_{r-2}},x_n)\right] dP_n(x_n)$$

In terms of II, applying the assumptions of the induction gives

$$\begin{aligned}
II &\le (r-1)\int \mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n-1,r-1}} \psi_{1,(i_1,\dots,i_{r-1},n)}(X_{i_1},\dots X_{i_{r-1}},x_n)\right] dP_n(x_n)\\
&= (r-1)\,\mathbb{E}\left[\max_{(i_1,\dots,i_{r-1})\in I_{n-1,r-1}} \psi_{1,(i_1,\dots,i_{r-1},n)}(X_{i_1},\dots X_{i_{r-1}},X_n)\right]\\
&\le (r-1)\,\mathbb{E}\left[\max_{(i_1,\dots,i_r)\in I_{n,r}} \psi_{1,(i_1,\dots,i_r)}(X_{i_1},\dots X_{i_r})\right].
\end{aligned}$$

Summing up, we have

$$I + II \le r\,\mathbb{E}\left[\max_{(i_1,\dots,i_r)\in I_{n,r}} \psi_{1,(i_1,\dots,i_r)}(X_{i_1},\dots X_{i_r})\right].$$

□

The following result is the extension of Lemma 6 in Imai and Koike (2025).

Lemma 5. *There exists a universal constant such that*

$$\begin{aligned}
&\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}\\
&\le C\Big(\|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m\otimes P_l)}\\
&\quad + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}\|\psi_{k,(i,l)}\|_{L^2(P_i\otimes P_l)} + \|\psi_{j,(i,m)}\|_{L^2(P_i\otimes P_m)}\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)}\Big).
\end{aligned}$$

Proof. From the definition of $\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k$, we have

$$\begin{aligned}
&\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\\
&= \left\|\int \pi_{2,im}\psi_j(x_i,x_m)\pi_{2,il}\psi_k(x_i,x_l)dP_i(x_i)\right\|_{L^2(P_m\otimes P_l)}.
\end{aligned}$$

From Eq. (1), $\pi_{2,im}\psi(X_i,X_m)$ we have

$$\begin{aligned}
\pi_{2,im}\psi_j(X_i,X_m) &= (\delta_{X_i}-P_i)\psi_j(X_i,X_m) - (P_m - P_mP_i)\psi_j(X_i,X_m),\\
\pi_{2,il}\psi_k(X_i,X_l) &= (\delta_{X_i}-P_i)\psi_k(X_i,X_l) - (P_l - P_lP_i)\psi_k(X_i,X_l).
\end{aligned}$$

This and the triangle inequality imply

$$\begin{aligned}
&\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\\
&\le \|(\delta_{X_i}-P_i)\psi_{j,(i,m)} \star_i^1 (\delta_{X_i}-P_i)\psi_{k,(i,l)}\|_{L^2(P_m\otimes P_l)}
\end{aligned}$$

$$
\begin{aligned}
&+ \|(\delta_{X_i} - P_i)\psi_{j,(i,m)} \star_i^1 (P_l - P_l P_i)\psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} \\
&+ \|(P_m - P_m P_i)\psi_{j,(i,m)} \star_i^1 (\delta_{X_i} - P_i)\psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} \\
&+ \|(P_m - P_m P_i)\psi_{j,(i,m)} \star_i^1 (P_l - P_l P_i)\psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)}.
\end{aligned}
$$

Then, the triangle inequality and Schwarz inequality give

$$
\begin{aligned}
&\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m \otimes P_l)} \\
&\le \|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}\|\psi_{k,(i,l)}\|_{L^2(P_i \otimes P_l)} \\
&\quad + \|\psi_{j,(i,m)}\|_{L^2(P_i \otimes P_m)}\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)} + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)} \\
&\quad + \|\psi_{j,(i,m)}\|_{L^2(P_i \otimes P_m)}\|P_l\psi_{k,(i,l)}\|_{L^2(P_i)} + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}\||P_l\psi_{k,(i,l)}\|_{L^2(P_i)} \\
&\quad + \|\psi_{j,(i,m)}\|_{L^2(P_i \otimes P_m)}|P_i P_l\psi_{k,(i,l)}| + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}|P_i P_l\psi_{k,(i,l)}| \\
&\quad + \|P_m\psi_{j,(i,m)}\|_{L^2(P_i)}\|\psi_{k,(i,l)}\|_{L^2(P_i \otimes P_l)} + |P_m P_i\psi_{j,(i,m)}|\|\psi_{k,(i,l)}\|_{L^2(P_i \otimes P_l)} \\
&\quad + \|P_m\psi_{j,(i,m)}\|_{L^2(P_i)}\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)} + |P_m P_i\psi_{j,(i,m)}|\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)} \\
&\quad + \|P_m\psi_{j,(i,m)}\|_{L^2(P_i)}\|P_l\psi_{k,(i,l)}\|_{L^2(P_i)} + |P_m P_i\psi_{j,(i,m)}|\|P_l\psi_{k,(i,l)}\|_{L^2(P_i)} \\
&\quad + \|P_m\psi_{j,(i,m)}\|_{L^2(P_i)}|P_i P_l\psi_{k,(i,l)}| + |P_m P_i\psi_{j,(i,m)}||P_i P_l\psi_{k,(i,l)}|.
\end{aligned}
$$

This result and Jensen's inequality give

$$
\begin{aligned}
&\|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m \otimes P_l)} \\
&\lesssim \|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} \\
&\quad + \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)}\|\psi_{k,(i,l)}\|_{L^2(P_i \otimes P_l)} + \|\psi_{j,(i,m)}\|_{L^2(P_i \otimes P_m)}\|P_i\psi_{k,(i,l)}\|_{L^2(P_l)}.
\end{aligned}
$$

□

Proof of Corollary 1. From Eq. (1), Lemma 4 and Jensen's inequality, we have $\Delta_{2,q}(1) \le \tilde{\Delta}_{2,q}(1)$, $\Delta_{2,q}(2) \le \tilde{\Delta}_{2,q}(2)$ and $\Delta_{2,*}^{(5)}(2) \le \tilde{\Delta}_{2,q}^{(5)}(2)$. Next, from Lemma 5, we have

$$
\begin{aligned}
\Delta_1^{(0)} \log^3 p &= n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3 \\ m,l\neq i}} \|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m \otimes P_l)} \log^3 p \\
&\lesssim n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3 \\ m,l\neq i}} \|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} \log^3 p \\
&\quad + n^2 \log^3 p \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_i\psi_{j,(i,m)}\|_{L^2(P_m)} \max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \|\psi_{k,(i,l)}\|_{L^2(P_i \otimes P_l)}.
\end{aligned}
$$

Therefore, AM-GM inequality and Lyapnov inequality give

$$
\begin{aligned}
\Delta_1^{(0)} \log^3 p &\lesssim n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3 \\ m,l\neq i}} \|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m \otimes P_l)} \log^3 p \\
&\quad + n^{5/2} \log^{5/2} p \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_i\psi_{j,(i,m)}\|^2_{L^2(P_m)} + n^{3/2} \log^{7/2} p \max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \|\psi_{k,(i,l)}\|^2_{L^2(P_i \otimes P_l)}
\end{aligned}
$$

$$
\begin{aligned}
&\le n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}} \|\psi_{j,(i,m)} \star_i^1 \psi_{k,(i,l)}\|_{L^2(P_m\otimes P_l)} \log^3 p \\
&\quad + n^{5/2}\log^{5/2} p \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_i\psi_{j,(i,m)}\|_{L^4(P_m)}^2 + n^{3/2}\log^{7/2} p \max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \|P_i(\psi_{k,(i,l)}^2)\|_{L^2(P_l)} \\
&\le \tilde{\Delta}_1^{(0)} \log^3 p + \sqrt{\left(\tilde{\Delta}_{2,*}^{(2)}(2) + \tilde{\Delta}_{2,*}^{(1)}(1)\right)\log^5 p}.
\end{aligned}
$$

Finally, we evaluate $\Delta_1^{(1)}$. From $\int \pi_{1,i}\psi_j(x_i)dP_i(x_i) = 0$ and Fubini's theorem, it holds that

$$
\begin{aligned}
&\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k(X_m) \\
&= \int \pi_{1,i}\psi_j(x_i)(\delta_{x_i} - P_i)(\delta_{X_m} - P_m)\psi_k(x_i, X_m)dP_i(x_i) \\
&= \int \pi_{1,i}\psi_j(x_i)\psi_k(x_i, X_m)dP_i(x_i) - \mathbb{E}\left[\int \pi_{1,i}\psi_j(x_i)\psi_k(x_i, X_m)dP_i(x_i)\right].
\end{aligned}
$$

This result and applying Schwarz inequality repeatedly give

$$
\begin{aligned}
\|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)} &\le \left\|\int \pi_{1,i}\psi_j(x_i)\psi_k(x_i, X_m)dP_i(x_i)\right\|_{L^2(P_m)} \\
&\le \left\|\int (\pi_{1,i}\psi_j)^2(x_i)dP_i(x) \cdot \int (\psi_k)^2(x_i, X_m)dP_i(x_i)\right\|_{L^2(P_m)} \\
&= \|\pi_{1,i}\psi_j\|_{L^2(P_i)} \left\|\int (\psi_k)^2(x_i, X_m)dP_i(x_i)\right\|_{L^2(P_m)} \\
&\le \|\pi_{1,i}\psi_j\|_{L^2(P_i)} \|\psi_{k,(i,m)} \star_i^1 \psi_{k,(i,m)}\|_{L^2(P_m)}^{1/2}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Delta_1^{(1)} &= n^{3/2} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)} \\
&\le n^{3/2} \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^2(P_i)} \max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \|\psi_{k,(i,m)} \star_i^1 \psi_{k,(i,m)}\|_{L^2(P_m)}^{1/2} \\
&= n^{1/2} \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^2(P_i)} \sqrt{\tilde{\Delta}_1^{(0)}}.
\end{aligned}
$$

Observe that

$$
\max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^2(P_i)}^2 \le n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \mathrm{Var}[P_m\psi_{j,(i,m)}(X_i)].
$$

Thus

$$
\Delta_1^{(1)} \le n^{3/2} \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \sqrt{\mathrm{Var}[P_m\psi_{j,(i,m)}(X_i)]}\sqrt{\tilde{\Delta}_1^{(0)}}.
$$

Now, the desired result follows by inserting obtained bounds into Eq. (2). □

B.3 Proof of Corollary 3

Proof of Corollary 3. As shown in the proof of Corollary 1, it holds that

$$\Delta_{2,q}(2) \le \tilde{\Delta}_{2,q}(2), \quad \Delta_{2,q}^{(5)}(2) \le \tilde{\Delta}_{2,q}^{(5)}(2)$$

and

$$\Delta_1^{(0)} \log^3 p \lesssim \tilde{\Delta}_1^{(0)} \log^3 p + \sqrt{\left(\tilde{\Delta}_{2,*}^{(2)}(2) + \tilde{\Delta}_{2,*}^{(1)}(1)\right) \log^5 p}.$$

Also, from Eq. (3), Lemma 4 and Jensen's inequality, we have $\Delta_{2,q}^V(1) \le \tilde{\Delta}_{2,q}^V(1)$. Finally, we evaluate $(\Delta_1^{(1)})^V$. In the same way as the evaluation $\Delta_1^{(1)}$, we have

$$(\Delta_1^{(1)})^V \le n^{1/2} \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}^V \psi_j\|_{L^2(P_i)} \sqrt{\tilde{\Delta}_1^{(0)}}.$$

Observe that

$$\begin{aligned}
&\max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}^V \psi_j\|_{L^2(P_i)}^2 \\
&\le n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \mathrm{Var}[P_m \psi_{j,(i,m)}(X_i)] + \max_{j\in[p]} \max_{i\in[n]} \mathrm{Var}[\psi_{j,(i,i)}(X_i, X_i)].
\end{aligned}$$

Thus

$$\begin{aligned}
\Delta_1^{(1)} \le\ & n^{3/2} \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \sqrt{\mathrm{Var}[P_m \psi_{j,(i,m)}(X_i)]} \sqrt{\tilde{\Delta}_1^{(0)}} \\
& + n^{1/2} \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \sqrt{\mathrm{Var}[\psi_{j,(i,i)}(X_i, X_i)]} \sqrt{\tilde{\Delta}_1^{(0)}}.
\end{aligned}$$

Now, the desired result follows by inserting obtained bounds into Eq. (4). □

C Proof of Auxiliary Result

C.1 Proof of Theorem 2 and Theorem 3

Lemma 6. *Let $q \ge 1$, $\boldsymbol{\psi}_j = \{\psi_{j,(i_1,\dots,i_r)}\}_{(i_1,\dots,i_r)\in I_{n,r}}$ and $\boldsymbol{\psi}_j^2 = \{\psi_{j,(i_1,\dots,i_r)}^2\}_{(i_1,\dots,i_r)\in I_{n,r}}$. Assume $\psi_j \in L^q(\otimes_{s=1}^r P_{i_s})$ $(j = 1,\dots,p)$ be degenerate, symmetric kernels of order $r \ge 1$ for all $(i_1,\dots,i_r) \in I_{n,r}$. Then, there exists a constant C_r depending only on r such that*

$$\left\| \max_{j\in[p]} |J_r(\boldsymbol{\psi}_j)| \right\|_{L^q(\mathbb{P})} \le C_r (q + \log p)^{r/2} \left\| \max_{j\in[p]} \sqrt{J_r(\boldsymbol{\psi}_j^2)} \right\|_{L^q(\mathbb{P})}.$$

Proof. The proof of Lemma 13 in Imai and Koike (2025) relies only on (i) the randomization theorem for U-statistics (Theorem 3.5.3 in de la Peña and Giné, 1999), (ii) hypercontractivity of Rademacher chaos (Theorem 3.5.2 in de la Peña and Giné, 1999), and (iii) the identity $\mathbb{E}_\varepsilon[|J_r^\varepsilon(\boldsymbol{\psi}_j)|^2 \mid X] = J_r(\boldsymbol{\psi}_j^2)$, none of which requires the assumption of identical distributions. Thus the same inequality holds even under i.n.i.d. setting. □

Remark 1. Although de la Peña and Giné (1999) state the hypercontractivity of Rademacher chaos and the randomization theorem for U-statistics with index-independent kernels, we can regard (i, X_i) as a single coordinate so we can utilize both theorems in our setting.

