Online Appendix to "On the Asymptotic Properties of Debiased Machine Learning Estimators"

Amilcar Velez
Department of Economics
Northwestern University
amilcare@u.northwestern.edu

This version: November 3, 2024.
Newest version here.

E Proofs of Auxiliary Results

Notation: Recall $\hat{\eta}_i = \hat{\eta}_k(X_i)$ for $i \in \mathcal{I}_k$ and $n_k = n/K$ is the number of observations on the fold $\mathcal{I}_k$. Denote $\psi_i^z = \psi^z(W_i, \eta_i)$ and $\hat{\psi}_i^z = \psi^z(W_i, \hat{\eta}_i)$ for $z = a, b$; $m_i = m(W_i, \theta_0, \eta_i)$, and $\hat{m}_i = m(W_i, \theta_0, \hat{\eta}_i)$; $\partial_\eta m_i = \partial_\eta m(W_i, \theta_0, \eta_i)$ and $\partial_\eta \hat{m}_i = \partial_\eta m(W_i, \theta_0, \hat{\eta}_i)$; $\partial_\eta^2 m_i = \partial_\eta^2 m(W_i, \theta_0, \eta_i)$ and $\partial_\eta^2 \hat{m}_i = \partial_\eta^2 m(W_i, \theta_0, \hat{\eta}_i)$. Here, $|| \cdot ||$ is the euclidean norm (ℓ_2 norm), $J_0 = E[\psi_i^a]$, CLT is for Central Limit Theorem, LLN is for Law of Large Numbers, LIE is for Law of Iterated Expectations, C-S is for Cauchy-Schwartz inequality, RHS is for right-hand side.

E.1 Proof of Theorem C.1

Proof. Using the notation of this section, the definitions of the DML1 estimator in (2.6) and the moment function m in (2.2), it follows

$$n^{1/2}\left(\hat{\theta}_{n,1} - \theta_0\right) = K^{-1/2} \sum_{k=1}^{K} \frac{n_k^{-1/2} \sum_{i \in \mathcal{I}_K} \hat{m}_i}{n_k^{-1} \sum_{i \in \mathcal{I}_K} \hat{\psi}_i^a} ,$$

and similarly for the oracle version defined in (2.8),

$$n^{1/2}\left(\hat{\theta}^*_{n,1}-\theta_0\right)=K^{-1/2}\sum_{k=1}^{K}\frac{n_k^{-1/2}\sum_{i\in\mathcal{I}_K} m_i}{n_k^{-1}\sum_{i\in\mathcal{I}_K}\psi_i^a}\ .$$

Using the previous two expressions,

$$n^{1/2}\left(\hat{\theta}_{n,1}-\hat{\theta}^*_{n,1}\right)=I_1+I_2$$

where

$$I_1=K^{-1/2}\sum_{k=1}^{K}\frac{n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{m}_i-m_i)}{n_k^{-1}\sum_{i\in\mathcal{I}_K}\hat{\psi}_i^a}I_2\quad=K^{-1/2}\sum_{k=1}^{K}\frac{\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_K} m_i\right)\left(n_k^{-1}\sum_{i\in\mathcal{I}_k}\psi_i^a-\hat{\psi}_i^a\right)}{\left(n_k^{-1}\sum_{i\in\mathcal{I}_K}\hat{\psi}_i^a\right)\left(n_k^{-1}\sum_{i\in\mathcal{I}_k}\psi_i^a\right)}$$

In what follows, I will show that both I_1 and I_2 are $o_p(1)$, which is sufficient to complete the proof of the theorem.

Claim 1: $I_1=o_p(1)$. I first rewrite I_1 using the identity $a(1+b)^{-1}=a-ab(1+b)^{-1}$ with $a=\hat{I}_{1,k}$ and $b=I_{1,k}$, where

$$\hat{I}_{1,k}=n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{m}_i-m_i)/J_0\ ,$$
$$I_{1,k}=n_k^{-1}\sum_{i\in\mathcal{I}_K}(\hat{\psi}_i^a-J_0)/J_0\ .$$

This implies

$$I_1=K^{-1/2}\sum_{k=1}^{K}\hat{I}_{1,k}-\hat{I}_{1,k}I_{1,k}\left(1+I_{1,k}\right)^{-1}\tag{E.1}$$

To show the claim, consider the following derivations

$$\begin{aligned}|I_1|&\overset{(1)}{\leq}\left|K^{-1/2}\sum_{k=1}^{K}\hat{I}_{1,k}\right|+\left(K^{-1/2}\sum_{k=1}^{K}|\hat{I}_{1,k}||I_{1,k}|\right)\times\max_{k=1,\ldots,K}\left|n_k^{-1}\sum_{i\in\mathcal{I}_K}\hat{\psi}_i^a/J_0\right|^{-1}\\&\overset{(2)}{=}\left|n^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{m}_i-m_i)/J_0\right|+\left(K^{-1/2}\sum_{k=1}^{K}|\hat{I}_{1,k}||I_{1,k}|\right)\times O_p(1)\ ,\\&\overset{(3)}{=}o_p(1)+o_p(1)\times O_p(1)\ ,\end{aligned}$$

where (1) holds by triangular inequality used on (E.1) and definition of $I_{1,k}$, (2) holds by definition of $\hat{I}_{1,k}$ and part 2 of Lemma E.1, and (3) hold by part 3 of Lemma C.2 and (E.2) presented below,

$$K^{-1/2}\sum_{k=1}^{K}|\hat{I}_{1,k}||I_{1,k}| = o_p(1) \ . \tag{E.2}$$

I use Taylor expansion and the mean value theorem to write $\hat{I}_{1,k} = \hat{I}_{1,1,k} + \hat{I}_{1,2,k}$ and $I_{1,k} = n_k^{-1/2}(I_{1,1,k} + I_{1,2,k} + I_{1,3,k})$, where

$$\begin{aligned}
\hat{I}_{1,1,k} &= n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top \partial_\eta m_i / J_0 \ , \\
\hat{I}_{1,2,k} &= n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top (\partial_\eta^2 \tilde{m}_i/(2J_0))(\hat{\eta}_i - \eta_i)/J_0 \ , \\
I_{1,1,k} &= n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\psi_i^a - J_0)/J_0 \ , \\
I_{1,2,k} &= n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top \partial_\eta \psi_i^a / J_0 \ , \\
I_{1,3,k} &= n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{\psi}_i^a/(2J_0)(\hat{\eta}_i - \eta_i) \ ,
\end{aligned}$$

with $\partial_\eta^2 \tilde{m}_i = \partial_\eta^2 m(W_i, \theta_0, \tilde{\eta}_i)$ for some $\tilde{\eta}_i$, due to mean value theorem, and similar for $\partial_\eta^2 \tilde{\psi}_i^a$. In what follows I prove $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,j_1,k}||I_{1,j_2,k}| = o_p(1)$ for $j_1 = 1,2$ and $j_2 = 1,2,3$, which is sufficient to prove (E.2).

Claim 1.1: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}||I_{1,1,k}| = o_p(1)$. Consider the following

$$\begin{aligned}
K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}||I_{1,1,k}| &\overset{(1)}{\le} \max_{k=1,\dots,K}\left| n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top \partial_\eta m_i/J_0 \right| \times n^{-1/2}\sum_{k=1}^{K}|I_{1,1,k}| \\
&\overset{(2)}{=} o_p(1) \times (n^{-1/2}K) \times O_p(1) \\
&\overset{(3)}{=} o_p(1) \ ,
\end{aligned}$$

where (1) holds by definition of $\hat{I}_{1,1,k}$, (2) holds by Lemma C.3 and the derivation presented below, and (3) holds since $K = O(n^{1/2})$.

$$E\left[n^{-1/2}\sum_{k=1}^{K}|I_{1,1,k}|\right] \overset{(1)}{=} n^{-1/2}KE[|I_{1,1,k}|]$$
$$\overset{(2)}{\leq} n^{-1/2}KE\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0\right)^2\right]^{1/2}$$
$$\overset{(3)}{\leq} n^{-1/2}KO(1)$$

where (1) holds since $I_{1,1,k}$ are i.i.d. random variables, (2) holds by Jensen's inequality and definition of $I_{1,1,k}$, and (3) holds since $\{\psi_i^a - J_0 : i \in \mathcal{I}_k\}$ are zero mean i.i.d. random variables and by parts (a) and (c) of Assumption 3.1.

Claim 1.2: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}| \times |I_{1,2,k}| = o_p(1)$. It follows by

$$K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}||I_{1,2,k}| \leq \max_{k=1,\dots,K}|\hat{I}_{1,1,k}| \times \max_{k=1,\dots,K}|I_{1,2,k}| \times n^{-1/2}K$$
$$\overset{(1)}{=} o_p(1) \times o_p(1) \times O(1)$$

where (1) holds by Lemma C.3 and because $K = O(n^{1/2})$.

Claim 1.3: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}||I_{1,3,k}| = o_p(1)$. It follows by

$$K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,1,k}||I_{1,3,k}| \leq \max_{k=1,\dots,K}|\hat{I}_{1,1,k}| \times n^{-1/2}\sum_{k=1}^{K}\left|n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i - \eta_i)^\top \partial_\eta^2\tilde{\psi}_i^a/(2J_0)(\hat{\eta}_i - \eta_i)\right|$$
$$\overset{(1)}{\leq} \max_{k=1,\dots,K}|\hat{I}_{1,1,k}| \times n^{-1/2}\sum_{k=1}^{K}(C_2 p/(2|J_0|)) \times n_k^{-1/2}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_i - \eta_i||^2$$
$$= \max_{k=1,\dots,K}|\hat{I}_{1,1,k}| \times (C_2 p/(2|J_0|))(Kn^{-1/2})^{1/2} n^{1/4} n^{-1}\sum_{i=1}^{n}||\hat{\eta}_i - \eta_i||^2$$
$$\overset{(2)}{=} o_p(1) \times O(1) \times n^{1/4} \times O_p(n^{-2\min\{\varphi_1,\varphi_2\}})$$
$$\overset{(3)}{=} o_p(1)$$

where (1) holds by part (e) of Assumption 3.1 and Loeve's inequality (Davidson (1994, Theorem 9.28)), (2) holds by Lemmas C.3 and C.4 and because $K = O(n^{1/2})$, and

(3) holds since $\min\{\varphi_1, \varphi_2\} > 1/4$.

Claim 1.4: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,2,k}||I_{1,1,k}| = o_p(1)$. Consider the following derivations,

$$
\begin{aligned}
& n^{-1/2}\sum_{k=1}^{K}|\hat{I}_{1,2,k}||I_{1,1,k}| \\
&\overset{(1)}{\leq} n^{-1/2}\left(\sum_{k=1}^{K}\left|n_k^{-1/2}\sum_{i\in\mathcal{I}_K}(\hat{\eta}_i-\eta_i)^\top(\partial_\eta^2\tilde{m}_i/(2J_0))(\hat{\eta}_i-\eta_i)/J_0\right|^2\right)^{1/2}\left(\sum_{k=1}^{K}|I_{1,1,k}|^2\right)^{1/2} \\
&\overset{(2)}{\leq} (C_2p/2)n_k^{-1/2}\left(\sum_{k=1}^{K}|n_k^{-1/2}\sum_{i\in\mathcal{I}_K}||\hat{\eta}_i-\eta_i||^2|^2\right)^{1/2} K^{-1/2}\left(\sum_{k=1}^{K}|I_{1,1,k}|^2\right)^{1/2} \\
&\overset{(3)}{\leq} (C_2p/2)K^{1/2}\left(\sum_{k=1}^{K}n^{-1}\sum_{i\in\mathcal{I}_K}||\hat{\eta}_i-\eta_i||^4\right)^{1/2}\left(K^{-1}\sum_{k=1}^{K}|I_{1,1,k}|^2\right)^{1/2} \\
&\overset{(4)}{=} (Kn^{-1/2})^{1/2}n^{1/4}\times O_p(n^{-2\min\{\varphi_1,\varphi_2\}})\times\left(K^{-1}\sum_{k=1}^{K}|I_{1,1,k}|^2\right)^{1/2} \\
&\overset{(5)}{=} O(1)\times n^{1/4}\times O_p(n^{-2\min\{\varphi_1,\varphi_2\}})\times O_p(1) \\
&\overset{(6)}{=} o_p(1)
\end{aligned}
$$

where (1) holds by Cauchy-Schwartz and definition of $\hat{I}_{1,2,k}$, (2) holds by part (e) of Assumption 3.1 and Loeve's inequality (Davidson (1994, Theorem 9.28)), (3) holds by Jensen's inequality, (4) holds by Lemma C.4, (5) holds because $K = O(n^{1/2})$, $E[K^{-1}\sum_{k=1}^{K}|I_{1,1,k}|^2] = O(1)$ by definition of $I_{1,1,k}$ and due to parts (a) and (c) of Assumption 3.1, and (6) holds since $\min\{\varphi_1, \varphi_2\} > 1/4$.

Claim 1.5: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,2,k}||I_{1,2,k}| = o_p(1)$. The proof is similar to the proof of Claim 1.3; therefore, it is omitted.

Claim 1.6: $K^{-1/2}\sum_{k=1}^{K} n_k^{-1/2}|\hat{I}_{1,2,k}|\times|I_{1,3,k}| = o_p(1)$. Consider the derivations,

$$
\begin{aligned}
K^{-1/2}\sum_{k=1}^{K}n_k^{-1/2}|\hat{I}_{1,2,k}||I_{1,3,k}| &\overset{(1)}{\leq} n^{-1/2}\sum_{k=1}^{K}(C_2p/(2|J_0|))^2\times\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_i-\eta_i||^2\right)^2 \\
&\overset{(2)}{\leq} (C_2p/(2J_0))^2\times n^{1/2}n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_i-\eta_i||^4 \\
&\overset{(3)}{=} (C_2p/(2J_0))^2\times n^{1/2}\times O_p(n^{-4\min\{\varphi_1,\varphi_2\}})
\end{aligned}
$$

$$\stackrel{(4)}{=} o_p(1) \ ,$$

where (1) holds by using the definition of $\hat{I}_{1,2,k}$ and $I_{1,3,k}$, part (e) of Assumption 3.1, and Loeve's inequality (Davidson (1994, Theorem 9.28)), (2) holds by Jensen's inequality, (3) holds by Lemma C.4, and (4) holds since $\min\{\varphi_1, \varphi_2\} > 1/4$.

Claim 2: $I_2 = o_p(1)$. Consider the following representation of I_2,

$$\begin{aligned} I_2 &= K^{-1/2} \sum_{k=1}^{K} \frac{\left(n_k^{-1/2} \sum_{i \in \mathcal{I}_K} m_i\right)\left(n_k^{-1} \sum_{i \in \mathcal{I}_k} \psi_i^a - \hat{\psi}_i^a\right)}{\left(n_k^{-1} \sum_{i \in \mathcal{I}_K} \hat{\psi}_i^a\right)\left(n_k^{-1} \sum_{i \in \mathcal{I}_k} \psi_i^a\right)} \\ &= K^{-1/2} \sum_{k=1}^{K} \frac{n_k^{-1/2} I_{2,k} \hat{I}_{2,k}}{\left(n_k^{-1} \sum_{i \in \mathcal{I}_K} \hat{\psi}_i^a / J_0\right)\left(n_k^{-1} \sum_{i \in \mathcal{I}_k} \psi_i^a / J_0\right)} \ , \end{aligned}$$

where

$$\begin{aligned} I_{2,k} &= n_k^{-1/2} \sum_{i \in \mathcal{I}_K} m_i / J_0 \\ \hat{I}_{2,k} &= n_k^{-1/2} \sum_{i \in \mathcal{I}_k} (\psi_i^a - \hat{\psi}_i^a) / J_0 \ . \end{aligned}$$

To show the claim, consider the following derivation

$$\begin{aligned} |I_2| &\stackrel{(1)}{\leq} \max_{k=1,\ldots,K} \left| n_k^{-1} \sum_{i \in \mathcal{I}_K} \hat{\psi}_i^a / J_0 \right|^{-1} \times \max_{k=1,\ldots,K} \left| n_k^{-1} \sum_{i \in \mathcal{I}_K} \psi_i^a / J_0 \right|^{-1} \times K^{-1} \sum_{k=1}^{K} n_k^{-1/2} |I_{2,k}| |\hat{I}_{2,k}| \\ &\stackrel{(2)}{=} O_p(1) \times O_p(1) \times o_p(1) \ , \end{aligned}$$

where (1) holds by triangular inequality and definition of I_2, and (2) by Lemma E.1 and (E.3) presented below,

$$K^{-1} \sum_{k=1}^{K} n_k^{-1/2} |\hat{I}_{2,k}| |I_{2,k}| = o_p(1) \ . \tag{E.3}$$

As in the proof of claim 1, I use Taylor approximation and mean value theorem to write $\hat{I}_{2,k} = \hat{I}_{2,1,k} + \hat{I}_{2,2,k}$, where

$$\hat{I}_{2,1,k} = n_k^{-1/2} \sum_{i \in \mathcal{I}_K} (\hat{\eta}_i - \eta_i)^\top \partial_\eta \psi_i^a / J_0 \ ,$$
$$\hat{I}_{2,2,k} = n_k^{-1/2} \sum_{i \in \mathcal{I}_K} (\hat{\eta}_i - \eta_i)^\top (\partial_\eta^2 \tilde{\psi}_i^a / (2J_0))(\hat{\eta}_i - \eta_i)/J_0 \ .$$

Finally, in what follows I prove $K^{-1/2} \sum_{k=1} n_k^{-1/2} |\hat{I}_{2,j,k}||I_{2,k}| = o_p(1)$ for $j = 1, 2$, which is sufficient to prove (E.3).

Claim 2.1: $K^{-1/2} \sum_{k=1}^K n_k^{-1/2} |\hat{I}_{2,1,k}||I_{2,k}| = o_p(1)$. The proof is similar to the one in Claim 1.1; therefore, it is omitted.

Claim 2.2: $K^{-1/2} \sum_{k=1}^K n_k^{-1/2} |\hat{I}_{2,2,k}||I_{2,k}| = o_p(1)$. The proof is similar to the one in Claim 1.4; therefore, it is omitted. □

E.2 Proof of Theorem C.2

Proof. Notation: In the proof of this theorem, $x_{n,K} = o_p(1)$ denotes a sequence of random variables $x_{n,K}$ converging to zero uniformly on $K \to \infty$ as $n \to \infty$ (equivalently, $\lim_{n\to\infty} \sup_{K \le n} P(|x_{n,K}| > \epsilon) = 0$ for any given $\epsilon > 0$).

Using the definitions of the DML2 estimator in (2.7) and the moment function m in (2.2), it follows

$$n^{1/2} \left(\hat{\theta}_{n,2} - \theta_0 \right) = \frac{n^{-1/2} \sum_{i=1}^n \hat{m}_i}{n^{-1} \sum_{i=1}^n \hat{\psi}_i^a} \ ,$$

and similarly for the oracle version defined in (2.9),

$$n^{1/2} \left(\hat{\theta}_{n,2}^* - \theta_0 \right) = \frac{n^{-1/2} \sum_{i=1}^n m_i}{n^{-1} \sum_{i=1}^n \psi_i^a} \ .$$

Using the previous two expressions, it follows

$$n^{1/2} \left(\hat{\theta}_{n,2} - \hat{\theta}_{n,2}^* \right) = I_1 + I_2$$

where

$$I_1 = \frac{n^{-1/2}\sum_{i=1}^{n}(\hat{m}_i - m_i)/J_0}{n^{-1}\sum_{i=1}^{n}\hat{\psi}_i^a/J_0} \tag{E.4}$$

$$I_2 = \frac{\left(n^{-1/2}\sum_{i=1}^{n} m_i/J_0\right)\left(n^{-1}\sum_{i=1}^{n}(\psi_i^a - \hat{\psi}_i^a)/J_0\right)}{\left(n^{-1}\sum_{i=1}^{n}\hat{\psi}_i^a/J_0\right)\left(n^{-1}\sum_{i=1}^{n}\psi_i^a/J_0\right)} \tag{E.5}$$

In what follows, I show that $I_1 = \mathcal{T}_{n,K}^{l} + \mathcal{T}_{n,K}^{nl} + o_p(n^{-\zeta})$ and $I_2 = o_p(n^{-\zeta})$, which is sufficient to complete the proof of the theorem since both $\mathcal{T}_{n,K}^{l}$ and $\mathcal{T}_{n,K}^{nl}$ are $O_p(n^{-\varphi_1})$ and $O_p(n^{1/2-2\varphi_1})$, respectively, under Assumptions 3.1 and 3.2 and by the proof of Propositions C.4 and C.5. Furthermore, if Assumption 3.3 holds, part 2 of Proposition C.5 implies $\text{Var}[n^{2\varphi_1-1/2}\mathcal{T}_{n,K}^{nl}] = G_\delta(K^2-3K+3)(K-1)^{-1-4\varphi_1}K^{4\varphi_1-1} + n^{4\varphi_1-1}r_{n,K}^{nl}$, which implies that $\lim_{n\to\infty}\inf_{K\leq n}\text{Var}[n^{2\varphi_1-1/2}\mathcal{T}_{n,K}^{nl}] > 0$. Part 1 of Proposition C.5 implies that $\sup_{K\leq n}|n^{2\varphi_1-1/2}E[\mathcal{T}_{n,K}^{nl}]| < \infty$; then, $\lim_{n\to\infty}\sup_{K\leq n}E[(n^{2\varphi_1-1/2}\mathcal{T}_{n,K}^{nl})^2] < \infty$. Similarly, the proof of Propositions C.4 guarantees $\lim_{n\to\infty}\sup_{K\leq n}E[(n^{\varphi_1}\mathcal{T}_{n,K}^{l})^2] < \infty$.

Claim 1: $I_1 = \mathcal{T}_{n,K}^{l} + \mathcal{T}_{n,K}^{nl} + o_p(n^{-\zeta})$. I first rewrite the RHS of (E.4) using the identity $a(1+b)^{-1} = a - ab(1+b)^{-1}$, where $a = n^{-1/2}\sum_{i=1}^{n}(\hat{m}_i - m_i)/J_0$ and $b = n^{-1}\sum_{i=1}^{n}(\hat{\psi}_i^a - J_0)/J_0$. That is

$$I_1 = a - ab(1+b)^{-1}$$

I then conclude the proof of the claim by using claims 1.1 and 1.2, stated below.

Claim 1.1: $a = \mathcal{T}_{n,K}^{l} + \mathcal{T}_{n,K}^{nl} + o_p(n^{-\zeta})$. This result holds by part 4 of Lemma C.2 since $a = n^{-1/2}\sum_{i=1}^{n}(\hat{m}_i - m_i)/J_0$.

Claim 1.2: $ab(1+b)^{-1} = o_p(n^{-\zeta})$. Note that part 4 of Lemma C.2 implies $a = O_p(n^{1/2-2\varphi_1})$. Note also that part 2 of Lemma C.2 and CLT imply $b = O_p(n^{-1/2})$, which guarantees that $(1+b)^{-1} = O_p(1)$; therefore, $ab(1+b)^{-1} = O_p(n^{-2\varphi_1})$, which is $o_p(n^{-\zeta})$ since $\varphi_1 < 1/2$.

