Supplement to "Detection of Multiple Structural Breaks in Large Covariance Matrices"

Yu-Ning Li* Degui Li† Piotr Fryzlewicz‡

This version: January 17, 2022

In this supplement, we provide the detailed proofs of the main theoretical results as well as additional simulation studies. Appendix B gives a simple motivating example for the factor model transformation stated in Proposition 2.1, Appendix C proves Propositions 2.1 and 3.1 for the transformed factor model, Appendix D proves the asymptotic properties of the WBS-Cov for the common components, Appendix E proves the asymptotic properties of the WSBS-Cov for the idiosyncratic error components, and Appendix F reports additional simulation results. Throughout the supplemental document, we let M be a generic positive constant whose value may change from line to line.

Appendix B: A motivating example of factor model transformation

In this appendix, we provide a simple motivating example to show how to transform breaks in factor loadings of a factor model to breaks in covariance of (transformed) factors, a transformation mechanism summarised in Proposition 2.1. Consider an approximate factor model with $K_1 = 2$:

$$\mathbf{X}_t = \mathbf{\Lambda}^0_{k+1}\mathbf{F}_t + \boldsymbol{\epsilon}_t,\ \ \eta^c_k + 1 \leqslant t \leqslant \eta^c_{k+1},$$

where $k = 0, 1, 2$, $\eta^c_0 = 0$ and $\eta^c_3 = n$. We assume that the number of factors and the column ranks of the factor loading matrices are all equal to r. Furthermore, we assume the column rank of $(\mathbf{\Lambda}^0_1, \mathbf{\Lambda}^0_2)$ is r, indicating that there exists an $r \times r$ matrix $\mathbf{T}$ such that $\mathbf{\Lambda}^0_2 = \mathbf{\Lambda}^0_1\mathbf{T}$; and the column

*School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China.
†The corresponding author, Department of Mathematics, University of York, YO10 5DD, UK. Email address: degui.li@york.ac.uk.
‡Department of Statistics, London School of Economics, WC2A 2AE, UK.

rank of $(\boldsymbol{\Lambda}_2^0, \boldsymbol{\Lambda}_3^0)$ is $2r$ (full column rank), indicating that $\boldsymbol{\Lambda}_2^0$ and $\boldsymbol{\Lambda}_3^0$ are linearly independent. Han and Inoue (2015) call the first break a "type 2 break" and the second break a "type 1 break"[1]. The transformed factor loadings and factors can be defined as $\boldsymbol{\Lambda}^\star = (\boldsymbol{\Lambda}_1^0, \boldsymbol{\Lambda}_3^0)$ and

$$\mathbf{F}_t^\star = \begin{cases} \left(\mathbf{F}_t^\intercal, \mathbf{0}^\intercal\right)^\intercal, & 1 \leqslant t \leqslant \eta_1^c, \\ \left(\mathbf{F}_t^\intercal \mathbf{T}^\intercal, \mathbf{0}^\intercal\right)^\intercal, & \eta_1^c + 1 \leqslant t \leqslant \eta_2^c, \\ \left(\mathbf{0}^\intercal, \mathbf{F}_t^\intercal\right)^\intercal, & \eta_2^c + 1 \leqslant t \leqslant n, \end{cases}$$

respectively. As a result, the original factor model can be equivalently written as

$$\mathbf{X}_t = \boldsymbol{\Lambda}^\star \mathbf{F}_t^\star + \boldsymbol{\epsilon}_t, \ \ t = 1, \cdots, n, \tag{B.1}$$

the same as (2.4) in Proposition 2.1. Note that the number of latent common factors has increased from r to $2r$ in model (B.1). Letting $\boldsymbol{\Sigma}(\mathbf{F}) = \mathsf{Cov}(\mathbf{F}_t)$, $\boldsymbol{\Sigma}_t(\boldsymbol{\Lambda}, \mathbf{F})$ in (1.3) can be re-formulated as

$$\boldsymbol{\Sigma}_t(\boldsymbol{\Lambda}, \mathbf{F}) = \begin{cases} \boldsymbol{\Lambda}^\star \mathsf{diag}\{\boldsymbol{\Sigma}(\mathbf{F}), \mathbf{O}\} (\boldsymbol{\Lambda}^\star)^\intercal, & 1 \leqslant t \leqslant \eta_1^c, \\ \boldsymbol{\Lambda}^\star \mathsf{diag}\left\{\mathbf{T}\boldsymbol{\Sigma}(\mathbf{F})\mathbf{T}^\intercal, \mathbf{O}\right\} (\boldsymbol{\Lambda}^\star)^\intercal, & \eta_1^c + 1 \leqslant t \leqslant \eta_2^c, \\ \boldsymbol{\Lambda}^\star \mathsf{diag}\{\mathbf{O}, \boldsymbol{\Sigma}(\mathbf{F})\} (\boldsymbol{\Lambda}^\star)^\intercal, & \eta_2^c + 1 \leqslant t \leqslant n. \end{cases} \tag{B.2}$$

where $\mathsf{diag}\{\mathbf{A}, \mathbf{B}\}$ denotes a block diagonal matrix with $\mathbf{A}$ and $\mathbf{B}$ being two square matrices and $\mathbf{O}$ denotes a null matrix whose size may change from one place to another. As the transformed factor loading matrix $\boldsymbol{\Lambda}^\star$ is time-invariant, structural breaks on $\boldsymbol{\Sigma}_t(\boldsymbol{\Lambda}, \mathbf{F})$ are purely caused by sudden changes in the covariance matrix for the transformed factors $\mathbf{F}_t^\star$.

Appendix C: Proofs of Propositions 2.1 and 3.1

PROOF OF PROPOSITION 2.1. Let $\mathcal{L}(\boldsymbol{\Lambda})$ be the space spanned by the column vectors of $\boldsymbol{\Lambda}_k^0$, $k = 1, \cdots, K_1 + 1$, and q_0 be its dimension. It is straightforward to show that

$$\max_{1 \leqslant k \leqslant K_1+1} \underline{r}_k \leqslant q_0 \leqslant \sum_{k=1}^{K_1+1} \underline{r}_k, \tag{C.1}$$

where $\underline{r}_k$ denotes the column rank of $\boldsymbol{\Lambda}_k^0$. As $\mathcal{L}(\boldsymbol{\Lambda})$ is a q_0-dimensional subspace of $\mathbb{R}^d$, we may construct a $d \times q_0$ matrix $\boldsymbol{\Lambda}^\star$ by stacking a group of basis for this vector space. Noting that the column vectors of $\boldsymbol{\Lambda}_k^0$ lie in the space $\mathcal{L}(\boldsymbol{\Lambda})$, there must exist a $q_0 \times r_k$ transformation matrix $\mathbf{T}_k$

[1]For the intermediate case with the rank of $(\boldsymbol{\Lambda}_k^0, \boldsymbol{\Lambda}_{k+1}^0)$ strictly between r and $2r$, Han and Inoue (2015) call it a "type 3 break". In this case, the factors and factor loadings can be similarly transformed by separating the linearly independent columns of $\boldsymbol{\Lambda}_k^0$ and $\boldsymbol{\Lambda}_{k+1}^0$ from the linearly dependent ones.

such that

$$\mathbf{\Lambda}_k^0 = \mathbf{\Lambda}^\star \mathbf{T}_k, \quad k = 1, \cdots, K_1 + 1. \tag{C.2}$$

Then the transformed factors can be defined as

$$\mathbf{F}_t^\star = \begin{cases} \mathbf{T}_1 \mathbf{F}_{t,1}, & 1 \leqslant t \leqslant \eta_1^c, \\ \mathbf{T}_2 \mathbf{F}_{t,2}, & \eta_1^c + 1 \leqslant t \leqslant \eta_2^c, \\ \vdots & \vdots \\ \mathbf{T}_{K_1+1} \mathbf{F}_{t,K_1+1}, & \eta_{K_1}^c + 1 \leqslant t \leqslant n. \end{cases} \tag{C.3}$$

With (2.2), (C.2) and (C.3), we readily have that, when $\eta_{k-1}^c + 1 \leqslant t \leqslant \eta_k^c$,

$$\mathbf{X}_t = \mathbf{\Lambda}_k^0 \mathbf{F}_{t,k} + \boldsymbol{\epsilon}_t = \mathbf{\Lambda}^\star \mathbf{T}_k \mathbf{F}_{t,k} + \boldsymbol{\epsilon}_t = \mathbf{\Lambda}^\star \mathbf{F}_t^\star + \boldsymbol{\epsilon}_t. \tag{C.4}$$

The inequalities in (2.5) can be proved by combining (C.1) and the fact of $\underline{r}_k \leqslant r_k$. □

PROOF OF PROPOSITION 3.1. Letting $\mathcal{L}(\mathbf{\Lambda})$ be defined as in the proof of Proposition 2.1, we may obtain a group of basis vectors for $\mathcal{L}(\mathbf{\Lambda})$ directly from the column vectors of $\mathbf{\Lambda}_k^0$, for $k = 1, \cdots K_1+1$. Specifically, define $\mathbf{\Lambda}^\star = \left[\mathbf{\Lambda}_1^0, \cdots, \mathbf{\Lambda}_{K_1+1}^0\right] \mathbf{S}$, where $\mathbf{S}$ is a $\sum_{k=1}^{K_1+1} r_k \times q_0$ selection matrix whose entries are either 1 or 0. By Assumption 2(ii) in Appendix A, $\mathbf{\Lambda}^\star$ is of full column rank and the smallest eigenvalue of $\frac{1}{d}\mathbf{\Lambda}^{\star\intercal}\mathbf{\Lambda}^\star$ is positive and bounded away from zero.