Proof of Theorem 3. The following argument is essentially same as that of proof of Theorem 8 in Imai and Koike (2025).

First, since $\max_{j\in[p]} J_r(\boldsymbol{\psi}_j) \le n^r \max_{j\in[p]} \max_{(i_1,\dots,i_r)\in I_{n,r}} |\psi_j(X_{i_1},\dots,X_{i_r})|$, the claim trivially holds if $q + \log p > n$; hence it suffices to consider the case $q + \log p \le n$.

We prove the claim by induction on r. It is trivial when $r = 0$. Next, suppose $r > 0$ and that the claim holds for all non-negative integers less than r. We are going to show that there exists a constant $c_1 \ge 1$ depending only on r such that the statement of Theorem 3 holds. By Hoeffding decomposition, we have

$$
\begin{aligned}
I &:= \left\| \max_{j\in[p]} J_r(\boldsymbol{\psi}_j) \right\|_{L^q(\mathbb{P})} \\
&\le \max_{j\in[p]} \mathbb{E}[J_r(\boldsymbol{\psi}_j)] + \sum_{s=1}^{r} \frac{1}{s!} \left\| \max_{j\in[p]} \sum_{(i_1,\dots,i_s)\in I_{n,s}} \pi_{s,(i_1,\dots,i_s)}\psi(X_{i_1},\dots,X_{i_s}) \right\|_{L^q(\mathbb{P})}
\end{aligned}
$$

For every s, Lemma 6 gives

$$
\begin{aligned}
&\left\| \max_{j\in[p]} \sum_{(i_1,\dots,i_s)\in I_{n,s}} \pi_{s,(i_1,\dots,i_s)}\psi_j(X_{i_1},\dots,X_{i_s}) \right\|_{L^q(\mathbb{P})} \\
&\le C_r(q+\log p)^{s/2} \left\| \max_{j\in[p]} \sqrt{\sum_{(i_1,\dots,i_s)\in I_{n,s}} (\pi_{s,(i_1,\dots,i_s)}\psi_j)^2(X_{i_1},\dots,X_{i_s})} \right\|_{L^q(\mathbb{P})}.
\end{aligned}
$$

Recall the definition of Hoeffding projections;

$$
\begin{aligned}
&\pi_{s,(i_1,\dots,i_s)}\psi(X_{i_1},\dots,X_{i_s}) \\
&= \sum_{(k_1,\dots,k_{r-s})\in K_{n,r-s}(i_1,\dots,i_s)} \left(\prod_{l=1}^{s} (\delta_{X_{i_l}} - P_{i_l}) \prod_{m=1}^{r-s} P_{k_m}\psi_j(X_{i_1},\dots,X_{i_s},X_{k_1},\dots,X_{k_{r-s}}) \right).
\end{aligned}
$$

Cauchy-Schwarz's inequality and so-called c_r-inequality

$$
\begin{aligned}
&(\pi_{s,(i_1,\dots,i_s)}\psi_j)^2(X_{i_1},\dots,X_{i_s}) \\
&\le C_r n^{r-s} \sum_{(k_1,\dots k_{r-s})\in K_{n,r-s}(i_1,\dots,i_s)} \left(\prod_{l=1}^{s} (\delta_{X_{i_l}} - P_{i_l}) \prod_{m=1}^{r-s} P_{k_m}\psi_j(X_{i_1},\dots,X_{i_s},X_{k_1},\dots,X_{k_{r-s}}) \right)^2 \\
&\le C_r n^{r-s} \sum_{m=0}^{s} \sum_{1\le l(1)<\dots<l(m)\le s} \sum_{(k_1,\dots k_{r-s})\in K_{n,r-s}(i_1,\dots,i_s)}
\end{aligned}
$$

$$\left(\prod_{(u_1,\ldots,u_{s-m})\in K_{s,s-m}(l(1),\ldots,l(m))} P_{i_u} \prod_{l=1}^{r-s} P_{k_l} \psi_j(X_{l(1)},\ldots,X_{l(m)},X_{u_1},\ldots,X_{u_{s-m}},X_{k_1},\ldots,X_{k_{r-s}})\right)^2$$

Thus

$$\sum_{(i_1,\ldots,i_s)\in I_{n,s}} (\pi_{s,(i_1,\ldots,i_s)}\psi_j)^2(X_{i_1},\ldots,X_{i_s})$$

$$\leq C_r n^{r-s} \sum_{m=0}^{s} \sum_{1\leq l(1)<\cdots<l(m)\leq s} \sum_{(i_1,\ldots,i_s)\in I_{n,s}} \sum_{(k_1,\ldots k_{r-s})\in K_{n,r-s}(i_1,\ldots,i_s)}$$

$$\left(\prod_{(u_1,\ldots,u_{s-m})\in K_{s,s-m}(l(1),\ldots,l(m))} P_{i_u} \prod_{l=1}^{r-s} P_{k_l} \psi_j(X_{l(1)},\ldots,X_{l(m)},X_{u_1},\ldots,X_{u_{s-m}},X_{k_1},\ldots,X_{k_{r-s}})\right)^2$$

$$\leq C_r n^{r-s} \sum_{m=0}^{s} \binom{s}{m} \sum_{1\leq i_1<\cdots<i_m\leq n} \sum_{(u_1,\ldots,u_{s-m})\in K_{s,s-m}(i_1,\ldots,i_m)} \sum_{(k_1,\ldots,k_{r-s})\in K_{n,r-s}(i_1,\ldots,i_s,u_1,\ldots,u_{s-m})}$$

$$\left(\prod_{t=1}^{s-m} P_{u_t} \prod_{l=1}^{r-s} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{u_1},\ldots,X_{u_{s-m}},X_{k_1},\ldots,X_{k_{r-s}})\right)^2$$

$$\leq C_r n^{2(r-s)} \sum_{m=0}^{s} \binom{s}{m} \sum_{1\leq i_1<\cdots<i_m\leq n} \sum_{(u_1,\ldots,u_{s-m})\in K_{s,s-m}(i_1,\ldots,i_m)} \max_{(k_1,\ldots k_{r-s})\in K_{n,r-s}(i_1,\ldots,i_s,u_1,\ldots,u_{s-m})}$$

$$\left(\prod_{t=1}^{s-m} P_{u_t} \prod_{l=1}^{r-s} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{u_1},\ldots,X_{u_{s-m}},X_{k_1},\ldots,X_{k_{r-s}})\right)^2$$

$$\leq C_r n^{2(r-s)} \sum_{m=0}^{s} n^{s-m} \binom{s}{m} \sum_{1\leq i_1<\cdots<i_m\leq n} \max_{(u_1,\ldots,u_{s-m})\in K_{s,s-m}(i_1,\ldots,i_m)} \max_{(k_1,\ldots k_{r-s})\in K_{n,r-s}(i_1,\ldots,i_s,u_1,\ldots,u_{s-m})}$$

$$\left(\prod_{t=1}^{s-m} P_{u_t} \prod_{l=1}^{r-s} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{u_1},\ldots,X_{u_{s-m}},X_{k_1},\ldots,X_{k_{r-s}})\right)^2$$

$$= C_r n^{2(r-s)} \sum_{m=0}^{s} n^{s-m} \binom{s}{m}$$

$$\times \sum_{1\leq i_1<\cdots<i_m\leq n} \max_{(k_1,\ldots,k_{r-m})\in K_{n,r-m}(i_1,\ldots,i_m)} \left(\prod_{l=1}^{r-m} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{k_1},\ldots,X_{k_{r-m}})\right)^2$$

$$\leq C_r n^{2(r-s)} \sum_{m=0}^{s} n^{s-m} \max_{1\leq i_1<\cdots<i_m\leq n} \max_{(k_1,\ldots,k_{r-m})\in K_{n,r-m}(i_1,\ldots,i_m)} \left(\prod_{l=1}^{r-m} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{k_1},\ldots,X_{k_{r-m}})\right)$$

$$\times \sum_{1\leq i_1<\cdots<i_m\leq n} \max_{(k_1,\ldots,k_{r-m})\in K_{n,r-m}(i_1,\ldots,i_m)} \left(\prod_{l=1}^{r-m} P_{k_l} \psi_j(X_{i_1},\ldots,X_{i_m},X_{k_1},\ldots,X_{k_{r-m}})\right)$$

From this result, the inequality $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ for any $x, y \geq 0$, Minkowski's inequality and

Cauchy-Schwarz's inequality, it holds that

$$\begin{aligned}
&\left\|\max_{j\in[p]}\sqrt{\sum_{(i_1,\dots,i_s)\in I_{n,s}}(\pi_{s,(i_1,\dots,i_s)}\psi_j)^2(X_{i_1},\dots,X_{i_s})}\right\|_{L^q(\mathbb{P})}\\
&\le C_r n^{r-s}\max_{0\le m\le s} n^{(s-m)/2}\\
&\times\left\|\max_{j\in[p]}\max_{(i_1,\dots,i_m)\in I_{n,m}}\max_{(k_1,\dots,k_{r-m})\in K_{n,r-m}(i_1,\dots,i_m)}\prod_{l=1}^{r-m}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}^{1/2}\\
&\times\left\|\max_{j\in[p]}\sum_{(i_1,\dots,i_m)\in I_{n,m}}\max_{(k_1,\dots,k_{r-m})\in K_{n,r-m}(i_1,\dots,i_m)}\prod_{l=1}^{r-m}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}^{1/2}.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
&\sum_{s=1}^{r}\left\|\max_{j\in[p]}\sum_{(i_1,\dots,i_s)\in I_{n,s}}\pi_{s,(i_1,\dots,i_s)}\psi(X_{i_1},\dots,X_{i_s})\right\|_{L^q(\mathbb{P})}\\
&\le C_r\sum_{s=1}^{r}n^{r-s}(q+\log p)^{s/2}\max_{0\le m\le s} n^{(s-m)/2}\\
&\times\left\|\max_{j\in[p]}\max_{(i_1,\dots,i_m)\in I_{n,m}}\max_{(k_1,\dots,k_{r-m})\in K_{n,r-m}(i_1,\dots,i_m)}\prod_{l=1}^{r-m}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}^{1/2}\\
&\times\left\|\max_{j\in[p]}\sum_{(i_1,\dots,i_m)\in I_{n,m}}\max_{(k_1,\dots,k_{r-m})\in K_{n,r-m}(i_1,\dots,i_m)}\prod_{l=1}^{r-m}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}^{1/2}.
\end{aligned}$$

Now, by the assumption of the induction, for every $0\le m<r$, there exists a comstant $c_m\ge 1$ depending only on m such that

$$\begin{aligned}
&\left\|\max_{j\in[p]}\sum_{(i_1,\dots,i_m)\in I_{n,m}}\max_{(k_1,\dots,k_{r-m})\in K_{n,r-m}(i_1,\dots,i_m)}\prod_{l=1}^{r-m}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}\\
&\le c_m\max_{0\le t\le m} n^{m-t}(q+\log p)^t\left\|\max_{j\in[p]}\max_{(i_1,\dots,i_t)\in I_{n,t}}\max_{(k_1,\dots,k_{r-t})\in K_{n,r-t}(i_1,\dots,i_t)}\prod_{l=1}^{r-t}P_{k_l}\psi_j\right\|_{L^q(\mathbb{P})}.
\end{aligned}$$

The remainder of the proof is identical to that of Imai and Koike (2025). □

Proof of Theorem 2. The theorem is the immediate consequence of Theorem 3 (cf. Imai and Koike, 2025). □

C.2 Proof of Lemma 2

Before the proof of Lemma 2, we restate Lemma 16 in Imai and Koike (2025) adapted to our i.n.i.d. setting.