Claim 2: $I_2 = o_p(n^{-\zeta})$. I first rewrite I_2 defined in (E.5) as follows,

$$I_2 = ab(1+c-b)^{-1}(1+c)^{-1} \ ,$$

where $a = n^{-1/2}\sum_{i=1}^{n} m_i/J_0$, $b = n^{-1}\sum_{i=1}^{n}(\psi_i^a - \hat{\psi}_i^a)/J_0$, and $c = n^{-1}\sum_{i=1}^{n}(\psi_i^a -$

$J_0)/J_0$. CLT implies that $a = O_p(1)$ and $c = O_p(n^{-1/2})$. Part 2 of Lemma C.2 implies $b = O_p(n^{-2\varphi_1})$. Therefore, $ab = O_p(n^{-2\varphi_1})$, and both $(1+c-b)^{-1}$ and $(1+c)^{-1}$ are $O_p(1)$. This implies I_2 is $O_p(n^{-2\varphi_1})$, which is $o_p(n^{-\zeta})$ since $\varphi_1 < 1/2$. □

E.3 Proof of Proposition C.1

Proof. **Part 1:** By the definition of the oracle version of the DML1 estimator in (2.8) and the moment function m in (2.2), it follows

$$n^{1/2}\left(\hat{\theta}^*_{n,1} - \theta_0\right) = K^{-1/2}\sum_{k=1}^{K}\frac{n_k^{-1/2}\sum_{i\in\mathcal{I}_K} m_i}{n_k^{-1}\sum_{i\in\mathcal{I}_K}\psi_i^a} \ . \tag{E.6}$$

I first rewrite the RHS of (E.6) using the identity $a_k(1+b_k)^{-1} = a_k - a_k b_k + a_k b_k^2(1+b_k)^{-1}$ with $a_k = n_k^{-1/2}\sum_{i\in\mathcal{I}_k} m_i/J_0$ and $b_k = n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0$. That is

$$n^{1/2}\left(\hat{\theta}^*_{n,1} - \theta_0\right) = I_1 + I_2 + I_3 \ ,$$

where

$$I_1 = K^{-1/2}\sum_{k=1}^{K} a_k$$
$$I_2 = -K^{-1/2}\sum_{k=1}^{K} a_k b_k$$
$$I_3 = K^{-1/2}\sum_{k=1}^{K} a_k b_k^2(1+b_k)^{-1}$$

By CLT, it follows $I_1 = n^{-1/2}\sum_{i=1}^{n} m_i/J_0 \xrightarrow{d} N(0,\sigma^2)$ as $n\to\infty$, where σ^2 is as in (2.11). Therefore, if $I_2 - K/\sqrt{n}\Lambda$ and I_3 are $o_p(1)$, then

$$n^{1/2}\left(\hat{\theta}^*_{n,1} - \theta_0\right) = n^{-1/2}\sum_{i=1}^{n} m_i/J_0 + K/\sqrt{n}\Lambda + o_p(1) \ ,$$

which is sufficient to complete the proof of part 1 since $K/\sqrt{n} \to c$ as $n\to\infty$. In what follows, Claim 1 shows $I_2 - K/\sqrt{n}\Lambda = o_p(1)$ and Claim 2 shows I_3 are $o_p(1)$.

Claim 1: $I_2 - K/\sqrt{n}\Lambda = o_p(1)$. First, note that $E[-a_k b_k] = K^{1/2}/\sqrt{n}\Lambda$ due to the following derivations,

$$
\begin{aligned}
E[a_k b_k] &\overset{(1)}{=} E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k} m_i/J_0\right)\left(n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0\right)\right] \\
&\overset{(2)}{=} n_k^{-1/2} E\left[(m_i/J_0)\left((\psi_i^a - J_0)/J_0\right)\right] \\
&\overset{(3)}{=} -n^{-1/2}K^{1/2}\Lambda
\end{aligned}
$$

where (1) holds by definition of a_k and b_k, (2) holds since $\{(m_i, \psi_i^a - J_0) : i \in \mathcal{I}_k\}$ are zero mean i.i.d. random vectors, and (3) holds by the definition of Λ in (3.6) and condition (2.1).

Therefore, $E[I_2] = -K^{-1/2}\sum_{k=1}^{K} E[a_k b_k] = K/\sqrt{n}\Lambda$, which implies that the claim is equivalent to show that $I_2 - E[I_2]$ is $o_p(1)$, which follows by the following derivations

$$
\begin{aligned}
E\left[(I_2 - E[I_2])^2\right] &\overset{(1)}{=} E\left[\left(K^{-1/2}\sum_{k=1}^{K}(a_k b_k - E[a_k b_k])\right)^2\right] \\
&\overset{(2)}{=} K^{-1}\sum_{k=1}^{K} E\left[(a_k b_k - E[a_k b_k])^2\right] \\
&\overset{(3)}{\leq} E\left[(a_k b_k)^2\right] \\
&\overset{(4)}{=} n_k^{-1} E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k} m_i/J_0\right)^2\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0\right)^2\right] \\
&\overset{(5)}{\leq} n_k^{-1} E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k} m_i/J_0\right)^4\right]^{1/2} \times E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0\right)^4\right]^{1/2} \\
&\overset{(6)}{=} n_k^{-1} \times O(1) \times O(1)
\end{aligned}
$$

where (1) holds by definition of I_2; (2) and (3) hold since $\{a_k b_k - E[a_k b_K] : 1 \leq k \leq K\}$ are zero mean i.i.d random variables due to the definition of a_k and b_k; (4) holds by definition of a_k and b_k; (5) holds by Cauchy-Schwartz; and (6) holds since $\{(m_i, \psi_i^a - J_0) : i \in \mathcal{I}_k\}$ are zero mean i.i.d. random vectors, parts (a) and (c) of Assumption 3.1, and $n_k \to \infty$. This completes the proof of Claim 1.

Claim 2: $I_3 = o_p(1)$. Consider the following derivation

$$|I_3| \overset{(1)}{\leq} \max_{k=1,\ldots,K} \left|1 + n_k^{-1} \sum_{i \in \mathcal{I}_k} (\psi_i^a - J_0)/J_0 \right|^{-1} \times K^{-1/2} \sum_{k=1}^{K} |a_k| b_k^2$$
$$= O_p(1) \times o_p(1) \ ,$$

where (1) holds by definition of I_3 and triangular inequality, and (2) holds by Lemma E.1 and (E.7) presented below,

$$K^{-1/2} \sum_{k=1}^{K} |a_k| b_k^2 = o_p(1) \ . \tag{E.7}$$

To prove (E.7), consider the following

$$E\left[K^{-1/2} \sum_{k=1}^{K} |a_k| b_k^2\right] \overset{(1)}{\leq} K^{-1/2} \sum_{k=1}^{K} E[|a_k|^2]^{1/2} E[|b_k|^4]^{1/2}$$
$$\overset{(2)}{=} K^{1/2} n_k^{-1} E\left[\left(n_k^{-1/2} \sum_{i \in \mathcal{I}_k} m_i/J_0\right)^2\right]^{1/2} \times E\left[\left(n_k^{-1/2} \sum_{i \in \mathcal{I}_k} (\psi_i^a - J_0)/J_0\right)^4\right]^{1/2}$$
$$\overset{(3)}{=} (Kn^{-1/2})^{3/2} n^{-1/4} O(1) \times O(1)$$
$$\overset{(4)}{=} o(1) \ ,$$

where (1) holds by Cauchy-Schwartz, (2) holds since $\{(a_k, b_k\} : 1 \leq k \leq K\}$ are i.i.d random vectors and the definition of a_k and b_k, (3) holds since $\{(m_i, \psi_i^a - J_0) : i \in \mathcal{I}_k\}$ are zero mean i.i.d. random vectors, and parts (a) and (c) of Assumption 3.1, and (4) holds since $K = O(n^{1/2})$. This completes the proof of Claim 2.

Part 2: By the definition of $\hat{\theta}^*_{n,2}$ in (2.9), and the moment function m in (2.2), it follows

$$n^{1/2}\left(\hat{\theta}^*_{n,2} - \theta_0\right) = \frac{n^{-1/2} \sum_{i=1}^{n} m_i/J_0}{n^{-1} \sum_{i=1}^{n} \psi_i^a/J_0}$$

Since the denominator converges to 1 in probability by the LLN and the numerator converges to $N(0, \sigma^2)$ in distribution due to the CLT, it follows that $n^{1/2}\left(\hat{\theta}^*_{n,2} - \theta_0\right)$ converges in distribution to $N(0, \sigma^2)$. This completes the proof of part 2 □

E.4 Proof of Proposition C.2

Proof. **Part 1:** By the definition of the oracle version of DML2 estimator in (2.9), and the moment function m in (2.2), it follows

$$n^{1/2}\left(\hat{\theta}^*_{n,2}-\theta_0\right)=\frac{n^{-1/2}\sum_{i=1}^n m_i}{n^{-1}\sum_{i=1}^n \psi_i^a}\ . \tag{E.8}$$

I rewrite the RHS of (E.8) using the identity $a(1+b)^{-1}=a-ab+ab^2(1+b)^{-1}$ with $a=n^{-1/2}\sum_{i=1}^n m_i/J_0$ and $b=n^{-1}\sum_{i=1}^n(\psi_i^a-J_0)/J_0$. That is

$$\begin{aligned} n^{1/2}\left(\hat{\theta}^*_{n,2}-\theta_0\right) &= a-ab+ab^2(1+b)^{-1}\\ &\overset{(1)}{=} \mathcal{T}_n^*+\mathcal{T}_n^{dml2}+ab^2(1+b)^{-1}\end{aligned}$$

where (1) holds by the definition of $\mathcal{T}_n^*$ and $\mathcal{T}_n^{dml2}$. It is sufficient to show $ab^2(1+b)^{-1}$ is $O_p(n^{-1})$ to complete the proof, which follows by CLT that implies $a=O_p(1)$ and $b=O_p(n^{-1/2})$, and $(1+b)^{-1}=O_p(1)$.

Finally, consider the following derivations

$$\begin{aligned} E[\mathcal{T}_n^{dml2}] &\overset{(1)}{=} -n^{-1/2}E\left[\left(n^{-1/2}\sum_{i=1}^n m_i/J_0\right)\left(n^{-1/2}\sum_{i=1}^n(\psi_i^a-J_0)/J_0\right)\right]\\ &\overset{(2)}{=} -n^{-1/2}\Lambda\end{aligned}$$

where (1) holds by the definition of $\mathcal{T}_n^{dml2}$, and (2) holds since $\{(m_i,(\psi_i^a-J_0)/J_0):1\le i\le n\}$ are zero mean i.i.d. random vectors and by the definition of Λ in (3.6).

Part 2: By definition of σ^2 and since $\{(m_i/J_0):1\le i\le n\}$ are zero mean i.i.d. random variables, it follows that $E[\mathcal{T}_n^*]=0$ and $E[(\mathcal{T}_n^*)^2]=\sigma^2$.

Part 3: First note that $\text{Cov}(\mathcal{T}_n^*,\mathcal{T}_n^{dml2})=E[(\mathcal{T}_n^*)(\mathcal{T}_n^{dml2})]$. Now, consider the following derivations,

$$\begin{aligned} E[(\mathcal{T}_n^*)(\mathcal{T}_n^{dml2})] &\overset{(1)}{=} -n^{-1/2}E\left[\left(n^{-1/2}\sum_{i=1}^n m_i/J_0\right)^2\left(n^{-1/2}\sum_{i=1}^n(\psi_i^a-J_0)/J_0\right)\right]\\ &\overset{(2)}{=} -n^{-1}E\left[(m_i/J_0)^2\left((\psi_i^a-J_0)/J_0\right)\right]\end{aligned}$$

$$\stackrel{(3)}{=} -n^{-1}\Xi_1$$

where (1) holds by definition of $\mathcal{T}_n^*$ and $\mathcal{T}_n^{dml2}$, (2) holds since $\{(m_i/J_0, (\psi_i^a - J_0)/J_0) : 1 \leq i \leq n\}$ are zero mean i.i.d. random vectors, and (3) holds by definition of Ξ_1 in (C-1).

Similarly, consider the following derivations

$$\begin{aligned}
Var[\mathcal{T}_n^{dml2}] &\stackrel{(1)}{=} n^{-1}E\left[\left(n^{-1/2}\sum_{i=1}^{n} m_i/J_0\right)^2 \left(n^{-1/2}\sum_{i=1}^{n}(\psi_i^a - J_0)/J_0\right)^2\right] - n^{-1}\Lambda^2 \\
&\stackrel{(2)}{=} n^{-1}\left(E[(m_i/J_0)^2]E[((\psi_i^a - J_0)/J_0)^2] + 2\Lambda^2 + O(n^{-1})\right) - n^{-1}\Lambda^2 \\
&\stackrel{(3)}{=} n^{-1}(\sigma^2\sigma_a^2 + \Lambda^2) + O(n^{-2})
\end{aligned}$$

where (1) holds by definition of $\mathcal{T}_n^{dml2}$ and Λ in (A-4) and (3.6), respectively, (2) holds since $\{(m_i/J_0, (\psi_i^a - J_0)/J_0) : 1 \leq i \leq n\}$ are zero mean i.i.d. random vectors and by definition of Λ, and (3) holds by definition of σ^2 and σ_a^2 in (2.11) and (C-2), respectively. □

E.5 Proof of Proposition C.3

Proof. For $i \in \mathcal{I}_k$, denote $\Delta_i = \Delta_i^b + \Delta_i^l$, where

$$\begin{aligned}
\Delta_i^l &= n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} \delta_{n_0,j,i} \ , \\
\Delta_i^b &= n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} b_{n_0,j,i} \ ,
\end{aligned}$$

Here, $\delta_{n_0,j,i} = \delta_{n_0}(W_j, X_i)$ and $b_{n_0,j,i} = b_{n_0}(X_j, X_i)$, and δ_{n_0} and b_{n_0} are functions satisfying Assumption 3.2.

Part 1: Using the previous notation, it holds that $E[(\Delta_i)^\top \partial_\eta m_i/J_0] = 0$. To see this, consider the following derivations

$$\begin{aligned}
E[(\Delta_i)^\top \partial_\eta m_i/J_0] &\stackrel{(1)}{=} E\left[(\Delta_i)^\top E\left[\partial_\eta m_i/J_0 \mid (W_j : j \notin \mathcal{I}_k), X_i\right]\right] \\
&\stackrel{(2)}{=} E\left[(\Delta_i)^\top E\left[\partial_\eta m_i/J_0 \mid X_i\right]\right]
\end{aligned}$$

$$\stackrel{(3)}{=} 0 \ ,$$

where (1) holds by the law of interactive expectations and because Δ_i is non-stochastic conditional on $(W_j : j \notin \mathcal{I}_k)$ and X_i, (2) holds since $\{W_j : 1 \leq j \leq n\}$ are i.i.d. random vectors and $i \in \mathcal{I}_k$, and (3) holds by the Neyman orthogonality condition (part (b) of Assumption 3.1).

Therefore, $E[\mathcal{T}^l_{n,K}] = 0$ holds due to the definition of $\mathcal{T}^l_{n,K}$ in (A-3) and the previous result,

$$\begin{aligned} E[\mathcal{T}^l_{n,K}] &= n^{-1/2} \sum_{i=1}^{n} E[(\Delta_i)^\top \partial_\eta m_i / J_0] \\ &= 0 \ . \end{aligned}$$

Part 2: By part 1, $\text{Var}[\mathcal{T}^l_{n,K}] = E[(\mathcal{T}^l_{n,K})^2]$. Now, consider the following decomposition:

$$\begin{aligned} E[(\mathcal{T}^l_{n,K})^2] &= E\left[\left(n^{-1/2} \sum_{i=1}^{n} (\Delta_i)^\top \partial_\eta m_i / J_0\right)^2\right] \\ &= n^{-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} E\left[\left((\Delta_{i_1})^\top \partial_\eta m_{i_1} / J_0\right)\left((\Delta_{i_2})^\top \partial_\eta m_{i_2} / J_0\right)\right] \\ &= I_1 + 2I_2 + I_3 \ , \end{aligned}$$

where I use $\Delta_i = \Delta^l_i + \Delta^b_i$ in the last equality, with I_1, I_2, and I_3 defined below,

$$\begin{aligned} I_1 &= n^{-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} E\left[\left((\Delta^l_{i_1})^\top \partial_\eta m_{i_1} / J_0\right)\left((\Delta^l_{i_2})^\top \partial_\eta m_{i_2} / J_0\right)\right] \\ I_2 &= n^{-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} E\left[\left((\Delta^l_{i_1})^\top \partial_\eta m_{i_1} / J_0\right)\left((\Delta^b_{i_2})^\top \partial_\eta m_{i_2} / J_0\right)\right] \\ I_3 &= n^{-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} E\left[\left((\Delta^b_{i_1})^\top \partial_\eta m_{i_1} / J_0\right)\left((\Delta^b_{i_2})^\top \partial_\eta m_{i_2} / J_0\right)\right] \end{aligned}$$

In what follows, I show $I_1 = n_0^{-2\varphi_1} G^l_\delta + o(n^{-2\varphi_1})$ with G^l_δ defined as in (A-6), $I_2 = 0$, and $I_3 = O(n^{-2\varphi_1})$, which is sufficient to complete the proof of Part 2.

Claim 1: $I_1 = n_0^{-2\varphi_1} G_\delta^l + o(n^{-2\varphi_1})$. Consider the following derivations,

$$
\begin{aligned}
& I_1 \\
&= n^{-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} E\left[\left((\Delta_{i_1}^l)^\top \partial_\eta m_{i_1}/J_0 \right) \left((\Delta_{i_2}^l)^\top \partial_\eta m_{i_2}/J_0 \right) \right] \\
&= n^{-1} \sum_{k_1,k_2=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \in \mathcal{I}_{k_2}} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_{k_1}} \sum_{j_2 \notin \mathcal{I}_{k_2}} E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] \\
&\overset{(1)}{=} n^{-1} \sum_{k_1,k_2=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \in \mathcal{I}_{k_2}} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_{k_1}} \sum_{j_2 \notin \mathcal{I}_{k_2}} E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] I\{k_1 \neq k_2\} \\
&\quad + n^{-1} \sum_{k=1}^{K} \sum_{i_1 \in \mathcal{I}_k} \sum_{i_2 \in \mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] I\{i_1 \neq i_2\} \\
&\quad + n^{-1} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E\left[\left(\delta_{n_0,j_1,i}^\top \partial_\eta m_i/J_0 \right) \left(\delta_{n_0,j_2,i}^\top \partial_\eta m_i/J_0 \right) \right] \\
&\overset{(2)}{=} n^{-1} \sum_{k_1,k_2=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \in \mathcal{I}_{k_2}} n_0^{-2\varphi_1} n_0^{-1} E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] I\{k_1 \neq k_2\} \\
&\quad + n^{-1} \sum_{k=1}^{K} \sum_{i_1 \in \mathcal{I}_k} \sum_{i_2 \in \mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] I\{i_1 \neq i_2\} \\
&\quad + n^{-1} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \right] \\
&\overset{(3)}{=} n^{-1} \sum_{k_1,k_2=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \in \mathcal{I}_{k_2}} n_0^{-2\varphi_1} n_0^{-1} E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] I\{k_1 \neq k_2\} \\
&\quad + n^{-1} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \right] \\
&\overset{(4)}{=} n^{-1} \sum_{k_1=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \notin \mathcal{I}_{k_1}} n_0^{-2\varphi_1} n_0^{-1} E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] \\
&\quad + n_0^{-2\varphi_1} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \right] \\
&\overset{(5)}{=} n_0^{-2\varphi_1} \left(E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0 \right) \left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0 \right) \right] + E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0 \right) \right] \right)
\end{aligned}
$$

$$\stackrel{(6)}{=} n_0^{-2\varphi_1} G_\delta^l + o(n^{-2\varphi_1}) \ ,$$

where (1) holds because there are 3 possible situations for $i_1 \in \mathcal{I}_{k_1}$ and $i_2 \in \mathcal{I}_{k_2}$: i) $k_1 \neq k_2$, ii) $k_1 = k_2$ but $i_1 \neq i_2$, and iii) $i_1 = i_2$, (2) holds by the law of iterative expectations and since

$$E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid X_{i_1}, W_{i_2}, W_{j_1}, W_{j_2}\right] = 0 \ , \ \text{when } i_1 \neq j_2$$

and

$$E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid X_{i_2}, W_{i_1}, W_{j_1}, W_{j_2}\right] = 0 \ , \ \text{when } i_2 \neq j_1 \ ,$$

(3) holds by the law of iterative expectations and since

$$E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(\delta_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid X_{i_2}, W_{i_1}, W_{j_1}, W_{j_2}\right] = 0 \ , \text{when } i_1 \neq i_2 \ ,$$

(4) holds since $\{\left(\delta_{n_0,j,i}^\top \partial_\eta m_i/J_0\right) : j \notin \mathcal{I}_k\}$ are i.i.d. random variables conditional on W_i (here I use $i \in \mathcal{I}_k$), by noting that $\sum_{k_2=1}^K I_{k_2 \neq k_1} \sum_{i_2 \in \mathcal{I}_{k_2}} (\cdot) = \sum_{i_2 \notin \mathcal{I}_{k_1}} (\cdot)$, and recalling that n_0 is the number of observations outside the fold $\mathcal{I}_k$, (5) holds because the random variables $\{\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(\delta_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0\right) : i_1 \neq i_2\}$ are identically distributed, and (6) holds by the definition of G_δ^l in (A-6) and $n/2 \leq n_0 \leq n$. This completes the proof of claim 1.

Claim 2: $I_2 = 0$. First, consider the following derivations

$$\begin{aligned}
&E\left[\left((\Delta_{i_1}^l)^\top \partial_\eta m_{i_1}/J_0\right)\left((\Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&\stackrel{(1)}{=} n_0^{-\varphi_1-\varphi_2} n_0^{-3/2} \sum_{j_1 \notin \mathcal{I}_{k_1}} \sum_{j_2 \notin \mathcal{I}_{k_2}} E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&\stackrel{(2)}{=} 0 \ ,
\end{aligned}$$

where (1) holds by definition of Δ_i^l and Δ_i^b, and (2) holds by considering 3 possible cases:

- If $j_1 \neq i_2$ and $j_1 \neq j_2$ (j_1 different than all other sub-indices), then

$$E\left[\left(\delta_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid W_{i_1}, W_{i_2}, W_{j_2}\right] = 0 \ ,$$

since $E[\delta_{n_0,j_1,i_1} \mid W_{i_1}, W_{i_2}, W_{j_2}] = 0$ due to part (a) of Assumption 3.2.

- If $j_1 = i_2$, then $i_2 \neq i_1$ (otherwise $j_1 \in \mathcal{I}_k$) and

$$E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid W_{i_2}, X_{i_1}, W_{j_2}\right] = 0 \ ,$$

  since $E[\delta_{n_0,i_2,i_1} \mid W_{i_1}, X_{i_2}, W_{j_2}] = 0$ due to part (a) of Assumption 3.2.

- If $j_1 = j_2 = j$, then

$$E\left[\left(\delta_{n_0,j,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,j,i_2}^\top \partial_\eta m_{i_2}/J_0\right) \mid X_j, W_{i_1}, W_{i_2}\right] = 0 \ ,$$

  since $E[\delta_{n_0,j,i_1} \mid X_j, W_{i_1}, W_{i_2}] = 0$ due to part (a) of Assumption 3.2.