By (C.2) and von Neumann's trace inequality (e.g., Marshall, Olkin and Arnold, 2011), we have

$$\mathsf{tr}\left(\frac{1}{d}\mathbf{\Lambda}_k^{0\intercal}\mathbf{\Lambda}_k^0\right) = \mathsf{tr}\left(\frac{1}{d}\mathbf{T}_k^\intercal\mathbf{\Lambda}^{\star\intercal}\mathbf{\Lambda}^\star\mathbf{T}_k\right) = \mathsf{tr}\left(\frac{1}{d}\mathbf{\Lambda}^{\star\intercal}\mathbf{\Lambda}^\star\mathbf{T}_k\mathbf{T}_k^\intercal\right) \geqslant \sum_{j=1}^{q_0}\mu_j\left(\frac{1}{d}\mathbf{\Lambda}^{\star\intercal}\mathbf{\Lambda}^\star\right)\mu_{q_0-j+1}\left(\mathbf{T}_k\mathbf{T}_k^\intercal\right),$$

where $\mathsf{tr}(\cdot)$ denotes trace of a square matrix. This indicates that

$$\mu_1\left(\mathbf{T}_k\mathbf{T}_k^\intercal\right) \leqslant \mathsf{tr}\left(\frac{1}{d}\mathbf{\Lambda}_k^{0\intercal}\mathbf{\Lambda}_k^0\right)/\mu_{q_0}\left(\frac{1}{d}\mathbf{\Lambda}^{\star\intercal}\mathbf{\Lambda}^\star\right),$$

which is bounded uniformly over $k = 1, \cdots, K_1 + 1$ by Assumption 2(ii), and thus

$$\max_{1\leqslant k\leqslant K_1+1}\|\mathbf{T}_k\|_F^2 = \max_{1\leqslant k\leqslant K_1+1}\mathsf{tr}\left(\mathbf{T}_k\mathbf{T}_k^\intercal\right) \leqslant \left(\max_{1\leqslant k\leqslant K_1+1} r_k\right)\cdot\max_{1\leqslant k\leqslant K_1+1}\mu_1\left(\mathbf{T}_k\mathbf{T}_k^\intercal\right) \leqslant M, \tag{C.5}$$

for some positive constant M, as $\max_{1\leqslant k\leqslant K_1+1} r_k$ is bounded by Assumption 2(i). Note that

$$\left\|\frac{1}{n}\sum_{t=1}^n \mathbf{F}_t^\star\mathbf{F}_t^{\star\intercal}\right\|_F \leqslant \sum_{k=1}^{K_1+1}\left\|\frac{1}{\eta_k^c - \eta_{k-1}^c}\sum_{t:\eta_{k-1}^c+1\leqslant t\leqslant\eta_k^c}\mathbf{T}_k\mathbf{F}_{t,k}\mathbf{F}_{t,k}^\intercal\mathbf{T}_k^\intercal\right\|_F$$

$$\leqslant \sum_{k=1}^{K_1+1} \mu_1 \left(\frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t:\eta_{k-1}^c+1 \leqslant t \leqslant \eta_k^c} \mathbf{F}_{t,k}\mathbf{F}_{t,k}^{\intercal} \right) \cdot \|\mathbf{T}_k\|_F^2 . \tag{C.6}$$

As $\eta_k^c - \eta_{k-1}^c \geqslant \kappa_n^c \to \infty$, by Assumption 2(i) and the Law of Large Numbers for the α-mixing sequence (e.g., Lin and Lu, 1996),

$$\frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t:\eta_k^c+1 \leqslant t \leqslant \eta_{k+1}^c} \mathbf{F}_{t,k}\mathbf{F}_{t,k}^{\intercal} \xrightarrow{P} \mathbf{\Sigma}_{\mathbf{F},k}, \quad k = 1, \cdots, K_1 + 1. \tag{C.7}$$

Combining (C.5)–(C.7), we have $\left\| \frac{1}{n} \sum_{t=1}^{n} \mathbf{F}_t^\star \mathbf{F}_t^{\star\intercal} \right\|_F = O_P(1)$.

From (C.7), we readily have that

$$\frac{1}{n} \sum_{t=1}^{n} \mathbf{F}_t^\star \mathbf{F}_t^{\star\intercal} = \sum_{k=1}^{K_1+1} \frac{\eta_k^c - \eta_{k-1}^c}{n} \cdot \frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t:\eta_{k-1}^c+1 \leqslant t \leqslant \eta_k^c} \mathbf{T}_k \mathbf{F}_{t,k}\mathbf{F}_{t,k}^{\intercal} \mathbf{T}_k^{\intercal} \xrightarrow{P} \mathbf{\Sigma}_{\mathbf{F}}, \tag{C.8}$$

where $\mathbf{\Sigma}_{\mathbf{F}}$ is a weighted average of $\mathbf{T}_k \mathbf{\Sigma}_{\mathbf{F},k} \mathbf{T}_k^{\intercal}$ over $k = 1, \cdots, K_1 + 1$, and the weights are strictly positive as $\kappa_n^c \asymp n$. We next only need to show that the smallest eigenvalue of $\mathbf{\Sigma}_{\mathbf{F}}$ is positive, which is to be proved by contradiction. Assume that there exists a q_0-dimensional vector $\boldsymbol{\nu} \neq \mathbf{0}$ such that $\boldsymbol{\nu}^{\intercal} \mathbf{\Sigma}_{\mathbf{F}} \boldsymbol{\nu} = 0$. This implies that $\boldsymbol{\nu}^{\intercal} \mathbf{T}_k \mathbf{\Sigma}_{\mathbf{F},k} \mathbf{T}_k^{\intercal} \boldsymbol{\nu} = 0$, and thus $\mathbf{T}_k^{\intercal} \boldsymbol{\nu} = \mathbf{0}$ for all $k = 1, \cdots, K_1+1$, since $\mathbf{\Sigma}_{\mathbf{F},k}$ is positive definite by Assumption 2(i). As the rank of $\mathbf{\Lambda}^\star$ is q_0, we may write $\boldsymbol{\nu} = (\mathbf{\Lambda}^\star)^{\intercal} \boldsymbol{\nu}^\star$ for some d-dimensional vector $\boldsymbol{\nu}^\star$. Then, by (C.2), we have $\mathbf{T}_k^{\intercal} \boldsymbol{\nu} = \mathbf{T}_k^{\intercal} (\mathbf{\Lambda}^\star)^{\intercal} \boldsymbol{\nu}^\star = (\mathbf{\Lambda}^\star \mathbf{T}_k)^{\intercal} \boldsymbol{\nu}^\star = \left(\mathbf{\Lambda}_k^0\right)^{\intercal} \boldsymbol{\nu}^\star = \mathbf{0}$. However, $\mathbf{\Lambda}^\star$ is constructed from the column vectors of $\mathbf{\Lambda}_k^0$, $k = 1, \cdots, K_1 + 1$, thus we must have $\boldsymbol{\nu} = (\mathbf{\Lambda}^\star)^{\intercal} \boldsymbol{\nu}^\star = \mathbf{0}$, leading to a contradiction. □

Appendix D: Proofs of the WBS-Cov theory for the common components

As construction of the CUSUM statistics relies on PCA estimates of the transformed common factors and idiosyncratic errors, we start with some uniform convergence results for the PCA estimation which are analogous to those derived in Bai and Ng (2002), Fan, Liao and Mincheva (2013) and Han and Inoue (2015).

LEMMA D.1. *Suppose that Assumptions 1, 2 and 3(i) in Appendix A are satisfied. Then, if $\kappa_n^c \asymp n$, we have (i)*

$$\max_{1 \leqslant t \leqslant n} \left\| \widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^\star \right\|_2 = O_P\left(\frac{1}{n^{1/2}} + \frac{n^{2/\delta}}{d^{1/2}} \right), \tag{D.1}$$

where $\delta = \delta_F \wedge \delta_\epsilon$*; and (ii)*

$$\max_{1\leqslant j\leqslant d}\left\|\widehat{\lambda}_j - \left(\mathbf{H}^{-1}\right)^{\intercal}\lambda_j^{\star}\right\|_2 = O_P\left(\left(\frac{\log d}{n}\right)^{1/2} + \frac{n^{2/\delta}}{d^{1/2}}\right), \tag{D.2}$$

if, in addition, Assumption 3(ii) is satisfied and $d = O\left(\exp\{n^{\nu}\}\right)$ *with* $0 \leqslant \nu < 1/5$*, where the rotation matrix* $\mathbf{H}$ *is defined in (3.3), and* $F_t^{\star}$ *and* $\lambda_j^{\star}$ *are the transformed factors and factor loadings.*

PROOF. (i) By the definition of PCA estimation, we may show that

$$\begin{aligned}\boldsymbol{\Omega}_{q_0}\left(\widehat{F}_t - \mathbf{H}F_t^{\star}\right) &= \frac{1}{nd}\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{F}_s F_s^{\star\intercal}\lambda_j^{\star}\epsilon_{tj} + \frac{1}{nd}\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{F}_s F_t^{\star\intercal}\lambda_j^{\star}\epsilon_{sj} + \frac{1}{nd}\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{F}_s \mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right] \\ &\quad + \frac{1}{nd}\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{F}_s\left\{\epsilon_{sj}\epsilon_{tj} - \mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right\} \\ &=: \mathbf{V}_{nt,1} + \mathbf{V}_{nt,2} + \mathbf{V}_{nt,3} + \mathbf{V}_{nt,4}\end{aligned} \tag{D.3}$$

for any $1 \leqslant t \leqslant n$, where $\boldsymbol{\Omega}_{q_0}$ is defined in Section 3.1.

We first consider $\mathbf{V}_{nt,1}$. As

$$\frac{1}{n}\sum_{t=1}^{n}\widehat{F}_t\widehat{F}_t^{\intercal} = \mathbf{I}_{q_0},\ \ \frac{1}{n}\sum_{t=1}^{n}F_t^{\star}F_t^{\star\intercal} = O_P(1),$$

by Proposition 3.1, using the Cauchy-Schwarz inequality, we have

$$\left\|\sum_{s=1}^{n}\widehat{F}_s F_s^{\star\intercal}\right\|_F = O_P(n). \tag{D.4}$$

By the C_r-inequality(e.g., Theorem 9.1.a in Lin and Bai, 2010), we have

$$\max_{1\leqslant t\leqslant n}\mathsf{E}\left[\left\|\sum_{j=1}^{d}\lambda_j^{\star}\epsilon_{tj}\right\|_2^{\delta_\epsilon}\right] \leqslant c_0\cdot\max_{1\leqslant t\leqslant n}\sum_{k=1}^{K_1+1}\mathsf{E}\left[\left\|\sum_{j=1}^{d}\lambda_{k,j}^{0}\epsilon_{tj}\right\|_2^{\delta_\epsilon}\right],$$

where c_0 is a positive constant. Then, by (A.2) in Assumption 3(i), the Bonferroni and Markov inequalities, we may prove that for any $\varepsilon > 0$,

$$\mathsf{P}\left(\max_{1\leqslant t\leqslant n}\left\|\sum_{j=1}^{d}\lambda_j^{\star}\epsilon_{tj}\right\|_2 > c_1 n^{1/\delta_\epsilon}d^{1/2}\right) \leqslant \sum_{t=1}^{n}\mathsf{P}\left(\left\|\sum_{j=1}^{d}\lambda_j^{\star}\epsilon_{tj}\right\|_2 > c_1 n^{1/\delta_\epsilon}d^{1/2}\right)$$

$$\leqslant \max_{1\leqslant t\leqslant n} \mathsf{E}\left[\left\|\sum_{j=1}^{d}\lambda_j^{\star}\epsilon_{tj}\right\|_2^{\delta_\epsilon}\right]/(c_1^{\delta_\epsilon}d^{\delta_\epsilon/2}) \leqslant \frac{c_0\iota_0(K_1+1)}{c_1^{\delta_\epsilon}} < \varepsilon \tag{D.5}$$

by letting $c_1 > [c_0\iota_0(K_1+1)/\varepsilon]^{1/\delta_\epsilon}$, where ι_0 is defined in Assumption 3(i). With (D.4) and (D.5), we readily have that

$$\max_{1\leqslant t\leqslant n}\|\mathbf{V}_{nt,1}\|_2 = O_P(n^{1/\delta_\epsilon}/d^{1/2}). \tag{D.6}$$