Lemma 7. *Let $t \geq 1$. Assume $\psi_{j,(i,m)} \in L^t(P_i \otimes P_m)$ for all $j \in [p]$ and $(i,m) \in I_{n,2}$. Then, there exists a universal constant C such that, for every $i \in [n]$*

$$\mathbb{E}\left[\max_{j\in[p]} \int \left| \sum_{m\in[n]:m\neq i} \{\psi_j(X_m, x_i) - (P_m\psi_j)(x_i)\} \right|^t dP_i(x_i)\right]$$
$$\leq C(t + \log p)^{t/2}\, \mathbb{E}\left[\max_{j\in[p]} \int \left(\sum_{m\in[n]:m\neq i} \psi_j^2(x_i, X_m) \right)^{t/2} dP_i(x_i)\right]. \tag{16}$$

Proof. With fixed i, the proof is almost identical to that of Lemma 16 in Imai and Koike (2025); we only replace the law of common distribution with the index-dependent laws. $\square$

Proof of Lemma 2. First, we prove Eq. (8). By Lemma 7,

$$\begin{aligned}
I &:= \max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \left| \sum_{m\in[n]:m>i} \psi_j(x_i, X_m) \right|^4 dP_i(x_i)\right] \\
&\lesssim \max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \left(\sum_{m\neq i} \psi_j^2(x_i, X_m) \right)^2 dP_i(x_i)\right] (\log p)^2 \\
&= \max_{i\in[n]} \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]} \int \left(\sum_{m\neq i} (P_m\psi_j^2)(x_i) + \psi_j^2(x_i, X_m) - (P_m\psi_j^2)(x_i) \right)^2 dP_i(x_i)\right] (\log p)^2 \\
&\lesssim \max_{i\in[n]} \left(\max_{i\in[n]}\max_{j\in[p]} \int \left(\sum_{m\neq i} (P_m\psi_j^2)(x_i) \right)^2 dP_i(x_i) \right. \\
&\quad \left. + \max_{i\in[n]} \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]} \int \left| \sum_{m\neq i} \{\psi_j^2(x_i, X_m) - (P_m\psi_j^2)(x_i)\} \right|^2 dP_i(x_i)\right] \right) (\log p)^2 \\
&=: I_1 + I_2.
\end{aligned}$$

By definition,

$$I_1 \leq n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(\psi_j^2)\|^2_{L^2(P_i)} \log^2 p.$$

Applying Lemma 7 to functions $(y,x) \to \psi_j(y,x)^2$, we have

$$I_2 \leq \max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \int \sum_{m\neq i} \psi_j^4(x_i, X_m) dP_i(x_i)\right] (\log p)^3 = \max_{i\in[n]} \mathbb{E}\left[\max_{j\in[p]} \sum_{m\neq i} P_i\psi_j^4(X_m)\right] (\log p)^3.$$

By Lemma 9 in Chernozhukov et al. (2015),

$$\mathbb{E}\left[\max_{j\in[p]} \sum_{m\neq i} P_i\psi_j^4(X_m)\right] \lesssim \max_{j\in[p]} \sum_{m\neq i} P_m P_i \psi_j^4 + \mathbb{E}\left[\max_{m:m\neq i}\max_{j\in[p]} P_i\psi_j^4(X_m)\right] \log p$$

$$\leq n \max_{j\in[p]} \max_{m:m\neq i} \|\psi_j\|^4_{L^4(P_i\otimes P_m)} + \mathbb{E}\left[\max_{m:m\neq i}\max_{j\in[p]} P_i\psi_j^4(X_m)\right]\log p.$$

Consequently, we have Eq. (8). The proof of Eq. (9) is almost same as that of Eq. (8).

Next, we prove Eq. (10).

$$\begin{aligned}
II &:= \mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n\left|\sum_{m\in[n]:m\neq i}\psi_j(X_i,X_m)\right|^4\right] \\
&\leq 8\left(\mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n\left|\sum_{m\in[n]:m<i}\psi_j(X_i,X_m)\right|^4\right] + \mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n\left|\sum_{m\in[n]:m>i}\psi_j(X_i,X_m)\right|^4\right]\right) \\
&=: 8(II_1 + II_2).
\end{aligned} \tag{17}$$

We can construct a bound on II_1 likewise the proof of Lemma 3 in Imai and Koike (2025), which uses Lemma 2 in Imai and Koike (2025). Define a filtration $(\mathcal{G}_i)_{i=1}^n$ as $\mathcal{G}_i := \sigma(X_1,\dots,X_i)$ for $i\in[n]$. Also, for every $i\in[n]$, define a random vector $\eta_i^< = (\eta_{i1}^<,\dots,\eta_{ip}^<)^\top$ as

$$\eta_{ij}^< := \left|\sum_{m\in[n]:m<i}\psi_j(X_i,X_m)\right|^4, \quad j=1,\dots,p.$$

Then, $(\eta_i^<)_{i=1}^n$ is adapted to the filtration $(\mathcal{G}_i)_{i=1}^n$. Hence Lemma 2 in Imai and Koike (2025) gives

$$II_1 \lesssim \mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n \mathbb{E}[\eta_{ij}^< \mid \mathcal{G}_{i-1}]\right] + \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\eta_{ij}^<\right]\log p =: III_1 + IV_1\log p,$$

where we set $\mathcal{G}_0 := \{\emptyset,\Omega\}$. Since X_i is independent of $\mathcal{G}_{i-1}$ for every i, by the same argument as that of the proof of Lemma 3 in Imai and Koike (2025), Eq. (9) gives

$$\begin{aligned}
III_1 \lesssim\ & n^3\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(\psi_j^2)\|^2_{L^2(P_i)}\log^2 p \\
&+ n^2\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|\psi_j\|^4_{L^4(P_i\otimes P_m)}\log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i\psi_j^4(X_m)\right]\log^4 p.
\end{aligned}$$

To bound IV_1, observe that

$$IV_1 = \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left|\sum_{m\in[n]:m<i}\psi_j(X_i,X_m)\right|^4\right] = \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left|\sum_{m=1}^n Y_{m,(i,j)}^<\right|^4\right],$$

where $Y_{m,(i,j)}^< = \psi_j(X_i,X_m)1\{m<i\}$. Unlike Imai and Koike (2025), we can not use time-reversal symmetry, so we define $\tilde{\mathcal{G}}_m^{(i)} := \sigma(X_1,\dots,X_m,X_i)$. Then, $(Y_{m,(i,j)}^<)_{m=1}^n$ is a martingale difference sequence with respect to $(\tilde{\mathcal{G}}_m^{(i)})_{m=0}^n$ for all $i\in[n]$ and $j\in[p]$. Hence Lemma 1 in Imai and Koike (2025) gives

$$IV_1 \lesssim \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left(\sum_{m=1}^n\mathbb{E}\left[(Y_{m,(i,j)}^<)^2 \mid \tilde{\mathcal{G}}_{m-1}^{(i)}\right]\right)^2\right]\log^2(np) + \mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}|Y_{m,(i,j)}^<|^4\right]\log^4(np)$$

$$\leq n^2\,\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\{P_m\psi_j^2(X_i)\}^2\right]\log^2(np)+\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\psi_j^4(X_i,X_m)\right]\log^4(np).$$

Define a filtration $(\mathcal{G}_i^>)_{i=1}^n$ as $\mathcal{G}_i^> := \sigma(X_i, X_{i+1},\dots,X_n)$. Also, for every $i\in[n]$, define a random vector $\eta_i^> = (\eta_{i1}^>,\dots,\eta_{ip}^>)^\top$ as

$$\eta_{ij}^> := \left|\sum_{m\in[n]:m>i}\psi_j(X_i,X_m)\right|^4, \quad j=1,\dots,p.$$

Then, $(\eta_i^>)_{i=1}^n$ is adapted to the filtration $(\mathcal{G}_i^>)_{i=1}^n$. Hence Lemma 2 in Imai and Koike (2025) gives

$$II_2 \lesssim \mathbb{E}\left[\max_{j\in[p]}\sum_{i=1}^n\mathbb{E}[\eta_{ij}^> \mid \mathcal{G}_{i+1}^>]\right]+\mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\eta_{ij}^>\right]\log p =: III_2 + IV_2\log p,$$

where we set $\mathcal{G}_{n+1}^> := \{\emptyset,\Omega\}$. Since X_i is independent of $\mathcal{G}_{i+1}^>$, for every i, by the same argument as that of the proof of Lemma 3 in Imai and Koike (2025), Eq. (8) gives

$$\begin{aligned}III_2 &\lesssim n^3\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(\psi_j^2)\|_{L^2(P_i)}^2\log^2 p\\ &\quad + n^2\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|\psi_j\|_{L^4(P_i\otimes P_m)}^4\log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i\psi_j^4(X_m)\right]\log^4 p.\end{aligned}$$

To bound IV_2, observe that

$$IV_2 = \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left|\sum_{m\in[n]:m>i}\psi_j(X_i,X_m)\right|^4\right] = \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left|\sum_{m=1}^n Y_{m,(i,j)}^>\right|^4\right],$$

where $Y_{m,(i,j)}^> = \psi_j(X_i,X_m)1\{m>i\}$. Observe that $(Y_{m,(i,j)}^>)_{m=1}^n$ is martingale difference sequence with respect to $(\mathcal{G}_m)_{m=0}^n$ for all $i\in[n]$ and $j\in[p]$. Hence Lemma 1 in Imai and Koike (2025) gives

$$\begin{aligned}IV_2 &\lesssim \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[p]}\left(\sum_{m=1}^n\mathbb{E}[(Y_{m,(i,j)})^2\mid\mathcal{G}_{m-1}]\right)^2\right]\log^2(np)+\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}|Y_{m,(i,j)}|^4\right]\log^4(np)\\ &\leq n^2\,\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\{P_m\psi_j^2(X_i)\}^2\right]\log^2(np)+\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\psi_j^4(X_i,X_m)\right]\log^4(np).\end{aligned}$$

Summing up,

$$\begin{aligned}II &\lesssim n^3\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(\psi_j^2)\|_{L^2(P_i)}^2\log^2 p\\ &\quad + n^2\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|\psi_j\|_{L^4(P_i\otimes P_m)}^4\log^3 p + n\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i\psi_j^4(X_m)\right]\log^4 p\\ &\quad + n^2\,\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\{P_m\psi_j^2(X_i)\}^2\right]\log^3(np)+\mathbb{E}\left[\max_{(i,m)\in I_{n,2}}\max_{j\in[p]}\psi_j^4(X_i,X_m)\right]\log^5(np).\end{aligned}$$

□

C.3 Proof of Lemma 3

Before the proof of Lemma 3, we introduce the following auxiliary results. The proof of these results is in Section C.3.1.

Lemma 8. *It holds that*

$$\max_{(j,k)\in[p]^2}\max_{i\in[n]}\int(\chi^{(1,1)}_{i,jk})^2(x_i)dP_i(x_i)\le\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|^4_{L^4(P_i)}, \tag{18}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{i\in[n]}(\chi^{(1,1)}_{i,jk})^2(X_i)\right\|_{L^1(\mathbb{P})}\le\mathbb{E}\left[\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^4(X_i)\right]. \tag{19}$$

Lemma 9. *It holds that*

$$\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde\chi^{(1,2)}_{im,jk})^2(x_i,x_m)dP_i(x_i)dP_m(x_m)\\
&\quad\le 2\left(\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|^2_{L^4(P_i)}\right)\left(\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(|\pi_{2,im}\psi_k|^2)\|_{L^2(P_i)}\right)\\
&\qquad+2\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\|\pi_{1,i}\psi_j\star^1_i\pi_{2,im}\psi_k\|^2_{L^2(P_m)},
\end{aligned} \tag{20}$$

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde\chi^{(1,2)}_{im,jk})^2(X_i,x_m)dP_m(x_m)\right\|_{L^1(\mathbb{P})}\\
&\quad\le\mathbb{E}\left[\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2}\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right]^{1/2}\\
&\qquad+\max_{j\in[p]}\max_{m\in[n]}\|\pi_{1,m}\psi_j\|^2_{L^4(P_m)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}P_m\{(\pi_{2,mi}\psi_k)^4(X_m,X_i)\}\right]^{1/2}\\
&\qquad+2\max_{j\in[p]}\max_{m\in[n]}\|\pi_{1,m}\psi_j\|^2_{L^2(P_m)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}(P_m\{\pi_{2,mi}\psi_k(X_m,X_i)\}^2)^2\right]^{1/2},
\end{aligned} \tag{21}$$

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\tilde\chi^{(1,2)}_{im,jk})^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}\\
&\quad\le 2\,\mathbb{E}\left[\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2}\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_k)^4(X_i,X_m)\right]^{1/2}\\
&\qquad+2\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|^2_{L^2(P_i)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right]^{1/2}.
\end{aligned} \tag{22}$$

Lemma 10. *It holds that*

$$\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{i\in[n]}\int(\varphi^{(1,2)}_{i,jk})^2(x_i)dP_i(x_i)\\
&\quad\le 4n^2\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\|\pi_{1,i}\psi_j\star^1_i\pi_{2,im}\psi_k\|^2_{L^2(P_m)},
\end{aligned} \tag{23}$$

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{i\in[n]}(\varphi^{(1,2)}_{i,jk})^2(X_i)\right\|_{L^1(\mathbb{P})}\\
&\quad\le 4n^2\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|^2_{L^2(P_i)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\{P_i(\pi_{2,im}\psi_k)^2(X_m)\}^2\right]^{1/2}.
\end{aligned} \tag{24}$$

Lemma 11. *It holds that*

$$\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde{\chi}^{(2,2),\texttt{diag}}_{im,jk})^2(x_i,x_m)dP_i(x_i)dP_m(x_m)$$
$$\leq \max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|\pi_{2,im}\psi_j\|^4_{L^4(P_i\otimes P_m)}+\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_i(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_m)}, \tag{25}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde{\chi}^{(2,2),\texttt{diag}}_{im,jk})^2(X_i,x_m)dP_m(x_m)\right\|_{L^1(\mathbb{P})}$$
$$\leq \frac{1}{4}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,im}\psi_j)^4(X_i)\right], \tag{26}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\tilde{\chi}^{(2,2),\texttt{diag}}_{im,jk})^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}$$
$$\leq \frac{1}{4}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right]. \tag{27}$$