Therefore,

$$\begin{aligned} I_{1,2} &= n^{-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} E\left[\left((\Delta_{i_1}^l)^\top \partial_\eta m_{i_1}/J_0\right)\left((\Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\ &= 0 \ , \end{aligned}$$

which completes the proof of claim 2.

Claim 3: $I_3 = O(n^{-2\varphi_1})$. Algebra shows

$$\begin{aligned} I_3 &= n^{-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} E\left[\left((\Delta_{i_1}^b)^\top \partial_\eta m_{i_1}/J_0\right)\left((\Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\ &= n^{-1}\sum_{k_1,k_2=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{i_2\in\mathcal{I}_{k_2}} n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1\notin\mathcal{I}_{k_1}}\sum_{j_2\notin\mathcal{I}_{k_2}} E\left[\left(b_{n_0,j_1,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,j_2,i_2}^\top \partial_\eta m_{i_2}/J_0\right)\right] \\ &\overset{(1)}{=} n^{-1}\sum_{k_1=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{i_2\notin\mathcal{I}_{k_1}} n_0^{-2\varphi_2} n_0^{-2} E\left[\left(b_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0\right)\left(b_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0\right)\right] \\ &\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2} n_0^{-2}\sum_{j\notin\mathcal{I}_k} E\left[\left(b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)\left(b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)\right] \\ &\overset{(2)}{=} n_0^{-1} E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)\left(n_0^{-\varphi_2} b_{n_0,i,j}^\top \partial_\eta m_j/J_0\right)\right] \\ &\quad + n_0^{-1} E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)\right] \ , \\ &\overset{(3)}{\le} n_0^{-1}\left(E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i/J_0\right)^2\right]\right)^{1/2}\left(E\left[\left(n_0^{-\varphi_2} b_{n_0,i,j}^\top \partial_\eta m_i/J_0\right)^2\right]\right)^{1/2} \end{aligned}$$

$$
\begin{aligned}
&+ n_0^{-1} E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i / J_0\right)^2\right] \\
&\overset{(4)}{\leq} 2(p\tilde{C}_1/C_0^2) n_0^{-1} E\left[||n_0^{-\varphi_2} b_{n_0,j,i}||^2\right] \\
&\overset{(5)}{\leq} 2(p\tilde{C}_1/C_0^2) n_0^{-1} n_0^{1-2\varphi_1} \tau_{n_0} \\
&\overset{(6)}{=} o(n^{-2\varphi_1}) \ ,
\end{aligned}
$$

where (1) uses the same argument to calculate I_1, (2) holds since the random vectors $\{\left(b_{n_0,i_2,i_1}^\top \partial_\eta m_{i_1}/J_0, b_{n_0,i_1,i_2}^\top \partial_\eta m_{i_2}/J_0\right) : i_1 \neq i_2\}$ are identically distributed, (3) holds by Cauchy-Schwartz inequality, (4) holds by the inequalities (E.9) and (E.10) presented below where $\tilde{C}_1$ is a constant depending only on (C_0, C_1, M), (5) holds by part (b.4) in Assumption 3.2, and (6) holds since $n/2 \leq n_0 \leq n$ and $\tau_{n_0} = o(1)$.

$$
E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i / J_0\right)^2\right] \leq (p\tilde{C}_1/C_0^2) E\left[||n_0^{-\varphi_2} b_{n_0,j,i}||^2\right] \tag{E.9}
$$

$$
E\left[\left(n_0^{-\varphi_2} b_{n_0,i,j}^\top \partial_\eta m_i / J_0\right)^2\right] \leq (p\tilde{C}_1/C_0^2) E\left[||n_0^{-\varphi_2} b_{n_0,j,i}||^2\right] \tag{E.10}
$$

To verify (E.9) consider the following derivation,

$$
\begin{aligned}
E\left[\left(n_0^{-\varphi_2} b_{n_0,j,i}^\top \partial_\eta m_i / J_0\right)^2\right] &\overset{(1)}{=} E\left[n_0^{-\varphi_2} b_{n_0,j,i}^\top E\left[(\partial_\eta m_i/J_0)(\partial_\eta m_i/J_0)^\top \mid X_j, X_i\right] n_0^{-\varphi_2} b_{n_0,j,i}\right] \\
&\overset{(2)}{\leq} (1/C_0^2) E\left[n_0^{-\varphi_2} b_{n_0,j,i}^\top E\left[(\partial_\eta m_i)(\partial_\eta m_i)^\top \mid X_i\right] n_0^{-\varphi_2} b_{n_0,j,i}\right] \\
&\overset{(3)}{\leq} p(\tilde{C}_1/C_0^2) E\left[||n_0^{-\varphi_2} b_{n_0,j,i}||^2\right]
\end{aligned}
$$

where (1) holds by LIE and since $b_{n_0,j,i}$ is non-random conditional on X_j and X_i, (2) holds by part (a) of Assumption 3.1 and independence between X_i and X_j since $i \neq j$, and (3) holds by definition of euclidean norm and since $||E[(\partial_\eta m_i)(\partial_\eta m_i)^\top \mid X_i]||_\infty \leq \tilde{C}_1 = C_1(1 + M^{1/4}/C_0)^2$ due to parts (d) in Assumption 3.1 and $|\theta_0| \leq M^{1/4}/C_0$ (which holds by definition of θ_0 and parts (a) and (c) of Assumption 3.1).

The verification of (E.10) follows the same previous derivations but reverting the role of i and j. Lastly, it uses that $E\left[||n_0^{-\varphi_2} b_{n_0,i,j}||^2\right] = E\left[||n_0^{-\varphi_2} b_{n_0,j,i}||^2\right]$ since $b_{n_0,j,i}$ and $b_{n_0,i,j}$ have the same distribution for $i \neq j$.

Part 3: By Cauchy-Schwartz, parts 3 of Proposition C.2, and part 2 of this proposition,

$$\mathrm{Cov}(\mathcal{T}_n^{dml}, \mathcal{T}_{n,K}^{l}) \leq \left((G_\delta^l)^{1/2}(K/(K-1))^{\varphi_1} + o(1)\right)^{1/2} \left(\sigma^2\sigma_a^2 + \Lambda^2 + o(1)\right)^{1/2} n^{-\varphi_1 - 1/2} ,$$

which implies the RHS is $O(n^{-\varphi_1 - 1/2})$, and this is $o(n^{-2\varphi_1})$ since $\varphi_1 < 1/2$. □

E.6 Proof of Proposition C.4

Proof. For $i \in \mathcal{I}_k$, denote $\Delta_i = \Delta_i^b + \Delta_i^l$, where

$$\Delta_i^l = n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} \delta_{n_0,j,i} ,$$
$$\Delta_i^b = n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} b_{n_0,j,i} ,$$

Here, $\delta_{n_0,j,i} = \delta_{n_0}(W_j, X_i)$ and $b_{n_0,j,i} = b_{n_0}(X_j, X_i)$, and δ_{n_0} and b_{n_0} are functions satisfying Assumption 3.2. Denote $H_i = \partial_\eta^2 m_i/(2J_0)$.

Part 1: Consider the following decomposition using the definition of $\mathcal{T}_{n,K}^{nl}$ in (3.12),

$$E[\mathcal{T}_{n,K}^{nl}] = n^{-1/2} \sum_{i=1}^{n} E\left[\Delta_i^\top H_i \Delta_i\right]$$
$$= I_1 + 2I_2 + I_3$$

where

$$I_1 = n^{-1/2} \sum_{i=1}^{n} E\left[(\Delta_i^l)^\top H_i \Delta_i^l\right]$$
$$I_2 = n^{-1/2} \sum_{i=1}^{n} E\left[(\Delta_i^b)^\top H_i \Delta_i^l\right]$$
$$I_3 = n^{-1/2} \sum_{i=1}^{n} E\left[(\Delta_i^b)^\top H_i \Delta_i^b\right]$$

In what follows, I show $I_1 = n^{1/2} n_0^{-2\varphi_1} F_\delta + o(n^{1/2-2\varphi_1})$, $I_2 = 0$, $I_3 = n^{1/2} n_0^{-2\varphi_2} F_b + o(n^{1/2-2\varphi_1})$, which is sufficient to complete the proof of Part 1 since $n_0 = ((K-1)/K)n$.

Claim 1: $I_1 = n^{1/2} n_0^{-2\varphi_1} F_\delta + o(n^{1/2-2\varphi_1})$. Consider the following derivations,

$$
\begin{aligned}
E[(\Delta_i^l)^\top H_i \Delta_i^l] &\overset{(1)}{=} n_0^{-2\varphi_1} n_0^{-1} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(\delta_{n_0,j_1,i})^\top H_i (\delta_{n_0,j_2,i})] \\
&\overset{(2)}{=} n_0^{-2\varphi_1} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E[(\delta_{n_0,j,i})^\top H_i (\delta_{n_0,j,i})] \\
&\overset{(3)}{=} n_0^{-2\varphi_1} E[(\delta_{n_0,j,i})^\top H_i (\delta_{n_0,j,i})] \ ,
\end{aligned}
$$

where (1) holds by definition of Δ_i^l, and (2) and (3) hold since $\{\delta_{n_0,j,i} : j \notin \mathcal{I}_k\}$ are zero mean i.i.d. random vectors conditional on W_i due to part (a) of Assumption 3.2 (here I use that $i \in \mathcal{I}_k$). Therefore,

$$
\begin{aligned}
I_1 &= n^{-1/2} \sum_{i=1}^{n} E\left[(\Delta_i^l)^\top H_i (\Delta_i^l)\right] \\
&= n^{1/2} n_0^{-2\varphi_1} E[(\delta_{n_0,j,i})^\top H_i (\delta_{n_0,j,i})] \\
&\overset{(1)}{=} n^{1/2} n_0^{-2\varphi_1} F_\delta + o(n^{1/2-2\varphi_1})
\end{aligned}
$$

where (1) holds by definition of F_δ in (3.3), Assumption A.1, and because $n/2 \leq n_0 \leq n$. This completes the proof of claim 1.

Claim 2: $I_2 = 0$. Consider the following derivations,

$$
\begin{aligned}
E[(\Delta_i^b)^\top H_i (\Delta_i^l)] &\overset{(1)}{=} n_0^{-\varphi_2-\varphi_1} n_0^{-3/2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(b_{n_0,j_1,i})^\top H_i (\delta_{n_0,j_2,i})] \\
&\overset{(2)}{=} 0
\end{aligned}
$$

where (1) holds by definition of Δ_i^b and Δ_i^l, and (2) holds since

$$
E[(b_{n_0,j_1,i})^\top (\partial_\eta^2 m_i/(2J_0))(\delta_{n_0,j_2,i}) \mid X_{j_2}, X_{j_1}, W_i] = 0
$$

due to part (a) of Assumption 3.2 ($E[\delta_{n_0,j_2,i} \mid X_{j_2}, W_i] = 0$). Therefore,

$$
I_2 = n^{-1/2} \sum_{i=1}^{n} E\left[(\Delta_i^b)^\top (\partial_\eta^2 m_i/(2J_0))(\Delta_i^l)\right]
$$

$$= 0 \ ,$$

which completes the proof of claim 2.

Claim 3: $I_3 = n^{1/2} n_0^{-2\varphi_2} F_b + o(n^{1/2-2\varphi_1})$. Denote $\tilde{b}_{n_0,i} = E[b_{n_0,j,i} \mid X_i]$ for $j \neq i$. Consider the following derivations,

$$\begin{aligned}
E[(\Delta_i^b)^\top H_i(\Delta_i^l)] &\overset{(1)}{=} n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(b_{n_0,j_1,i})^\top H_i(b_{n_0,j_2,i})] \\
&\overset{(2)}{=} n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(b_{n_0,j_1,i} - \tilde{b}_{n_0,i})^\top H_i(b_{n_0,j_2,i} - \tilde{b}_{n_0,i})] \\
&\quad + n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(b_{n_0,j_1,i} - \tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})] \\
&\quad + n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(\tilde{b}_{n_0,i})^\top H_i(b_{n_0,j_2,i} - \tilde{b}_{n_0,i})] \\
&\quad + n_0^{-2\varphi_2} n_0^{-2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E[(\tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})] \\
&\overset{(3)}{=} n_0^{-2\varphi_2} n_0^{-1} E[(b_{n_0,j,i} - \tilde{b}_{n_0,i})^\top H_i(b_{n_0,j,i} - \tilde{b}_{n_0,i})] \\
&\quad + n_0^{-2\varphi_2} E[(\tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})]
\end{aligned}$$

where (1) holds by definition of Δ_i^b, (2) holds by adding and subtracting $\tilde{b}_{n_0,i}$, and (3) holds since $\{b_{n_0,j,i} - \tilde{b}_{n_0,i} : j \notin \mathcal{I}_k\}$ are zero mean i.i.d. random vectors conditional on W_i, which implies $E[(\tilde{b}_{n_0,i})^\top H_i(b_{n_0,j,i} - \tilde{b}_{n_0,i}) \mid W_i] = 0$. Therefore,

$$\begin{aligned}
I_3 &= n^{-1/2} \sum_{i=1}^n E\left[(\Delta_i^b)^\top H_i(\Delta_i^b)\right] \\
&= n^{1/2} n_0^{-1} n_0^{-2\varphi_2} E[(b_{n_0,j,i} - \tilde{b}_{n_0,i})^\top H_i(b_{n_0,j,i} - \tilde{b}_{n_0,i})] + n^{1/2} n_0^{-2\varphi_2} E[(\tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})] \\
&\overset{(1)}{=} o(n^{1/2-2\varphi_1}) + n^{1/2} n_0^{-2\varphi_2} E[(\tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})] \\
&\overset{(2)}{=} o(n^{1/2-2\varphi_1}) + n^{1/2} n_0^{-2\varphi_2} F_b + o(n^{1/2-2\varphi_2}) \\
&\overset{(3)}{=} n^{1/2} n_0^{-2\varphi_2} F_b + o(n^{1/2-2\varphi_1})
\end{aligned}$$

where (1) holds by (E.11) presented below and since $n/2 \leq n_0 \leq n$, (2) holds by definition of F_b in (3.4) and Assumption A.1, and (3) since $\varphi_1 \leq \varphi_2$ and $n/2 \leq n_0 \leq n$.

$$n_0^{-2\varphi_2} E[(b_{n_0,j,i} - \tilde{b}_{n_0,i})^\top H_i(b_{n_0,j,i} - \tilde{b}_{n_0,i})] = o(n^{1-2\varphi_1}) \ . \tag{E.11}$$

To verify (E.11) consider the following derivations,

$$
\begin{aligned}
|n_0^{-2\varphi_2}E[(b_{n_0,j,i}-\tilde{b}_{n_0,i})^\top H_i(b_{n_0,j,i}-\tilde{b}_{n_0,i})]| &\overset{(1)}{\leq} n_0^{-2\varphi_2}(2C_0)^{-1}E[|(b_{n_0,j,i}-\tilde{b}_{n_0,i})^\top \partial_\eta^2 m_i(b_{n_0,j,i}-\tilde{b}_{n_0,i})|] \\
&\overset{(2)}{\leq} (2C_0)^{-1}(p\tilde{C}_2)E[||n_0^{-\varphi_2}b_{n_0,j,i}-n_0^{-\varphi_2}\tilde{b}_{n_0,i}||^2] \\
&\overset{(3)}{\leq} 2(2C_0)^{-1}(p\tilde{C}_2)\left(E[||n_0^{-\varphi_2}b_{n_0,j,i}||^2]+E[||n_0^{-\varphi_2}\tilde{b}_{n_0,i}||^2]\right) \\
&\overset{(4)}{\leq} (p\tilde{C}_2/C_0)\left(n_0^{1-2\varphi_1}\tau_{n_0}+n_0^{-2\varphi_2}M_1^{1/2}\right) \\
&\overset{(5)}{=} o(n^{1-2\varphi_1}) \ ,
\end{aligned}
$$

where (1) holds by triangular inequality and part (a) of Assumption 3.1, (2) holds by definition of euclidean norm and since $||E[\partial_\eta^2 m_i \mid X_i]||_\infty \leq \tilde{C}_2 = C_2(1+M^{1/4}/C_0)$ due to part (e) of Assumption 3.1 and $|\theta_0| \leq M^{1/4}/C_0$ (which holds by definition of θ_0 and parts (a) and (c) of Assumption 3.1), (3) holds by standard properties of euclidean norm, (4) holds by parts (b.3) and (b.4) of Assumption 3.2 with $\tau_{n_0} = o(1)$, and (5) holds since $\varphi_1 \leq \varphi_2$ and $n/2 \leq n_0 \leq n$.

Part 2: Consider the following decomposition,

$$
\mathcal{T}_{n,K}^{nl} - E[\mathcal{T}_{n,K}^{nl}] = I_{l,l} + 2I_{l,b} + I_{b,b}
$$

where

$$
\begin{aligned}
I_{l,l} &= n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\left(\delta_{n_0,j_1,i}^\top H_i\delta_{n_0,j_2,i} - E[\delta_{n_0,j_1,i}^\top H_i\delta_{n_0,j_2,i}]\right) \\
I_{l,b} &= n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\left(\delta_{n_0,j_1,i}^\top H_i b_{n_0,j_2,i} - E[\delta_{n_0,j_1,i}^\top H_i b_{n_0,j_2,i}]\right) \\
I_{b,b} &= n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\left(b_{n_0,j_1,i}^\top H_i b_{n_0,j_2,i} - E[b_{n_0,j_1,i}^\top H_i b_{n_0,j_2,i}]\right)
\end{aligned}
$$

which implies

$$
\begin{aligned}
Var[\mathcal{T}_{n,K}^{nl}] &= E\left[(I_{l,l}+2I_{l,b}+I_{b,b})^2\right] \\
&= E[I_{l,l}^2] + E[I_{b,b}^2] + 4E[I_{l,b}^2] + 2E[I_{l,l}I_{b,b}] + 4E[(I_{l,l}+I_{b,b})I_{l,b}]
\end{aligned}
$$

In what follows, I show $E[I_{l,l}^2] = G_\delta(K^2-3K+3)(K-1)^{-2}n_0^{1-4\varphi_1} + o(n^{-\zeta})$, $E[I_{b,b}^2] = o(n^{-\zeta})$, and $E[I_{l,b}^2] = o(n^{-\zeta})$, which is sufficient to complete the proof of Part 2 since $n_0 = ((K-1)/K)n$ and by Cauchy-Schwartz it holds $E[I_{l,l}I_{b,b}] = o(n^{-\zeta})$, and $E[(I_{l,l}+I_{b,b})I_{l,b}] = o(n^{-\zeta})$.

Claim 1: $E[I_{l,l}^2] = G_\delta(K^2-3K+3)(K-1)^{-2}n_0^{1-4\varphi_1} + o(n^{-\zeta})$. Consider the following notation

$$\Gamma^{l,l}_{j_1,j_2,i} = \left(\delta^\top_{n_0,j_1,i}H_i\delta_{n_0,j_2,i} - E[\delta^\top_{n_0,j_1,i}H_i\delta_{n_0,j_2,i}]\right) \ .$$

Note that $E[\Gamma^{l,l}_{j_1,j_2,i}] = 0$ by construction, and $j_1 \neq j_2$ implies

$$E[\delta^\top_{n_0,j_1,i}H_i\delta_{n_0,j_2,i}] = 0 \quad \text{and} \quad \Gamma^{l,l}_{j_1,j_2,i} = \delta^\top_{n_0,j_1,i}H_i\delta_{n_0,j_2,i} \ .$$

Therefore, $E[\Gamma^{l,l}_{j_1,j_2,i} \mid W_i, W_{j_1}, X_{j_2}] = 0$ and $E[\Gamma^{l,l}_{j_1,j_2,i} \mid W_i, W_{j_2}, X_{j_1}] = 0$ when $j_1 \neq j_2$ due to part (a) of Assumption 3.2. Furthermore,

$$\left|E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]\right| \leq (p\tilde{C}_2/C_0)^2 n_0^{1-2\varphi_1}M_1 \ , \tag{E.12}$$

which follows by Cauchy-Schwartz, part (e) of Assumption 3.1, and part (b.2) of Assumption 3.2, with $\tilde{C}_2 = C_2(1+M^{1/4}/C_0)$.

Using the previous notation, $I_{l,l}$ can be written as follows

$$I_{l,l} = n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\Gamma^{l,l}_{j_1,j_2,i} \tag{E.13}$$

and $E[I_{l,l}^2]$ can be decompose in three terms

$$\begin{aligned}
E[I_{l,l}^2] &= E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\Gamma^{l,l}_{j_1,j_2,i}\right)^2\right] \\
&= E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j\notin\mathcal{I}_k}\Gamma^{l,l}_{j,j,i} + n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\Gamma^{l,l}_{j_1,j_2,i}I\{j_1\neq j_2\}\right)^2\right] \\
&= I_1 + I_2 + 2I_3 \ ,
\end{aligned}$$

where

$$I_1 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j\notin\mathcal{I}_k}\Gamma^{l,l}_{j,j,i}\right)^2\right]$$
$$I_2 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\Gamma^{l,l}_{j_1,j_2,i}I\{j_1\neq j_2\}\right)^2\right]$$
$$I_3 = E\left[\left(n^{-1/2}\sum_{k_1=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}} n_0^{-2\varphi_1-1}\sum_{j_1\notin\mathcal{I}_{k_1}}\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(n^{-1/2}\sum_{k_2=1}^{K}\sum_{i_2\in\mathcal{I}_{k_2}} n_0^{-2\varphi_1-1}\sum_{j_3\notin\mathcal{I}_{k_2}}\sum_{j_4\notin\mathcal{I}_{k_2}}\Gamma^{l,l}_{j_3,j_4,i_2}I\{j_3\neq j_4\}\right.\right.$$

In what follows, I show that $I_1 = o(n^{-\zeta})$, $I_2 = G_\delta(K^2-3K+3)(K-1)^{-2}n_0^{1-4\varphi_1} + o(n^{-\zeta})$, and $I_3 = o(n^{-\zeta})$, which is sufficient to complete the proof of Claim 1.

Claim 1.1 $I_1 = o(n^{-\zeta})$. Consider the following expansion

$$\begin{aligned} I_1 &= n^{-1}n_0^{-4\varphi_1-2}\sum_{k_1=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{j_1\notin\mathcal{I}_{k_1}}\sum_{k_2=1}^{K}\sum_{i_2\in\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_{k_2}} E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right] \\ &= n^{-1}n_0^{-4\varphi_1-2}\sum_{(i_1,i_2,j_1,i_2)\in\mathcal{E}} E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right] \ , \end{aligned}$$

where $\mathcal{E} = \{(i_1,i_2,j_1,j_2)\in[n]^4 : i_1\in\mathcal{I}_{k_1}, i_2\in\mathcal{I}_{k_2}, j_1\notin\mathcal{I}_{k_1}, j_2\notin\mathcal{I}_{k_2}, k_1\in[K], k_2\in[K]\}$, with $[n]$ denoting $\{1,\ldots,n\}$. Let $\mathcal{E}_4\subseteq[n]^4$ be the subset of indices with distinct entries. (e.g., $i_1\notin\{i_2,j_1,j_2\}$, $i_2\notin\{j_1,j_2\}$, $j_1\neq j_2$). Let $\mathcal{E}_{\leq 3}\subset[n]^4$ be the subset of indices with at most three distinct entries.