By (C.5) and

$$\max_{1\leqslant k\leqslant K_1+1}\max_{\eta_{k-1}^c+1\leqslant t\leqslant \eta_k^c}\mathsf{E}\left[\|\mathbf{F}_{t,k}\|_2^{\delta_F}\right] < \infty$$

in Assumption 1(ii), we can prove that $\max_{1\leqslant t\leqslant n}\|\mathbf{F}_t^{\star}\|_2 = O_P\left(n^{1/\delta_F}\right)$, which together with $\frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{F}}_t\widehat{\mathbf{F}}_t^{\intercal} = \mathbf{I}_{q_0}$, (D.5) and the Cauchy-Schwarz inequality, implies that

$$\begin{aligned}
\max_{1\leqslant t\leqslant n}\|\mathbf{V}_{nt,2}\|_2 &= \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\left\|\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{\mathbf{F}}_s\mathbf{F}_t^{\star\intercal}\lambda_j^{\star}\epsilon_{sj}\right\|_2 \\
&\leqslant \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\|\mathbf{F}_t^{\star}\|_2\left(\sum_{s=1}^{n}\left\|\widehat{\mathbf{F}}_s\right\|_2^2\right)^{1/2}\left(\sum_{s=1}^{n}\left\|\sum_{j=1}^{d}\lambda_j^{\star}\epsilon_{sj}\right\|_2^2\right)^{1/2} \\
&= \frac{1}{nd}\cdot O_P\left(n^{1/\delta_F}\right)\cdot O_P\left(n^{1/2}\right)\cdot O_P\left(n^{1/2+1/\delta_\epsilon}d^{1/2}\right) \\
&= O_P\left(n^{2/\delta}/d^{1/2}\right).
\end{aligned} \tag{D.7}$$

By a basic inequality on the covariance bound for the α-mixing sequence (e.g., Lemma 1.2.4 in Lin and Lu, 1996), we have

$$\sum_{j=1}^{d}\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right] \leqslant 10\cdot\alpha_{|s-t|}^{1-2/\delta_\epsilon}\sum_{j=1}^{d}\left\{\mathsf{E}\left[|\epsilon_{sj}|^{\delta_\epsilon}\right]\right\}^{1/\delta_\epsilon}\left\{\mathsf{E}\left[|\epsilon_{tj}|^{\delta_\epsilon}\right]\right\}^{1/\delta_\epsilon} = O\left(d\cdot[\alpha(|s-t|)]^{1-2/\delta_\epsilon}\right),$$

where $\alpha(s) = \max_{1\leqslant k\leqslant K_1+1}\alpha_k(s)$, indicating that

$$\begin{aligned}
\max_{1\leqslant t\leqslant n}\|\mathbf{V}_{nt,3}\|_2 &= \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\left\|\sum_{s=1}^{n}\sum_{j=1}^{d}\widehat{\mathbf{F}}_s\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right\|_2 \\
&\leqslant \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\left(\sum_{s=1}^{n}\|\widehat{\mathbf{F}}_s\|_2^2\right)^{1/2}\left(\sum_{s=1}^{n}\left(\sum_{j=1}^{d}\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)^2\right)^{1/2} \\
&= \frac{1}{nd}\cdot O_P\left(n^{1/2}\right)\cdot O\left(d\cdot\left[\sum_{k=1}^{n}[\alpha(k)]^{2(1-2/\delta_\epsilon)}\right]^{1/2}\right)
\end{aligned}$$

$$= \quad O_P\left(n^{-1/2}\right) \tag{D.8}$$

as $\sum_{k=1}^{n}[\alpha(k)]^{2(1-2/\delta_\epsilon)} < \infty$ when $\alpha(k)$ decays to zero at a geometric rate.

By (A.3) in Assumption 3(i) and using the Bonferroni and Markov inequalities again, we may show that for any $\varepsilon > 0$,

$$\begin{aligned}
& \mathsf{P}\left(\max_{1\leqslant s,t\leqslant n}\left|\sum_{j=1}^{d}\left(\epsilon_{sj}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)\right| > c_2 n^{2/\delta_\epsilon} d^{1/2}\right) \\
\leqslant\ & \sum_{s=1}^{n}\sum_{t=1}^{n}\mathsf{P}\left(\left|\sum_{j=1}^{d}\left(\epsilon_{sj}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)\right| > c_2 n^{2/\delta_\epsilon} d^{1/2}\right) \\
\leqslant\ & \sum_{s=1}^{n}\sum_{t=1}^{n}\mathsf{E}\left[\left|\sum_{j=1}^{d}\left(\epsilon_{sj}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)\right|^{\delta_\epsilon}\right]/(c_2^{\delta_\epsilon} n^2 d^{\delta_\epsilon/2}) \\
\leqslant\ & \iota_0/c_2^{\delta_\epsilon} < \varepsilon,
\end{aligned}$$

where c_2 is chosen to be larger than $(\iota_0/\varepsilon)^{1/\delta_\epsilon}$. As a result, we have

$$\begin{aligned}
\max_{1\leqslant t\leqslant n}\|\mathbf{V}_{nt,4}\|_2 &= \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\left\|\sum_{s=1}^{n}\widehat{\mathsf{F}}_s\sum_{j=1}^{d}\left(\epsilon_{sj}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)\right\|_2 \\
&\leqslant \frac{1}{nd}\cdot\max_{1\leqslant t\leqslant n}\left(\sum_{s=1}^{n}\|\widehat{\mathsf{F}}_s\|_2^2\right)^{1/2}\left(\sum_{s=1}^{n}\left(\sum_{j=1}^{d}\left(\epsilon_{sj}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{sj}\epsilon_{tj}\right]\right)\right)^2\right)^{1/2} \\
&= \frac{1}{nd}\cdot O_P\left(n^{1/2}\right)\cdot O_P\left(n^{1/2}n^{2/\delta_\epsilon}d^{1/2}\right) \\
&= O_P\left(n^{2/\delta_\epsilon}/d^{1/2}\right).
\end{aligned} \tag{D.9}$$

By (D.3) and (D.6)–(D.9), we can prove (D.1) if $\mathbf{\Omega}_{q_0}$ is asymptotically invertible. The latter can be proved by following the proof of Theorem 3(i) in Chen *et al* (2018). The proof of Lemma D.1(i) is thus completed.

(ii) From Proposition 2.1 in Section 2.2 and by the fact of $\frac{1}{n}\sum_{t=1}^{n}\widehat{\mathsf{F}}_t\widehat{\mathsf{F}}_t^{\intercal} = \mathbf{I}_{q_0}$, we have

$$\begin{aligned}
\widehat{\lambda}_j &= \frac{1}{n}\sum_{t=1}^{n}X_{tj}\widehat{\mathsf{F}}_t = \frac{1}{n}\sum_{t=1}^{n}\left(\lambda_j^{\star\intercal}\mathsf{F}_t^{\star}+\epsilon_{tj}\right)\widehat{\mathsf{F}}_t \\
&= \frac{1}{n}\sum_{t=1}^{n}\widehat{\mathsf{F}}_t\mathsf{F}_t^{\star\intercal}\lambda_j^{\star} + \frac{1}{n}\sum_{t=1}^{n}\epsilon_{tj}\widehat{\mathsf{F}}_t
\end{aligned}$$

$$= (\mathbf{H}^{-1})^{\intercal}\lambda_j^{\star} + \frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{F}}_t\left(\mathbf{F}_t^{\star} - \mathbf{H}^{-1}\widehat{\mathbf{F}}_t\right)^{\intercal}\lambda_j^{\star}$$

$$+\mathbf{H}\cdot\frac{1}{n}\sum_{t=1}^{n}\epsilon_{tj}\mathbf{F}_t^{\star} + \frac{1}{n}\sum_{t=1}^{n}\epsilon_{tj}\left(\widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^{\star}\right). \tag{D.10}$$

By Lemma D.1(i), we readily have

$$\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{F}}_t\left(\mathbf{F}_t^{\star} - \mathbf{H}^{-1}\widehat{\mathbf{F}}_t\right)^{\intercal}\lambda_j^{\star}\right\|_2 = O_P\left(\frac{1}{n^{1/2}} + \frac{n^{2/\delta}}{d^{1/2}}\right), \tag{D.11}$$

and

$$\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\epsilon_{tj}\left(\widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^{\star}\right)\right\|_2 = O_P\left(\frac{1}{n^{1/2}} + \frac{n^{2/\delta}}{d^{1/2}}\right). \tag{D.12}$$

By (D.10)–(D.12) and noting that $\mathbf{H} = O_P(1)$, to complete the proof of (D.2), we only need to show that

$$\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\epsilon_{tj}\mathbf{F}_t^{\star}\right\|_2 = O_P\left(\sqrt{(\log d)/n}\right). \tag{D.13}$$

The proof of (D.13) is standard. Let $\zeta_{tj} = \epsilon_{tj}\mathbf{F}_t^{\star}$ for notational simplicity. From $\mathsf{E}[\epsilon_{tj}\mathbf{F}_t] = \mathbf{0}$ in Assumption 1(ii), we have $\mathsf{E}[\zeta_{tj}] = \mathsf{E}[\epsilon_{tj}\mathbf{F}_t^{\star}] = \mathbf{0}$, indicating that

$$\zeta_{tj} = \zeta_{tj} - \mathsf{E}[\zeta_{tj}] = \overline{\zeta}_{tj} - \mathsf{E}\left[\overline{\zeta}_{tj}\right] + \widetilde{\zeta}_{tj} - \mathsf{E}\left[\widetilde{\zeta}_{tj}\right],$$

where

$$\overline{\zeta}_{tj} = \zeta_{tj}\cdot\mathcal{I}\left(\|\zeta_{tj}\|_2 \leqslant c_3\log(dn)\right),\ \ \widetilde{\zeta}_{tj} = \zeta_{tj}\cdot\mathcal{I}\left(\|\zeta_{tj}\|_2 > c_3\log(dn)\right),$$

and c_3 is a positive constant to be determined later. Hence, in order to prove (D.13), we only have to show that

$$\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(\overline{\zeta}_{tj} - \mathsf{E}\left[\overline{\zeta}_{tj}\right]\right)\right\|_2 = O_P\left(\sqrt{(\log d)/n}\right) \tag{D.14}$$

and

$$\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(\widetilde{\zeta}_{tj} - \mathsf{E}\left[\widetilde{\zeta}_{tj}\right]\right)\right\|_2 = O_P\left(\sqrt{(\log d)/n}\right). \tag{D.15}$$