Lemma 12. *It holds that*

$$\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(x_i,x_m,x_l)dP_i(x_i)dP_m(x_m)dP_l(x_l)$$
$$\leq \frac{1}{2}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_i)}+\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star^1_i\pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}, \tag{28}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$\leq \frac{1}{6}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\big(P_m(\pi_{2,im}\psi_j)^2(X_i)\big)^2\right]+\frac{1}{6}\max_{j\in[p]}\max_{(i,m\in I_{n,2}}\big\|P_m(|\pi_{2,im}\psi_j|^2)\big\|^2_{L^2(P_i)}$$
$$+\frac{2}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/2}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|_{L^2(P_m)}, \tag{29}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$\leq \frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right]^{1/2}\mathbb{E}\left[\max_{k\in[p]}\max_{(i,l)\in I_{n,2}}\big(P_l(\pi_{2,il}\psi_k)^2(X_i)\big)^2\right]^{1/2}$$
$$+\frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i(\pi_{2,im}\psi_j)^4(X_m)\right]^{1/2}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|\pi_{2,il}\psi_k\|^2_{L^4(P_i\otimes P_l)}$$
$$+\frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_l(\pi_{2,li}\psi_j)^4(X_i)\right], \tag{30}$$

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,X_l)\right\|_{L^1(\mathbb{P})}\leq \mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i(\pi_{2,im}\psi_j)^4(X_m)\right]. \tag{31}$$

Lemma 13. *It holds that*

$$\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(x_m,x_l)dP_m(x_m)dP_l(x_l)$$

$$\leq n^2 \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}, \tag{32}$$

$$\left\| \max_{(j,k)\in[p]^2} \max_{(m,l)\in I_{n,2}} \int (\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m, x_l) dP_l(x_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}} \|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)} \, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_k)^2(X_m)\right)^2\right]^{1/2}, \tag{33}$$

$$\left\| \max_{(j,k)\in[p]^2} \max_{(m,l)\in I_{n,2}} (\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m, X_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq n^2 \, \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]. \tag{34}$$

Lemma 14. *It holds that*

$$\max_{(j,k)\in[p]^2} \max_{m\in[n]} \int (\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(x_m) dP_m(x_m) \leq 4n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\|P_i(|\pi_{2,im}\psi_j|^2)\right\|^2_{L^2(P_m)}, \tag{35}$$

$$\left\| \max_{(j,k)\in[p]^2} \max_{m\in[n]} (\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(X_m) \right\|_{L^1(\mathbb{P})} \leq 4n^2 \, \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]. \tag{36}$$

Proof of Lemma 3

Proof of Eq. (12): Theorem 2 and Lemma 8 give

$$\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_1(\boldsymbol{\chi}^{(1,1)}_{jk})|\right] (\log p)^{3/2}$$

$$\leq C\left(n^{\frac{1}{2}} \log^2 p \left(\max_{(j,k)\in[p]^2} \max_{i\in[n]} \int (\chi^{(1,1)}_{i,jk})^2(x_i) dP_i(x_i)\right)^{1/2} \vee \log^{5/2} p \left\| \max_{(j,k)\in[p]^2} \max_{i\in[n]} (\chi^{(1,1)}_{i,jk})^2(X_i) \right\|^{1/2}_{L^1(\mathbb{P})}\right)$$

$$\leq C n^{1/2} \log^2 p \sqrt{\max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|^4_{L^4(P_i)}} + \log^{5/2} p \, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2}$$

$$\leq C\sqrt{\left(\Delta^{(1)}_{2,*}(1) + \Delta^{(2)}_{2,*}(1)\right) \log^5 p}.$$

Proof of Eq. (13): Theorem 2 and Lemma 9 give

$$\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\chi}}^{(1,2)}_{jk})|\right] (\log p)^{3/2}$$

$$\leq C\left(n \log^{5/2} p \left(\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}^{(1,2)}_{im,jk})^2(x_i, x_m) dP_i(x_i) dP_m(x_m)\right)^{1/2}\right.$$

$$\vee n^{1/2} \log^3 p \left\| \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}^{(1,2)}_{im,jk})^2(X_i, x_m) dP_m(x_m) \right\|^{1/2}_{L^1(\mathbb{P})}$$

$$
\begin{aligned}
&\qquad \vee \log^{7/2} p \left\| \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (\tilde{\chi}^{(1,2)}_{im,jk})^2(X_i, X_m) \right\|^{1/2}_{L^1(\mathbb{P})} \Bigg) \\
&\le Cn \log^{5/2} p \left(\max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j(X_i)\|^2_{L^4(P_i)} \right)^{1/2} \left(\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(|\pi_{2,im}\psi_k|^2)\|_{L^2(P_i)} \right)^{1/2} \\
&\quad + Cn \log^{5/2} p \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star^1_i \pi_{2,im}\psi_k(X_m)\|_{L^2(P_m)} \\
&\quad + Cn^{1/2} \log^3 p\, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i) \right]^{1/4} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \big(P_m(\pi_{2,im}\psi_k)^2(X_i)\big)^2 \right]^{1/4} \\
&\quad + Cn^{1/2} \log^3 p \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^4(P_m)}\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_m(\pi_{2,mi}\psi_k)^4(X_i) \right]^{1/4} \\
&\quad + Cn^{1/2} \log^3 p \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^2(P_m)}\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} (P_m(\pi_{2,mi}\psi_k)^2(X_i))^2 \right] \\
&\quad + C \log^{7/2} p\, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i) \right]^{1/4} \mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (\pi_{2,im}\psi_k)^4(X_i, X_m) \right]^{1/4} \\
&\quad + C \log^{7/2} p \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^2(P_i)}\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \big(P_m(\pi_{2,im}\psi_k)^2(X_i)\big)^2 \right]^{1/4}.
\end{aligned}
$$

From AM-GM inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\chi}}^{(1,2)}_{jk})| \right] (\log p)^{3/2} \\
&\le C \log^{5/2} p \left(n^{1/2} \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j(X_i)\|^2_{L^4(P_i)} + n^{3/2} \max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(|\pi_{2,im}\psi_k|^2)\|_{L^2(P_i)} \right) \\
&\quad + C \left(n \log^{5/2} p \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star^1_i \pi_{2,im}\psi_k(X_m)\|_{L^2(P_m)} \right) \\
&\quad + C \log^2 p \left(\log^{1/2} p\, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i) \right]^{1/2} + n \log^{3/2} p\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \big(P_m(\pi_{2,im}\psi_k)^2(X_i)\big)^2 \right]^{1/2} \right) \\
&\quad + C \log^2 p \left(n^{1/2} \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|^2_{L^4(P_m)} + n^{1/2} \log^2 p\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_m(\pi_{2,mi}\psi_k)^4(X_i) \right]^{1/2} \right) \\
&\quad + C \log^{9/4} \left(n^{1/2} \log^{3/4} p \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^2(P_m)}\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} (P_m(\pi_{2,mi}\psi_k)^2(X_i))^2 \right]^{1/4} \right) \\
&\quad + C \log^2 p \left(\log^{1/2} p\, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i) \right]^{1/2} + \log^{5/2} p\, \mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (\pi_{2,im}\psi_k)^4(X_i, X_m) \right]^{1/2} \right) \\
&\quad + C \log^{5/2} p \left(\log^{1/2} p\, \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4 \right]^{1/2} + \log^{3/2} p\, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \big(P_m(\pi_{2,im}\psi_k)^2(X_i)\big)^2 \right]^{1/2} \right).
\end{aligned}
$$

Then, it holds that

$$
\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_2(\tilde{\boldsymbol{\chi}}^{(1,2)}_{jk})| \right] (\log p)^{3/2}
$$

$$
\begin{aligned}
&\le C\Bigg[\log^{5/2} p\left(\sqrt{\Delta_{2,*}^{(1)}(1)}+\sqrt{\Delta_{2,*}^{(2)}(2)}\right)+\frac{1}{\sqrt{n}}\Delta_1^{(1)}\log^{5/2} p \\
&\quad+\log^2 p\left(\sqrt{\Delta_{2,*}^{(2)}(1)}+\sqrt{\Delta_{2,*}^{(5)}(2)}\right)+\log^2 p\left(\sqrt{\Delta_{2,*}^{(1)}(1)}+\sqrt{\Delta_{2,*}^{(3)}(2)}\right) \\
&\quad+\max_{j\in[p]}\max_{m\in[n]}\|\pi_{1,m}\psi_j\|_{L^2(P_m)}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4}+\log^{5/2} p\left(\sqrt{\Delta_{2,*}^{(2)}(1)}+\sqrt{\Delta_{2,*}^{(5)}(2)}\right)\Bigg] \\
&\le C\Bigg(\frac{1}{\sqrt{n}}\Delta_1^{(1)}\log^{5/2} p+\max_{j\in[p]}\max_{m\in[n]}\|\pi_{1,m}\psi_j\|_{L^2(P_m)}\left(\Delta_{2,*}^{(5)}(2)\log^9 p\right)^{1/4} \\
&\qquad\qquad+\sqrt{(\Delta_{2,*}(1)+\Delta_{2,*}(2))\log^5 p}\Bigg). \qquad (37)
\end{aligned}
$$

Also, Theorem 2 and Lemma 10 give

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_1(\boldsymbol{\varphi}_{jk}^{(1,2)})|\right](\log p)^{3/2} \\
&\le C\left(n^{\frac{1}{2}}\log^2 p\left(\max_{(j,k)\in[p]^2}\max_{m\in[n]}\int(\varphi_{m,jk}^{(1,2)})^2(x_m)dP_m(x_m)\right)^{1/2}\vee\log^{5/2} p\left\|\max_{(j,k)\in[p]^2}\max_{m\in[n]}(\varphi_{m,jk}^{(1,2)})^2(X_m)\right\|_{L^1(\mathbb{P})}^{1/2}\right) \\
&\le C\Bigg(n^{3/2}\log^2 p\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\|\pi_{1,i}\psi_j\star_i^1\pi_{2,im}\psi_k\|_{L^2(P_m)} \\
&\qquad+n\log^{5/2} p\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|_{L^2(P_i)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\{P_i(\pi_{2,im}\psi_k)^2(X_m)\}^2\right]^{1/4}\Bigg) \\
&\le C\left(\Delta_1^{(1)}\log^2 p+n^{1/2}\max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|_{L^2(P_i)}\left(\Delta_{2,*}^{(5)}(2)\log^7 p\right)^{1/4}\right). \qquad (38)
\end{aligned}
$$

From Eq. (37) and Eq. (38), we have Eq. (13).

Proof of Eq. (14): From Theorem 2 and Lemma 11,

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\chi}}_{jk}^{(2,2),\mathtt{diag}})|\right](\log p)^{3/2} \\
&\le C\Bigg(n\log^{5/2} p\left(\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde{\chi}_{im,jk}^{(2,2),\mathtt{diag}})^2(x_i,x_m)dP_i(x_i)dP_m(x_m)\right)^{1/2} \\
&\qquad\vee n^{1/2}\log^3 p\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde{\chi}_{im,jk}^{(2,2),\mathtt{diag}})^2(X_i,x_m)dP_m(x_m)\right\|_{L^1(\mathbb{P})}^{1/2} \\
&\qquad\vee\log^{7/2} p\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\tilde{\chi}_{im,jk}^{(2,2),\mathtt{diag}})^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}^{1/2}\Bigg) \\
&\le Cn\log^{5/2} p\left(\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|\pi_{2,im}\psi_j\|_{L^4(P_i\otimes P_m)}^2+\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_i(|\pi_{2,im}\psi_j|^2)\|_{L^2(P_m)}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ Cn^{1/2}\log^3 p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_m(\pi_{2,im}\psi_j)^4(X_i)\right]^{1/2} \\
&+ C\log^{7/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} (\pi_{2,im}\psi_j)^4(X_i,X_m)\right]^{1/2} \\
&\le C\sqrt{\Delta_{2,*}(2)\log^5 p}. \qquad (39)
\end{aligned}
$$

From Theorem 2 and Lemma 12,

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_3(\tilde{\boldsymbol{\chi}}^{(2,2)}_{jk})|\right](\log p)^{3/2} \\
&\le C\Bigg(n^{3/2}\log^3 p\left(\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(x_i,x_m,x_l)dP_i(x_i)dP_m(x_m)dP_l(x_l)\right)^{1/2} \\
&\qquad \vee n\log^{7/2} p\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|^{1/2}_{L^1(\mathbb{P})} \\
&\qquad \vee n^{1/2}\log^4 p\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,x_l)dP_l(x_l)\right\|^{1/2}_{L^1(\mathbb{P})} \\
&\qquad \vee \log^{9/2}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,X_l)\right\|^{1/2}_{L^1(\mathbb{P})}\Bigg) \\
&\le Cn^{3/2}\log^3 p\left(\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(|\pi_{2,im}\psi_j|^2)\|_{L^2(P_i)} + \max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\right) \\
&\quad + Cn\log^{7/2} p\Bigg(\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_m(\pi_{2,im}\psi_j)^2(X_i)\right)^2\right]^{1/2} + \max_{j\in[p]}\max_{(i,m\in I_{n,2}}\left\|P_m(|\pi_{2,im}\psi_j|^2)\right\|_{L^2(P_i)} \\
&\quad + \mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/4}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|^{1/2}_{L^2(P_m)}\Bigg) \\
&\quad + Cn^{1/2}\log^4 p\Bigg(\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right]^{1/4}\mathbb{E}\left[\max_{k\in[p]}\max_{(i,l)\in I_{n,2}}\left(P_l(\pi_{2,il}\psi_k)^2(X_i)\right)^2\right]^{1/4} \\
&\quad + \mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right]^{1/4}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|\pi_{2,il}\psi_k\|_{L^4(P_i\otimes P_l)} \\
&\quad + \mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_l(\pi_{2,li}\psi_j)^4(X_i)\right]^{1/2}\Bigg) \\
&\quad + C\log^{9/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4\right]^{1/2}.
\end{aligned}
$$