Now, take $(i_1,i_2,j_1,j_2)\in\mathcal{E}\cap\mathcal{E}_4$. It follows that

$$E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right] = E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\right]E\left[\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right] = 0 \ ,$$

due to independence and by definition of $\Gamma^{l,l}_{j_1,j_1,i_1}$.

Therefore,

$$|I_1| = \left|n^{-1}n_0^{-4\varphi_1-2}\sum_{(i_1,i_2,j_1,i_2)\in\mathcal{E}\setminus\mathcal{E}_4} E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right]\right|$$

$$\overset{(1)}{\leq} n^{-1} n_0^{-4\varphi_1 - 2} \sum_{(i_1,i_2,j_1,i_2)\in\mathcal{E}_{\leq 3}} \left| E\left[\left(\Gamma^{l,l}_{j_1,j_1,i_1}\right)\left(\Gamma^{l,l}_{j_2,j_2,i_2}\right)\right]\right|$$

$$\overset{(2)}{\leq} n^{-1} n_0^{-4\varphi_1 - 2} \sum_{(i_1,i_2,j_1,i_2)\in\mathcal{E}_{\leq 3}} (p\tilde{C}_2/C_0)^2 n_0^{1-2\varphi_1} M_1$$

$$\overset{(3)}{\leq} n^{-1} n_0^{-4\varphi_1 - 2} \times 3^4 n^3 \times (p\tilde{C}_2/C_0)^2 n_0^{1-2\varphi_1} M_1$$

$$\overset{(4)}{=} O(n^{1-6\varphi_1}) ,$$

where (1) holds by triangular inequality and since $\mathcal{E}\backslash\mathcal{E}_4 \subset \mathcal{E}_{\leq 3}$, (2) holds by (E.12), (3) holds since the number of elements of $\mathcal{E}_3$ is at most $3^4 n^3$ (for each 3-tuple $(a,b,c) \in [n]^3$ consider the functions from the positions $\{1,2,3,4\}$ into the possible values $\{a,b,c\}$, the number of all these functions is 3^4 and there number of 3-tuple is n^3), and (4) hold since $n/2 \leq n \leq n$. Therefore, I_1 is $O(n^{1-6\varphi_1})$, which is $o(n^{-\zeta})$ since $6\varphi_1 - 1 > 4\varphi_1 - 1 \geq \zeta$. This completes the proof of Claim 1.1.

Claim 1.2: $I_2 = G_\delta(K^2 - 3K + 3)(K-1)^{-2} n_0^{1-4\varphi_1} + o(n^{-\zeta})$. Consider the following expansion

$$I_2 = n^{-1} n_0^{-4\varphi_1-2} \sum_{k_1=1}^{K} \sum_{i_1\in\mathcal{I}_{k_1}} \sum_{j_1,j_2\notin\mathcal{I}_{k_1}} \sum_{k_2=1}^{K} \sum_{i_2\in\mathcal{I}_{k_2}} \sum_{j_3,j_4\notin\mathcal{I}_{k_2}} E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1} I\{j_1 \neq j_2\}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2} I\{j_3 \neq j_4\}\right)\right]$$

$$= n^{-1} n_0^{-4\varphi_1-2} \sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}} E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right] ,$$

where $\mathcal{E} = \{(i_1,i_2,j_1,j_2,j_3,j_4) \in [n]^6 : i_1 \in \mathcal{I}_{k_1}, i_2 \in \mathcal{I}_{k_2}; j_1, j_2 \notin \mathcal{I}_{k_1}; j_3, j_4 \notin \mathcal{I}_{k_1}; j_1 \neq j_2; j_3 \neq j_4; k_1, k_2 \in [K]\}$. Let $\mathcal{E}_6 \subseteq [n]^6$ be the subset of indices with distinct entries. Let $\mathcal{E}_5 \subseteq [n]^6$ be the subset of indices with exactly five distinct entries, meaning that two entries are identical while the remaining entries are distinct. Let $\mathcal{E}_4 \subseteq [n]^6$ be the subset of indices with exactly four distinct entries. Let $\mathcal{E}_{\leq 3} \subset [n]^6$ be the subset of indices with at most three distinct entries. Note that $[n]^6 = \mathcal{E}_{\leq 3} \cup \mathcal{E}_4 \cup \mathcal{E}_5 \cup \mathcal{E}_6$.

Now, take $(i_1,i_2,j_1,j_2,j_3,j_4) \in \mathcal{E} \cap \mathcal{E}_6$. It follows that

$$E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right] = 0 , \tag{E.14}$$

since $\Gamma^{l,l}_{j_1,j_2,i_1}$ and $\Gamma^{l,l}_{j_3,j_4,i_2}$ are independent zero mean random variables.

Now take $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$. Without loss of generality, assume that j_1 is different than all the other indices (otherwise, this statement holds with j_2 or j_3 or j_4). Then,

$$\begin{aligned} E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right] &\overset{(1)}{=} E\left[\left(\delta^\top_{n_0,j_1,i_1} H_{i_1} \delta_{n_0,j_2,i_1}\right)\left(\delta^\top_{n_0,j_3,i_2} H_{i_2} \delta_{n_0,j_4,i_2}\right)\right] \\ &\overset{(2)}{=} E\left[E\left[\delta^\top_{n_0,j_1,i_1} \mid W_{i_1}, W_{i_2}, W_{j_2}, W_{j_3}, W_{j_4}\right]\left(H_{i_1} \delta_{n_0,j_2,i_1}\right)\left(\delta^\top_{n_0,j_3,i_2} H_{i_2} \delta_{n_0,j_4,i_2}\right)\right] \\ &\overset{(3)}{=} 0 \end{aligned} \tag{E.15}$$

where (1) holds by definition of $\Gamma^{l,l}_{j_1,j_2,i_1}$ and $\Gamma^{l,l}_{j_3,j_4,i_2}$ since $j_1 \neq j_2$ and $j_3 \neq j_4$, (2) holds by LIE, and (3) holds by part (a) of Assumption 3.2. Note that this argument can be used whenever one j_s is different than all the other indices, for some $s \in \{1, 2, 3, 4\}$.

Now take $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_4$. Suppose $\{a, b, c, d\}$ are four different indices, then there are two possible distributions for the 6-tuples: (i) two pairs, e.g., (a, a, b, b, c, d), or (ii) one triple, e.g., (a, a, a, b, c, d). Notice that for 6-tuples in (ii), there exists one j_s different than all the other indices, for some $s \in \{1, 2, 3, 4\}$. In this case, $E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]$ equals zero due to the argument described above. Therefore, in what follows, I consider only 6-tuples in (i), specifically, the cases where j_s appears in a pair for all $s = 1, 2, 3, 4$.

- Case 1: $j_1 = j_3$, $j_2 = j_4$, and $i_1 \neq i_2$. Then,

$$E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right] = E\left[\left(\delta^\top_{n_0,j_1,i_1} H_{i_1} \delta_{n_0,j_2,i_1}\right)\left(\delta^\top_{n_0,j_1,i_2} H_{i_2} \delta_{n_0,j_2,i_2}\right)\right] \tag{E.16}$$

  To compute the number of indices $(i_1, i_2, j_1, j_2, j_1, j_2) \in \mathcal{E}$ in this case, recall that $i_1 \in \mathcal{I}_{k_1}$ and $i_2 \in \mathcal{I}_{k_2}$, therefore, there are two situations (i) $k_1 = k_2$ or (ii) $k_1 \neq k_2$. For the first situation, i_1 can take n values, i_2 can take $n_k - 1$ values (since it is different than i_1 but is in the same fold $\mathcal{I}_k$), and j_1 and j_2 can take n_0 and $n_0 - 1$ values (since they are different but not in $\mathcal{I}_k$). That is $n(n_k - 1)n_0(n_0 - 1)$ combinations. For the second situation, i_1 can take n values, i_2 can take n_0 values, then j_1 and j_2 take values in all the data except into the two folds that contain i_1 and i_2 (since $j_1, j_2 \notin \mathcal{I}_{k_1}$ and $j_1 = j_3, j_2 = j_4 \notin \mathcal{I}_{k_2}$), that is $(n_0 - n_k)(n_0 - n_k - 1)$. That is $n(n - n_k)(n_0 - n_k)(n_0 - n_k - 1)$ combinations.

Therefore, the total number of indices is equal to

$$nn_0^3\left(\frac{K^2-3K+3}{(K-1)^2}-2n_0^{-1}+n_0^{-2}\right) \tag{E.17}$$

- Case 2: $j_1=j_4$, $j_2=j_3$, and $i_1\neq i_2$. Then,

$$E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]=E\left[\left(\delta^\top_{n_0,j_1,i_1}H_{i_1}\delta_{n_0,j_2,i_1}\right)\left(\delta^\top_{n_0,j_1,i_2}H_{i_2}\delta_{n_0,j_2,i_2}\right)\right]\ .$$

  The number of indices $(i_1,i_2,j_1,j_2,j_1,j_2)\in\mathcal{E}$ in this case is exactly the same as in the previous case, which is presented in (E.17).

Finally, note that

$$\left|\sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cup\mathcal{E}_{\leq 3}}E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]\right|\leq 3^6n^3(p\tilde{C}_2/C_0)n_0^{1-2\varphi_1}M_1\ , \tag{E.18}$$

which follows by triangular inequality, (E.12), and by using that the number of elements in $\mathcal{E}_{\leq 3}$ is lower or equal to 3^6n^3 (for each 3-tuple $(a,b,c)\in[n]^3$, consider the functions from the positions $\{1,2,3,4,5,6\}$ into the possible values $\{a,b,c\}$, the number of all these functions is 3^4, while the number of 3-tuple is n^3).

In what follows, I use the preliminary findings to calculate I_2 up to an error of size $o(n^{-\zeta})$,

$$\begin{aligned}
I_2 &= n^{-1}n_0^{-4\varphi_1-2}\sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}}E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]\\
&\overset{(1)}{=} n^{-1}n_0^{-4\varphi_1-2}\sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cap\mathcal{E}_4}E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]\\
&\quad + n^{-1}n_0^{-4\varphi_1-2}\sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cup\mathcal{E}_{\leq 3}}E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]\\
&\overset{(2)}{=} n^{-1}n_0^{-4\varphi_1-2}\sum_{k_1=1}^{K}\sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cap\mathcal{E}_4}E\left[\left(\Gamma^{l,l}_{j_1,j_2,i_1}\right)\left(\Gamma^{l,l}_{j_3,j_4,i_2}\right)\right]+O(n^{1-6\varphi_1})\\
&\overset{(3)}{=} n_0^{1-4\varphi_1}\left(\frac{K^2-3K+3}{(K-1)^2}-2n_0^{-1}+n_0^{-2}\right)2E\left[\left(\delta^\top_{n_0,j_1,i_1}H_{i_1}\delta_{n_0,j_2,i_1}\right)\left(\delta^\top_{n_0,j_1,i_2}H_{i_2}\delta_{n_0,j_2,i_2}\right)\right]+O(n^{1-6\varphi_1})\\
&\overset{(4)}{=} n_0^{1-4\varphi_1}\left(\frac{K^2-3K+3}{(K-1)^2}\right)G_\delta+o(n^{-\zeta})\ ,
\end{aligned}$$

where (1) holds by the derivations in (E.14) and (E.15), (2) holds by (E.18), (3) holds by (E.16) that computes the expected value and (E.17) that calculates the number of indices to consider, and (4) holds by definition of G_δ in (3.2), Assumption A.1, $n/2 \le n_0 \le n$, and since $6\varphi_1 - 1 > \zeta$. This completes the proof of Claim 1.2.

Claim 1.3: $I_3 = o(n^{-\zeta})$. First, claim 1.1 implies I_1 is $o(n^{-\zeta})$. Second, claim 1.2 implies I_2 is $O(n^{-\zeta})$ since $4\varphi_1 - 1 \ge \zeta$. Finally, then I_3 is $o(n^{-\zeta})$ due to Cauchy-Schwartz ($|I_3| \le |I_1|^{1/2}|I_2|^{1/2}$). This completes the proof of Claim 1.3.

Claim 2: $E[I_{b,b}^2] = o(n^{-\zeta})$. Consider the following notation,

$$\Gamma^{b,b}_{j_1,j_2,i} = b^\top_{n_0,j_1,i} H_i b_{n_0,j_2,i} - E[b^\top_{n_0,j_1,i} H_i b_{n_0,j_2,i}] \ ,$$

where by construction $E[\Gamma^{b,b}_{j_1,j_2,i}] = 0$. Denote $\tilde{b}_{n_0,i} = E[b_{n_0,j,i} \mid X_i]$. Note that if $j_1 \ne j_2$, then

$$\Gamma^{b,b}_{j_1,j_2,i} = b^\top_{n_0,j_1,i} H_i b_{n_0,j_2,i} - E[\tilde{b}^\top_{n_0,i} H_i \tilde{b}_{n_0,i}] \ ,$$

Furthermore,

$$n_0^{-4\varphi_2} E\left[|\Gamma^{b,b}_{j_1,j_1,i}|^2\right] \le (p\tilde{C}_2/C_0) n_0^{3(1-2\varphi_1)} \tau_{n_0} \ , \tag{E.19}$$

which follows by C-S, part (e) of Assumption 3.1, and part (b.4) of Assumption 3.2, with $\tilde{C}_2 = C_2(1 + M^{1/4}/C_0)$. And, if $j_1 \ne j_2$,

$$n_0^{-4\varphi_2} E\left[|\Gamma^{b,b}_{j_1,j_2,i}|^2\right] \le (p\tilde{C}_2/C_0) n_0^{2(1-2\varphi_1)} \tau_{n_0} \ , \tag{E.20}$$

which holds by C-S, part (e) of Assumption 3.1, and part (b.1) of Assumption 3.2.

The previous notation can be used to rewrite $I_{b,b}$ as follows

$$\begin{aligned} E[I_{b,b}^2] &= E\left[\left(n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2} \sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} \Gamma^{b,b}_{j_1,j_2,i}\right)^2\right] \\ &= E\left[\left(n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2} \sum_{j\notin\mathcal{I}_k} \Gamma^{b,b}_{j,j,i} + n^{-1/2}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2} \sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} \Gamma^{b,b}_{j_1,j_2,i} I\{j_1 \ne j_2\}\right)^2\right] \end{aligned}$$

$$= I_1 + I_2 + 2I_3$$

where

$$I_1 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j\notin\mathcal{I}_k}\Gamma^{b,b}_{j,j,i}\right)^2\right]$$

$$I_2 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k}\Gamma^{b,b}_{j_1,j_2,i}I\{j_1\neq j_2\}\right)^2\right]$$

$$I_3 = E\left[\left(n^{-1/2}\sum_{k_1=1}^{K}\sum_{i\in\mathcal{I}_{k_1}} n_0^{-2\varphi_2-2}\sum_{j_1\notin\mathcal{I}_{k_1}}\Gamma^{b,b}_{j_1,j_1,i_1}\right)\left(n^{-1/2}\sum_{k_2=1}^{K}\sum_{i_2\in\mathcal{I}_{k_2}} n_0^{-2\varphi_2-2}\sum_{j_3\notin\mathcal{I}_{k_2}}\sum_{j_4\notin\mathcal{I}_{k_2}}\Gamma^{b,b}_{j_3,j_4,i}I\{j_3\neq j_4\}\right)\right.$$

In what follows, I show that $I_1 = o(n^{-\zeta})$, $I_2 = o(n^{-\zeta})$, and $I_3 = o(n^{-\zeta})$, which is sufficient to complete the proof of Claim 2.

Claim 2.1: $I_1 = o(n^{-\zeta})$. Consider the following expansion,

$$\begin{aligned} I_1 &= n^{-1}n_0^{-4\varphi_2-4}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{i_2\in\mathcal{I}_{k_2}}\sum_{j_1\notin\mathcal{I}_{k_1}}\sum_{j_2\notin\mathcal{I}_{k_2}} E\left[\Gamma^{b,b}_{j_1,j_1,i_1}\Gamma^{b,b}_{j_2,j_2,i_2}\right] \\ &= n^{-1}n_0^{-4\varphi_2-4}\sum_{(i_1,i_2,j_1,j_2)\in\mathcal{E}} E\left[\Gamma^{b,b}_{j_1,j_1,i_1}\Gamma^{b,b}_{j_2,j_2,i_2}\right] ,\end{aligned}$$

where $\mathcal{E} = \{(i_1,i_2,j_1,j_2)\in[n]^4 : i_1\in\mathcal{I}_{k_1}, i_2\in\mathcal{I}_{k_2}, j_1\notin\mathcal{I}_{k_1}, j_2\notin\mathcal{I}_{k_2}, k_1\in[K], k_2\in[K]\}$, with $[n]$ denoting $\{1,\ldots,n\}$. Let $\mathcal{E}_4\subseteq[n]^4$ be the subset of indices with distinct entries. Let $\mathcal{E}_{\leq 3}\subset[n]^4$ be the subset of indices with at most three distinct entries.

Now, take $(i_1,i_2,j_1,j_2)\in\mathcal{E}\cap\mathcal{E}_4$. It follows that

$$E\left[\Gamma^{b,b}_{j_1,j_1,i_1}\Gamma^{b,b}_{j_2,j_2,i_2}\right] = 0 ,$$

since $\Gamma^{b,b}_{j_1,j_1,i_1}$ and $\Gamma^{b,b}_{j_2,j_2,i_2}$ are zero mean independent random variables.

Now, take $(i_1,i_2,j_1,j_2)\in\mathcal{E}\cap\mathcal{E}_{\leq 3}$. Consider the following derivation,

$$n_0^{-4\varphi_2}\left|E\left[\Gamma^{b,b}_{j_1,j_1,i_1}\Gamma^{b,b}_{j_2,j_2,i_2}\right]\right| \leq (p\tilde{C}_2/C_0)n_0^{3(1-2\varphi_1)}\tau_{n_0} ,$$

which follows by C-S and (E.19).

Therefore,

$$\left| n^{-1} n_0^{-4\varphi_2-4} \sum_{(i_1,i_2,j_1,j_2)\in\mathcal{E}\cap\mathcal{E}_{\leq 3}} E\left[\Gamma^{b,b}_{j_1,j_1,i_1}\Gamma^{b,b}_{j_2,j_2,i_2}\right]\right| \leq n^{-1} n_0^{-4} 3^4 n^3 (p\tilde{C}_2/C_0) n_0^{3(1-2\varphi_1)} \tau_{n_0} \ ,$$

which uses that $\mathcal{E}_{\leq 3}$ has at most $3^4 n^3$ elements (as in the proof of claim 1.1).

Using these two preliminary results, it follows that

$$I_1 = o(n^{1-6\varphi_1}) \ ,$$

since $\tilde{\tau}_{n_0} = o(1)$ and $n/2 \leq n_0 \leq n$. This completes the proof of Claim 2.2 since $6\varphi_1 - 1 > \zeta$.

Claim 2.2: $I_2 = o(n^{-\zeta})$. Consider the following expansion,

$$\begin{aligned} I_2 &= n^{-1} n_0^{-4\varphi_2-4} \sum_{k_1=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{k_2=1}^{K}\sum_{i_2\in\mathcal{I}_{k_2}}\sum_{j_1,j_2\notin\mathcal{I}_{k_1}}\sum_{j_3,j_4\notin\mathcal{I}_{k_2}} E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,i_2}\right] I\{j_1\neq j_2\} I\{j_3\neq j_4\} \\ &= n^{-1} n_0^{-4\varphi_2-4} \sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}} E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,i_2}\right] \ , \end{aligned}$$

where $\mathcal{E} = \{(i_1,i_2,j_1,j_2,j_3,j_4)\in[n]^6 : i_1\in\mathcal{I}_{k_1}, i_2\in\mathcal{I}_{k_2}; j_1,j_2\notin\mathcal{I}_{k_1}; j_3,j_4\notin\mathcal{I}_{k_1}; j_1\neq j_2; j_3\neq j_4; k_1,k_2\in[K]\}$. Let $\mathcal{E}_6\subseteq[n]^6$ be the subset of indices with distinct entries. Let $\mathcal{E}_5\subseteq[n]^6$ be the subset of indices with exactly five distinct entries, meaning that two entries are identical while the remaining entries are distinct. Let $\mathcal{E}_4\subseteq[n]^6$ be the subset of indices with exactly four distinct entries. Let $\mathcal{E}_{\leq 3}\subset[n]^6$ be the subset of indices with at most three distinct entries. Note that $[n]^6 = \mathcal{E}_{\leq 3}\cup\mathcal{E}_4\cup\mathcal{E}_5\cup\mathcal{E}_6$.

Note that for all $(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}_6$, it follows $E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,i_2}\right] = 0$ since $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ are independent zero mean random variables.

Now, take $(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cap\mathcal{E}_5$. There are three possible cases:

- Case 1: $j_s = j_r$ for some $s\in\{1,2\}$ and $r\in\{3,4\}$. Since all the sub-cases are similar, without loss of generality, take $(s, r) = (1, 3)$. It follows that

$$
\begin{aligned}
n_0^{-2\varphi_2} \left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,i_2}\right]\right| &\overset{(1)}{\le} \left| E\left[(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1}\tilde{b}_{n_0,i_1}(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_2}\tilde{b}_{n_0,i_2}\right]\right| \\
&\quad + n_0^{-2\varphi_2}\left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,i_2} H_{i_2}\tilde{b}_{n_0,i_2}\right]\right| \\
&\overset{(2)}{\le} \tilde{C} E\left[|(n_0^{-\varphi_2} b_{n_0,j_1,i_1})|^2 |\tilde{b}_{n_0,i_1}|^2\right] + n_0^{-2\varphi_2}\tilde{C}M_1 \\
&\overset{(3)}{\le} \tilde{C} E\left[E\left[|(n_0^{-\varphi_2} b_{n_0,j_1,i_1})|^2 \mid X_{i_1}\right]^2\right]^{1/2} E\left[|\tilde{b}_{n_0,i_1}|^4\right]^{1/2} + n_0^{-2\varphi_2}\tilde{C}M_1 \\
&\overset{(4)}{\le} \tilde{C} n_0^{1-2\varphi_1}\tau_{n_0} M_1^{1/2} + n_0^{-2\varphi_2}\tilde{C}M_1
\end{aligned}
$$

where (1) holds by triangular inequality, LIE and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $j_1 = j_3$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$, (2) holds by part (e) of Assumption 3.1 with $\tilde{C}$ as a function of (C_0, C_2, M, p) and part (b.3) of Assumption 3.2 with C-S, (3) holds by LIE and C-S, and (4) holds by parts (b.1) and (b.3) of Assumption 3.2.