We first consider proving (D.15). From (A.4) in Assumption 3(ii) and the arguments in the proof of Proposition 3.1, there exists a positive constant $\iota_1^{\diamond}$ (which may be different from ι_1) such that

$$\max_{1\leqslant j\leqslant d}\max_{1\leqslant t\leqslant n}\mathsf{E}\left[\exp\{\iota_1^{\diamond}\|\epsilon_{tj}\mathbf{F}_t^{\star}\|_2\}\right] < \infty.$$

Choosing c_3 such that $c_3\iota_1^\diamond > 1$, we have

$$
\begin{aligned}
\mathsf{E}\left[\left\|\widetilde{\zeta}_{tj}\right\|_2\right] &\leqslant \left\{\mathsf{E}\left[\|\zeta_{tj}\|_2^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\|\zeta_{tj}\|_2 > c_3\log(dn)\right)\right\}^{1/2} \\
&= \left\{\mathsf{E}\left[\|\zeta_{tj}\|_2^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\exp\{\iota_1^\diamond\|\zeta_{tj}\|_2\} > \exp\{\iota_1^\diamond c_3\log(dn)\}\right)\right\}^{1/2} \\
&\leqslant O\left((dn)^{-\iota_1^\diamond c_3/2}\right) = o(n^{-1/2})
\end{aligned}
$$

uniformly over j and t. Then, for any $M > 0$, we can show that

$$
\begin{aligned}
& \mathsf{P}\left(\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(\widetilde{\zeta}_{tj} - \mathsf{E}\left[\widetilde{\zeta}_{tj}\right]\right)\right\|_2 > M\cdot\sqrt{(\log d)/n}\right) \\
\leqslant\ & \mathsf{P}\left(\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\widetilde{\zeta}_{tj}\right\|_2 > \frac{M}{2}\cdot\sqrt{(\log d)/n}\right) \\
\leqslant\ & \mathsf{P}\left(\max_{1\leqslant j\leqslant d}\max_{1\leqslant t\leqslant n}\|\zeta_{tj}\|_2 > c_3\log(dn)\right) \\
\leqslant\ & \sum_{j=1}^{d}\sum_{t=1}^{n}\frac{\mathsf{E}[\exp\{\iota_1^\diamond\|\zeta_{tj}\|_2\}]}{\exp\{\iota_1^\diamond c_3\log(dn)\}} \\
=\ & O\left((dn)^{1-\iota_1^\diamond c_3}\right) = o(1),
\end{aligned}
$$

leading to (D.15).

We next turn to the proof of (D.14). Using an exponential inequality for the α-mixing sequence (e.g., Theorem 1.3(2) in Bosq, 1998) and noting that $d = O\left(\exp\{n^\nu\}\right)$ with $0 \leqslant \nu < 1/5$, we may show that by taking $M > 0$ sufficiently large,

$$
\begin{aligned}
& \mathsf{P}\left(\max_{1\leqslant j\leqslant d}\left\|\frac{1}{n}\sum_{t=1}^{n}\left(\overline{\zeta}_{tj} - \mathsf{E}\left[\overline{\zeta}_{tj}\right]\right)\right\|_2 > M\cdot\sqrt{(\log d)/n}\right) \\
=\ & O\left(d\exp\{-c_M\log d\}\right) + O\left(d(\log d)^{1/4}(\log d + \log n)^{3/2}n^{3/2}\rho^{\frac{\sqrt{n/\log d}}{c_M(\log d+\log n)}}\right) \\
=\ & O\left(d^{1-c_M} + n^{(7\nu/4)+3/2}\exp\left\{n^\nu - (\log\rho/c_M)n^{(1/2)-(3\nu/2)}\right\}\right) = o(1),
\end{aligned}
$$

where $c_M > 0$ is a sufficiently large constant when M is large enough, completing the proof of (D.14). $\square$

With the uniform convergence result given in Lemma D.1(i), we can easily prove Proposition 3.2.

PROOF OF PROPOSITION 3.2. Note that

$$
\begin{aligned}
\mathbf{C}_{l,u}^{\widehat{\mathbf{F}}}(s)-\mathbf{C}_{l,u}^{\mathbf{HF}^\star}(s) &= \sqrt{\frac{u-s}{(u-l+1)(s-l+1)}}\sum_{t=l}^{s}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right] \\
&\quad +\sqrt{\frac{u-s}{(u-l+1)(s-l+1)}}\sum_{t=l}^{s}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal+\mathbf{HF}_t^\star\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right] \\
&\quad -\sqrt{\frac{s-l+1}{(u-l+1)(u-s)}}\sum_{t=s+1}^{u}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right] \\
&\quad -\sqrt{\frac{s-l+1}{(u-l+1)(u-s)}}\sum_{t=s+1}^{u}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal+\mathbf{HF}_t^\star\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right].
\end{aligned}
$$

By (D.1) in Lemma D.1 and noting that $n=O\left(d^{\delta/(\delta+4)}\right)$, we readily have

$$
\begin{aligned}
& \max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\sqrt{\frac{u-s}{(u-l+1)(s-l+1)}}\left\|\sum_{t=l}^{s}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right]\right\|_2 \\
= & \max_{(l,u):\,1\leqslant l<u\leqslant n}(u-l+1)^{-1/2}\cdot O_P\left(n^{-1}\right)\max_{s:l\leqslant s<u}\sqrt{(u-s)(s-l+1)} \\
= & \max_{(l,u):\,1\leqslant l<u\leqslant n}(u-l+1)^{1/2}\cdot O_P\left(n^{-1}\right)=O_P\left(n^{-1/2}\right), \qquad\qquad (D.16)
\end{aligned}
$$

and similarly

$$
\begin{aligned}
& \max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\sqrt{\frac{s-l+1}{(u-l+1)(u-s)}}\left\|\sum_{t=s+1}^{u}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)^\intercal\right]\right\|_2 \\
= & \max_{(l,u):\,1\leqslant l<u\leqslant n}(u-l+1)^{1/2}\cdot O_P\left(n^{-1}\right)=O_P\left(n^{-1/2}\right). \qquad\qquad (D.17)
\end{aligned}
$$

On the other hand, by the Cauchy-Schwarz inequality, Lemma D.1(i) and Proposition 3.1, we can prove that

$$
\begin{aligned}
& \left\|\sum_{t=l}^{s}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal\right]\right\|_2 \\
\leqslant & \left(\sum_{t=l}^{s}\left\|\widehat{\mathbf{F}}_t-\mathbf{HF}_t^\star\right\|_2^2\right)^{1/2}\left(\sum_{t=l}^{s}\left\|\mathbf{F}_t^\star\right\|_2^2\right)^{1/2}\cdot O_P(1) \\
= & O_P\left((s-l+1)/n^{1/2}\right),
\end{aligned}
$$

and similarly

$$\left\| \sum_{t=s+1}^{u} \mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal\right]\right\|_2 = O_P\left((u-s)/n^{1/2}\right).$$

Consequently, we have

$$\begin{aligned}
&\max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u} \sqrt{\frac{u-s}{(u-l+1)(s-l+1)}}\left\|\sum_{t=l}^{s}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal\right]\right\|_2 \\
=\ & \max_{(l,u):\,1\leqslant l<u\leqslant n} (u-l+1)^{-1/2}\cdot O_P\left(n^{-1/2}\right)\max_{s:l\leqslant s<u}\sqrt{(u-s)(s-l+1)} \\
=\ & \max_{(l,u):\,1\leqslant l<u\leqslant n} (u-l+1)^{1/2}\cdot O_P\left(n^{-1/2}\right) = O_P(1) \qquad \text{(D.18)}
\end{aligned}$$

and

$$\begin{aligned}
&\max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u} \sqrt{\frac{s-l+1}{(u-l+1)(u-s)}}\left\|\sum_{t=s+1}^{u}\mathsf{vech}\left[\left(\widehat{\mathbf{F}}_t - \mathbf{H}\mathbf{F}_t^\star\right)\mathbf{F}_t^{\star\intercal}\mathbf{H}^\intercal\right]\right\|_2 \\
=\ & \max_{(l,u):\,1\leqslant l<u\leqslant n} (u-l+1)^{1/2}\cdot O_P\left(n^{-1/2}\right) = O_P(1). \qquad \text{(D.19)}
\end{aligned}$$

By (D.16)–(D.19), we can complete the proof of (3.4). □

We next turn to the proof of Theorem 3.1. In order to facilitate the proof, we first introduce some additional notation. Let

$$\begin{aligned}
\mathbf{Z}_t^{\mathbf{F}^\star} &= \mathsf{vech}\left(\mathbf{F}_t^\star\mathbf{F}_t^{\star\intercal}\right) = \left(Z_{t,1}^{\mathbf{F}^\star},\cdots,Z_{t,q_0(q_0+1)/2}^{\mathbf{F}^\star}\right)^\intercal, \\
\mathbf{G}_t^{\mathbf{F}^\star} &= \mathsf{E}\left[\mathsf{vech}\left(\mathbf{F}_t^\star\mathbf{F}_t^{\star\intercal}\right)\right] = \left(G_{t,1}^{\mathbf{F}^\star},\cdots,G_{t,q_0(q_0+1)/2}^{\mathbf{F}^\star}\right)^\intercal, \\
z_t^{\mathbf{F}^\star} &= \mathbf{Z}_t^{\mathbf{F}^\star} - \mathbf{G}_t^{\mathbf{F}^\star} = \left(z_{t,1}^{\mathbf{F}^\star},\cdots,z_{t,q_0(q_0+1)/2}^{\mathbf{F}^\star}\right)^\intercal.
\end{aligned}$$

Define

$$\mathbf{C}_{l,u}^{\mathbf{F}^\star}(s) = \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}\mathbf{Z}_t^{\mathbf{F}^\star} - \frac{1}{u-s}\sum_{t=s+1}^{u}\mathbf{Z}_t^{\mathbf{F}^\star}\right).$$

Then

$$\begin{aligned}
\mathbf{C}_{l,u}^{\mathbf{F}^\star}(s) &= \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}\mathbf{G}_t^{\mathbf{F}^\star} - \frac{1}{u-s}\sum_{t=s+1}^{u}\mathbf{G}_t^{\mathbf{F}^\star}\right) \\
&\quad + \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}z_t^{\mathbf{F}^\star} - \frac{1}{u-s}\sum_{t=s+1}^{u}z_t^{\mathbf{F}^\star}\right) \\
&=: \mathbf{C}_{l,u}^{\mathbf{G},\mathbf{F}^\star}(s) + \mathbf{C}_{l,u}^{z,\mathbf{F}^\star}(s). \qquad \text{(D.20)}
\end{aligned}$$