From AM-GM inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2} |J_3(\tilde{\boldsymbol{\chi}}^{(2,2)}_{jk})|\right](\log p)^{3/2} \\
&\le Cn^{3/2}\log^3 p\left(\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(|\pi_{2,im}\psi_j|^2)\|_{L^2(P_i)} + \max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ C\log^{5/2} p\Bigg(n\log p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\big(P_m(\pi_{2,im}\psi_j)^2(X_i)\big)^2\right]^{1/2} \\
&\qquad + n\log p\max_{j\in[p]}\max_{(i,m\in I_{n,2}}\big\|P_m(|\pi_{2,im}\psi_j|^2)\big\|_{L^2(P_i)}\Bigg) \\
&+ C\log^{9/4} p\Bigg(n^{1/2}\log^{3/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/2} \\
&\qquad + n^{3/2}\log p\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|^{1/2}_{L^2(P_m)}\Bigg) \\
&+ C\log^{9/4} p\Bigg(\log^{5/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right]^{1/2} \\
&\qquad + n\log p\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,l)\in I_{n,2}}\big(P_l(\pi_{2,il}\psi_k)^2(X_i)\big)^2\right]^{1/2}\Bigg) \\
&\quad + C\log^{9/4} p\left(\log^2 p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right]^{1/2} + n\log^{3/2} p\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|\pi_{2,il}\psi_k\|^2_{L^4(P_i\otimes P_l)}\right) \\
&\quad + C\log^2 p\left(n^{1/2}\log^2 p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_l(\pi_{2,li}\psi_j)^4(X_i)\right]^{1/2}\right) \\
&+ C\log^{9/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4\right]^{1/2} \\
&\le C\left(\frac{1}{\sqrt{n}}\Delta_1^{(0)}\log^3 p + \sqrt{\Delta_{2,*}(2)\log^5 p}\right).
\end{aligned}
\tag{40}
$$

Also, from Theorem 2 and Lemma 13, we have

$$
\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\varphi}}^{(2,2)}_{jk})|\right](\log p)^{3/2} \\
&\le C\Bigg(n\log^{5/2} p\left(\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(x_m,x_l)dP_m(x_m)dP_l(x_l)\right)^{1/2} \\
&\qquad \vee n^{1/2}\log^3 p\left\|\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m,x_l)dP_l(x_l)\right\|^{1/2}_{L^1(\mathbb{P})} \\
&\qquad \vee \log^{7/2} p\left\|\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m,X_l)\right\|^{1/2}_{L^1(\mathbb{P})}\Bigg) \\
&\le C\Bigg(n^2\log^{5/2} p\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)} \\
&\quad + n^{3/2}\log^3 p\max_{(j,k)\in[p]^2}\max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|^{1/2}_{L^2(P_m\otimes P_l)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\big(P_i(\pi_{2,im}\psi_k)^2(X_m)\big)^2\right]^{1/4}
\end{aligned}
$$

$$+ n\log^{7/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]^{1/2}\Bigg).$$

From AM-GM inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_2(\tilde{\boldsymbol{\varphi}}^{(2,2)}_{jk})|\right](\log p)^{3/2}\\
&\le C\Bigg(n^2\log^{5/2} p\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\Bigg)\\
&\quad+C\Bigg(n^2\log^3 p\max_{(j,k)\in[p]^2}\max_{\substack{(i,m,l)\in[n]^3\\ m,l\neq i}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)}\\
&\qquad\qquad+n\log^3 p\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_i(\pi_{2,im}\psi_k)^2(X_m)\right)^2\right]^{1/2}\Bigg)\\
&\quad+C\log^2 p\left(n\log^{3/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]^{1/2}\right)\\
&\le C\left(\Delta_1^{(0)}\log^3 p+\sqrt{\Delta_{2,*}(2)\log^5 p}\right).
\end{aligned}\tag{41}$$

Also, from Theorem 2 and Lemma 14, we have

$$\begin{aligned}
&\mathbb{E}\left[\max_{(j,k)\in[p]^2}|J_1(\boldsymbol{\varphi}^{(2,2),\mathtt{diag}}_{jk})|\right](\log p)^{3/2}\\
&\le C\Bigg(n^{\frac12}\log^2 p\left(\max_{(j,k)\in[p]^2}\max_{m\in[n]}\int(\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(x_m)dP_m(x_m)\right)^{1/2}\\
&\qquad\vee\log^{5/2} p\left\|\max_{(j,k)\in[p]^2}\max_{m\in[n]}(\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(X_m)\right\|^{1/2}_{L^1(\mathbb{P})}\Bigg)\\
&\le C\left(n^{3/2}\log^2 p\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left\|P_i(|\pi_{2,im}\psi_j|^2)\right\|_{L^2(P_m)}+n\log^{5/2} p\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]^{1/2}\right)\\
&\le C\sqrt{\Delta_{2,*}(2)\log^5 p}.
\end{aligned}\tag{42}$$

From Eq. (39),Eq. (40), Eq. (41) and Eq. (42), we have Eq. (14).

C.3.1 Proof of Lemma 8-14

Proof of Eq. (18). Recall that

$$(\chi^{(1,1)}_{i,jk})^2(X_i)=\left(\pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)-P_i\{\pi_{1,i}\psi_j(X_i)\pi_{1,i}\psi_k(X_i)\}\right)^2.\tag{43}$$

then, from Jensen inequality, we have

$$\max_{(j,k)\in[p]^2}\max_{i\in[n]}\int(\chi^{(1,1)}_{i,jk})^2(x_i)dP_i(x_i)$$

$$
\begin{aligned}
&\leq \max_{(j,k)\in[p]^2} \max_{i\in[n]} \left(P_i\{(\pi_{1,i}\psi_j)^2(X_i)(\pi_{1,i}\psi_k)^2(X_i)\}\right) \\
&\leq \max_{(j,k)\in[p]^2} \max_{i\in[n]} \left(P_i\{(\pi_{1,i}\psi_j)^4(X_i)\}\right)^{1/2} \left(P_i\{(\pi_{1,i}\psi_k)^4(X_i)\}\right)^{1/2} \\
&\leq \max_{j\in[p]} \max_{i\in[n]} \left(P_i\{(\pi_{1,i}\psi_j)^4(X_i)\}\right),
\end{aligned}
$$

where the first inequality follows from Jensen's inequality and the second inequality follows from Cauchy-Schwarz inequality. □

Proof of Eq. (19). From Eq. (43) and Jensen inequality,

$$
\begin{aligned}
&\left\| \max_{(j,k)\in[p]^2} \max_{i\in[n]} (\chi^{(1,1)}_{i,jk})^2(X_i) \right\|_{L^1(\mathbb{P})} \\
&\leq \left\| \max_{(j,k)\in[p]^2} \max_{i\in[n]} \{(\pi_{1,i}\psi_j)^2(X_i)(\pi_{1,i}\psi_k)^2(X_i)\} \right\|_{L^1(\mathbb{P})}.
\end{aligned}
$$

Since,

$$
\max_{(j,k)\in[p]^2} \max_{i\in[n]} (\pi_{1,i}\psi_j)^2(X_i)(\pi_{1,i}\psi_k)^2(X_i) = \left(\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^2(X_i) \right)^2 = \max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i),
$$

we have

$$
\left\| \max_{(j,k)\in[p]^2} \max_{i\in[n]} (\chi^{(1,1)}_{i,jk})^2(X_i) \right\|_{L^1(\mathbb{P})} \leq \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i) \right]
$$

□

Proof of Eq. (20). Observe that c_r inequality gives

$$
\begin{aligned}
(\bar{\chi}^{(1,2)}_{im,jk})^2(X_i, X_m) &\leq 2(\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i, X_m) + 2\big(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i, X_m)\}\big)^2 \\
(\bar{\chi}^{(1,2)}_{mi,jk})^2(X_m, X_i) &\leq 2(\pi_{1,m}\psi_j)^2(X_m)(\pi_{2,mi}\psi_k)^2(X_m, X_i) + 2\big(P_m\{\pi_{1,m}\psi_j(X_m)\pi_{2,mi}\psi_k(X_m, X_i)\}\big)^2.
\end{aligned}
\tag{44}
$$

From Eq.(5) in Imai and Koike (2025) and Eq. (44), we can see that

$$
\begin{aligned}
&\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}^{(1,2)}_{im,jk})^2(x_i, x_m) dP_i(x_i) dP_m(x_m) \\
&\leq \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\bar{\chi}^{(1,2)}_{im,jk})^2(x_i, x_m) dP_i(x_i) dP_m(x_m) \\
&\leq 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(P_i P_m\{(\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i, X_m)\}\right) \\
&\quad + 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \left(\int \pi_{1,i}\psi_j(x_i)\pi_{2,im}\psi_k(x_i, x_m) dP_i(x_i)\right)^2 dP_m(x_m) \\
&= 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(P_i\{(\pi_{1,i}\psi_j)^2(X_i) P_m(\pi_{2,im}\psi_k)^2(X_i, X_m)\}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}^2 \\
&\le 2 \left(\max_{j\in[p]} \max_{i\in[n]} P_i(\pi_{1,i}\psi_j)^4(X_i)\right)^{1/2} \left(\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_i\{P_m(\pi_{2,im}\psi_k)^2(X_i, X_m)\}^2\right)^{1/2} \\
&\quad + 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}^2 \\
&= 2 \left(\max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j(X_i)\|_{L^4(P_i)}^2\right) \left(\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \|P_m(|\pi_{2,im}\psi_k|^2)\|_{L^2(P_i)}\right) \\
&\quad + 2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}^2.
\end{aligned}
$$

□

Proof of Eq. (21). From Eq. (44), we can see that

$$
\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}_{im,jk}^{(1,2)})^2(X_i, x_m) dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{4} \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \{(\bar{\chi}_{im,jk}^{(1,2)})^2(X_i, x_m) + (\chi_{mi,jk}^{(1,2)})^2(x_m, X_i)\} dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\le \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i, x_m) dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\quad + \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i, x_m)\})^2 dP_m(x_m) \\
&\quad + \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\pi_{1,m}\psi_j)^2(x_m)(\pi_{2,mi}\psi_k)^2(x_m, X_i) dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\quad + \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (P_m\{\pi_{1,m}\psi_j(X_m)\pi_{2,mi}\psi_k(X_m, X_i)\})^2\right\|_{L^1(\mathbb{P})}.
\end{aligned}
$$

In terms of the first term,

$$
\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i, x_m) dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\le \left\|\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^2(X_i) \max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \int (\pi_{2,im}\psi_k)^2(X_i, x_m) dP_m(x_m)\right\|_{L^1(\mathbb{P})} \\
&\le \left\|\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i)\right\|_{L^1(\mathbb{P})}^{1/2} \left\|\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right\|_{L^1(\mathbb{P})}^{1/2} \\
&= \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right]^{1/2}.
\end{aligned}
$$

In terms of the second term, Schwarz inequality gives

$$
\begin{aligned}
&\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i, x_m)\})^2 dP_m(x_m) \\
&= \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k)^2(x_m) dP_m(x_m)
\end{aligned}
$$

$$= \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{1,i}\psi_j \star_i^1 \pi_{2,im}\psi_k\|_{L^2(P_m)}^2$$

$$\le \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|_{L^2(P_i)}^2 \, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right]^{1/2}.$$

In terms of the third term, form Schwarz inequality and Jensen inequality, we have

$$\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\pi_{1,m}\psi_j)^2(x_m)(\pi_{2,mi}\psi_k)^2(x_m, X_i)dP_m(x_m)\right\|_{L^1(\mathbb{P})}$$

$$\le \max_{j\in[p]} \max_{m\in[n]} \left(\int (\pi_{1,m}\psi_j)^4(x_m)dP_m(x_m)\right)^{1/2} \left\|\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(\int (\pi_{2,mi}\psi_k)^4(x_m, X_i)dP_m(x_m)\right)^{1/2}\right\|_{L^1(\mathbb{P})}$$

$$\le \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^4(P_m)}^2 \, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_m\{(\pi_{2,mi}\psi_k)^4(X_m, X_i)\}\right]^{1/2}.$$

Similarly, in terms of the fourth term, we have

$$\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (P_m\{\pi_{1,m}\psi_j(X_m)\pi_{2,mi}\psi_k(X_m, X_i)\})^2\right\|_{L^1(\mathbb{P})}$$

$$\le \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} P_m(\{\pi_{1,m}\psi_j(X_m)\})^2 P_m(\{\pi_{2,mi}\psi_k(X_m, X_i)\})^2\right\|_{L^1(\mathbb{P})}$$

$$\le \max_{j\in[p]} \max_{m\in[n]} P_m(\{\pi_{1,m}\psi_j(X_m)\})^2 \left\|\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_m(\{\pi_{2,mi}\psi_k(X_m, X_i)\})^2\right\|_{L^1(\mathbb{P})}$$

$$\le \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^2(P_m)}^2 \left\|\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} (P_m\{\pi_{2,mi}\psi_k(X_m, X_i)\}^2)^2\right\|_{L^1(\mathbb{P})}^{1/2}.$$

Summing up

$$\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}_{im,jk}^{(1,2)})^2(X_i, x_m)dP_m(x_m)\right\|_{L^1(\mathbb{P})}$$

$$\le \mathbb{E}\left[\max_{j\in[p]} \max_{i\in[n]} (\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_m(\pi_{2,im}\psi_k)^2(X_i)\right)^2\right]^{1/2}$$

$$+ \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^4(P_m)}^2 \, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} P_m\{(\pi_{2,mi}\psi_k)^4(X_m, X_i)\}\right]^{1/2}$$

$$+ 2 \max_{j\in[p]} \max_{m\in[n]} \|\pi_{1,m}\psi_j\|_{L^2(P_m)}^2 \, \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_m(\pi_{2,mi}\psi_k)^2(X_i)\right)^2\right]^{1/2}.$$