- Case 2: $j_s = i_2$ for some $s \in \{1, 2\}$ or $j_r = i_1$ for some $r \in \{3, 4\}$. Since all the sub-cases are similar, take $s = 1$. It follows that

$$
\begin{aligned}
n_0^{-\varphi_2} \left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,j_1}\right]\right| &\overset{(1)}{\le} \left| E\left[(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1}\tilde{b}_{n_0,i_1}\tilde{b}^\top_{n_0,j_1} H_{j_1}\tilde{b}_{n_0,j_1}\right]\right| \\
&\quad + n_0^{-\varphi_2}\left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,j_1} H_{j_1}\tilde{b}_{n_0,j_1}\right]\right| \\
&\overset{(2)}{\le} \tilde{C} E\left[|n_0^{-\varphi_2} b_{n_0,j_1,i_1}|^4\right]^{1/4} E\left[|\tilde{b}_{n_0,j_1}|^4\right]^{3/4} + n_0^{-\varphi_2}\tilde{C}M_1 \\
&\overset{(3)}{\le} \tilde{C} n_0^{3(1-2\varphi_1)/4}\tau_{n_0}^{1/4} M_1^{3/4} + n_0^{-\varphi_2}\tilde{C}M_1
\end{aligned}
$$

where (1) holds by triangular inequality, LIE, and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $j_1 = i_2$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$; (2) holds by part (e) of Assumption 3.1 with $\tilde{C}$ as a function of (C_0, C_2, M, p), C-S, and part (b.3) of Assumption 3.2 with C-S; (3) holds by parts (b.3) and (b.4) of Assumption 3.2.

- Case 3: $i_1 = i_2$. It follows that

$$
\begin{aligned}
\left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,j_1}\right]\right| &\overset{(1)}{\le} \left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right]\right| \\
&\quad + \left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right]\right|
\end{aligned}
$$

$$\overset{(2)}{\leq} 2\tilde{C}M_1$$

where (1) holds by triangular inequality, LIE, and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $i_1 = i_2$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$; and (2) holds by part (e) of Assumption 3.1 with $\tilde{C}$ as a function of (C_0, C_2, M, p), C-S, and part (b.3) of Assumption 3.2 with C-S.

Therefore, for the indices on $\mathcal{E} \cap \mathcal{E}_5$, it follows that

$$\begin{aligned} n^{-1} n_0^{-4\varphi_2 - 4} \sum_{(i_1,i_2,j_1,j_2,j_3,j_4) \in \mathcal{E} \cap \mathcal{E}_5} E\left[\Gamma^{b,b}_{j_1,j_2,i_1} \Gamma^{b,b}_{j_3,j_4,i_2}\right] &\overset{(1)}{=} o(n^{1-2\varphi_1-2\varphi_2}) + o(n^{3/4-3\varphi_1/2-3\varphi_2}) + O(n^{-4\varphi_2}) \\ &\overset{(2)}{=} o(n^{-\zeta}) \end{aligned} \tag{E.21}$$

where (1) holds since the number of elements of $\mathcal{E} \cap \mathcal{E}_5$ is lower than $5^6 n^5$ and the preliminary findings in cases 1, 2, and 3, and (2) holds since $2\varphi_1 + 2\varphi_2 - 1 > \varphi_1 + \varphi_2 - 1/2 \geq \zeta$, $3\varphi_1/2 + 3\varphi_2 - 3/4 > \varphi_1 + \varphi_2 - 1/2 \geq \zeta$, and $4\varphi_2 > 4\varphi_1 - 1 \geq \zeta$.

Now, take $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_4$. There are three cases.

- Case 1: $j_1 = j_s$ and $j_2 = j_r$ for $\{r, s\} = \{3, 4\}$. Without loss of generality, consider $(s, r) = (3, 4)$. It follows

$$\begin{aligned} & n_0^{-4\varphi_2} \left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1} \Gamma^{b,b}_{j_1,j_2,i_2}\right]\right| \\ &\overset{(1)}{\leq} \left| E\left[(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1} (n_0^{-\varphi_2} b_{n_0,j_2,i_1})(n_0^{-\varphi_2} b_{n_0,j_1,i_2})^\top H_{i_2} (n_0^{-\varphi_2} b_{n_0,j_2,i_2})\right]\right| \\ &\quad + n_0^{-4\varphi_2} \left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1} \tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,i_2} H_{i_2} \tilde{b}_{n_0,i_2}\right]\right| \\ &\overset{(2)}{\leq} \tilde{C} E\left[|n_0^{-\varphi_2} b_{n_0,j_1,i_1})||n_0^{-\varphi_2} b_{n_0,j_2,i_1}||n_0^{-\varphi_2} b_{n_0,j_1,i_2}||n_0^{-\varphi_2} b_{n_0,j_2,i_2}|\right] + n_0^{-4\varphi_2} \tilde{C} M_1 \\ &\overset{(3)}{=} \tilde{C} E\left[E\left[|n_0^{-\varphi_2} b_{n_0,j_1,i_1})||n_0^{-\varphi_2} b_{n_0,j_1,i_2}| \mid X_{i_1}, X_{i_2}\right] E\left[|n_0^{-\varphi_2} b_{n_0,j_2,i_1}||n_0^{-\varphi_2} b_{n_0,j_2,i_2}| \mid X_{i_1}, X_{i_2}\right]\right] \\ &\quad + n_0^{-4\varphi_2} \tilde{C} M_1 \\ &\overset{(4)}{\leq} \tilde{C} E\left[E\left[|n_0^{-\varphi_2} b_{n_0,j_1,i_1})|^2 \mid X_{i_1}\right] E\left[|n_0^{-\varphi_2} b_{n_0,j_2,i_2}|^2 \mid X_{i_2}\right]\right] + n_0^{-4\varphi_2} \tilde{C} M_1 \\ &\overset{(5)}{\leq} \tilde{C} E\left[E\left[|n_0^{-\varphi_2} b_{n_0,j_1,i_1})|^2 \mid X_{i_1}\right]^2\right]^{1/2} E\left[E\left[|n_0^{-\varphi_2} b_{n_0,j_2,i_2}|^2 \mid X_{i_2}\right]^2\right]^{1/2} + n_0^{-4\varphi_2} \tilde{C} M_1 \\ &\overset{(6)}{\leq} \tilde{C} n_0^{2(1-2\varphi_1)} \tau_{n_0} + n_0^{-4\varphi_2} \tilde{C} M_1 \end{aligned}$$

where (1) holds by triangular inequality, LIE, and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $j_1 = j_3$, $j_2 = j_4$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$; (2) holds by part (e) of Assumption 3.1 with $\tilde{C}$ as a function of (C_0, C_2, M, p), C-S, and part (b.3) of Assumption 3.2 with C-S; (3) holds by LIE; (4) and (5) holds by C-S; and (6) holds by part (b.1) of Assumption 3.2.

- Case 2: $j_1 = j_s$ for $s \in \{3, 4\}$ and $j_2 = i_2$. Without loss of generality, $s = 3$. It follows

$$\begin{aligned} n_0^{-3\varphi_2} \left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_1,j_2,i_2}\right]\right| &\overset{(1)}{\leq} \left| E\left[(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1}(n_0^{-\varphi_2} b_{n_0,j_2,i_1})(n_0^{-\varphi_2} b_{n_0,j_1,j_2})^\top H_{j_2}\tilde{b}_{n_0,j_2})\right]\right| \\ &\quad + n_0^{-3\varphi_2}\left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,j_2} H_{j_2}\tilde{b}_{n_0,j_2}\right]\right| \\ &\overset{(3)}{\leq} \tilde{C} n_0^{9(1-2\varphi_1)/4}\tau_{n_0}^{3/4} M_1^{1/4} + \tilde{C} n_0^{-3\varphi_2} M_1 \end{aligned}$$

  where (1) holds by triangular inequality, LIE, and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $j_1 = j_3$, $j_2 = i_2$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$; and (2) holds by C-S and parts (b.3) and (b.4) of Assumption 3.2.

- Case 3: $j_1 = j_s$ for $s \in \{3, 4\}$ and $i_1 = i_2$. Without loss of generality, consider $s = 3$. It follows that

$$\begin{aligned} n_0^{-2\varphi_2} \left| E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_1,j_4,i_1}\right]\right| &\overset{(1)}{\leq} \left| E\left[(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1}\tilde{b}_{n_0,i_1}(n_0^{-\varphi_2} b_{n_0,j_1,i_1})^\top H_{i_1}\tilde{b}_{n_0,i_1}\right]\right| \\ &\quad + n_0^{-2\varphi_2}\left| E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right] E\left[\tilde{b}^\top_{n_0,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\right]\right| \\ &\overset{(2)}{\leq} \tilde{C} n_0^{1-2\varphi_1}\tau_{n_0} M_1^{1/2} + n_0^{-2\varphi_2}\tilde{C} M_1 \end{aligned}$$

  where (1) holds by triangular inequality, LIE and definition of $\Gamma^{b,b}_{j_1,j_2,i_1}$ and $\Gamma^{b,b}_{j_3,j_4,i_2}$ when $j_1 = j_3$ and $i_1 = i_2$ and $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_4$; and (2) holds by the same derivations presented in Case 1 when $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_5$; therefore, it is omitted.

Therefore, for the indices on $\mathcal{E} \cap \mathcal{E}_4$, it follows that

$$\begin{aligned} n^{-1} n_0^{-4\varphi_2 - 4} \sum_{(i_1,i_2,j_1,j_2,j_3,j_4)\in\mathcal{E}\cap\mathcal{E}_4} E\left[\Gamma^{b,b}_{j_1,j_2,i_1}\Gamma^{b,b}_{j_3,j_4,i_2}\right] &\overset{(1)}{=} o(n^{1-4\varphi_1}) + o(n^{5/4-9\varphi_1/2-\varphi_2}) + O(n^{-2\varphi_1-2\varphi_2}) \\ &\overset{(2)}{=} o(n^{-\zeta}) \end{aligned} \tag{E.22}$$

where (1) holds since the number of elements of $\mathcal{E} \cap \mathcal{E}_4$ is lower than $4^6 n^4$ and the preliminary findings in cases 1, 2, and 3; and (2) holds since $4\varphi_1 - 1 \geq \zeta$, $9\varphi_1/2 + \varphi_2 - 5/4 > \varphi_1 + \varphi_2 - 1/2 \geq \zeta$, and $2\varphi_1 + 2\varphi_2 > \varphi_1 + \varphi_2 - 1/2 \geq \zeta$.

Now, take $(i_1, i_2, j_1, j_2, j_3, j_4) \in \mathcal{E} \cap \mathcal{E}_{\leq 3}$. Similar to the proof of Claim 2.1 but using (E.20) instead of (E.19), it follows that

$$\left| n^{-1} n_0^{-4\varphi_2 - 4} \sum_{(i_1, i_2, j_1, j_2) \in \mathcal{E} \cap \mathcal{E}_{\leq 3}} E\left[\Gamma^{b,b}_{j_1, j_1, i_1} \Gamma^{b,b}_{j_2, j_2, i_2}\right] \right| = o(n^{-\zeta}) \tag{E.23}$$

Finally, using (E.21), (E.22), and (E.23), it follows that $I_2 = o(n^{-\zeta})$. This completes the proof of Claim 2.2.

Claim 2.3: $I_3 = o(n^{-\zeta})$. This result is a consequence of C-S and Claims 2.1 and 2.2.

Claim 3: $E[I^2_{l,b}] = o(n^{-\zeta})$. Consider the following notation,

$$\Gamma^{l,b}_{j_1, j_2, i} = \delta^{\top}_{n_0, j_1, i} H_i b_{n_0, j_2, i}$$

where it holds $E[\Gamma^{l,b}_{j_1, j_2, i}] = 0$ due to part (a) of Assumption 3.2.

Furthermore,

$$n_0^{-2\varphi_2} E\left[|\Gamma^{l,b}_{j_1, j_2, i}|^2\right] \leq \tilde{C} n_0^{2(1-2\varphi_1)} \tau^{1/2}_{n_0} M_1^{1/2} \ , \tag{E.24}$$

which follows by C-S, part (e) of Assumption 3.1, and parts (b.2) and (b.4) of Assumption 3.2, with $\tilde{C}$ function of (C_2, M, C_0, p).

The previous notation can be used to rewrite $E[I^2_{l,b}]$ as follows

$$\begin{aligned}
&= E\left[\left(n^{-1/2} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-\varphi_1 - \varphi_2 - 3/2} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} \Gamma^{l,b}_{j_1, j_2, i}\right)^2\right] \\
&= E\left[\left(n^{-1/2} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-\varphi_1 - \varphi_2 - 3/2} \sum_{j \notin \mathcal{I}_k} \Gamma^{l,b}_{j,i} + n^{-1/2} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-\varphi_1 - \varphi_2 - 3/2} \sum_{j_1, j_2 \notin \mathcal{I}_k} \Gamma^{l,b}_{j_1, j_2, i} I\{j_1 \neq j_2\}\right)^2\right] \\
&= I_1 + I_2 + 2I_3
\end{aligned}$$

where

$$I_1 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2}\sum_{j\notin\mathcal{I}_k}\Gamma^{l,b}_{j,j,i}\right)^2\right]$$

$$I_2 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2}\sum_{j_1,j_2\notin\mathcal{I}_k}\Gamma^{l,b}_{j_1,j_2,i}I\{j_1\neq j_2\}\right)^2\right]$$

$$I_3 = E\left[\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2}\sum_{j\notin\mathcal{I}_k}\Gamma^{l,b}_{j,j,i}\right)\left(n^{-1/2}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2}\sum_{j_1,j_2\notin\mathcal{I}_k}\Gamma^{l,b}_{j_1,j_2,i}I\{j_1\neq j_2\}\right)\right]$$

In what follows, I show that $I_1 = o(n^{-\zeta})$, $I_2 = o(n^{-\zeta})$, and $I_3 = o(n^{-\zeta})$, which is sufficient to complete the proof of Claim 2.

Claim 3.1: $I_1 = o(n^{-\zeta})$. Consider the following expansion,

$$I_1 = n^{-1}n_0^{-2\varphi_1-2\varphi_2-3}\sum_{k_1,k_2=1}^{K}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{i_1\in\mathcal{I}_{k_1}}\sum_{j_1\notin\mathcal{I}_{k_1}}\sum_{j_2\notin\mathcal{I}_{k_2}} E[\Gamma^{l,b}_{j_1,j_1,i_1}\Gamma^{l,b}_{j_2,j_2,i_2}]\ .$$

Now, $E[\Gamma^{l,b}_{j_1,j_1,i_1}\Gamma^{l,b}_{j_2,j_2,i_2}]$ is calculated under the two possible cases based on the indices (i_1, j_1, i_2, j_2).

- Case 1: (i_1, j_1) and (i_2, j_2) have no element in common. Then $E[\Gamma^{l,b}_{j_1,j_1,i_1}\Gamma^{l,b}_{j_2,j_2,i_2}]$ is zero since $\Gamma^{l,b}_{j_1,j_1,i_1}$ and $\Gamma^{l,b}_{j_2,j_2,i_2}$ are independent zero mean random variables.
- Case 2: (i_1, j_1) and (i_2, j_2) have at least one element in common. In this case, there are at most 3^4n^3 possible indices. Moreover, due to (E.24) and C-S, it follows

$$|E[\Gamma^{l,b}_{j_1,j_1,i_1}\Gamma^{l,b}_{j_2,j_2,i_2}]| \leq \tilde{C}n_0^{3(1-2\varphi_1)/2}\tau_{n_0}^{1/2}n_0^{(1-2\varphi_1)}M_1^{1/2}\ .$$

Therefore,

$$\begin{aligned}|I_1| &\leq n^{-1}n_0^{-2\varphi_1-3}3^4n^3\tilde{C}n_0^{2(1-2\varphi_1)}\tau_{n_0}^{1/2}M_1^{1/2}\\ &= o(n^{1-6\varphi_1})\ ,\end{aligned}$$

which is sufficient to conclude that I_1 is $o(n^{-\zeta})$ since $6\varphi_1 - 1 > 4\varphi_1 - 1 \geq \zeta$. This completes the proof of Claim 3.1.

Claim 3.2: $I_2 = o(n^{-\zeta})$. Consider the following expansion,

$$I_2 = n^{-1} \sum_{k_1,k_2=1}^{K} \sum_{i_1 \in \mathcal{I}_{k_1}} \sum_{i_2 \in \mathcal{I}_{k_2}} n_0^{-2\varphi_1-2\varphi_2-3} \sum_{j_1,j_2 \notin \mathcal{I}_{k_1}} \sum_{j_3,j_4 \notin \mathcal{I}_{k_2}} E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,i_2}\right] I\{j_1 \neq j_2\} I\{j_3 \neq j_4\}$$

Now, $E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,i_2}\right]$ is calculated under four possible cases based on the indices.

- Case 1: all indices are different. Then , $E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,i_2}\right]$ equals zero since $\Gamma^{l,b}_{j_1,j_1,i_1}$ and $\Gamma^{l,b}_{j_2,j_2,i_2}$ are independent zero mean random variables.
- Case 2: there are exactly five different indices. Then, consider the four different sub-cases:
  - $j_1 = j_3$, then

$$\begin{aligned}\left|E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_1,j_4,i_2}\right]\right| &= \left|E\left[\delta^{\top}_{n_0,j_1,i_1} H_{i_1}\tilde{b}_{n_0,i_1}\delta^{\top}_{n_0,j_1,i_2} H_{i_2}\tilde{b}_{n_0,i_2}\right]\right| \\ &\leq \tilde{C}E\left[|\delta_{n_0,j_1,i_1}||\tilde{b}_{n_0,i_1}||\delta_{n_0,j_1,i_2}||\tilde{b}_{n_0,i_2}|\right] \\ &\leq \tilde{C}E\left[E\left[|\delta_{n_0,j_1,i_1}|^2 \mid X_{i_1}\right]|\tilde{b}_{n_0,i_1}|^2\right] \\ &= O(1)\end{aligned}$$

    which holds by C-S, LIE, and part (b.1) and (b.3) of Assumption 3.2. Since there are at most $5^6 n^5$ terms, it follows these terms contributed to I_2 with $O(n^{1-2\varphi_1-2\varphi_2})$ which is larger than $o(n^{-\zeta})$ since $2\varphi_1 + 2\varphi_1 - 1 > \varphi_1 + \varphi_2 - 1/2 \geq \zeta$.
  - $j_1 = j_4$, then

$$\begin{aligned}E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_1,i_2}\right] &= E\left[\delta^{\top}_{n_0,j_1,i_1} H_{i_1} b_{n_0,j_2,i_1}\delta^{\top}_{n_0,j_3,i_2} H_{i_2} b_{n_0,j_1,i_2}\right] \\ &= E[E[\delta^{\top}_{n_0,j_1,i_1} \mid X_{j_1}, W_{i_1}, W_{i_2}, W_{j_2}, W_{j_3}] H_{i_1} b_{n_0,j_2,i_1}\delta^{\top}_{n_0,j_3,i_2} H_{i_2} b_{n_0,j_1,i_2}] \\ &= 0\ ,\end{aligned}$$

    which holds due to part (a) of Assumption 3.2.

– $j_1 = i_2$, then

$$\begin{aligned}E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,j_1}\right] &= E\left[\delta^\top_{n_0,j_1,i_1} H_{i_1} b_{n_0,j_2,i_1} \delta^\top_{n_0,j_3,j_1} H_{j_1} b_{n_0,j_4,j_1}\right] \\ &= E\left[\delta^\top_{n_0,j_1,i_1} H_{i_1} b_{n_0,j_2,i_1} E\left[\delta^\top_{n_0,j_3,j_1} \mid W_{i_1}, W_{j_1}, W_{j_2}, W_{j_4}\right] H_{j_1} b_{n_0,j_4,j_1}\right] \\ &= 0.\end{aligned}$$

– $i_1 = i_2$, then j_3 is different than all and the previous argument used for $j_1 = i_2$ applies and implies

$$E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,i_2}\right] = 0$$

- Case 3: there are exactly four different indices. Then
  – if j_2 or j_4 is different than all, then

$$\begin{aligned}\left|E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_3,j_4,i_2}\right]\right| &= \left|E\left[\delta^\top_{n_0,j_1,i_1} H_{i_1} \tilde{b}_{n_0,i_1} \delta^\top_{n_0,j_3,i_2} H_{i_2} \tilde{b}_{n_0,i_2}\right]\right| \\ &= \tilde{C} E\left[|\delta_{n_0,j_1,i_1}||\tilde{b}_{n_0,i_1}||\delta_{n_0,j_3,i_2}||\tilde{b}_{n_0,i_2}|\right] \\ &= O(1)\end{aligned}$$

  which holds by C-S, LIE, and parts (b.1) and (b.3) of Assumption 3.2. Since there are at most $4^6 n^4$ terms, it follows these terms, in this case, contributed to I_2 with $O(n^{-2\varphi_1-2\varphi_2})$, which is $o(n^{-\zeta})$ since $2\varphi_1 + 2\varphi_2 > 4\varphi_1 - 1$.

  – if $j_1 = j_4$ and $j_2 = j_3$, then

$$\begin{aligned}E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_2,j_1,i_2}\right] &= E\left[\delta^\top_{n_0,j_1,i_1} H_{i_1} b_{n_0,j_2,i_1} \delta^\top_{n_0,j_2,i_2} H_{i_2} b_{n_0,j_1,i_2}\right] \\ &= E[E[\delta^\top_{n_0,j_1,i_1} \mid X_{j_1}, W_{i_1}, W_{j_2}, W_{i_2}] H_{i_1} b_{n_0,j_2,i_1} \delta^\top_{n_0,j_2,i_2} H_{i_2} b_{n_0,j_1,i_2}] \\ &= 0\ ,\end{aligned}$$

  which follows by part (a) of Assumption 3.2.

  – if $j_2 = j_3$ and $i_1 = j_4$, then

$$E\left[\Gamma^{l,b}_{j_1,j_2,i_1}\Gamma^{l,b}_{j_2,i_1,i_2}\right] = E\left[\delta^\top_{n_0,j_1,i_1} H_{i_1} b_{n_0,j_2,i_1} \delta^\top_{n_0,j_2,i_2} H_{i_2} b_{n_0,i_1,i_2}\right]$$

$$= E[E[\delta_{n_0,j_1,i_1}^\top \mid X_{j_1}, W_{i_1}, W_{j_2}, W_{i_2}]H_{i_1}b_{n_0,j_2,i_1}\delta_{n_0,j_2,i_2}^\top H_{i_2}b_{n_0,i_1,i_2}]$$
$$= 0 ,$$

which follows by part (a) of Assumption 3.2.

– if $j_1 = j_3$ and $j_2 = j_4$, then

$$\begin{aligned} n_0^{-2\varphi_2}\left|E\left[\Gamma_{j_1,j_2,i_1}^{l,b}\Gamma_{j_1,j_2,i_2}^{l,b}\right]\right| &= \left|E\left[\delta_{n_0,j_1,i_1}^\top H_{i_1}b_{n_0,j_2,i_1}\delta_{n_0,j_1,i_2}^\top H_{i_2}b_{n_0,j_2,i_2}\right]\right| \\ &\leq \tilde{C}E\left[E\left[|\delta_{n_0,j_1,i_1}|^2 \mid X_{i_1}\right]E\left[|n_0^{-\varphi_2}b_{n_0,j_2,i_1}|^2 \mid X_{i_1}\right]\right] \\ &\leq \tilde{C}E\left[E\left[|\delta_{n_0,j_1,i_1}|^2 \mid X_{i_1}\right]^2\right]^{1/2}E\left[E\left[|n_0^{-\varphi_2}b_{n_0,j_2,i_1}|^2 \mid X_{i_1}\right]^2\right]^{1/2} \\ &\leq \tilde{C}M_1^{1/2}n_0^{(1-2\varphi_1)}\tau_{n_0}^{1/2} \end{aligned}$$

which holds due to C-S, LIE, part (b.1) of Assumption 3.2. Since there are at most 4^6n^4 terms, it follows these terms contributed to I_2 with $o(n^{1-4\varphi_1})$, which is $o(n^{-\zeta})$ since $4\varphi_1 - 1 \geq \zeta$.