Recall that the two positive integers l and u denote the "lower" and "upper" bounds of a segment. We assume that

$$\eta^c_{k_0} \leqslant l < \eta^c_{k_0+1} < \cdots < \eta^c_{k_0+k_1} < u \leqslant \eta^c_{k_0+k_1+1}, \tag{D.21}$$

where $k_0 \in \{0, \cdots, K_1 - k_1\}$ and $k_1 \in \{1, \cdots, K_1 - k_0\}$. The following two conditions are key to the WBS-Cov asymptotic analysis: for some $1 \leqslant k \leqslant k_1$,

$$l < \eta^c_{k_0+k} - c_4\kappa^c_n < \eta^c_{k_0+k} + c_4\kappa^c_n < u \tag{D.22}$$

and

$$\left\{(l - \eta^c_{k_0}) \wedge (\eta^c_{k_0+1} - l)\right\} \vee \left\{(u - \eta^c_{k_0+k_1}) \wedge (\eta^c_{k_0+k_1+1} - u)\right\} \leqslant c_5\varphi^c_n, \tag{D.23}$$

where c_4 and c_5 are two positive constants, κ^c_n is defined in Assumption 4(ii), and φ^c_n is defined in Theorem 3.1. Define the intervals

$$\mathcal{I}^c_k = \left[\eta^c_{k-1} + (\eta^c_k - \eta^c_{k-1})/3,\ \eta^c_{k-1} + 2(\eta^c_k - \eta^c_{k-1})/3\right],\ \ k = 1, \cdots, K_1 + 1,$$

and the event

$$\mathcal{D}^c_n = \left\{\forall\, k = 1, \cdots, K_1,\ \exists\, m = 1, \cdots, M^c_n \text{ such that } l_m \in \mathcal{I}^c_k \text{ and } u_m \in \mathcal{I}^c_{k+1}\right\},$$

where M^c_n is defined in Section 2.3.

LEMMA D.2. *Letting $\overline{\mathcal{D}}^c_n$ be the complement of $\mathcal{D}^c_n$, we have*

$$\mathsf{P}\left(\overline{\mathcal{D}}^c_n\right) \leqslant K_1\left[1 - (\kappa^c_n/(3n))^2\right]^{M^c_n}, \tag{D.24}$$

where κ^c_n is defined in Assumption 4(ii).

PROOF. From the definition of $\overline{\mathcal{D}}^c_n$ and noting that the two random points l_m and u_m are drawn uniformly from the set $\{l, l+1, \cdots, u-1, u\}$ with $1 \leqslant l < u \leqslant n$, we readily have that

$$\begin{aligned}
\mathsf{P}\left(\overline{\mathcal{D}}^c_n\right) &\leqslant \sum_{k=1}^{K_1}\prod_{m=1}^{M^c_n}\left[1 - \mathsf{P}\left(l_m \in \mathcal{I}^c_k \text{ and } u_m \in \mathcal{I}^c_{k+1}\right)\right] \\
&\leqslant K_1\prod_{m=1}^{M^c_n}\left(1 - \frac{\eta^c_k - \eta^c_{k-1}}{3n}\cdot\frac{\eta^c_{k+1} - \eta^c_k}{3n}\right) \\
&\leqslant K_1\left[1 - (\kappa^c_n/(3n))^2\right]^{M^c_n},
\end{aligned} \tag{D.25}$$

completing the proof of Lemma D.2. □

The following lemma derives an asymptotic order for $\mathbf{C}_{l,u}^{z,\mathbf{F}^\star}(s)$ uniformly over l, u and s.

LEMMA D.3. *Suppose that Assumptions 1, 2 and 3(ii) are satisfied. If $\kappa_n^c \asymp n$, there exists a positive constant c_6 such that*

$$\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\left\|\mathbf{C}_{l,u}^{z,\mathbf{F}^\star}(s)\right\|_2 > c_6\cdot\log^2 n\right)\to 0, \tag{D.26}$$

as $n\to\infty$.

PROOF. Note that $\mathbf{C}_{l,u}^{z,\mathbf{F}^\star}(s)$ is a column vector with dimension $q_0(q_0+1)/2$. Let $C_{l,u,k}^{z,\mathbf{F}^\star}(s)$ be the k-th element of $\mathbf{C}_{l,u}^{z,\mathbf{F}^\star}(s)$, i.e.,

$$C_{l,u,k}^{z,\mathbf{F}^\star}(s) = \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s} z_{t,k}^{\mathbf{F}^\star} - \frac{1}{u-s}\sum_{t=s+1}^{u} z_{t,k}^{\mathbf{F}^\star}\right),\ \ k=1,\cdots,q_0(q_0+1)/2.$$

By the Bonferroni inequality and noting that q_0 is assumed to be bounded, in order to prove (D.26), we only need to show that

$$\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\left|C_{l,u,k}^{z,\mathbf{F}^\star}(s)\right| > \frac{2c_6}{q_0(q_0+1)}\cdot\log^2 n\right)\to 0 \tag{D.27}$$

for each $k=1,\cdots,q_0(q_0+1)/2$. Letting

$$C_{l,u,k}^{z,\mathbf{F}^\star}(s;1) = \sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s} z_{t,k}^{\mathbf{F}^\star}$$

and

$$C_{l,u,k}^{z,\mathbf{F}^\star}(s;2) = \sqrt{\frac{s-l+1}{u-l+1}}\cdot\frac{1}{\sqrt{u-s}}\cdot\sum_{t=s+1}^{u} z_{t,k}^{\mathbf{F}^\star},$$

it suffices to prove that

$$\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\left|C_{l,u,k}^{z,\mathbf{F}^\star}(s;j)\right| > \bar{c}(q_0)\cdot\log^2 n\right)\to 0 \tag{D.28}$$

for $j=1$ and 2, where $\bar{c}(q_0)=\frac{c_6}{q_0(q_0+1)}$.

The proof of (D.28) is similar to the proof of (D.13) in Lemma D.1(ii). Define

$$\overline{z}_{t,k}^{\mathbf{F}^\star} = z_{t,k}^{\mathbf{F}^\star}\cdot\mathcal{I}\left(\left|z_{t,k}^{\mathbf{F}^\star}\right|\leqslant c_7\log n\right),\ \ \widetilde{z}_{t,k}^{\mathbf{F}^\star} = z_{t,k}^{\mathbf{F}^\star}\cdot\mathcal{I}\left(\left|z_{t,k}^{\mathbf{F}^\star}\right| > c_7\log n\right),$$

where $c_7>0$ is a sufficiently large constant. Letting $\overline{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)$ and $\widetilde{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)$ be defined similarly to $C_{l,u,k}^{z,\mathbf{F}^\star}(s;1)$ but with $z_{t,k}^{\mathbf{F}^\star}$ replaced by $\overline{z}_{t,k}^{\mathbf{F}^\star}-\mathsf{E}\left[\overline{z}_{t,k}^{\mathbf{F}^\star}\right]$ and $\widetilde{z}_{t,k}^{\mathbf{F}^\star}-\mathsf{E}\left[\widetilde{z}_{t,k}^{\mathbf{F}^\star}\right]$, respectively. From Assumption 3(ii) and Proposition 3.1, there exists a positive constant $\iota_6 > 0$ (which may be different from ι_1) such that

$$\max_{1\leqslant t\leqslant n}\max_{1\leqslant k\leqslant q_0(q_0+1)/2}\mathsf{E}\left[\exp\left\{\iota_6\left|z_{t,k}^{\mathbf{F}^\star}\right|\right\}\right]<\infty.$$

Consequently, we can show that

$$\begin{aligned}\mathsf{E}\left[\left|\widetilde{z}_{t,k}^{\mathbf{F}^\star}\right|\right] &\leqslant \left\{\mathsf{E}\left[\left|z_{t,k}^{\mathbf{F}^\star}\right|^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\left|z_{t,k}^{\mathbf{F}^\star}\right|>c_7\log n\right)\right\}^{1/2}\\ &= \left\{\mathsf{E}\left[\left|z_{t,k}^{\mathbf{F}^\star}\right|^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\exp\left\{\iota_6\left|z_{t,k}^{\mathbf{F}^\star}\right|\right\}>\exp\{\iota_6 c_7\log n\}\right)\right\}^{1/2}\\ &\leqslant O\left(n^{-\iota_6 c_7/2}\right)=o\left(n^{-1/2}\log^2 n\right)\end{aligned}$$

uniformly over k and t, where the constant c_7 is chosen so that $c_7\iota_6 > 1$. Therefore, we can prove that

$$\begin{aligned}&\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\widetilde{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)\right|>\frac{\bar{c}(q_0)}{2}\cdot\log^2 n\right)\\ \leqslant\ &\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s}\widetilde{z}_{t,k}^{\mathbf{F}^\star}\right|>\frac{\bar{c}(q_0)}{3}\cdot\log^2 n\right)\\ \leqslant\ &\mathsf{P}\left(\max_{1\leqslant t\leqslant n}\left|z_{t,k}^{\mathbf{F}^\star}\right|>c_7\log n\right)\leqslant\sum_{t=1}^{n}\frac{\mathsf{E}[\exp\{\iota_6\left|z_{t,k}^{\mathbf{F}}\right|\}]}{\exp\{\iota_6 c_7\log n\}}=o(1).\end{aligned}\tag{D.29}$$

We next prove

$$\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\overline{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)\right|>\frac{\bar{c}(q_0)}{2}\cdot\log^2 n\right)\to 0.\tag{D.30}$$

Consider the following two scenarios: (i) $s-l+1\leqslant c_8\log^2 n$, and (ii) $s-l+1>c_8\log^2 n$, where c_8 is a sufficiently large positive constant. For scenario (i), it is easy to see that

$$\begin{aligned}\left|\overline{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)\right| &\leqslant \sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s}\left(\left|\overline{z}_{t,k}^{\mathbf{F}^\star}\right|+\mathsf{E}\left[\left|\overline{z}_{t,k}^{\mathbf{F}^\star}\right|\right]\right)\\ &\leqslant \sqrt{s-l+1}\cdot(2c_7\log n)\leqslant(2c_7c_8)\cdot\log^2 n.\end{aligned}$$

For scenario (ii), by Theorem 1.3(2) in Bosq (1998) (choosing $p=\sqrt{s-l}$), we then have

$$\begin{aligned}&\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\overline{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)\right|>\frac{\bar{c}(q_0)}{2}\cdot\log^2 n\right)\\ \leqslant\ &\mathsf{P}\left(\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l+c_8\log^2 n-1\leqslant s<u}\left|\overline{C}_{l,u,k}^{z,\mathbf{F}^\star}(s;1)\right|>\left[\frac{\bar{c}(q_0)}{2}-2c_7c_8\right]\cdot\log^2 n\right)\end{aligned}$$

$$\leqslant \quad O\left(n^3 \exp\{-M \log n\} + n^{3+3/4}\rho^{\sqrt{c_8}\log n}\right) = o(1),$$

where c_6 is chosen to be sufficiently large such that $\frac{\bar{c}(q_0)}{2} - 2c_7c_8$ is strictly larger than zero and the constant M is larger than 3, and the constant c_8 is chosen to be larger than $(-15/(4\log\rho))^2$. This proves (D.30).