$\square$

Proof of Eq. (22). From Eq. (44), we have

$$\left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (\tilde{\chi}_{im,jk}^{(1,2)})^2(X_i, X_m)\right\|_{L^1(\mathbb{P})}$$

$$\le \left\|\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \{(\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i, X_m)\}\right\|_{L^1(\mathbb{P})}$$

$$+\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\})^2\right\|_{L^1(\mathbb{P})}.$$

In terms of the first term,

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\{(\pi_{1,i}\psi_j)^2(X_i)(\pi_{2,im}\psi_k)^2(X_i,X_m)\}\right\|_{L^1(\mathbb{P})}\\
&\leq \left\|\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^2(X_i)\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_k)^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}\\
&\leq \left\|\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^4(X_i)\right\|_{L^1(\mathbb{P})}^{1/2}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_k)^4(X_i,X_m)\right\|_{L^1(\mathbb{P})}^{1/2}\\
&= \mathbb{E}\left[\max_{j\in[p]}\max_{i\in[n]}(\pi_{1,i}\psi_j)^4(X_i)\right]^{1/2}\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_k)^4(X_i,X_m)\right]^{1/2}.
\end{aligned}$$

In terms of the second term

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\})^2\right\|_{L^1(\mathbb{P})}\\
&\leq \left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}P_i(\pi_{1,i}\psi_j)^2P_i(\pi_{2,im}\psi_k)^2(X_m)\}\right\|_{L^1(\mathbb{P})}\\
&\leq \max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|_{L^2(P_i)}^2\left\|\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}P_i(\pi_{2,im}\psi_k)^2(X_m)\right\|_{L^1(\mathbb{P})}\\
&\leq \max_{j\in[p]}\max_{i\in[n]}\|\pi_{1,i}\psi_j\|_{L^2(P_i)}^2\left\|\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}\{P_i(\pi_{2,im}\psi_k)^2(X_m)\}^2\right\|_{L^1(\mathbb{P})}^{1/2},
\end{aligned}$$

where the second inequality follows from Schwarz inequality. □

Proof of Eq. (23). Since $\varphi_{m,jk}^{(1,2)}(X_m) \coloneqq \sum_{i:m\neq i}2P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\}$, we have

$$\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{m\in[n]}\int(\varphi_{m,jk}^{(1,2)})^2(x_m)dP_m(x_m)\\
&\leq 4n^2\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int\left(P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,x_m)\}\right)^2dP_m(x_m)\\
&= 4n^2\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\pi_{1,i}\psi_j\star_i^1\pi_{2,im}\psi_k)^2(x_m)dP_m(x_m)\\
&= 4n^2\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\|\pi_{1,i}\psi_j\star_i^1\pi_{2,im}\psi_k\|_{L^2(P_m)}^2.
\end{aligned}$$

□

Proof of Eq. (24).

$$\left\|\max_{(j,k)\in[p]^2}\max_{m\in[n]}(\varphi_{m,jk}^{(1,2)})^2(X_m)\right\|_{L^1(\mathbb{P})}$$

$$\leq \left\| \max_{(j,k)\in[p]^2} \max_{m\in[n]} \left(\sum_{i:m\neq i} 2P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\} \right)^2 \right\|_{L^1(\mathbb{P})}$$
$$\leq 4n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} (P_i\{\pi_{1,i}\psi_j(X_i)\pi_{2,im}\psi_k(X_i,X_m)\})^2 \right\|_{L^1(\mathbb{P})}$$
$$\leq 4n^2 \max_{j\in[p]} \max_{i\in[n]} \|\pi_{1,i}\psi_j\|^2_{L^2(P_i)} \left\| \max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \{P_i(\pi_{2,im}\psi_k)^2(X_m)\}^2 \right\|^{1/2}_{L^1(\mathbb{P})}.$$

where the third inequality follows from Schwarz inequality. □

Proof of Eq. (25). Observe that

$$\begin{aligned}
&(\bar{\chi}^{(2,2)}_{imm,jk})^2(X_i,X_m) \\
&= \frac{1}{4}\big(\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m) - P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\}\big)^2, \\
&(\bar{\chi}^{(2,2)}_{mii,jk})^2(X_m,X_i) \\
&= \frac{1}{4}\big(\pi_{2,mi}\psi_j(X_m,X_i)\pi_{2,mi}\psi_k(X_m,X_i) - P_m\{\pi_{2,mi}\psi_j(X_m,X_i)\pi_{2,mi}\psi_k(X_m,X_i)\}\big)^2.
\end{aligned} \tag{45}$$

In conjunction with Eq.(5) in Imai and Koike (2025), we have

$$\begin{aligned}
&\max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\tilde{\chi}^{(2,2),\mathtt{diag}}_{im,jk})^2(x_i,x_m)dP_i(x_i)dP_m(x_m) \\
&\leq \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int (\bar{\chi}^{(2,2)}_{imm,jk})^2(x_i,x_m)dP_i(x_i)dP_m(x_m) \\
&\leq \frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \Big(\pi_{2,im}\psi_j(x_i,x_m)\pi_{2,im}\psi_k(x_i,x_m)\Big)^2 dP_i(x_i)dP_m(x_m) \\
&\quad + \frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \Big(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,im}\psi_k(X_i,x_m)\}\Big)^2 dP_m(x_m).
\end{aligned}$$

In terms of the first term, Schwarz inequality gives

$$\begin{aligned}
&\frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \Big(\pi_{2,im}\psi_j(x_i,x_m)\pi_{2,im}\psi_k(x_i,x_m)\Big)^2 dP_i(x_i)dP_m(x_m) \\
&\leq \frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \|\pi_{2,im}\psi_j\|^2_{L^4(P_i\otimes P_m)} \|\pi_{2,im}\psi_k\|^2_{L^4(P_i\otimes P_m)} \\
&= \frac{1}{4} \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \|\pi_{2,im}\psi_j\|^4_{L^4(P_i\otimes P_m)}.
\end{aligned}$$

In terms of the second term, Schwarz and Jensen inequality give

$$\begin{aligned}
&\frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \Big(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,im}\psi_k(X_i,x_m)\}\Big)^2 dP_m(x_m) \\
&\leq \frac{1}{4} \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(\int \Big(P_i(\pi_{2,im}\psi_j)^2(X_i,x_m)\Big)^2 dP_m(x_m)\right)^{1/2} \left(\int \Big(P_i(\pi_{2,im}\psi_k)^2(X_i,x_m)\Big)^2 dP_m(x_m)\right)^{1/2}
\end{aligned}$$

$$= \frac{1}{4} \max_{j \in [p]} \max_{(i,m) \in I_{n,2}} \|P_i(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_m)}.$$

Summing up,

$$\begin{aligned}
&\max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int (\tilde{\chi}^{(2,2),\mathtt{diag}}_{im,jk})^2 (x_i, x_m) dP_i(x_i) dP_m(x_m) \\
&\le \frac{1}{4} \max_{j \in [p]} \max_{(i,m) \in I_{n,2}} \|\pi_{2,im}\psi_j\|^4_{L^4(P_i \otimes P_m)} + \frac{1}{4} \max_{j \in [p]} \max_{(i,m) \in I_{n,2}} \|P_i(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_m)}.
\end{aligned}$$

$\square$

Proof of Eq. (26). From Eq. (45), we have

$$\begin{aligned}
&\left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int (\tilde{\chi}^{(2,2),\mathtt{diag}}_{im,jk})^2 (X_i, x_m) dP_m(x_m) \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{8} \left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int \Big(\pi_{2,im}\psi_j(X_i, x_m) \pi_{2,im}\psi_k(X_i, x_m) \right. \\
&\qquad\qquad \left. - P_i\{\pi_{2,im}\psi_j(X_i, x_m)\pi_{2,im}\psi_k(X_i, x_m)\} \Big)^2 dP_m(x_m) \right\|_{L^1(\mathbb{P})} \\
&\quad + \frac{1}{8} \left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int \Big(\pi_{2,mi}\psi_j(x_m, X_i) \pi_{2,mi}\psi_k(x_m, X_i) \right. \\
&\qquad\qquad \left. - P_m\{\pi_{2,mi}\psi_j(X_m, X_i)\pi_{2,mi}\psi_k(X_m, X_i)\} \Big)^2 dP_m(x_m) \right\|_{L^1(\mathbb{P})}.
\end{aligned}$$

In terms of the first term, from the Jensen inequality

$$\begin{aligned}
&\frac{1}{8} \left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int \Big(\pi_{2,im}\psi_j(X_i, x_m) \pi_{2,im}\psi_k(X_i, x_m) \right. \\
&\qquad\qquad \left. - P_i\{\pi_{2,im}\psi_j(X_i, x_m)\pi_{2,im}\psi_k(X_i, x_m)\} \Big)^2 dP_m(x_m)\} \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{8} \left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int (\pi_{2,im}\psi_j)^2 (X_i, x_m) (\pi_{2,im}\psi_k)^2 (X_i, x_m) dP_m(x_m) \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{8} \mathbb{E} \left[\max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \left(P_m (\pi_{2,im}\psi_j)^4 (X_i) \right)^{1/2} \left(P_m (\pi_{2,im}\psi_k)^4 (X_i) \right)^{1/2} \right] \\
&\le \frac{1}{8} \mathbb{E} \left[\max_{j \in [p]} \max_{(i,m) \in I_{n,2}} P_m (\pi_{2,im}\psi_j)^4 (X_i) \right],
\end{aligned}$$

where the second inequality follows from Schwarz inequality. In terms of the second term, similar evaluation gives

$$\begin{aligned}
&\frac{1}{8} \left\| \max_{(j,k) \in [p]^2} \max_{(i,m) \in I_{n,2}} \int \Big(\pi_{2,mi}\psi_j(x_m, X_i) \pi_{2,mi}\psi_k(x_m, X_i) \right. \\
&\qquad\qquad \left. - P_m\{\pi_{2,mi}\psi_j(x_m, X_i)\pi_{2,mi}\psi_k(x_m, X_i)\} \Big)^2 dP_m(x_m) \right\|_{L^1(\mathbb{P})}
\end{aligned}$$

$$\leq \frac{1}{8}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_m(\pi_{2,mi}\psi_j)^4(X_i)\right].$$

Summing up

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\int(\tilde{\chi}^{(2,2)}_{im,jk})^2(X_i,x_m)dP_m(x_m)\right\|_{L^1(\mathbb{P})}\leq\frac{1}{4}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}} P_m(\pi_{2,im}\psi_j)^4(X_i)\right].$$

□

Proof of Eq. (27). From Eq. (45), we have

$$\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\tilde{\chi}^{(2,2)}_{im,jk})^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}\\
&\leq\frac{1}{4}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}\Big(\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\right.\\
&\qquad\qquad\left.-P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\}\Big)^2\right\|_{L^1(\mathbb{P})}\\
&\leq\frac{1}{4}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^2(X_i,X_m)(\pi_{2,im}\psi_k)^2(X_i,X_m)\right\|_{L^1(\mathbb{P})}\\
&\leq\frac{1}{4}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right],
\end{aligned}$$

where the second inequality follows from Jensen's inequality and the final inequality follows from the AM-GM inequality and Schwarz inequality. □

Proof of Eq. (28). Recall that

$$\bar{\chi}^{(2,2)}_{iml,jk}(X_i,X_m,X_l):=\frac{1}{2}\big(\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)-P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\big). \tag{46}$$

Then, from Eq.(5) in Imai and Koike (2025) and c_r inequality, we have, we have

$$\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(x_i,x_m,x_l)dP_i(x_i)dP_m(x_m)dP_l(x_l)\\
&\leq\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{iml,jk})^2(x_i,x_m,x_l)dP_i(x_i)dP_m(x_m)dP_l(x_l)\\
&\leq\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\Big(\pi_{2,im}\psi_j(x_i,x_m)\pi_{2,il}\psi_k(x_i,x_l)\Big)^2dP_i(x_i)dP_m(x_m)dP_l(x_l)\\
&\quad+\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\Big(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,il}\psi_k(X_i,x_l)\}\Big)^2dP_m(x_m)dP_l(x_l)\\
&\leq\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int P_m\{(\pi_{2,im}\psi_j)^2(x_i,X_m)\}P_l\{(\pi_{2,il}\psi_k)^2(x_i,X_l)\}dP_i(x_i)\\
&\quad+\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\Big(\int\{\pi_{2,im}\psi_j(x_i,x_m)\pi_{2,il}\psi_k(x_i,x_l)\}dP_i(x_i)\Big)^2dP_m(x_m)dP_l(x_l)
\end{aligned}$$

$$\le \frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|P_m(|\pi_{2,im}\psi_j|^2)\|_{L^2(P_i)}\|P_l(|\pi_{2,il}\psi_k|^2)\|_{L^2(P_l)}$$
$$+\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}$$
$$\le \frac{1}{2}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\|P_m(|\pi_{2,im}\psi_j|^2)\|^2_{L^2(P_i)}+\frac{1}{2}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}.$$

□

Proof of Eq. (29). From the c_r inequality, we have,

$$\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$\le\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$+\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mil,jk})^2(x_m,X_i,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$+\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mli,jk})^2(x_m,x_l,X_i)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}.$$