– $j_1 = i_2$ and $j_2 = j_4$, then

$$\begin{aligned} E\left[\Gamma_{j_1,j_2,i_1}^{l,b}\Gamma_{j_3,j_2,j_1}^{l,b}\right] &= E\left[\delta_{n_0,j_1,i_1}^\top H_{i_1}b_{n_0,j_2,i_1}\delta_{n_0,j_3,j_1}^\top H_{j_1}b_{n_0,j_2,j_1}\right] \\ &= E[\delta_{n_0,j_1,i_1}^\top H_{i_1}b_{n_0,j_2,i_1}E[\delta_{n_0,j_3,j_1}^\top \mid X_{j_3}, W_{i_1}, W_{j_2}, W_{j_1}]H_{j_1}b_{n_0,j_2,j_1}] \\ &= 0 , \end{aligned}$$

which holds due to part (a) of Assumption 3.2.

- Case 4: there are exactly three different indices. All the terms in this case contributed to I_2 with $o(n^{-\zeta})$ by a similar argument as Case 2 in the proof of Claim 3.1.

All the previous cases imply that $I_2 = o(n^{-\zeta})$, which completes the proof of Claim 3.2.

Claim 3.3: $I_3 = o(n^{-\zeta})$. This result is a consequence of C-S and Claims 3.1 and 3.2.

Part 3: It follows by Cauchy-Schwartz, using part 2 of this proposition and part 1 of Proposition C.2.

Part 4: It follows by Cauchy-Schwartz, using part 2 of this proposition and part 2 of Proposition C.3. □

E.7 Proof of Proposition C.5

Proof. For $i \in \mathcal{I}_k$, denote $\Delta_i = \Delta_i^b + \Delta_i^l$, where

$$\Delta_i^l = n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} \delta_{n_0,j,i} \ ,$$

$$\Delta_i^b = n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} b_{n_0,j,i} \ ,$$

Here, $\delta_{n_0,j,i} = \delta_{n_0}(W_j, X_i)$ and $b_{n_0,j,i} = b_{n_0}(X_j, X_i)$, and δ_{n_0} and b_{n_0} are functions satisfying Assumption 3.2.

Part 1: Consider the following decomposition

$$\begin{aligned}
E[\mathcal{T}_n^* \mathcal{T}_{n,K}^l] &= E\left[\left(n^{-1/2}\sum_{i_1=1}^{n} m_{i_1}/J_0\right)\left(n^{-1/2}\sum_{i_2=1}^{n}(\Delta_{i_2})^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&\overset{(1)}{=} n^{-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} E\left[(m_{i_1}/J_0)\left((\Delta_{i_2}^l + \Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&= n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((\Delta_{i_2}^l)^\top \partial_\eta m_{i_2}/J_0\right)\right] + E\left[(m_{i_1}/J_0)\left((\Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&= I_1 + I_2 \ ,
\end{aligned}$$

where (1) holds since $\Delta_i = \Delta_i^l + \Delta_i^b$. Claim 1.1 below shows that $I_1 = 0$, while Claim 1.2 shows $I_2 = (G_b^l/2) n_0^{-\varphi_2} + o(n^{-\varphi_2})$.

Claim 1: $I_1 = 0$. To see this, consider the following derivations

$$\begin{aligned}
I_1 &= n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((\Delta_{i_2}^l)^\top \partial_\eta m_{i_2}/J_0\right)\right] \\
&\overset{(1)}{=} n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} n_0^{-\varphi_1} n_0^{-1/2}\sum_{j\notin\mathcal{I}_k} E\left[(m_{i_1}/J_0)\, \delta_{n,j,i_2}^\top \partial_\eta m_{i_2}/J_0\right]
\end{aligned}$$

$$\overset{(2)}{=} n^{-1} \sum_{i_1=1}^{n} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} E\left[(m_{i_1}/J_0)\, E\left[\delta_{n,j,i_2}^\top \mid W_{i_1}, W_{i_2}\right] \partial_\eta m_{i_2}/J_0\right]$$

$$\overset{(3)}{=} n^{-1} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} E\left[(m_j/J_0)\, \delta_{n,j,i_2}^\top E\left[\partial_\eta m_{i_2}/J_0 \mid X_{i_2}, W_j\right]\right]$$

$$\overset{(4)}{=} 0$$

where (1) holds by definition of Δ_i^l, (2) holds by the law of iterative expectations, (3) holds since $E\left[\delta_{n,j,i_2}^\top \mid W_{i_1}, W_{i_2}\right] = 0$ when $i_1 \neq j$ due to part (a) of Assumption 3.2 and by the law of iterative expectations, and (4) holds by the Neyman orthogonality condition implied by part (b) of Assumption 3.1.

Claim 2: $I_2 = (G_b^l/2) n_0^{-\varphi_2} + o(n^{-\varphi_2})$. To see this, consider the following derivations

$$I_2 = n^{-1} \sum_{i_1=1}^{n} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((\Delta_{i_2}^b)^\top \partial_\eta m_{i_2}/J_0\right)\right]$$

$$\overset{(1)}{=} n^{-1} \sum_{i_1=1}^{n} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[(m_{i_1}/J_0)\, b_{n_0,j,i_2}^\top \partial_\eta m_{i_2}/J_0\right]$$

$$\overset{(2)}{=} n^{-1} \sum_{i_1=1}^{n} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[(m_{i_1}/J_0)\, b_{n_0,j,i_2}^\top E\left[\partial_\eta m_{i_2}/J_0 \mid X_{i_2}, W_{i_1}, X_j\right]\right]$$

$$\overset{(3)}{=} n^{-1} \sum_{k=1}^{K} \sum_{i_2 \in \mathcal{I}_k} n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[(m_{i_2}/J_0)\, E\left[b_{n_0,j,i_2}^\top \mid W_{i_2}\right] \partial_\eta m_{i_2}/J_0\right]$$

$$\overset{(4)}{=} n_0^{-\varphi_2} E\left[(m_{i_2}/J_0)\, \tilde{b}_{n_0}(X_{i_2}) \partial_\eta m_{i_2}/J_0\right]$$

$$\overset{(4)}{=} n_0^{-\varphi_2}(G_b^l/2) + o(n_0^{-\varphi_2})\ ,$$

where (1) holds by definition of Δ_i^b, (2) holds by the law of iterative expectations, (3) holds since $E\left[\partial_\eta m_{i_2}/J_0 \mid X_{i_2}, W_{i_1}, X_j\right] = 0$ when $i_1 \neq i_2$ due to the Neyman orthogonality condition implied by part (b) of Assumption 3.1 and the law of iterative expectations, (4) holds by definitions of $\tilde{b}_{n_0,i} = E[b_{n_0,j,i} \mid X_i]$ which is equal to $E\left[b_{n_0,j,i} \mid W_i\right]$, and (5) holds by definition of G_b^l in (A-7) and Assumption A.1.

Part 2: Consider the following decomposition,

$$
\begin{aligned}
E[\mathcal{T}_n^* \mathcal{T}_{n,K}^{nl}] &= E\left[\left(n^{-1/2}\sum_{i_1=1}^{n} m_{i_1}/J_0\right)\left(n^{-1/2}\sum_{i_2=1}^{n}(\Delta_{i_2})^\top H_{i_2}(\Delta_{i_2})\right)\right] \\
&= I_1 + 2I_2 + I_3
\end{aligned}
$$

where

$$
\begin{aligned}
I_1 &= E\left[\left(n^{-1/2}\sum_{i_1=1}^{n} m_{i_1}/J_0\right)\left(n^{-1/2}\sum_{i_2=1}^{n}(\Delta_{i_2}^l)^\top H_{i_2}(\Delta_{i_2}^l)\right)\right] \\
I_2 &= E\left[\left(n^{-1/2}\sum_{i_1=1}^{n} m_{i_1}/J_0\right)\left(n^{-1/2}\sum_{i_2=1}^{n}(\Delta_{i_2}^l)^\top H_{i_2}(\Delta_{i_2}^b)\right)\right] \\
I_3 &= E\left[\left(n^{-1/2}\sum_{i_1=1}^{n} m_{i_1}/J_0\right)\left(n^{-1/2}\sum_{i_2=1}^{n}(\Delta_{i_2}^b)^\top H_{i_2}(\Delta_{i_2}^b)\right)\right]
\end{aligned}
$$

In what follows, I show that $I_1 = o(n^{-\zeta})$, $I_2 = (G_b/2)n_0^{1/2-\varphi_1-\varphi_2} + o(n^{-\zeta})$, and $I_3 = o(n^{-\zeta})$.

Claim 1: $I_1 = o(n^{-\zeta})$. Consider the following derivations,

$$
\begin{aligned}
I_1 &\overset{(1)}{=} n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((\delta_{n,j_1,i_2})^\top H_{i_2}(\delta_{n,j_2,i_2})\right)\right] \\
&\overset{(2)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_i/J_0)\,(\delta_{n,j_1,i})^\top H_i(\delta_{n,j_2,i})\right] \\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_{j_1}/J_0)\,(\delta_{n,j_1,i})^\top H_i(\delta_{n,j_2,i})\right] \\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_{j_2}/J_0)\,(\delta_{n,j_1,i})^\top H_i(\delta_{n,j_2,i})\right] \\
&\overset{(3)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j\notin\mathcal{I}_k} E\left[(m_i/J_0)\,(\delta_{n,j,i})^\top H_i(\delta_{n,j,i})\right] \\
&\quad + 2n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_1} n_0^{-1}\sum_{j\notin\mathcal{I}_k} E\left[(m_j/J_0)\,(\delta_{n,j,i})^\top H_i(\delta_{n,j,i})\right]
\end{aligned}
$$

$$= n_0^{-2\varphi_1} E\left[(m_i/J_0)\left(\delta_{n,j,i}\right)^\top H_i(\delta_{n,j,i})\right] + 2n_0^{-2\varphi_1} E\left[(m_j/J_0)\left(\delta_{n,j,i}\right)^\top H_i(\delta_{n,j,i})\right]$$
$$\stackrel{(4)}{=} O(n^{-2\varphi_1})$$

where (1) holds by definition of Δ_i^l, (2) holds since $i_1 \notin \{i_2, j_1, j_2\}$ implies

$$E\left[(m_{i_1}/J_0)\left((\delta_{n,j_1,i_2})^\top H_{i_2}(\delta_{n,j_2,i_2})\right)\right] = 0 \ ,$$

which follows since m_{i_1} is a zero mean random variable independent of W_{i_2}, W_{j_1} and W_{j_2}, (3) holds since $j_1 \neq j_2$ implies

$$E\left[(m_i/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(\delta_{n,j_2,i})\right] = 0$$
$$E\left[(m_{j_1}/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(\delta_{n,j_2,i})\right] = 0$$
$$E\left[(m_{j_2}/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(\delta_{n,j_2,i})\right] = 0 \ ,$$

which follows by the law of iterative expectations and noting $E[\delta_{n,j_2,i} \mid W_i, W_{j_1}] = 0$ and $E[\delta_{n,j_1,i} \mid W_i, W_{j_2}] = 0$ (due to part (a) of Assumption 3.2), and (4) holds by Holder's inequality, part (e) of Assumption 3.1, and part (a) of Assumption 3.2.

Claim 2: $I_2 = (G_b/2)n_0^{1/2-\varphi_1-\varphi_2} + o(n^{-\zeta})$. Consider the following derivations,

$$I_2 \stackrel{(1)}{=} n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2} n_0^{-3/2} \sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((\delta_{n,j_1,i_2})^\top H_{i_2}(b_{n_0,j_2,i_2})\right)\right]$$
$$\stackrel{(2)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_i/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(b_{n_0,j_2,i})\right]$$
$$+ n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_1}/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(b_{n_0,j_2,i})\right] I\{j_1 \neq j_2\}$$
$$+ n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_2}/J_0)\left(\delta_{n,j_1,i}\right)^\top H_i(b_{n_0,j_2,i})\right] I\{j_1 \neq j_2\}$$
$$+ n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j\notin\mathcal{I}_k} E\left[(m_j/J_0)\left(\delta_{n,j,i}\right)^\top H_i(b_{n_0,j,i})\right]$$
$$\stackrel{(3)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j\notin\mathcal{I}_k} E\left[(m_i/J_0)\left(\delta_{n,j,i}\right)^\top H_i(b_{n_0,j,i})\right]$$

$$
\begin{aligned}
&+ n^{-1} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j_1, j_2 \notin \mathcal{I}_k} E\left[(m_{j_1}/J_0) \, (\delta_{n,j_1,i})^\top H_i(\tilde{b}_{n_0,i}) \right] I\{j_1 \neq j_2\} \\
&+ n^{-1} \sum_{k=1}^{K} \sum_{i \in \mathcal{I}_k} n_0^{-\varphi_1-\varphi_2-3/2} \sum_{j \notin \mathcal{I}_k} E\left[(m_j/J_0) \, (\delta_{n,j,i})^\top H_i(b_{n_0,j,i}) \right] \\
&\overset{(4)}{=} n_0^{-\varphi_1-\varphi_2+1/2} E\left[(m_{j_1}/J_0) \, (\delta_{n,j_1,i})^\top H_i(\tilde{b}_{n_0,i}) \right] \\
&- n_0^{-\varphi_1-\varphi_2-1/2} E\left[(m_{j_1}/J_0) \, (\delta_{n,j_1,i})^\top H_i(\tilde{b}_{n_0,i}) \right] \\
&+ n_0^{-\varphi_1-1/2} E\left[(m_j/J_0) \, (\delta_{n,j,i})^\top H_i(n_0^{-\varphi_2} b_{n_0,j,i}) \right] \\
&\overset{(5)}{=} (G_b/2) n_0^{1/2-\varphi_1-\varphi_2} + o(n^{-\zeta}) \ ,
\end{aligned}
$$

where (1) holds by definition of $\Delta_{1,i}^l$ and $\Delta_{1,i}^b$, (2) holds since $i_1 \notin \{i_2, j_1, j_2\}$ implies

$$
E\left[(m_{i_1}/J_0) \left((\delta_{n,j_1,i_2})^\top H_{i_2}(b_{n_0,j_2,i_2}) \right) \right]
$$

by using that m_{i_1} is zero mean and independent of W_{i_2}, W_{j_1}, and W_{j_2}, (3) holds by definition of $\tilde{b}_{n_0,i} = E[b_{n_0,j,i} \mid X_i]$, and since $j_1 \neq j_2$ and the law of iterative expectations implies

$$
(m_\ell/J_0) \, E\left[(\delta_{n,j_1,i})^\top \mid W_i, W_{j_2} \right] H_{i_2}(b_{n_0,j_2,i}) = 0
$$

for $\ell = i, j_2$ by using part (a) of Assumption 3.2, (4) holds by the law of the iterative expectations,

$$
(m_i/J_0) \, E\left[(\delta_{n,j,i})^\top \mid W_i, X_j \right] H_i(b_{n_0,j,i}) = 0 \ ,
$$

and by parts (a) of Assumption 3.2, and (5) holds by definition of G_b in (A-6) and Assumption A.1 and because

$$
\begin{aligned}
& n_0^{-\varphi_1-1/2} E\left[(m_j/J_0) \, (\delta_{n,j,i})^\top H_i(n_0^{-\varphi_2} b_{n_0,j,i}) \right] \\
& \leq C n_0^{-\varphi_1-1/2} E[|m_j/J_0|^2]^{1/2} E[||\delta_{n,j,i}||^4]^{1/4} E[||n_0^{-\varphi_2} b_{n_0,j,i}||^4] \\
& = O(n^{-\varphi_1-1/2}) \times O(1) \times O(n_0^{1/4-\varphi_1/2}) \times o(n_0^{3/4-3\varphi_1/2}) \\
& = o(n^{1/2-3\varphi_1})
\end{aligned}
$$

where the inequality uses part (d) of Assumption 3.1 and Cauchy-Schwartz inequality, and the equalities follows by part (b) of Assumption 3.2. The proof of the claim is completed since $\zeta < 1/2 - 3\varphi_1$ whenever $\varphi_1 < 1/2$.

Claim 3: $I_3 = o(n^{-\zeta})$. Consider the following derivations,

$$
\begin{aligned}
I_4 &\overset{(1)}{=} n^{-1}\sum_{i_1=1}^{n}\sum_{k=1}^{K}\sum_{i_2\in\mathcal{I}_k} n_0^{-2\varphi_2} n_0^{-2}\sum_{j_1\notin\mathcal{I}_k}\sum_{j_2\notin\mathcal{I}_k} E\left[(m_{i_1}/J_0)\left((b_{n_0,j_1,i_2})^\top H_{i_2}(b_{n_0,j_2,i_2})\right)\right]\\
&\overset{(2)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_i/J_0)\,(b_{n_0,j_1,i})^\top H_i(b_{n_0,j_2,i})\right]\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_1}/J_0)\,(b_{n_0,j_1,i})^\top H_i(b_{n_0,j_2,i})\right] I\{j_1\neq j_2\}\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_2}/J_0)\,(b_{n_0,j_1,i})^\top H_i(b_{n_0,j_2,i})\right] I\{j_1\neq j_2\}\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j\notin\mathcal{I}_k} E\left[(m_j/J_0)\,(b_{n_0,j,i})^\top H_i(b_{n_0,j,i})\right]\\
&\overset{(3)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_i/J_0)\,(\tilde{b}_{n_0,i})^\top H_i(\tilde{b}_{n_0,i})\right] I\{j_1\neq j_2\}\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2}\sum_{j\notin\mathcal{I}_k} E\left[(m_i/J_0)\,(n_0^{-\varphi_2}b_{n_0,j,i})^\top H_i(n_0^{-\varphi_2}b_{n_0,j,i})\right]\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_1}/J_0)\,(n_0^{-\varphi_2}b_{n_0,j_1,i})^\top H_i(\tilde{b}_{n_0,i})\right] I\{j_1\neq j_2\}\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-\varphi_2-2}\sum_{j_1,j_2\notin\mathcal{I}_k} E\left[(m_{j_2}/J_0)\,(\tilde{b}_{n_0,i})^\top H_i(n_0^{-\varphi_2}b_{n_0,j_2,i})\right] I\{j_1\neq j_2\}\\
&\quad + n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} n_0^{-2}\sum_{j\notin\mathcal{I}_k} E\left[(m_j/J_0)\,(n_0^{-\varphi_2}b_{n_0,j,i})^\top H_i(n_0^{-\varphi_2}b_{n_0,j,i})\right]\\
&\overset{(4)}{=} O(n^{-2\varphi_2}) + o(n^{1/2-3\varphi_1}) + o(n^{1/2-\varphi_1-\varphi_2}) + o(n^{1/2-3\varphi_1})\\
&\overset{(5)}{=} o(n^{-\zeta})\ ,
\end{aligned}
$$

where (1) holds by definition of $\Delta^b_{1,i}$, (2) holds since $i_1 \notin \{j_1, j_2, i_2\}$ implies

$$E\left[(m_{i_1}/J_0)\left((b_{n,j_1,i_2})^\top(\partial^2_\eta m_{i_2}/(2J_0))(b_{n,j_2,i_2})\right)\right] = 0$$

by using that m_{i_1} is zero mean and independent of W_{i_2}, W_{j_1}, and W_{j_2}; (3) holds by the law of iterative expectations; (4) holds by parts (b) and (c) of Assumption 3.2, parts (c) and (e) of Assumption 3.1, part (b.1) of Assumption 3.2, and Holder's inequality; and (5) holds since $\zeta < 1/2 - 3\varphi_1$ and $\zeta < 2\varphi_2$ (because $\varphi_1 < 1/2$ and $\varphi_1 \leq \varphi_2$). $\square$

E.8 Proof of Lemma C.1

Proof. It is sufficient to show the result for the case when δ_{n_0} and b_{n_0} are real-valued functions since for any $x = (x_1, \ldots, x_p) \in \mathbf{R}^p$ it holds $||x||^4 = \left(\sum_{\ell=1}^p x_\ell^2\right)^2 \leq p\sum_{\ell=1}^p |x_\ell|^4$. In the proof, I use that $E[(\sum_{\ell \notin \mathcal{I}_k} Z_\ell)^4] \leq n_0 E[Z_\ell^4] + 3n_0^2 E[Z_\ell^2]^2$, which holds for zero mean i.i.d. random variables Z_ℓ.

Part 1: Fix $i \in \mathcal{I}_k$ and denote $Z_\ell = \delta_{n_0}(W_\ell, X_i)$ for any $\ell \notin \mathcal{I}_k$. Conditional on X_i, it holds that $\{Z_\ell : \ell \notin \mathcal{I}_k\}$ is a zero mean i.i.d. sequence of random variables due to part (a) of Assumption 3.2. Therefore,

$$E\left[|n_0^{-1/2}\sum_{\ell \notin \mathcal{I}_k} n_0^{-\varphi_1} Z_\ell|^4 \mid X_i\right] \leq n_0^{-2-4\varphi_1}\left(n_0 E[Z_\ell^4 \mid X_i] + 3n_0^2 E[Z_\ell^2 \mid X_i]^2\right)$$

Using the previous inequality and the law of iterative expectations, it follows

$$\begin{aligned} E\left[|n_0^{-1/2}\sum_{\ell \notin \mathcal{I}_k} n_0^{-\varphi_1} Z_\ell|^4\right] &\leq n_0^{-4\varphi_1}\left(n_0^{-1} E[Z_\ell^4] + 3E[E[Z_\ell^2 \mid X_i]^2]\right) \\ &\overset{(1)}{\leq} n_0^{-4\varphi_1}\left(n_0^{-2\varphi_1} M_1 + 3M_1\right)\ , \end{aligned}$$

where (1) holds by parts (b.1) and (b.2) of Assumption 3.2, and the definition of Z_ℓ. Taking $C \geq 4M_1$ completes the proof of part 1.

Part 2: Fix $i \in \mathcal{I}_k$ and denote $Z_\ell = n_0^{-\varphi_2}(b_{n_0}(X_\ell, X_i) - \tilde{b}_{n_0}(X_i))$ for any $\ell \notin \mathcal{I}_k$, where $\tilde{b}_{n_0}(X_i) = E[b_{n_0}(X_\ell, X_i) \mid X_i]$. As in part 1, $\{Z_\ell : \ell \notin \mathcal{I}\}$ conditional on X_i are zero mean i.i.d. random variables. Therefore,

$$E\left[|n_0^{-1}\sum_{\ell \notin \mathcal{I}_k}(Z_\ell + n_0^{-\varphi_2}\tilde{b}_{n_0}(X_i))|^4\right] \overset{(1)}{\leq} 2^3 E\left[|n_0^{-1}\sum_{\ell \notin \mathcal{I}_k} Z_\ell|^4\right] + 2^3 E\left[|n_0^{-1}\sum_{\ell \notin \mathcal{I}_k} n_0^{-\varphi_2}\tilde{b}_{n_0}(X_i)|^4\right]$$

$$\overset{(2)}{\leq} 8n_0^{-2}\left(n_0^{-1}E[Z_\ell^4] + 3E[E[Z_\ell^2 \mid X_i]^2]\right) + 8n_0^{-4\varphi_2}E[|\tilde{b}_{n_0}(X_i)|^4]$$
$$\overset{(3)}{\leq} 8n_0^{-6\varphi_1}\tau_{n_0} + 24n_0^{-4\varphi_1}\tau_{n_0} + 8n_0^{-4\varphi_2}M_1$$

where (1) holds by Loeve's inequality (Davidson (1994, Theorem 9.28)), (2) holds by the same arguments as in part 1, and (3) holds by part (b.1), (b.3) and (b.4) of Assumption 3.2. Taking $C \geq 32 + 8M_1$ completes the proof of part 2. □

E.9 Proof of Lemma C.2

Proof. For $i \in \mathcal{I}_k$, denote $\Delta_i = \Delta_i^b + \Delta_i^l$, where

$$\Delta_i^l = n_0^{-\varphi_1} n_0^{-1/2} \sum_{j \notin \mathcal{I}_k} \delta_{n_0,j,i} \ ,$$
$$\Delta_i^b = n_0^{-\varphi_2} n_0^{-1} \sum_{j \notin \mathcal{I}_k} b_{n_0,j,i} \ ,$$

Here, $\delta_{n_0,j,i} = \delta_{n_0}(W_j, X_i)$ and $b_{n_0,j,i} = b_{n_0}(X_j, X_i)$, and δ_{n_0} and b_{n_0} are functions satisfying Assumption 3.2. In what follows, the results are proved for any given sequence K that diverges to infinity as n diverges to infinity, which is sufficient to guarantee the results of this lemma.