With (D.29) and (D.30), we can show (D.28), completing the proof of the lemma. □

The following lemma derives a lower bound for the CUSUM statistic in the WBS-Cov when l and u satisfy (D.22) and (D.23).

LEMMA D.4. *Suppose that the assumptions in Lemma D.3 and Assumption 4(ii) are satisfied, and let l and u (the lower and upper bound of the segment, respectively) satisfy the conditions (D.22) and (D.23). Conditional on that the rotation matrix $\mathbf{H}$ is non-singular, we have*

$$\mathsf{P}\left(\left\|\mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_0^c}, u_{m_0^c}}(s_0^c)\right\|_2 \geqslant c_9 \cdot (\kappa_n^c \underline{\omega}_n^c)^{1/2}\right) \to 1 \tag{D.31}$$

as $n \to \infty$, where c_9 is a positive constant, m_0^c and s_0^c are defined as in Algorithm 2 of Section 2.3.

PROOF. From the definition of $\mathbf{F}_t^\star$ given in (C.3), we readily have the following time-varying covariance structure for $\mathbf{F}_t^\star$:

$$\boldsymbol{\Sigma}_t(\mathbf{F}^\star) = \begin{cases} \boldsymbol{\Sigma}_1^0(\mathbf{F}^\star), & 1 \leqslant t \leqslant \eta_1^c, \\ \boldsymbol{\Sigma}_2^0(\mathbf{F}^\star), & \eta_1^c + 1 \leqslant t \leqslant \eta_2^c, \\ \vdots & \vdots \\ \boldsymbol{\Sigma}_{K_1+1}^0(\mathbf{F}^\star), & \eta_{K_1}^c + 1 \leqslant t \leqslant n. \end{cases}$$

Hence, we have

$$\boldsymbol{\Sigma}_{k+1}^0(\boldsymbol{\Lambda}, \mathbf{F}) - \boldsymbol{\Sigma}_k^0(\boldsymbol{\Lambda}, \mathbf{F}) = \boldsymbol{\Lambda}^\star\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\boldsymbol{\Lambda}^{\star\intercal},$$

indicating that

$$\begin{aligned}
& \frac{1}{d^2} \cdot \left\|\boldsymbol{\Sigma}_{k+1}^0(\boldsymbol{\Lambda}, \mathbf{F}) - \boldsymbol{\Sigma}_k^0(\boldsymbol{\Lambda}, \mathbf{F})\right\|_F^2 \\
= \; & \frac{1}{d^2} \cdot \mathsf{Trace}\left\{\boldsymbol{\Lambda}^\star\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\boldsymbol{\Lambda}^{\star\intercal}\boldsymbol{\Lambda}^\star\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\boldsymbol{\Lambda}^{\star\intercal}\right\} \\
= \; & \mathsf{Trace}\left\{\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\left[\boldsymbol{\Lambda}^{\star\intercal}\boldsymbol{\Lambda}^\star/d\right]\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\left[\boldsymbol{\Lambda}^{\star\intercal}\boldsymbol{\Lambda}^\star/d\right]\right\} \\
= \; & \left\|\left[\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^\star) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^\star)\right]\left[\boldsymbol{\Lambda}^{\star\intercal}\boldsymbol{\Lambda}^\star/d\right]\right\|_F^2.
\end{aligned}$$

From the proof of Proposition 3.1, all the eigenvalues of $\boldsymbol{\Lambda}^{\star\intercal}\boldsymbol{\Lambda}^\star/d$ are bounded and strictly positive.

Using the inequality

$$\left\| \left[\boldsymbol{\Sigma}^0_{k+1}(\mathbf{F}^\star) - \boldsymbol{\Sigma}^0_k(\mathbf{F}^\star)\right] \left[\boldsymbol{\Lambda}^{\star^\intercal}\boldsymbol{\Lambda}^\star/d\right] \right\|_F^2 \leqslant \left\|\boldsymbol{\Sigma}^0_{k+1}(\mathbf{F}^\star) - \boldsymbol{\Sigma}^0_k(\mathbf{F}^\star)\right\|_F^2 \cdot \mu_1^2 \left(\boldsymbol{\Lambda}^{\star^\intercal}\boldsymbol{\Lambda}^\star/d\right)$$

with $\mu_1\left(\boldsymbol{\Lambda}^{\star^\intercal}\boldsymbol{\Lambda}^\star/d\right)$ being the maximum eigenvalue of $\boldsymbol{\Lambda}^{\star^\intercal}\boldsymbol{\Lambda}^\star/d$, we then have

$$\underline{\omega}^c_n \leqslant c_{10} \left\|\boldsymbol{\Sigma}^0_{k+1}(\mathbf{F}^\star) - \boldsymbol{\Sigma}^0_k(\mathbf{F}^\star)\right\|_F^2, \tag{D.32}$$

where c_{10} is a positive constant.

Consider that l and u satisfy the two conditions: (D.22) and (D.23). These conditions imply that l and u are close to the previously detected break points and bounded away from the previously undetected break points. Without loss of generality, we let η^c_k be one of these break points within $[l, u]$ satisfying $l + c_5\varphi^c_n < \eta^c_k < u - c_5\varphi^c_n$. On the set $\mathcal{D}^c_n$, there exists $1 \leqslant m_k \leqslant M^c_n$ such that $l_{m_k} \in \mathcal{I}^c_k$ and $u_{m_k} \in \mathcal{I}^c_{k+1}$, indicating that both $\eta^c_k - l_{m_k}$ and $u_{m_k} - \eta^c_k$ are larger than $\kappa^c_n/3$. Define

$$\boldsymbol{\varpi}^{\mathbf{F}^\star}_k = \mathsf{vech}\left(\boldsymbol{\Sigma}^0_{k+1}(\mathbf{F}^\star) - \boldsymbol{\Sigma}^0_k(\mathbf{F}^\star)\right) =: \left(\varpi^{\mathbf{F}^\star}_{k,1}, \cdots, \varpi^{\mathbf{F}^\star}_{k,q_0(q_0+1)/2}\right)^\intercal. \tag{D.33}$$

For $i = 1, \cdots, q_0(q_0+1)/2$, we have

$$\left|C^{\mathbf{G},\mathbf{F}^\star}_{l_{m_k},u_{m_k},i}(\eta^c_k)\right| = \sqrt{\frac{(\eta^c_k - l_{m_k}+1)(u_{m_k}-\eta^c_k)}{u_{m_k}-l_{m_k}+1}}\left|\varpi^{\mathbf{F}^\star}_{k,i}\right| \geqslant \left(\frac{\kappa^c_n}{6}\right)^{1/2}\left|\varpi^{\mathbf{F}^\star}_{k,i}\right|, \tag{D.34}$$

where $C^{\mathbf{G},\mathbf{F}^\star}_{l,u,i}(\cdot)$ is the i-th element of $\mathbf{C}^{\mathbf{G},\mathbf{F}^\star}_{l,u}(\cdot)$ defined in (D.20). Thus

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{F}^\star}_{l_{m_k},u_{m_k}}(\eta^c_k)\right\|_2 \geqslant c_{11}\left(\kappa^c_n\right)^{1/2}\left\|\boldsymbol{\varpi}^{\mathbf{F}^\star}_k\right\|_2, \tag{D.35}$$

where c_{11} is a positive constant. Let $\mathbf{L}_q$ and $\mathbf{D}_q$ be the $q(q+1)/2 \times q^2$ elimination matrix and the $q^2 \times q(q+1)/2$ duplication matrix, transforming the vectorisation of a matrix to its half vectorisation and vice versa, respectively. We have

$$\mathbf{C}^{\mathbf{H}\mathbf{F}^\star}_{l,u}(s) = \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{C}^{\mathbf{F}^\star}_{l,u}(s).$$

Noting that $\|\mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\|_F^2 = O_P(1)$, a combination of (D.32) and (D.35) leads to

$$\left\|\mathbf{L}_{q_0}(\mathbf{H}\otimes\mathbf{H})\mathbf{D}_{q_0}\cdot\mathbf{C}^{\mathbf{G},\mathbf{F}^\star}_{l_{m_k},u_{m_k}}(\eta^c_k)\right\|_2 \geqslant 2c_9\left(\kappa^c_n\underline{\omega}^c_n\right)^{1/2}. \tag{D.36}$$

By the definitions of m^c_0 and s^c_0 in Algorithm 2 and using Proposition 3.2 and Lemma D.3, we may show that conditional on that $\mathbf{H}$ is non-singular,

$$
\begin{aligned}
\left\|\mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_0^c},u_{m_0^c}}(s_0^c)\right\|_2 &\geqslant \left\|\mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_k},u_{m_k}}(\eta_k^c)\right\|_2 \\
&= \left\|\mathbf{C}^{\mathbf{H}\mathbf{F}^\star}_{l_{m_k},u_{m_k}}(\eta_k^c)\right\|_2 + O_P(1) \\
&= \left\|\mathbf{L}_{q_0}(\mathbf{H}\otimes\mathbf{H})\mathbf{D}_{q_0}\cdot \mathbf{C}^{\mathbf{G},\mathbf{F}^\star}_{l_{m_k},u_{m_k}}(\eta_k^c)\right\|_2 + O_P\left(\log^2 n\right) \\
&\geqslant 2c_9\left(\kappa_n^c\underline{\omega}_n^c\right)^{1/2} + O_P\left(\log^2 n\right).
\end{aligned}
$$

We then complete the proof of (D.31) by noting that $(\kappa_n^c\underline{\omega}_n^c)/\log^4 n\to\infty$ by Assumption 4(ii). □

Define the function $g(\cdot)$ as

$$
g(x)=\frac{|ax+b|}{[x(1-x)]^{1/2}},\quad 0<x<1,
$$

where a and b are two constants which do not depend on x. Lemma 2.2 in Venkatraman (1992) proves that $g(x)$ is a strictly quasi-convex function on $[c,d]$ with $0<c<d<1$, and

$$
g(x)<\max\{g(c),g(d)\},\ \forall\, c<x<d.
$$

As the CUSUM statistics proposed in the present paper are multi-dimensional vectors, we next provide an extension of Lemma 2.2 in Venkatraman (1992) (from the univariate binary segmentation to the multi-dimensional binary segmentation).