In terms of the first term, from Eq. (46), we can see that

$$\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$\le\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\pi_{2,im}\psi_j)^2(X_i,x_m)(\pi_{2,il}\psi_k)^2(X_i,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$+\frac{1}{6}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(\int\pi_{2,im}\psi_j(x_i,x_m)\pi_{2,il}\psi_k(x_i,x_l)dP_i(x_i)\right)^2dP_m(x_m)dP_l(x_l)$$
$$\le\frac{1}{6}\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left(P_m(\pi_{2,im}\psi_j)^2(X_i)\right)\left(P_l(\pi_{2,il}\psi_k)^2(X_i)\right)\right]$$
$$+\frac{1}{6}\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left\|P_i(|\pi_{2,im}\psi_j|^2)\right\|_{L^2(P_m)}\left\|P_i(|\pi_{2,il}\psi_k|^2)\right\|_{L^2(P_l)}$$
$$\le\frac{1}{6}\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_m(\pi_{2,im}\psi_j)^2(X_i)\right)^2\right]+\frac{1}{6}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left\|P_m(|\pi_{2,im}\psi_j|^2)\right\|^2_{L^2(P_i)},$$

where the second inequality follows from Schwarz inequality. In terms of the second term, from Eq. (46), we can see that

$$\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mil,jk})^2(x_m,X_i,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$\le\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\pi_{2,mi}\psi_j)^2(x_m,X_i)(\pi_{2,ml}\psi_k)^2(x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$
$$+\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(\int\pi_{2,mi}\psi_j(x_m,X_i)\pi_{2,ml}\psi_k(x_m,x_l)dP_m(x_m)\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}$$

$$
\begin{aligned}
&\le \frac{1}{6}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left(P_m(\pi_{2,mi}\psi_j)^4(X_i)\right)^{1/2}\left(P_m\{P_l(\pi_{2,ml}\psi_k)^2(X_m)\}^2\right)^{1/2}\right]\\
&\quad+\frac{1}{6}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}P_m(\pi_{2,mi}\psi_j)^2(X_i)P_lP_m(\pi_{2,ml}\psi_k)^2\right]\\
&\le \frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/2}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|_{L^2(P_m)},
\end{aligned}
$$

where the second inequality follows from Schwarz inequality. In terms of the third term, from Eq. (46), we can see that

$$
\begin{aligned}
&\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mli,jk})^2(x_m,x_l,X_i)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le \frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\pi_{2,ml}\psi_j)^2(x_m,x_l)(\pi_{2,mi}\psi_k)^2(x_m,X_i)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(\int\pi_{2,ml}\psi_j(x_m,x_l)\pi_{2,mi}\psi_k(x_m,X_i)dP_m(x_m)\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le \frac{1}{6}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left(P_m\{P_l(\pi_{2,ml}\psi_j)^2(X_m)\}^2\right)^{1/2}\left(P_m(\pi_{2,mi}\psi_k)^4(X_i)\right)^{1/2}\right]\\
&\quad+\frac{1}{6}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}P_lP_m(\pi_{2,ml}\psi_j)^2P_m(\pi_{2,mi}\psi_k)^2(X_i)\right]\\
&\le \frac{1}{3}\max_{j\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|_{L^2(P_m)}\,\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/2}.
\end{aligned}
$$

Summing up

$$
\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,x_m,x_l)dP_m(x_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le \frac{1}{6}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left(P_m(\pi_{2,im}\psi_j)^2(X_i)\right)^2\right]+\frac{1}{6}\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}\left\|P_m(|\pi_{2,im}\psi_j|^2)\right\|^2_{L^2(P_i)}\\
&\quad+\frac{2}{3}\,\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,mi}\psi_j)^4(X_i)\right]^{1/2}\max_{k\in[p]}\max_{(m,l)\in I_{n,2}}\|P_l(|\pi_{2,ml}\psi_k|^2)\|_{L^2(P_m)}.
\end{aligned}
$$

□

Proof of Eq. (30). From c_r inequality, we have

$$
\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le \frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{lim,jk})^2(x_l,X_i,X_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}
\end{aligned}
$$

$$+\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mli,jk})^2(X_m,x_l,X_i)dP_l(x_l)\right\|_{L^1(\mathbb{P})}.$$

In terms of the first term, from Eq. (46), we can see that

$$\begin{aligned}
&\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{iml,jk})^2(X_i,X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}(\pi_{2,im}\psi_j)^2(X_i,X_m)\int(\pi_{2,il}\psi_k)^2(X_i,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(\int\pi_{2,im}\psi_j(x_i,X_m)\pi_{2,il}\psi_k(x_i,x_l)dP_i(x_i)\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{6}\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}(\pi_{2,im}\psi_j)^4(X_i,X_m)\right]^{1/2}\mathbb{E}\left[\max_{k\in[p]}\max_{(i,l)\in I_{n,2}}\left(P_l(\pi_{2,il}\psi_k)^2(X_i)\right)^2\right]^{1/2}\\
&\quad+\frac{1}{6}\mathbb{E}\left[\max_{j\in[p]}\max_{(i,m)\in I_{n,2}}P_i(\pi_{2,im}\psi_j)^4(X_m)\right]^{1/2}\max_{k\in[p]}\max_{(i,l)\in I_{n,2}}\|\pi_{2,il}\psi_k\|^2_{L^4(P_i\otimes P_l)}.
\end{aligned}$$

In terms of the second term, from Eq. (46), we can see that

$$\begin{aligned}
&\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{lim,jk})^2(x_l,X_i,X_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\pi_{2,li}\psi_j)^2(x_l,X_i)(\pi_{2,lm}\psi_k)^2(x_l,X_m)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left(\int\pi_{2,li}\psi_j(x_l,X_i)\pi_{2,lm}\psi_k(x_l,X_m)dP_l(x_l)\right)^2\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{3}\mathbb{E}\left[\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\left(P_l(\pi_{2,li}\psi_j)^4(X_i)\right)^{1/2}\left(P_l(\pi_{2,lm}\psi_k)^4(X_m)\right)^{1/2}\right]\\
&\le\frac{1}{3}\mathbb{E}\left[\max_{j\in[p]}\max_{(i,l)\in I_{n,2}}P_l(\pi_{2,li}\psi_j)^4(X_i)\right].
\end{aligned}$$

In terms of the third term, from Eq. (46), we can see that

$$\begin{aligned}
&\frac{1}{3}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int(\bar{\chi}^{(2,2)}_{mli,jk})^2(X_m,x_l,X_i)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}(\pi_{2,mi}\psi_k)^2(X_m,X_i)\int(\pi_{2,ml}\psi_j)^2(X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{6}\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(\int\pi_{2,ml}\psi_j(x_m,x_l)\pi_{2,mi}\psi_k(x_m,X_i)dP_m(x_m)\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{6}\mathbb{E}\left[\max_{j\in[p]}\max_{(m,l)\in I_{n,2}}\left(P_l(\pi_{2,ml}\psi_j)^2(X_m)\right)^2\right]^{1/2}\mathbb{E}\left[\max_{k\in[p]}\max_{(m,i)\in I_{n,2}}(\pi_{2,mi}\psi_k)^4(X_m,X_i)\right]^{1/2}\\
&\quad+\frac{1}{6}\max_{j\in[p]}\max_{(m,l)\in I_{n,2}}\|\pi_{2,ml}\psi_j\|^2_{L^4(P_m\otimes P_l)}\mathbb{E}\left[\max_{k\in[p]}\max_{(i,m)\in I_{n,2}}P_m(\pi_{2,mi}\psi_k)^4(X_i)\right]^{1/2}.
\end{aligned}$$

Summing up,

$$
\begin{aligned}
&\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \int (\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i, X_m, x_l) dP_l(x_l) \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} (\pi_{2,im}\psi_j)^4(X_i, X_m)\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \left(P_l(\pi_{2,il}\psi_k)^2(X_i)\right)^2\right]^{1/2} \\
&\quad + \frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right]^{1/2} \max_{k\in[p]} \max_{(m,l)\in I_{n,2}} \|\pi_{2,il}\psi_k\|^2_{L^4(P_i\otimes P_l)} \\
&\quad + \frac{1}{3}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_l(\pi_{2,li}\psi_j)^4(X_i)\right].
\end{aligned}
$$

□

Proof of Eq. (31). From Eq.(5) in Imai and Koike (2025), we have

$$
\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} (\tilde{\chi}^{(2,2)}_{iml,jk})^2(X_i, X_m, X_l) \right\|_{L^1(\mathbb{P})} \le \left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} (\bar{\chi}^{(2,2)}_{iml,jk})^2(X_i, X_m, X_l) \right\|_{L^1(\mathbb{P})}.
$$

From Eq. (46) and Schwarz inequality, we can see that

$$
\begin{aligned}
&\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} (\chi^{(2,2)}_{iml,jk})^2(X_i, X_m, X_l) \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{2}\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} (\pi_{2,im}\psi_j)^2(X_i, X_m)(\pi_{2,il}\psi_k)^2(X_i, X_l) \right\|_{L^1(\mathbb{P})} \\
&\quad + \frac{1}{2}\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \left(\int \pi_{2,im}\psi_j(x_i, X_m)\pi_{2,il}\psi_k(x_i, X_l) dP_i(x_i)\right)^2 \right\|_{L^1(\mathbb{P})} \\
&\le \frac{1}{2}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} (\pi_{2,im}\psi_j)^4(X_i, X_m)\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,l)\in I_{n,2}} (\pi_{2,il}\psi_k)^4(X_i, X_l)\right]^{1/2} \\
&\quad + \frac{1}{2}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} P_i(\pi_{2,im}\psi_j)^2(X_m) P_i(\pi_{2,il}\psi_k)^2(X_l)\right].
\end{aligned}
$$

In terms of the second term

$$
\begin{aligned}
&\frac{1}{2}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} P_i(\pi_{2,im}\psi_j)^2(X_m) P_i(\pi_{2,il}\psi_k)^2(X_l)\right] \\
&\le \frac{1}{4}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \left\{\left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2 + \left(P_i(\pi_{2,il}\psi_k)^2(X_l)\right)^2\right\}\right] \\
&\le \frac{1}{4}\,\mathbb{E}\left[\max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \left\{P_i(\pi_{2,im}\psi_j)^4(X_m) + P_i(\pi_{2,il}\psi_k)^4(X_l)\right\}\right] \\
&\le \frac{1}{2}\,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right],
\end{aligned}
$$

where the first inequality follows from AM-GM inequality and the second inequality follows from Jensen's inequality. Therefore

$$
\left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} (\chi^{(2,2)}_{iml,jk})^2(X_i, X_m, X_l) \right\|_{L^1(\mathbb{P})} \le \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^4(X_m)\right].
$$

□

Proof of Eq. (32). From Eq.(5) in Imai and Koike (2025), we have

$$
\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(x_m,x_l)dP_m(x_m)dP_l(x_l)\\
&\le \max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\varphi^{(2,2)}_{ml,jk})^2(x_m,x_l)dP_m(x_m)dP_l(x_l).
\end{aligned}
$$

Since we consider the case where $i\neq m\neq l$, it holds that $P_iP_mP_l\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}=0$. Therefore, we can see that

$$
\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\varphi^{(2,2)}_{ml,jk})^2(x_m,x_l)dP_m(x_m)dP_l(x_l)\\
&\le n^2\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\Big(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,il}\psi_k(X_i,x_l)\}\Big)^2dP_m(x_m)dP_l(x_l)\\
&= n^2\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,im}\psi_j\star_i^1\pi_{2,il}\psi_k\|^2_{L^2(P_m\otimes P_l)}.
\end{aligned}
$$

□

Proof of Eq. (33). From c_r inequality, we have

$$
\begin{aligned}
&\left\|\max_{(j,k)\in[p]^2}\max_{(m,l)\in I_{n,2}}\int(\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m,x_l)dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\le\frac{1}{2}n^2\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,x_l)\}\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}\\
&\quad+\frac{1}{2}n^2\left\|\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(P_i\{\pi_{2,il}\psi_j(X_i,x_l)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2dP_l(x_l)\right\|_{L^1(\mathbb{P})}.
\end{aligned}
$$

Likewise the proof of Lemma 17(c) in Imai and Koike (2025), from Fubini's theorem, we have

$$
\begin{aligned}
&\int\left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,x_l)\}\right)^2dP_l(x_l)\\
&=\int\left(\int\pi_{2,im}\psi_j(x_i,X_m)\pi_{2,il}\psi_k(x_i,x_l)dP_i(x_i)\right)^2dP_l(x_l)\\
&=\int\int\int\pi_{2,im}\psi_j(x_i,X_m)\pi_{2,il}\psi_k(x_i,x_l)\pi_{2,im}\psi_j(x_i',X_m)\pi_{2,il}\psi_k(x_i',x_l)dP_i(x_i)dP_i(x_i')dP_l(x_l)\\
&=\int\int(\pi_{2,il}\psi_k\star_l^1\pi_{2,il}\psi_k)(x_i,x_i')\pi_{2,im}\psi_j(x_i,X_m)\pi_{2,im}\psi_j(x_i',X_m)dP_i(x_i)dP_i(x_i').
\end{aligned}
$$

Then, Schwarz inequality gives

$$
\begin{aligned}
&\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\int\left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,x_l)\}\right)^2dP_l(x_l)\\
&\le\max_{(j,k)\in[p]^2}\max_{(i,m,l)\in I_{n,3}}\|\pi_{2,il}\psi_k\star_l^1\pi_{2,il}\psi_k\|_{L^2(P_i\otimes P_i)}\times P_i(\pi_{2,im}\psi_j)^2(X_m)
\end{aligned}
$$

$$= \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \|\pi_{2,il}\psi_k \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_l\otimes P_l)} \times P_i(\pi_{2,im}\psi_j)^2(X_m).$$