Part 1: Using Assumption 3.2, it follows

$$n^{-1} \sum_{i=1}^{n} (\hat{\eta}_i - \eta_i)^\top \partial_\eta \psi_i^z = I_1 + I_2 + I_3 \ ,$$

where

$$I_1 = n^{-1} \sum_{i=1}^{n} (\Delta_i^l)^\top \partial_\eta \psi_i^z$$
$$I_2 = n^{-1} \sum_{i=1}^{n} (\Delta_i^b)^\top \partial_\eta \psi_i^z$$
$$I_3 = n^{-1} n_0^{-2\min\{\varphi_1,\varphi_2\}} \sum_{i=1}^{n} \hat{R}_1(X_i)^\top \partial_\eta \psi_i^z$$

and $n_0 = ((K-1)/K)n$.

Claim 1: $I_1 = O_p(n^{-1/2-\min\{\varphi_1,\varphi_2\}})$. I first show that $E[I_1] = 0$ (claim 1.1). I then show that $E[I_1^2] = O(n^{-1-2\min\{\varphi_1,\varphi_2\}})$ (claim 1.2), which is sufficient to conclude the claim.

Claim 1.1: $E[I_1] = 0$. Consider the following derivations,

$$\begin{aligned} E[I_1] &= n^{-1}\sum_{i=1}^{n} E[(\Delta_i^l)^\top \partial_\eta \psi_i^z] \\ &= n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} E\left[E\left[(\Delta_i^l)^\top \partial_\eta \psi_i^z \mid X_i, (W_j : j \notin \mathcal{I}_k)\right]\right] \\ &\overset{(1)}{=} n^{-1}\sum_{k=1}^{K}\sum_{i\in\mathcal{I}_k} E\left[(\Delta_i^l)^\top E\left[\partial_\eta \psi_i^z \mid X_i, (W_j : j \notin \mathcal{I}_k)\right]\right] \\ &\overset{(2)}{=} 0 , \end{aligned}$$

where (1) holds since $\Delta_i^l = \Delta_1^l(X_i)$ is a function of X_i and the data $(W_j : j \notin \mathcal{I}_k)$ used to estimate $\hat{\eta}_k(\cdot)$, and (2) holds by part (b) in Assumption 3.1.

Claim 1.2: $E[I_1^2] = O(n^{-1-2\min\{\varphi_1,\varphi_2\}})$. Recall that I use the following notation $\delta_{n_0,j,i} = \delta_{n_0}(W_j, X_i)$ and $\Delta_i^l = \Delta_1^l(X_i) = n_0^{-\varphi_1} n_0^{-1/2} \sum_{j\notin\mathcal{I}_k} \delta_{n_0,j,i}$ for $i \in \mathcal{I}_k$. To show $E[I_1^2] = O(n^{-1-2\min\{\varphi_1,\varphi_2\}})$, consider the following derivations,

$$\begin{aligned} E[I_1^2] &\overset{(1)}{=} n^{-2}\sum_{i_1,i_2=1}^{n} E\left[(\Delta_{i_1}^l)^\top \partial_\eta \psi_{i_1}^z (\Delta_{i_2}^l)^\top \partial_\eta \psi_{i_2}^z\right] \\ &\overset{(2)}{\leq} n_0^{-2\varphi_1} n^{-2}\sum_{i=1}^{n} E\left[((\Delta_i^l)^\top \partial_\eta \psi_i^z)^2\right] + n_0^{-2\varphi_1} n^{-2}\sum_{i_1\neq i_2}^{n} \left|E\left[(\Delta_{i_1}^l)^\top \partial_\eta \psi_{i_1}^z (\Delta_{i_2}^l)^\top \partial_\eta \psi_{i_2}^z\right]\right| \\ &\overset{(3)}{\leq} n_0^{-2\varphi_1-1} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta \psi_i^z\right)^2\right] + \frac{(n-1)n^{-1}}{n_0^{1+2\varphi_1}} \left|E\left[\delta_{n_0,i_2,i_1}^\top \partial_\eta \psi_{i_1}^z \delta_{n_0,i_1,i_2}^\top \partial_\eta \psi_{i_2}^z\right]\right| \\ &\overset{(4)}{\leq} n_0^{-2\varphi_1-1} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta \psi_i^z\right)^2\right] + (n-1)n^{-1} n_0^{-(1+2\varphi_1)} E\left[\left(\delta_{n_0,i_2,i_1}^\top \partial_\eta \psi_{i_1}^z\right)^2\right]^{1/2} E\left[\left(\delta_{n_0,i_1,i_2}^\top \partial_\eta \psi_{i_2}^z\right)^2\right]^{1/2} \\ &= n_0^{-2\varphi_1-1} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta \psi_i^z\right)^2\right] + (n-1)n^{-1} n_0^{-(1+2\varphi_1)} E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta \psi_i^z\right)^2\right] \\ &\overset{(5)}{\leq} n_0^{-2\varphi_1-1} M_1^{1/2} C_1 \times p\left(1 + (n-1)n^{-1}\right) , \end{aligned}$$

where (1) holds by definition of I_1, (2) holds by triangular inequality, (3) holds by

(E.25) and (E.26) presented below, (4) holds by Cauchy-Schwartz inequality, and (5) holds by the derivations presented next,

$$
\begin{aligned}
E\left[\left(\delta_{n_0,j,i}^\top \partial_\eta \psi_i^z\right)^2\right] &= E\left[\delta_{n_0,j,i}^\top E\left[\left(\partial_\eta \psi_i^z (\partial_\eta \psi_i^z)^\top\right) \mid X_i, W_j\right] \delta_{n_0,j,i}\right] \\
&\overset{(1)}{=} E\left[\delta_{n_0,j,i}^\top E\left[\left(\partial_\eta \psi_i^z (\partial_\eta \psi_i^z)^\top\right) \mid X_i\right] \delta_{n_0,j,i}\right] \\
&\overset{(2)}{\leq} E\left[||\delta_{n_0,j,i}||^2\right] C_1 \times p \\
&\overset{(3)}{\leq} M_1^{1/2} C_1 \times p
\end{aligned}
$$

where (1) holds since $i \neq j$, (2) holds by part (d) of Assumption 3.1 and Loeve's inequality (Davidson (1994, Theorem 9.28)), and (3) holds by Jensen's inequality (e.g., $E\left[||\delta(W_j, X_i)||^2\right]^{1/2} \leq E\left[||\delta(W_j, X_i)||^4\right]^{1/4}$) and part (b) of Assumption 3.2. Note these derivations complete the proof of claim 1.2.

The previous derivations used the following claims:

$$
E\left[((\Delta_i^l)^\top \partial_\eta \psi_i^z)^2\right] = n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[\delta(W_j, X_i)^\top \partial_\eta \psi_i^z \delta(W_j, X_i)^\top \partial_\eta \psi_i^z\right] \quad \text{(E.25)}
$$

$$
E\left[(\Delta_{i_1}^l)^\top \partial_\eta \psi_{i_1}^z (\Delta_{i_2}^l)^\top \partial_\eta \psi_{i_2}^z\right] = n_0^{-1} E\left[\delta_{n_0,i_2,i_1}^\top \partial_\eta \psi_{i_1}^z \delta_{n_0,i_1,i_2}^\top \partial_\eta \psi_{i_2}^z\right] I\{k_1 \neq k_2\} \quad \text{(E.26)}
$$

To show (E.25), consider the following derivations.

$$
\begin{aligned}
E\left[(\Delta_i^l)^\top \partial_\eta \psi_i^z (\Delta_i^l)^\top \partial_\eta \psi_i^z\right] &= n_0^{-1} \sum_{j_1 \notin \mathcal{I}_k} \sum_{j_2 \notin \mathcal{I}_k} E\left[\delta_{n_0,j_1,i}^\top \partial_\eta \psi_i^z \delta_{n_0,j_2,i}^\top \partial_\eta \psi_i^z\right] \\
&\overset{(1)}{=} n_0^{-1} \sum_{j \notin \mathcal{I}_k} E\left[\delta(W_j, X_i)^\top \partial_\eta \psi_i^z \delta(W_j, X_i)^\top \partial_\eta \psi_i^z\right]
\end{aligned}
$$

where (1) holds due to the following: if $j_1 \neq j_2$, then

$$
\begin{aligned}
E\left[\delta_{n_0,j_1,i}^\top \partial_\eta \psi_i^z \delta_{n_0,j_1,i}^\top \partial_\eta \psi_i^z\right] &= E\left[E\left[\delta_{n_0,j_1,i}^\top \mid W_i, W_{j_2}\right] \partial_\eta \psi_i^z \delta_{n_0,j_2,i}^\top \partial_\eta \psi_i^z\right] \\
&= E\left[E\left[\delta_{n_0,j_1,i}^\top \mid X_i\right] \partial_\eta \psi_i^z \delta_{n_0,j_2,i}^\top \partial_\eta \psi_i^z\right] \\
&\overset{(1)}{=} 0 \ ,
\end{aligned}
$$

where (1) holds by definition of δ_{n_0} in part (a) of Assumption 3.2.

To show (E.26), consider $i_1 \neq i_2$ where $i_1 \in \mathcal{I}_{k_1}$ and $i_2 \in \mathcal{I}_{k_2}$, therefore

$$E\left[(\Delta^l_{i_1})^\top \partial_\eta \psi^z_{i_1} (\Delta^l_{i_2})^\top \partial_\eta \psi^z_{i_2}\right] = n_0^{-1} \sum_{j_1 \notin \mathcal{I}_{k_1}} \sum_{j_2 \notin \mathcal{I}_{k_2}} E\left[\delta^\top_{n_0,j_1,i_1} \partial_\eta \psi^z_{i_1} \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2}\right]$$
$$\stackrel{(1)}{=} n_0^{-1} E\left[\delta^\top_{n_0,i_2,i_1} \partial_\eta \psi^z_{i_1} \delta^\top_{n_0,i_1,i_2} \partial_\eta \psi^z_{i_2}\right] I\{k_1 \neq k_2\}$$

where (1) holds since $k_1 = k_2$ implies $j_2 \neq i_1$ and $j_1 \neq i_2$, and because the conditions $j_2 \neq i_1$ or $j_1 \neq i_2$ imply that $E\left[\delta^\top_{n_0,j_1,i_1} \partial_\eta \psi^z_{i_1} \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2}\right]$ is zero. To see this, suppose $j_2 \neq i_1$ and consider the following derivations,

$$\begin{aligned} E\left[\delta^\top_{n_0,j_1,i_1} \partial_\eta \psi^z_{i_1} \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2}\right] &= E\left[E\left[\delta^\top_{n_0,j_1,i_1} \partial_\eta \psi^z_{i_1} \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2} \mid X_{i_1}, W_{i_2}, W_{j_1}, W_{j_2}\right]\right] \\ &= E\left[\delta^\top_{n_0,j_1,i_1} E\left[\partial_\eta \psi^z_{i_1} \mid X_{i_1}, W_{i_2}, W_{j_1}, W_{j_2}\right] \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2}\right] \\ &\stackrel{(1)}{=} E\left[\delta^\top_{n_0,j_1,i_1} E\left[\partial_\eta \psi^z_{i_1} \mid X_{i_1}\right] \delta^\top_{n_0,j_2,i_2} \partial_\eta \psi^z_{i_2}\right] \\ &\stackrel{(2)}{=} 0 \ , \end{aligned}$$

where (1) holds since $i_1 \notin \{i_2, j_1, j_2\}$ (since $i_1 \neq j_2$) and (2) holds by part (b) in Assumption 3.1. Similar derivations conclude the same for $j_1 \neq i_2$.

Claim 2: $I_2 = O(n^{-1/2-\min\{\varphi_1,\varphi_2\}})$. Define $X^{(n)} = \{X_i : 1 \leq i \leq n\}$. I first show $E[I_2 \mid X^{(n)}] = 0$. I then show $E[I_2^2] \leq n^{-1} E[||\Delta^b_i||^2] C_1 p$, which is sufficient to conclude due to Lemma C.1 that implies that $E[||\Delta^b_i||^2]$ is $O(n^{-2\min\{\varphi_1,\varphi_2\}})$ due to Cauchy-Schwartz.

The first part holds due to the following derivations,

$$\begin{aligned} E[I_2 \mid X^{(n)}] &= E\left[n^{-1} \sum_{i=1}^n (\Delta^b_i)^\top \partial_\eta \psi^z_i \mid X^{(n)}\right] \\ &\stackrel{(1)}{=} n^{-1} \sum_{i=1}^n (\Delta^b_i)^\top E\left[\partial_\eta \psi^z_i \mid X^{(n)}\right] \\ &\stackrel{(2)}{=} n^{-1} \sum_{i=1}^n (\Delta^b_i)^\top E\left[\partial_\eta \psi^z_i \mid X_i\right] \\ &\stackrel{(3)}{=} 0 \ , \end{aligned}$$

where (1) holds since Δ^b_i is function of $X^{(n)}$ and $\Delta^b_i = n_0^{-\varphi_2} n_0^{-1} \sum_{i_0 \notin \mathcal{I}_k} b(X_{i_0}, X_i)$ for

$i \in \mathcal{I}_k$ due to part (a) of Assumption 3.2, (2) holds since the observations are i.i.d., and (3) follows due to part (b) of Assumption 3.1.

To prove that $E[I_2^2] \leq n^{-1}E[||\Delta_i^b||^2]C_1 p$, first note that

$$\begin{aligned}
E[I_2^2 \mid X^{(n)}] &= E\left[\left(n^{-1}\sum_{i=1}^{n}(\Delta_i^b)^\top \partial_\eta \psi_i^z\right)^2 \mid X^{(n)}\right] \\
&\overset{(1)}{=} E\left[n^{-2}\sum_{i=1}^{n}\left((\Delta_i^b)^\top \partial_\eta \psi_i^z\right)^2 \mid X^{(n)}\right] \\
&\overset{(2)}{=} n^{-2}\sum_{i=1}^{n}(\Delta_i^b)^\top E\left[(\partial_\eta \psi_i^z)(\partial_\eta \psi_i^z)^\top \mid X_i\right]\Delta_i^b \\
&\overset{(1)}{\leq} n^{-2}\sum_{i=1}^{n}||\Delta_i^b||^2 C_1 \times p
\end{aligned}$$

where (1) holds because $E\left[\left((\Delta_i^b)^\top \partial_\eta \psi_i^z\right)\left((\Delta_j^b)^\top \partial_\eta \psi_j^z\right) \mid X^{(n)}\right] = 0$ when $i \neq j$ (since Δ_i^b and Δ_j^b are functions of $X^{(n)}$, and part (b) of Assumption 3.1), (2) holds since Δ_i^b and Δ_j^b are functions of $X^{(n)}$ and the observations are i.i.d., and (3) holds by part (d) of Assumption 3.1. Then,

$$E[E[I_2^2 \mid X^{(n)}]] \leq E[n^{-2}\sum_{i=1}^{n}||\Delta_i^b||^2 C_1 \times p] = n^{-1}E[||\Delta_i^b||^2]C_1 p$$

which completes the proof of this claim.

Claim 3: $I_3 = O_p(n_0^{-2\min\{\varphi_1,\varphi_2\}})$. Algebra shows

$$\begin{aligned}
|I_3| &= |n_0^{-2\min\{\varphi_1,\varphi_2\}} n^{-1}\sum_{i=1}^{n}\hat{R}_1(X_i)^\top \partial_\eta \psi_i^z| \\
&\leq n_0^{-2\min\{\varphi_1,\varphi_2\}}\left(n^{-1}\sum_{i=1}^{n}||\hat{R}_1(X_i)||^2\right)^{1/2}\left(n^{-1}\sum_{i=1}^{n}||\partial_\eta \psi_i^z||^2\right)^{1/2} \\
&\overset{(1)}{=} n_0^{-2\min\{\varphi_1,\varphi_2\}} \times O_p(1) \times \left(n^{-1}\sum_{i=1}^{n}||\partial_\eta \psi_i^z||^2\right)^{1/2}, \\
&\overset{(2)}{=} n_0^{-2\min\{\varphi_1,\varphi_2\}} \times O_p(1) \times O_p(1),
\end{aligned}$$

$$\overset{(3)}{=} O_p(n_0^{-2\min\{\varphi_1,\varphi_2\}}) \ ,$$

where (1) holds by part (c) of Assumption 3.2, (2) holds by the law of large numbers, Jensen's inequality (e.g., $(E[||\partial_\eta \psi_i^z||^2] \leq E[||\partial_\eta \psi_i^z||^4]^{1/2})$, and part (c) of Assumption 3.1, and (3) holds since $n/2 \leq n \leq n$

Part 2: By Taylor approximation and mean-value theorem (since $\psi^z(w,\eta)$ is twice continuously differentiable on η by Assumption 3.1), it follows

$$\hat{\psi}_i^z - \psi_i^z = (\hat{\eta}_i - \eta_i)^\top \partial_\eta \psi_i^z + \frac{1}{2}(\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{\psi}_i^z (\hat{\eta}_i - \eta_i)$$

where $\partial_\eta^2 \tilde{\psi}_i^z = \partial_\eta^2 \psi^z(W_i,\eta)|_{\eta=\tilde{\eta}_i}$ for some $\hat{\eta}_i$ (due to mean-value theorem). Using this

$$\begin{aligned} n^{-1}\sum_{i=1}^n (\hat{\psi}_i^z - \psi_i^z) &= n^{-1}\sum_{i=1}^n (\hat{\eta}_i - \eta_i)^\top \partial_\eta \psi_i^z + \frac{1}{2} n^{-1}\sum_{i=1}^n (\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{\psi}_i^z (\hat{\eta}_i - \eta_i) \\ &\overset{(1)}{=} O_p(n^{-\min\{\varphi_1,\varphi_2\}-1/2}) + \frac{1}{2} n^{-1}\sum_{i=1}^n (\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{\psi}_i^z (\hat{\eta}_i - \eta_i) \\ &\overset{(2)}{=} O_p(n^{-2\min\{\varphi_1,\varphi_2\}}) \ , \end{aligned}$$

where (1) holds due to Part 1, and (2) holds due to the derivations presented next,

$$\begin{aligned} |n^{-1}\sum_{i=1}^n (\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{m}_i (\hat{\eta}_i - \eta_i)| &\leq n^{-1}\sum_{i=1}^n |(\hat{\eta}_i - \eta_i)^\top \partial_\eta^2 \tilde{m}_i (\hat{\eta}_i - \eta_i)| \\ &\overset{(1)}{\leq} n^{-1}\sum_{i=1}^n ||\hat{\eta}_i - \eta_i||^2 C_2 \times p \ , \\ &\overset{(2)}{=} O_p(n^{-2\min\{\varphi_1,\varphi_2\}}) \ , \end{aligned}$$

where (1) holds due to part (e) of Assumption 3.1 and Loeve's inequality (Davidson (1994, Theorem 9.28)), and (2) holds due to part 4 of Lemma C.4.

Part 3: It follows from part 1, by using that $\partial_\eta m_i = \partial_\eta \psi_i^b - \partial_\eta \psi_i^a \theta_0$ and $|\theta_0| \leq M_1^{1/4}/C_0$ (due to parts (a) and (c) of Assumptions 3.1 and the representation of θ_0 as a ratio of expected values in (2.3)).

Part 4: By Taylor expansion and mean value theorem,

$$\hat{m}_i - m_i = (\hat{\eta}_i - \eta_i)^\top \partial_\eta m_i + (\hat{\eta}_i - \eta_i)^\top (\partial_\eta^2 m_i/2)(\hat{\eta}_i - \eta_i) + \tilde{r}_i \ ,$$

where $\tilde{r}_i$ is the Lagrange's remainder error term (since m is three-times continuous differentiable on η by assumption on ψ^z). Therefore,

$$|\tilde{r}_i| \leq (1/6)p^{3/2}C_3||\hat{\eta}_i - \eta_i||^3 \ , \tag{E.27}$$

where the bound follows by part (e) of Assumption 3.1, Jensen's inequality, and the definition of Euclidean norm. It follows

$$n^{-1/2}\sum_{i=1}^{n}(\hat{m}_i - m_i)/J_0 = I_1 + I_2 + I_3 \ ,$$

where

$$\begin{aligned} I_1 &= n^{-1/2}\sum_{i=1}^{n}(\hat{\eta}_i - \eta_i)^\top \partial_\eta m_i/J_0 \\ I_2 &= n^{-1/2}\sum_{i=1}^{n}(\hat{\eta}_i - \eta_i)^\top (\partial_\eta^2 m_i/(2J_0))(\hat{\eta}_i - \eta_i) \\ I_3 &= n^{-1/2}\sum_{i=1}^{n}\tilde{r}_i/J_0 \end{aligned}$$

In the claims below I show that $I_1 = \mathcal{T}^l_{n,K} + o_p(n^{-\zeta})$, $I_2 = \mathcal{T}^{nl}_{n,K} + o_p(n^{-\zeta})$, and $I_3 = o_p(n^{-\zeta})$, which is sufficient to complete the proof of part 4. Furthermore, if Assumption 3.3 holds, then Proposition C.5 implies $\lim_{n\to\infty}\inf_{K\leq n} Var[n^{2\varphi_1-1}\mathcal{T}^{nl}_{n,K}] > 0$; and if Assumption A.1 holds, then Proposition C.3 implies $\lim_{n\to\infty}\inf_{K\leq n} Var[n^{\varphi_1}\mathcal{T}^{l}_{n,K}] > 0$.

Claim 1: $I_1 = \mathcal{T}^l_{n,K} + O_p(n^{-2\min\{\varphi_1,\varphi_2\}})$. By part (a) of Assumption 3.2, it follows

$$I_1 = I_{1,1} + I_{1,2} \ ,$$

where $\hat{R}_i = \hat{R}(X_i)$ for $i \in \mathcal{I}_k$ and

$$I_{1,1} = n^{-1/2} \sum_{i=1}^{n} (\Delta_i)^\top \partial_\eta m_i / J_0$$
$$I_{1,2} = n^{-1/2} \sum_{i=1}^{n} (n_0^{-2\varphi_1} \hat{R}_i)^\top \partial_\eta m_i / J_0$$

By definition of $\mathcal{T}^l_{n,K}$ in (A-3), it follows that $I_{1,1} = \mathcal{T}^l_{n,K}$. Since $\partial_\eta m_i = \partial_\eta \psi_i^b - \theta_0 \partial_\eta \psi_i^a$ and $|\theta_0| \leq M^{1/4}/C_0$, it follows that $I_{1,2}$ is $O_p(n^{-2\min\{\varphi_1,\varphi_2\}})$ due to proof of Claim 3 in Part 1 of this lemma.