LEMMA D.5. *Define*

$$
G(x)=\frac{\left(\sum_{i=1}^m|a_ix+b_i|^p\right)^{1/p}}{[x(1-x)]^{1/2}},\quad 0<c\leqslant x\leqslant d<1, \tag{D.37}
$$

where a_i *and* b_i, $i=1,\cdots,m$, *are numbers independent of* x, m *is a positive integer and* $1\leqslant p\leqslant 2$. *The function* $G(x)$ *is quasi-convex over the interval* $[c,d]$.

PROOF. We first show that, for any positive convex function $G^\star(x)$ on $[c,d]$ and $\gamma\in(0,1]$, $G^\star(x)/[x(1-x)]^\gamma$ is a quasi-convex function over $[c,d]$. To prove this, it is sufficient to show that each sub-level set defined as

$$
\mathcal{S}_\alpha=\{x\mid G^\star(x)/[x(1-x)]^\gamma\leqslant\alpha\}
$$

is a convex set. Note that the sub-level set $\mathcal{S}_\alpha$ can be written as

$$
\mathcal{S}_\alpha=\{x\mid G^\star(x)-\alpha[x(1-x)]^\gamma\leqslant 0\}.
$$

As both $G^\star(x)$ and $-\alpha[x(1-x)]^\gamma$ are convex, we readily prove that $\mathcal{S}_\alpha$ is a convex set. Choosing $G^\star(x) = \sum_{i=1}^m |a_i x + b_i|^p$ which is positive and convex, we can then show that the function $\sum_{i=1}^m |a_i x + b_i|^p / [x(1-x)]^\gamma$ is quasi-convex. As a non-decreasing functional transformation preserves the quasi-convexity, the function $\left(\sum_{i=1}^m |a_i x + b_i|^p\right)^{1/p} / [x(1-x)]^{\gamma/p}$ is also quasi-convex. Letting $\gamma = p/2$, we prove that $G(x)$ is quasi-convex, completing the proof of the lemma. $\square$

Similarly to $\mathbf{Z}_t^{\mathbf{F}^\star}$, $\mathbf{G}_t^{\mathbf{F}^\star}$ and $\boldsymbol{z}_t^{\mathbf{F}^\star}$, we define

$$\begin{aligned}
\mathbf{Z}_t^{\mathbf{HF}^\star} &= \mathsf{vech}\left(\mathbf{H}\mathbf{F}_t^\star \mathbf{F}_t^{\star\intercal}\mathbf{H}\right) = \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathsf{vech}\left(\mathbf{F}_t^\star\mathbf{F}_t^{\star\intercal}\right) = \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{Z}_t^{\mathbf{F}^\star}, \\
\mathbf{G}_t^{\mathbf{HF}^\star} &= \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathsf{E}[\mathbf{Z}_t^{\mathbf{F}^\star}] = \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{G}_t^{\mathbf{F}^\star}, \\
\boldsymbol{z}_t^{\mathbf{HF}^\star} &= \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\left(\mathbf{Z}_t^{\mathbf{F}^\star} - \mathbf{G}_t^{\mathbf{F}^\star}\right) = \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\boldsymbol{z}_t^{\mathbf{F}^\star},
\end{aligned}$$

and then

$$\begin{aligned}
\mathbf{C}_{l,u}^{\mathbf{HF}^\star}(s) &= \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{C}_{l,u}^{\mathbf{F}^\star}(s) \\
&= \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{C}_{l,u}^{\mathbf{G},\mathbf{F}^\star}(s) + \mathbf{L}_{q_0}\left(\mathbf{H}\otimes\mathbf{H}\right)\mathbf{D}_{q_0}\mathbf{C}_{l,u}^{\boldsymbol{z},\mathbf{F}^\star}(s) \\
&=: \mathbf{C}_{l,u}^{\mathbf{G},\mathbf{HF}^\star}(s) + \mathbf{C}_{l,u}^{\boldsymbol{z},\mathbf{HF}^\star}(s).
\end{aligned}$$

We next give an extension of Lemma 2.6 in Venkatraman (1992) to the case of multi-dimensional WBS-Cov. In the following lemma and its proof, we use the notation ν with appropriate subscript to highlight the difference and similarity between Lemma 2.6 in Venkatraman (1992) and our lemma. For example, ν_h, ν_i, ν_j and ν_l in the following lemma correspond to h, i, j and l in Venkatraman (1992).

LEMMA D.6. *Suppose that the assumptions of Lemma D.4 and (D.21)–(D.23) are satisfied. Let* $s_\star^c \in [l_{m_0^c}, u_{m_0^c}]$ *be the point of maximising* $\left\|\mathbf{C}_{l_{m_0^c},u_{m_0^c}}^{\mathbf{G},\mathbf{HF}^\star}(s)\right\|_2$ *with respect to* s, *i.e.,*

$$s_\star^c = \arg\max_{l_{m_0^c}\leqslant s< u_{m_0^c}} \left\|\mathbf{C}_{l_{m_0^c},u_{m_0^c}}^{\mathbf{G},\mathbf{HF}^\star}(s)\right\|_2, \tag{D.38}$$

and define $\eta_{k_\diamond}^c$ *as a change point satisfying*

$$\left\|\mathbf{C}_{l_{m_0^c},u_{m_0^c}}^{\mathbf{G},\mathbf{HF}^\star}(\eta_{k_\diamond}^c)\right\|_2 > \left\|\mathbf{C}_{l_{m_0^c},u_{m_0^c}}^{\mathbf{G},\mathbf{HF}^\star}(s_\star^c)\right\|_2 - 3c_6\log^2 n, \tag{D.39}$$

where c_6 *is a positive constant defined in Lemma D.3. Then there exists* $c_{12} > 0$ *such that*

$$(\eta_{k_\diamond}^c - l_{m_0^c} + 1)\wedge(u_{m_0^c} - \eta_{k_\diamond}^c) \geqslant c_{12}\kappa_n^c, \tag{D.40}$$

and we further have

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_l)\right\|_2 + (c_{13}\nu_l\kappa^c_n)\cdot\frac{\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2}{(u_{m_0^c}-l_{m_0^c}+1)^2}, \tag{D.41}$$

where $0<\nu_l<c_{14}\gamma^c_n$ *with* $\gamma^c_n=(\kappa^c_n/\underline{\omega}^c_n)^{1/2}\log^2 n$, *and* c_{13} *and* c_{14} *are two positive constants.*

PROOF. Using Lemma D.5 with $p=2$ and $m=q_0(q_0+1)/2$, and noting

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l,u}(s)\right\|_2 = G\left(\frac{s-l+1}{u-l+1}\right)\sqrt{u-l+1}$$

(by appropriately choosing a_i and b_i in the definition of G), we may show that there exists a positive integer $k_\star$ such that $s^c_\star=\eta^c_{k_\star}$. From the conditions (D.22) and (D.23), we have that $(\eta^c_{k_\star}-l+1)\wedge(u-\eta^c_{k_\star})$ is either smaller than $c_5\varphi^c_n$ or larger than $\kappa^c_n-c_5\varphi^c_n$, where c_5 is defined in (D.23). Note that

$$\begin{aligned}\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l,u}(s)\right\|_2 &= \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left\|\frac{1}{s-l+1}\sum_{t=l}^{s}\mathbf{G}^{\mathbf{HF}^\star}_t-\frac{1}{u-s}\sum_{t=s+1}^{u}\mathbf{G}^{\mathbf{HF}^\star}_t\right\|_2\\ &\leqslant 2b_{l,u}\sqrt{(s-l+1)\wedge(u-s)},\end{aligned} \tag{D.42}$$

where

$$b_{l,u}=\sup_{l\leqslant s\leqslant u}\left\|\mathbf{G}^{\mathbf{HF}^\star}_s-\frac{1}{u-l+1}\sum_{t=l}^{u}\mathbf{G}^{\mathbf{HF}^\star}_t\right\|_2.$$

If $(\eta^c_{k_\star}-l+1)\wedge(u-\eta^c_{k_\star})\leqslant c_5\varphi^c_n$ holds, we have $(\eta^c_{k_\star}-l_{m_0^c}+1)\wedge(u_{m_0^c}-\eta^c_{k_\star})\leqslant c_5\varphi^c_n$ as $[l_{m_0^c},u_{m_0^c}]$ is a random sub-interval of $[l,u]$. By Assumption 4(ii), we have

$$b_{l_{m_0^c},u_{m_0^c}}\leqslant c_{15}\left(\overline{\omega}^c_{l_{m_0^c},u_{m_0^c}}\right)^{1/2}\leqslant c_{15}\left(\overline{\omega}^c_{l,u}\right)^{1/2}\leqslant c_{15}\left(\overline{\omega}^c_n\right)^{1/2}, \tag{D.43}$$

where c_{15} is a positive constant and

$$\overline{\omega}^c_{l,u}=\frac{1}{d^2}\cdot\max_{k:l+c_5\varphi^c_n\leqslant\eta^c_k\leqslant u-c_5\varphi^c_n}\left\|\mathbf{\Sigma}^0_{k+1}(\mathbf{\Lambda},\mathbf{F})-\mathbf{\Sigma}^0_k(\mathbf{\Lambda},\mathbf{F})\right\|_F^2. \tag{D.44}$$

With (D.42) and (D.43), we have

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\star})\right\|_2\leqslant 2b_{l_{m_0^c},u_{m_0^c}}(c_5\varphi^c_n)^{1/2}\leqslant 2c_{15}(c_5\overline{\omega}^c_n\varphi^c_n)^{1/2}. \tag{D.45}$$

Combining (D.36) with (D.45), we readily have that $\frac{\underline{\omega}^c_n}{\overline{\omega}^c_n}\cdot\frac{\kappa^c_n\underline{\omega}^c_n}{\log^4 n}$ is bounded. However, this leads to contradiction with the condition $\frac{\omega_n^c}{\overline{\omega}_n^c} \cdot \frac{\kappa_n^c \omega_n^c}{\log^4 n} \to \infty$ in Assumption 4(ii). Therefore, $(\eta_{k_\star}^c - l + 1) \wedge (u - \eta_{k_\star}^c)$ cannot be smaller than $c_5 \varphi_n^c$, and we must have

$$(\eta_{k_\star}^c - l + 1) \wedge (u - \eta_{k_\star}^c) \geqslant \kappa_n^c - c_5 \varphi_n^c, \tag{D.46}$$

which further indicates that there exists $m_\star^c \in \mathcal{M}_{l,u}^c$ such that $l_{m_\star^c} \in \mathcal{I}_{k_\star}^c$ and $u_{m_\star^c} \in \mathcal{I}_{k_\star+1}^c$.