Therefore

$$\frac{1}{2}n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \int \left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,x_l)\}\right)^2 dP_l(x_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq \frac{1}{2}n^2 \max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \|\pi_{2,il}\psi_k \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_l\otimes P_l)} \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_j)^2(X_m)\right)^2\right]^{1/2}.$$

Similarly,

$$\frac{1}{2}n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \int \left(P_i\{\pi_{2,il}\psi_j(X_i,x_l)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2 dP_l(x_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq \frac{1}{2}n^2 \max_{j\in[p]} \max_{(i,l)\in I_{n,2}} \|\pi_{2,il}\psi_j \star_i^1 \pi_{2,il}\psi_j\|_{L^2(P_l\otimes P_l)} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_k)^2(X_m)\right)^2\right]^{1/2}.$$

Summing up

$$\left\| \max_{(j,k)\in[p]^2} \max_{(m,l)\in I_{n,2}} \int (\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m,x_l)dP_l(x_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq n^2 \max_{(j,k)\in[p]^2} \max_{\substack{(i,m,l)\in[n]^3 \\ m,l\neq i}} \|\pi_{2,im}\psi_j \star_i^1 \pi_{2,il}\psi_k\|_{L^2(P_m\otimes P_l)} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left(P_i(\pi_{2,im}\psi_k)^2(X_m)\right)^2\right]^{1/2}.$$

□

Proof of Eq. (34). From c_r inequality, we have

$$\left\| \max_{(j,k)\in[p]^2} \max_{(m,l)\in I_{n,2}} (\tilde{\varphi}^{(2,2)}_{ml,jk})^2(X_m,X_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq \frac{1}{2}n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,il}\psi_k(X_i,X_l)\}\right)^2 \right\|_{L^1(\mathbb{P})}$$

$$+ \frac{1}{2}n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m,l)\in I_{n,3}} \left(P_i\{\pi_{2,il}\psi_j(X_i,X_l)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2 \right\|_{L^1(\mathbb{P})}$$

$$\leq n^2 \left\| \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^2(X_m) \max_{k\in[p]} \max_{(i,l)\in I_{n,2}} P_i(\pi_{2,il}\psi_k)^2(X_l) \right\|_{L^1(\mathbb{P})}$$

$$\leq n^2 \,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\{P_i(\pi_{2,im}\psi_j)^2(X_m)\right\}^2\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,l)\in I_{n,2}} \left\{P_i(\pi_{2,il}\psi_k)^2(X_l)\right\}^2\right]^{1/2}$$

$$\leq n^2 \,\mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\{P_i(\pi_{2,im}\psi_j)^2(X_m)\right\}^2\right],$$

where the second and third inequality follows from Schwarz inequality. □

Proof of Eq. (35). From c_r inequality, we have

$$\max_{(j,k)\in[p]^2} \max_{m\in[n]} \int (\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(x_m)dP_m(x_m)$$

$$
\begin{aligned}
&\le 2n^2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \left(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,im}\psi_k(X_i,x_m)\}\right)^2 dP_m(x_m) \\
&\quad + 2n^2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(P_iP_m\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2 \\
&\le 4n^2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int \left(P_i\{\pi_{2,im}\psi_j(X_i,x_m)\pi_{2,im}\psi_k(X_i,x_m)\}\right)^2 dP_m(x_m) \\
&\le 4n^2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \int P_i(\pi_{2,im}\psi_j)^2(x_m)P_i(\pi_{2,im}\psi_k)^2(x_m)dP_m(x_m) \\
&\le 4n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\|P_i(|\pi_{2,im}\psi_j|^2)\right\|_{L^2(P_m)} \max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left\|P_i(|\pi_{2,im}\psi_k|^2)\right\|_{L^2(P_m)} \\
&\le 4n^2 \max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\|P_i(|\pi_{2,im}\psi_j|^2)\right\|^2_{L^2(P_m)},
\end{aligned}
$$

where the second inequality follows from Jensen's inequality, the third and the final inequality follows from Schwarz inequality. □

Proof of Eq. (36). From c_r inequality, we have

$$
\begin{aligned}
&\left\| \max_{(j,k)\in[p]^2} \max_{m\in[n]} (\varphi^{(2,2),\mathtt{diag}}_{m,jk})^2(X_m) \right\|_{L^1(\mathbb{P})} \\
&\le 2n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(P_i\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2 \right\|_{L^1(\mathbb{P})} \\
&\quad + 2n^2 \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} \left(P_iP_m\{\pi_{2,im}\psi_j(X_i,X_m)\pi_{2,im}\psi_k(X_i,X_m)\}\right)^2 \\
&\le 4n^2 \left\| \max_{(j,k)\in[p]^2} \max_{(i,m)\in I_{n,2}} P_i(\pi_{2,im}\psi_j)^2(X_i,X_m)P_i(\pi_{2,im}\psi_k)^2(X_i,X_m) \right\|_{L^1(\mathbb{P})} \\
&\le 4n^2\, \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\{P_i(\pi_{2,im}\psi_j)^2(X_m)\right\}^2\right]^{1/2} \mathbb{E}\left[\max_{k\in[p]} \max_{(i,m)\in I_{n,2}} \left\{P_i(\pi_{2,im}\psi_k)^2(X_m)\right\}^2\right]^{1/2} \\
&\le 4n^2\, \mathbb{E}\left[\max_{j\in[p]} \max_{(i,m)\in I_{n,2}} \left\{P_i(\pi_{2,im}\psi_j)^2(X_m)\right\}^2\right],
\end{aligned}
$$

where the second inequality follows from Jensen's inequality and Schwarz inequality and the final inequality follows from Schwarz inequality. □

References

Adao, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010. 3

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11(1):727–753. 3

Angrist, J. D., Imbens, G. W., and Krueger, A. B. (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics*, 14(1):57–67. 11

Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The quarterly journal of economics*, 106(4):979–1014. 3

Armstrong, T. B. and Kolesár, M. (2018). A simple adjustment for bandwidth snooping. *The Review of Economic Studies*, 85(2):732–765. 2

Belloni, A., Chernozhukov, V., Chetverikov, D., Hansen, C., and Kato, K. (2018). High-dimensional econometrics and regularized GMM. *arXiv preprint arXiv:1806.01888*. 2, 4, 10

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890. 3

Bondy, M., Roth, S., and Sager, L. (2020). Crime is in the air: The contemporaneous relationship between air pollution and crime. *Journal of the Association of Environmental and Resource Economists*, 7(3):555–585. 3

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of economic studies*, 89(1):181–213. 3

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204. 3

Cattaneo, M. D., Crump, R. K., and Jansson, M. (2014). Small bandwidth asymptotics for density-weighted average derivatives. *Econometric Theory*, 30(1):176–200. 4

Cattaneo, M. D., Farrell, M. H., Jansson, M., and Masini, R. P. (2024). Higher-order refinements of small bandwidth asymptotics for density-weighted average derivative estimators. *Journal of Econometrics*, In press:105855. 3, 4

Cattaneo, M. D., Jansson, M., and Newey, W. K. (2018a). Alternative asymptotics and the partially linear model with many regressors. *Econometric Theory*, 34(2):277–301. 3, 4, 10, 11, 12, 13

Cattaneo, M. D., Jansson, M., and Newey, W. K. (2018b). Inference in linear regression models with many covariates and heteroscedasticity. *Journal of the American Statistical Association*, 113(523):1350–1361. 3, 10

Chao, J. C., Swanson, N. R., Hausman, J. A., Newey, W. K., and Woutersen, T. (2012). Asymptotic distribution of JIVE in a heteroskedastic IV regression with many instruments. *Econometric Theory*, 28(1):42–86. 3, 10, 11, 13

Chen, X. (2018). Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications. *Annals of Statistics*, 46(2):642–678. 2

Chen, X. and Kato, K. (2019). Randomized incomplete U-statistics in high dimensions. *Annals of Statistics*, 47(6):3127–3156. 2

Chen, X. and Kato, K. (2020). Jackknife multiplier bootstrap: finite sample approximations to the U-process supremum with applications. *Probability Theory and Related Fields*, 176:1097–1163. 2

Cheng, G., Liu, Z., and Peng, L. (2022). Gaussian approximations for high-dimensional non-degenerate U-statistics via exchangeable pairs. *Statistics & Probability Letters*, 182:109295. 2

Chernozhukov, V., Chetverikov, D., and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Annals of Statistics*, 41(6):2786–2819. 2

Chernozhukov, V., Chetverikov, D., and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probability Theory and Related Fields*, 162:47–70. 23, 34

Chernozhukov, V., Chetverikov, D., and Kato, K. (2019). Inference on causal and structural parameters using many moment inequalities. *The Review of Economic Studies*, 86(5):1867–1900. 2

Chernozhukov, V., Chetverikov, D., Kato, K., and Koike, Y. (2022). Improved central limit theorem and bootstrap approximation in high dimensions. *Annals of Statistics*, 50(5):2562–2586. 2

Chernozhukov, V., Chetverikov, D., Kato, K., and Koike, Y. (2023). High-dimensional data bootstrap. *Annual Review of Statistics and Its Application*, 10(1):427–449. 2

Chetverikov, D., Wilhelm, D., and Kim, D. (2021). An adaptive test of stochastic monotonicity. *Econometric Theory*, 37(3):495–536. 3

Chiang, H. D., Kato, K., and Sasaki, Y. (2023). Inference for high-dimensional exchangeable arrays. *Journal of the American Statistical Association*, 118(543):1595–1605. 2

Chiang, H. D., Matsushita, Y., and Otsu, T. (2025). Regression adjustment, cross-fitting, and randomized experimens with many controls. *arXiv preprint arXiv:2302.00469*. 3

Dahl, G. B., Kostøl, A. R., and Mogstad, M. (2014). Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752. 3

Davies, N. M., von Hinke Kessler Scholder, S., Farbmacher, H., Burgess, S., Windmeijer, F., and Smith, G. D. (2015). The many weak instruments problem and mendelian randomization. *Statistics in medicine*, 34(3):454–468. 3

de la Peña, V. and Giné, E. (1999). *Decoupling: From Dependence to Independence*. Springer Science & Business Media. 30, 31

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12):4178–4219. 3

Dobbie, W., Goldin, J., and Yang, C. S. (2018). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2):201–240. 3

Döbler, C. and Peccati, G. (2017). Quantitative de Jong theorems in any dimension. *Electronic Journal of Probability*, 22(2):1–35. 10, 17

Döbler, C. and Peccati, G. (2019). Quantitative CLTs for symmetric U-statistics using contractions. *Electronic Journal of Probability*, 24(5):1–43. 4

Frandsen, B., Lefgren, L., and Leslie, E. (2023). Judging judge fixed effects. *American Economic Review*, 113(1):253–277. 3

Gabaix, X. and Koijen, R. S. (2024). Granular instrumental variables. *Journal of Political Economy*, 132(7):2274–2303. 3

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624. 3

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. (2012). A kernel two-sample test. *The journal of machine learning research*, 13(1):723–773. 3, 14

Horowitz, J. L. and Spokoiny, V. G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, 69(3):599–631. 2

Imai, S. and Koike, Y. (2025). Gaussian approximation for high-dimensional U-statistics with size-dependent kernels. *arXiv preprint arXiv:2504.10866*. 2, 3, 4, 7, 13, 14, 15, 16, 17, 25, 26, 27, 30, 31, 33, 34, 35, 36, 45, 49, 51, 55, 56

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press. 3

Jochmans, K. (2022). Heteroscedasticity-robust inference in linear regression models with many covariates. *Journal of the American Statistical Association*, 117(538):887–896. 3, 10

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review*, 96(3):863–876. 3

Koike, Y. (2023). High-dimensional central limit theorems for homogeneous sums. *Journal of Theoretical Probability*, 36(1):1–45. 2

Kuchibhotla, A. K., Kolassa, J. E., and Kuffner, T. A. (2022). Post-selection inference. *Annual Review of Statistics and Its Application*, 9:505–527. 2

Lee, D. S., McCrary, J., Moreira, M. J., and Porter, J. (2022). Valid t-ratio inference for iv. *American Economic Review*, 112(10):3260–3290. 3

Lei, L. and Ding, P. (2021). Regression adjustment in completely randomized experiments with a diverging number of covariates. *Biometrika*, 108(4):815–828. 3

Li, T. and Yuan, M. (2024). On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. *Journal of Machine Learning Research*, 25(334):1–62. 13, 14

Mikusheva, A. and Sun, L. (2022). Inference with many weak instruments. *The Review of Economic Studies*, 89(5):2663–2686. 3

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342. 3

Powell, J. L., Stock, J. H., and Stoker, T. M. (1989). Semiparametric estimation of index coefficients. *Econometrica*, 57(6):1403–1430. 4

Song, Y., Chen, X., and Kato, K. (2019). Approximating high-dimensional infinite-order U-statistics: Statistical and computational guarantees. *Electronic Journal of Statistics*, 13(2):4794. 2

Song, Y., Chen, X., and Kato, K. (2023). Stratified incomplete local simplex tests for curvature of nonparametric multiple regression. *Bernoulli*, 29(1):323–349. 2

Sriperumbudur, B. K., Gretton, A., Fukumizu, K., Schölkopf, B., and Lanckriet, G. R. (2010). Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11:1517–1561. 14