Claim 2: $I_2 = \mathcal{T}^{nl}_{n,K} + O_p(n^{1/2-3\min\{\varphi_1,\varphi_2\}})$. By part (a) of Assumption 3.2, it follows

$$I_2 = I_{2,1} + 2I_{2,2} + I_{2,3}$$

where $\hat{R}_i = \hat{R}(X_i)$ for $i \in \mathcal{I}_k$ and

$$I_{2,1} = n^{-1/2} \sum_{i=1}^{n} (\Delta_i)^\top (\partial^2_\eta m_i / (2J_0))(\Delta_i)$$
$$I_{2,2} = n^{-1/2} \sum_{i=1}^{n} (\Delta_i)^\top (\partial^2_\eta m_i / (2J_0))(n_0^{-2\varphi_1} \hat{R}_i)$$
$$I_{2,3} = n^{-1/2} \sum_{i=1}^{n} (n_0^{-2\varphi_1} \hat{R}_i)^\top (\partial^2_\eta m_i / (2J_0))(n_0^{-2\varphi_1} \hat{R}_i)$$

By definition of $\mathcal{T}^{nl}_{n,K}$ in (3.12), it follows that $I_{2,1} = \mathcal{T}^{nl}_{n,K}$. In what follows, I prove claims that imply $I_{2,j} = o_p(n^{-\zeta})$ for $j = 2, 3$ using that $\zeta < 3\varphi_1 - 1/2$ since $\varphi_1 < 1/2$, which is sufficient to complete the proof of claim 2.

Claim 2.1: $I_{2,2} = O_p(n^{1/2-3\min\{\varphi_1,\varphi_2\}})$. To see this, consider the following derivations,

$$|I_{2,2}| \overset{(1)}{\leq} n^{1/2} \times pC_2 \left(n^{-1} \sum_{i=1}^{n} ||\Delta_i|| \times ||n_0^{-2\min\{\varphi_1,\varphi_2\}} \hat{R}_i|| \right)$$
$$\overset{(2)}{\leq} n^{1/2} n_0^{-2\min\{\varphi_1,\varphi_2\}} \times pC_2 \left(n^{-1} \sum_{i=1}^{n} ||\Delta_i||^2 \right)^{1/2} \times \left(n^{-1} \sum_{i=1}^{n} ||\hat{R}_i||^2 \right)^{1/2}$$

$$\overset{(3)}{=} n^{1/2} n_0^{-2\min\{\varphi_1,\varphi_2\}} \times O_p(n^{-\min\{\varphi_1,\varphi_2\}}) \times O_p(1)$$
$$= O_p(n^{1/2-3\min\{\varphi_1,\varphi_2\}})$$

where (1) holds by triangle inequality, part (e) of Assumption 3.1, Jensen's inequality, and definition of Euclidean norm, (2) holds by Cauchy-Schwartz inequality, (3) holds by Lemma C.1 and by part (c) of Assumption 3.2 and Markov's inequality.

Claim 2.2: $I_{2,3} = O_p(n^{1/2-4\min\{\varphi_1,\varphi_2\}})$. The proof is similar to Claim 2.1; therefore, it is omitted.

Claim 3: $I_3 = O_p(n^{1/2-3\min\{\varphi_1,\varphi_2\}})$. Using (E.27), it follows

$$|I_3| \leq (1/6) p^{3/2} C_3 / J_0 n^{-1/2} \sum_{i=1}^{n} ||\hat{\eta}_i - \eta_i||^3 .$$

In what follows, I prove that $n^{-1/2} \sum_{i=1}^{n} ||\hat{\eta}_i - \eta_i||^3$ is $O_p(n^{1/2-3\varphi_1})$.

By part (a) of Assumption 3.2 and since $\varphi_1 \leq \varphi_2$, it follows

$$\hat{\eta}_i - \eta_i = \Delta_i + n_0^{-2\min\{\varphi_1,\varphi_2\}} \hat{R}_i$$

where $\Delta_i = \Delta_i^l + \Delta_i^b$ and $\hat{R}_i = \hat{R}(X_i)$. Using triangle inequality and Loeve's inequality (Davidson (1994, Theorem 9.28)) in the previous expression, it follows

$$||\hat{\eta}_i - \eta_i||^3 \leq 2^2 \left(||\Delta_i||^3 + n_0^{-6\min\{\varphi_1,\varphi_2\}} ||\hat{R}_i||^3 \right)$$

which implies

$$n^{-1} \sum_{i=1}^{n} ||\hat{\eta}_i - \eta_i||^3 \leq 2^2 (I_{3,1} + I_{3,2})$$

where

$$I_{3,1} = n^{-1} \sum_{i=1}^{n} ||\Delta_i||^3$$
$$I_{3,2} = n^{-1} \sum_{i=1}^{n} n_0^{-6\varphi_1} ||\hat{R}_i||^3$$

To complete the proof of claim 3, it is sufficient to show $I_{3,1} = O_p(n^{-3\varphi_1})$ and $I_{3,2} = O_p(n^{1/2-6\varphi_1})$, since they imply $n^{-1/2}\sum_{i=1}^{n}||\hat{\eta}_i - \eta_i||^3$ is $O_p(n^{1/2-3\varphi_1})$.

Claim 3.1: $I_{3,1} = O_p(n^{-3\min\{\varphi_1,\varphi_2\}})$. The proof is a direct result of Lemma C.1 and Markov's inequality; therefore, it is omitted.

Claim 3.2: $I_{3,2} = O_p(n^{1/2-6\min\{\varphi_1,\varphi_2\}})$. Consider the following derivations,

$$\begin{aligned} I_{3,2} &\overset{(1)}{\leq} n^{1/2}n^{-6\min\{\varphi_1,\varphi_2\}}\left(n^{-1}\sum_{i=1}^{n}||\hat{R}_i||^2\right)^{3/2} \\ &\overset{(2)}{=} n^{1/2}\times n^{-6\min\{\varphi_1,\varphi_2\}}\times O_p(1) \end{aligned}$$

where (1) holds by Loeve's inequality (Davidson (1994, Theorem 9.28)), and (2) holds by part (c) of Assumption 3.2 with Markov's inequality. This completes the proof of claim 3.3.

Claim 1 and Claim 2 in the proof of Part 1 imply that $\mathcal{T}^{l}_{n,K}$ is $O_p(n^{-\min\{\varphi_1,\varphi_2\}})$. By the same argument used in the proof of Part 2 to bound the non-linear expression (but using Lemma E.8 instead of part 4 of Lemma C.4), it follows that $\mathcal{T}^{nl}_{n,K}$ is $O_p(n^{1/2-2\min\{\varphi_1,\varphi_2\}})$. □

E.10 Proof of Lemma C.3

Proof. In what follows, the results are proved for any given sequence K that diverges to infinity as n diverges to infinity, which is sufficient to guarantee the result of this lemma.

Part 1: The proof of (C-3) has two steps. The first step shows

$$E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right)^2\right]\leq Cn^{-2\min\{\varphi_1,\varphi_2\}}\ , \qquad \text{(E.28)}$$

for some positive constant $C = C(p, C_1, M_1, M_2)$. The second step shows

$$E\left[\max_{k=1,\dots,K}\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right)^2\right]\leq CKn^{-1/2}n^{-2\min\{\varphi_1,\varphi_2\}+1/2} \tag{E.29}$$

which is sufficient to prove (C-3) by using Markov's inequality and $1/2 < 2\min\{\varphi_1,\varphi_2\}$.

Step 1: Consider the following derivation

$$E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right)^2 \mid (W_j : j\notin\mathcal{I}_k)\right]$$

which is equal to

$$\begin{aligned}
&\overset{(1)}{=} E\left[n_k^{-1}\sum_{i\in\mathcal{I}_k}\left((\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right)^2 \mid (W_j : j\notin\mathcal{I}_k)\right]\\
&\overset{(2)}{=} n_k^{-1}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top E\left[(\partial_\eta\psi^z)(\partial_\eta\psi^z)^\top \mid (W_j : j\notin\mathcal{I}_k)\right](\hat{\eta}_i-\eta_i)\\
&\overset{(3)}{\leq} C_1\times p\times n_k^{-1}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_k(X_i)-\eta_0(X_i)||^2
\end{aligned}$$

where (1) by the i.i.d zero mean of the random vectors $\{(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z : i\in\mathcal{I}_k\}$ conditional on $(W_j : j\notin\mathcal{I}_k)$ that holds by part (b) of Assumption 3.1, (2) holds since $\hat{\eta}_i-\eta_i$ are not random conditional on $(W_j : j\notin\mathcal{I}_k)$, and (3) by part (d) of Assumption 3.1 and Loeve's inequality (Davidson (1994, Theorem 9.28)).

Using the previous derivations, it follows

$$\begin{aligned}
E\left[\left(n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right)^2\right] &\leq E\left[C_1\times p\times n_k^{-1}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_k(X_i)-\eta_0(X_i)||^2\right]\\
&\overset{(1)}{=} C_1\times p\times E\left[||\hat{\eta}_k(X_i)-\eta_0(X_i)||^2\right]\\
&\overset{(2)}{\leq} Cn^{-2\min\{\varphi_1,\varphi_2\}}
\end{aligned}$$

for some positive constant $C = C(p, C_1, M_1, M_2)$, where (1) holds since $\hat{\eta}_k(X_i)-\eta_0(X_i)$ are i.i.d. for $i\in\mathcal{I}_k$, and (2) by part 2 of Lemma C.4 and (E.31), which defines the constant C. This completes the proof of step 1.

Step 2: Note that the maximum of K positive number is bounded by their sum. Using this observation and (E.28), it follows (E.29).

Part 2: The proof of (C-4) is similar to the proof part 1. It follows from the following inequality:

$$
\begin{aligned}
E\left[\max_{k=1,\dots,K} n_k^{-1}\sum_{i\in\mathcal{I}_K}||\hat{\eta}_i-\eta_i||^2\right] &\le E\left[\sum_{k=1}^{K}\left(n_k^{-1}\sum_{i\in\mathcal{I}_K}||\hat{\eta}_i-\eta_i||^2\right)\right] \\
&= KE\left[||\hat{\eta}_i-\eta_i||^2\right] \\
&\le Kn^{-1/2}O(n^{-2\min\{\varphi_1,\varphi_2\}+1/2}) \ ,
\end{aligned}
$$

which goes to zero since $1/2 < 2\min\{\varphi_1,\varphi_2\}$, and this is sufficient to prove (C-4) by using Markov's inequality.

Part 3: The proof of (C-5) follows from (C-3), by using $\partial_\eta m_i = \partial_\eta\psi^b - \theta_0\partial_\eta\psi^a$ and $|\theta_0| \le M_1^{1/4}/C_0$ (due to parts (a) and (c) of Assumptions 3.1 and (2.3)).

Part 4: The proof of (C-6) follows from (C-3) and (C-4) and the following inequality

$$
\left|n_k^{-1}\sum_{i\in\mathcal{I}_k}\hat{\psi}_i^a-\psi_i^a\right| \le n_k^{-1/2}\left|n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\hat{\eta}_i-\eta_i)^\top\partial_\eta\psi_i^z\right| + C_2pn_k^{-1}\sum_{i\in\mathcal{I}_k}||\hat{\eta}_i-\eta_i||^2
$$

which holds due to Taylor expansion and mean valued theorem, part (e) of Assumption 3.1, and Loeves' inequality. □

E.11 Proof of Lemma C.4

Proof. For $i\in\mathcal{I}_k$, denote $\Delta_i = \Delta_i^b + \Delta_i^l$, where

$$
\begin{aligned}
\Delta_i^l &= n_0^{-\varphi_1}n_0^{-1/2}\sum_{j\notin\mathcal{I}_k}\delta_{n_0,j,i} \ , \\
\Delta_i^b &= n_0^{-\varphi_2}n_0^{-1}\sum_{j\notin\mathcal{I}_k}b_{n_0,j,i} \ ,
\end{aligned}
$$

Here, $\delta_{n_0,j,i} = \delta_{n_0}(W_j,X_i)$ and $b_{n_0,j,i} = b_{n_0}(X_j,X_i)$, and δ_{n_0} and b_{n_0} are functions satisfying Assumption 3.2. In what follows, the results are proved for any given sequence K that diverges to infinity as n diverges to infinity, which is sufficient to guarantee the result of this lemma.

Part 1: By part (a) of Assumption 3.2,

$$\hat{\eta}_i - \eta_i = \Delta_i^l + \Delta_i^b + n_0^{-2\min\{\varphi_1,\varphi_2\}}\hat{R}_i$$

and by Loeve's inequality (Davidson (1994, Theorem 9.28)),

$$||\hat{\eta}_i - \eta_i||^4 \leq 3^3 \left(||\Delta_i^l||^4 + ||\Delta_i^b||^4 + ||n_0^{-2\min\{\varphi_1,\varphi_2\}}\hat{R}_i||^4 \right) .$$

Using the previous inequality, it follows

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}||\hat{\eta}_i - \eta_i||^4 &\leq 3^3 \left(n^{-1}\sum_{i=1}^{n}||\Delta_i^l||^4 + n^{-1}\sum_{i=1}^{n}||\Delta_i^b||^4 + n_0^{-8\min\{\varphi_1,\varphi_2\}}n^{-1}\sum_{i=1}^{n}||\hat{R}_i||^4 \right) \\ &\overset{(1)}{\leq} 3^3 \left(O_p(n^{-4\min\{\varphi_1,\varphi_2\}}) + n_0^{-8\min\{\varphi_1,\varphi_2\}}n^{-1}\sum_{i=1}^{n}||\hat{R}_i||^4 \right) \\ &\overset{(2)}{=} O_p(n^{-4\min\{\varphi_1,\varphi_2\}}) , \end{aligned} \tag{E.30}$$

where (1) holds by Markov's inequality and Lemma C.1, and (2) holds by part (d) of Assumption 3.2 and since $n_0 = ((K-1)/K)n$, which completes the proof of part 1.

Part 2: By part (a) of Assumption 3.2,

$$\hat{\eta}_i - \eta_i = \Delta_i^l + \Delta_i^b + n_0^{-2\min\{\varphi_1,\varphi_2\}}\hat{R}_i$$

and by Loeve's inequality (Davidson (1994, Theorem 9.28)),

$$||\hat{\eta}_i - \eta_i||^2 \leq 3||\Delta_i^l||^2 + 3||\Delta_i^b||^2 + 3||n_0^{-2\varphi_1}\hat{R}_i||^2 .$$

Using the previous inequality, it follows

$$\begin{aligned} E[||\hat{\eta}_i - \eta_i||^2] &\leq 3E[||\Delta_i^l||^2] + 3E[||\Delta_i^b||^2] + 3E[||n_0^{-2\varphi_1}\hat{R}_i||^2] \\ &\overset{(1)}{\leq} 3E[||\Delta_i^l||^4]^{1/2} + 3E[||\Delta_i^b||^4]^{1/2} + 3n_0^{-2\min\{\varphi_1,\varphi_2\}}O(1) \\ &\overset{(2)}{=} O(n^{-2\min\{\varphi_1,\varphi_2\}}) , \end{aligned} \tag{E.31}$$

where (1) holds by Jensen's inequality and part (c) of Assumption 3.2, and (2) holds by Lemma C.1 and since $n_0 = ((K-1)/K)n$. This completes the proof of part 2.

Part 3: It follows from parts 2 and Markov's inequality.

Part 4: It follows from part 2 and by using that $n^{1/2-\min\{\varphi_1,\varphi_2\}} = o(1)$. □

Lemma E.1. *Suppose Assumptions 3.1 and 3.2 hold. In addition, assume K is such that $K \le n$, $K \to \infty$ and $K/n^\gamma \to c \in [0,+\infty)$ as $n \to \infty$.*

1. *If $\gamma = 1/2$, then*

$$\max_{k=1,\ldots,K}\left|1 + n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi^a(W_i,\eta_i) - J_0)/J_0\right|^{-1} = O_p(1)$$

2. *If $\gamma = 1/2$, $1/4 < \min\{\varphi_1,\varphi_2\}$ and $\varphi_1 \le 1/2$, then*

$$\max_{k=1,\ldots,K}\left|1 + n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi^a(W_i,\hat\eta_i) - J_0)/J_0\right|^{-1} = O_p(1)$$

3. *If $\gamma = 1$, then*

$$\left|n^{-1}\sum_{i=1}^{n}\psi^a(W_i,\eta_i)/J_0\right|^{-1} = O_p(1)$$

4. *If $\gamma = 1$ and $\varphi_1 > 1/4$, then*

$$\left|n^{-1}\sum_{i=1}^{n}\psi^a(W_i,\hat\eta_i)/J_0\right|^{-1} = O_p(1)$$

where $\eta_i = \eta_0(X_i)$, $\hat\eta_i$ is as in (2.5), and $J_0 = E[\psi^a(W_i,\eta_i)]$.

Proof. **Part 1:** Consider $M > 1$ and the following derivations

$$P\left(\max_{k=1,\ldots,K}\left|1 + n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a - J_0)/J_0\right|^{-1} \le M\right)$$

$$
\begin{aligned}
&\overset{(1)}{=} P\left(\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|^{-1}\leq M\right)^K\\
&= P\left(1/M\leq\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|\right)^K\\
&\geq P\left(1/M\leq 1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right)^K\\
&=\left\{1-P\left(n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0<-(M-1)/M\right)\right\}^K\\
&\overset{(2)}{\geq}\left\{1-P\left(\left|n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|>(M-1)/M\right)\right\}^K\\
&\overset{(3)}{\geq}\left\{1-((M-1)/M)^4 n_k^{-2}E\left[\left|n_k^{-1/2}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|^4\right]\right\}^K\\
&\overset{(4)}{\geq}\left\{1-((M-1)/M)^4 n_k^{-2}10E\left[|(\psi_i^a-J_0)/J_0|^4\right]\right\}^K\\
&\overset{(5)}{\geq} 1-((M-1)/M)^4 n^{-2}K^2 10E\left[|(\psi_i^a-J_0)/J_0|^4\right]K\\
&\overset{(6)}{\geq} 1-n^{-2+3/2}(Kn^{-1/2})^3((M-1)/M)^4 O(1)\\
&\overset{(7)}{\geq} 1-o(1)
\end{aligned}
$$

(1) holds since $\{\psi_i^a : 1\leq i\leq n\}$ are i.i.d. random variables and because $\{\mathcal{I}_k : 1\leq k\leq K\}$ defines a partition of $\{1,\ldots,n\}$, (2) holds since $M>1$, (3) holds by Markov's inequality, (4) holds since $\{\psi_i^a-J_0 : i\in\mathcal{I}_k\}$ are zero mean i.i.d. random variables, (5) holds by Bernoulli's inequality, (6) holds by parts (a) and (c) of Assumption 3.1, and (7) holds since $K=O(n^{1/2})$.

Part 2: Define the event $E_{n,\epsilon}=\{\max_{k=1,\ldots,K}\left|n_k^{-1}\sum_{i\in\mathcal{I}_k}(\hat{\psi}_i^a-\psi_i^a)/J_0\right|<\epsilon\}$. Now, consider an small $\epsilon>0$ and $M>1$ such that $(1/M+\epsilon)^{-1}>1$ and the following derivations

$$
P\left(\max_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0+(\hat{\psi}_i^a-\psi_i^a)/J_0\right|^{-1}\leq M\right)
$$

$$
\begin{aligned}
&= P\left(\min_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0+(\hat{\psi}_i^a-\psi_i^a)/J_0\right|\geq 1/M\right)\\
&\geq P\left(\min_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0+(\hat{\psi}_i^a-\psi_i^a)/J_0\right|\geq 1/M,\ E_{n,\epsilon}\right)\\
&\overset{(1)}{\geq} P\left(\min_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|\geq 1/M+\epsilon,\ E_{n,\epsilon}\right)\\
&\geq P\left(\min_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|\geq 1/M+\epsilon\right)-P(E_{n,\epsilon}^c)\\
&\overset{(2)}{\geq} P\left(\max_{k=1,\ldots,K}\left|1+n_k^{-1}\sum_{i\in\mathcal{I}_k}(\psi_i^a-J_0)/J_0\right|^{-1}\leq (1/M+\epsilon)^{-1}\right)-o(1)\\
&\overset{(3)}{\geq} 1-o(1)-o(1)
\end{aligned}
$$

where (1) holds because $\min_{k=1,\ldots,K}-\left|n_k^{-1}\sum_{i\in\mathcal{I}_k}(\hat{\psi}_i^a-\psi_i^a)/J_0\right|>-\epsilon$ conditional on the event $E_{n,\epsilon}$ and triangular inequality, (2) holds since $P(E_{n,\epsilon}^c)=o(1)$ due to Lemma C.3 (here I use $1/2<2\min\{\varphi_1,\varphi_2\}$ and $\varphi_1\leq 1/2$), and (3) holds by the same arguments presented in the proof of Part 1 by using $(1/M+\epsilon)^{-1}$ instead of M; therefore, it is omitted.

Part 3: Consider $M>1$ and $\tilde{M}>1$ such that $\tilde{M}<((M-1)/M)n^{1/2}$ and the following derivations,

$$
\begin{aligned}
P\left(\left|n^{-1}\sum_{i=1}^{n}\psi_i^a/J_0\right|^{-1}>M\right) &= P\left(\left|1+n^{-1}\sum_{i=1}^{n}(\psi_i^a-J_0)/J_0\right|<1/M\right)\\
&\overset{(1)}{\leq} P\left(n^{-1/2}\sum_{i=1}^{n}(\psi_i^a-J_0)/J_0<-((M-1)/M)n^{1/2}\right)\\
&\overset{(2)}{\leq} P\left(n^{-1/2}\sum_{i=1}^{n}(\psi_i^a-J_0)/J_0<-\tilde{M}\right)\\
&\overset{(3)}{=} \Phi(-\tilde{M}/\sigma_a)+o(1)
\end{aligned}
$$

where (1) holds since $M>1$, (2) holds by definition of $\tilde{M}$, and (3) holds by CLT as $n\to\infty$ (here, σ_a^2 is as in (C-2)). To complete the proof, note that $\Phi(-\tilde{M}/\sigma_a)\to 0$

as $\tilde{M} \to \infty$.

Part 4: Define the event $E_{n,\epsilon} = \{\left|n^{-1}\sum_{i=1}^{n}(\hat{\psi}_i^a - \psi_i^a)/J_0\right| < \epsilon\}$. Now consider an small $\epsilon > 0$ and $M > 1$ such that $(1/M + \epsilon)^{-1} > 1$. Note that $P(E_{n,\epsilon}^c) = o(1)$ due to Lemma C.2 (here I use $\min\{\varphi_1, \varphi_2\} > 1/4$). The proof is completed by similar arguments presented in part 2 and part 3; therefore, it is omitted. □

References

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, Advanced Texts in Econometrics, Oxford University Press.