We next strengthen (D.46) to

$$(\eta_{k_\star}^c - l_{m_0^c} + 1) \wedge (u_{m_0^c} - \eta_{k_\star}^c) \geqslant c_{12} \kappa_n^c. \tag{D.47}$$

Suppose that (D.47) fails, i.e., for any $c_\star$ and N, there exists some $n > N$ such that

$$(\eta_{k_\star}^c - l_{m_0^c} + 1) \wedge (u_{m_0^c} - \eta_{k_\star}^c) < c_\star \kappa_n^c. \tag{D.48}$$

Without loss of generality, we let $\eta_{k_\star}^c - l_{m_0^c} + 1 < c_\star \kappa_n^c$ and consider the following two cases of $u_{m_0^c}$:

(i) $\eta_{k_\star}^c \leqslant u_{m_0^c} < \eta_{k_0+k_1}^c$, or $\eta_{k_0+k_1+1}^c - c_5 \varphi_n^c \leqslant u \leqslant \eta_{k_0+k_1+1}^c$ and $\eta_{k_0+k_1}^c < u_{m_0^c} \leqslant u$;

(ii) $\eta_{k_0+k_1}^c \leqslant u \leqslant \eta_{k_0+k_1}^c + c_5 \varphi_n^c$ and $\eta_{k_\star}^c < \eta_{k_0+k_1}^c < u_{m_0^c} \leqslant u$.

The main difference between cases (i) and (ii) is that in case (ii) there does not exist any $m \in \mathcal{M}_{l,u}^c$ such that $l_m \in \mathcal{I}_{k_0+k_1}^c$ and $u_m \in \mathcal{I}_{k_0+k_1+1}^c$.

We first consider case (i). Following the proof of (D.36), we have

$$\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathrm{HF}^\star}(\eta_{k_\star}^c) \right\|_2 \geqslant 2c_9 \left(\kappa_n^c \overline{\omega}_{l,u}^c \right)^{1/2}, \tag{D.49}$$

where c_9 is defined in Lemma D.4. On the other hand, if (D.48) holds, using (D.42) and (D.43), we have

$$\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathrm{HF}^\star}(\eta_{k_\star}^c) \right\|_2 \leqslant 2 b_{l_{m_0^c}, u_{m_0^c}} (c_\star \kappa_n^c)^{1/2} \leqslant 2c_{15} \left(c_\star \kappa_n^c \overline{\omega}_{l,u}^c \right)^{1/2}. \tag{D.50}$$

Letting $c_\star$ be sufficiently close to zero, (D.49) and (D.50) would lead to a contradiction. As a result, case (i) would not occur when n is sufficiently large. We next turn to case (ii). By (D.36) in the proof of Lemma D.4, (D.49) still holds. On the other hand, since $\eta_{k_0+k_1}^c \leqslant u_{m_0^c} \leqslant u \leqslant \eta_{k_0+k_1}^c + c_5 \varphi_n^c$, by the triangle inequality, we have

$$\begin{aligned}
&\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathrm{HF}^\star}(\eta_{k_\star}^c) \right\|_2 \\
= &\sqrt{\frac{(\eta_{k_\star}^c - l_{m_0^c} + 1)(u_{m_0^c} - \eta_{k_\star}^c)}{u_{m_0^c} - l_{m_0^c} + 1}} \left\| \frac{1}{\eta_{k_\star}^c - l_{m_0^c} + 1} \sum_{t=l_{m_0^c}}^{\eta_{k_\star}^c} \mathbf{G}_t^{\mathrm{HF}^\star} - \frac{1}{u_{m_0^c} - \eta_{k_\star}^c} \sum_{t=\eta_{k_\star}^c+1}^{u_{m_0^c}} \mathbf{G}_t^{\mathrm{HF}^\star} \right\|_2
\end{aligned}$$

$$\leqslant \sqrt{\frac{(\eta^c_{k_\star}-l_{m^c_0}+1)(u_{m^c_0}-\eta^c_{k_\star})}{u_{m^c_0}-l_{m^c_0}+1}}\cdot\left\|\frac{1}{\eta^c_{k_\star}-l_{m^c_0}+1}\sum_{t=l_{m^c_0}}^{\eta^c_{k_\star}}\mathbf{G}^{\mathbf{HF}\star}_t-\frac{1}{u_{m^c_0}-\eta^c_{k_\star}}\sum_{t=\eta^c_{k_\star}+1}^{u_{m^c_0}}\mathbf{G}^{\mathbf{HF}\star}_{t\wedge(u_{m^c_0}-c_5\varphi^c_n)}\right\|_2$$
$$+\sqrt{\frac{(\eta^c_{k_\star}-l_{m^c_0}+1)(u_{m^c_0}-\eta^c_{k_\star})}{u_{m^c_0}-l_{m^c_0}+1}}\cdot\frac{c_5\varphi^c_n b_{u_{m^c_0}-c_5\varphi^c_n,u_{m^c_0}}}{u_{m^c_0}-\eta^c_{k_\star}}$$
$$\leqslant \left(2b_{l_{m^c_0},u_{m^c_0}-c_5\varphi^c_n}+c_5\varphi^c_n b_{u_{m^c_0}-c_5\varphi^c_n,u_{m^c_0}}/\kappa^c_n\right)\sqrt{(\eta^c_{k_\star}-l_{m^c_0}+1)\wedge(u_{m^c_0}-\eta^c_{k_\star})}. \tag{D.51}$$

Noting that

$$\varphi^c_n b_{u_{m^c_0}-c_5\varphi^c_n,u_{m^c_0}}/\kappa^c_n = O\left((\overline{\omega}^c_n)^{1/2}\log^4 n/(\underline{\omega}^c_n\kappa^c_n)\right) \text{ and } 2b_{l_{m^c_0},u_{m^c_0}-c_5\varphi^c_n}\geqslant(\underline{\omega}^c_n)^{1/2},$$

as $\left(\frac{\underline{\omega}^c_n}{\overline{\omega}^c_n}\right)^{1/2}\cdot\frac{\kappa^c_n\underline{\omega}^c_n}{\log^4 n}\to\infty$ from assumption 4(ii), we have

$$\varphi^c_n b_{u_{m^c_0}-c_5\varphi^c_n,u_{m^c_0}}/\kappa^c_n = o\left(b_{l_{m^c_0},u_{m^c_0}}\right),$$

which, together with (D.51), indicates that (D.50) holds as well. However, by letting $c_\star$ approach zero, (D.49) and (D.50) would lead to a contradiction. Hence, case (ii) would not occur when n is sufficiently large. Combining the above arguments, we may complete the proof of (D.47). Furthermore, following the similar arguments and using (D.39), we may prove (D.40).

We finally turn to the proof of (D.41). Consider two cases: (i) $u_{m^c_0}\leqslant\eta^c_{k_\diamond+1}$ and (ii) $\eta^c_{k_\diamond+1}<u_{m^c_0}$. We start with case (i) of $u_{m^c_0}\leqslant\eta^c_{k_\diamond+1}$. For notational simplicity, we let $\nu_i=\eta^c_{k_\diamond}-l_{m^c_0}+1$ and $\nu_h=u_{m^c_0}-\eta^c_{k_\diamond}$, and define $\boldsymbol{\beta}=\left(\beta_1,\cdots,\beta_{q_0(q_0+1)/2}\right)^\intercal$ with

$$\beta_k = C^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0},u_{m_0},k}(\eta^c_{k_\diamond})\left(\frac{\nu_i\nu_h}{\nu_i+\nu_h}\right)^{1/2},$$

where $C^{\mathbf{G},\mathbf{HF}\star}_{l,u,k}(\cdot)$ is the k-th element of $\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l,u,k}(\cdot)$. As $u_{m^c_0}\leqslant\eta^c_{k_\diamond+1}$ in this case, it is easy to verify that

$$C^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0},u_{m_0},k}(\eta^c_{k_\diamond})=\beta_k\left(\frac{\nu_i\nu_h}{\nu_i+\nu_h}\right)^{-1/2},\quad C^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0},u_{m_0},k}(\eta^c_{k_\diamond}+\nu_l)=\beta_k\cdot\frac{\nu_h-\nu_l}{\nu_h}\cdot\left[\frac{\nu_i+\nu_h}{(\nu_i+\nu_l)(\nu_h-\nu_l)}\right]^{1/2}$$

and

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m^c_0},u_{m^c_0}}(\eta^c_{k_\diamond})\right\|_2=\|\boldsymbol{\beta}\|_2\left(\frac{\nu_i\nu_h}{\nu_i+\nu_h}\right)^{-1/2}.$$

As $(\kappa^c_n\underline{\omega}^c_n)/\log^4 n\to\infty$ by Assumption 4(ii), we have $\gamma^c_n=o(\kappa^c_n)$ and consequently $\nu_l<\nu_i$ when

n is large enough. Hence, we have

$$
\begin{aligned}
& \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 - \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_l)\right\|_2 \\
= \ & \|\boldsymbol{\beta}\|_2 \cdot \frac{\sqrt{\nu_i+\nu_h}}{\nu_h}\left(\sqrt{\frac{\nu_h}{\nu_i}} - \sqrt{\frac{\nu_h-\nu_l}{\nu_i+\nu_l}}\right) \\
= \ & \nu_l \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \cdot \frac{\nu_i+\nu_h}{\sqrt{\nu_h}\sqrt{\nu_i+\nu_l}\left(\sqrt{\nu_h}\sqrt{\nu_i+\nu_l}+\sqrt{\nu_h-\nu_l}\sqrt{\nu_i}\right)} \\
\geqslant \ & \nu_l \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \cdot \frac{\nu_i+\nu_h}{\sqrt{\nu_h}\sqrt{2\nu_i}\left(\sqrt{\nu_h}\sqrt{2\nu_i}+\sqrt{\nu_h}\sqrt{\nu_i}\right)} \\
\geqslant \ & \nu_l \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \Big/ \left[2(u_{m_0^c}-l_{m_0^c}+1)\right], \qquad \text{(D.52)}
\end{aligned}
$$

which, together with (D.40), proves (D.41).

We next consider case (ii). Let $\nu_i = \eta^c_{k_\diamond} - l_{m_0^c} + 1$, $\nu_h = (c_{12} \wedge 1)\kappa^c_n/3$, $\nu_j = u_{m_0^c} - \eta^c_{k_\diamond} - \nu_h$ and

$$
\mathbf{V}^c_{\mathbf{G}} = \mathbf{G}^{\mathbf{HF}^\star}_{\eta^c_{k_\diamond}+1} - \frac{1}{u_{m_0^c}-l_{m_0^c}+1}\sum_{t=l_{m_0^c}}^{u_{m_0^c}} \mathbf{G}^{\mathbf{HF}^\star}_t.
$$

From the condition $(\kappa^c_n \underline{\omega}^c_n)/\log^4 n \to \infty$, we may show that $0 \leqslant \nu_l \leqslant \nu_h$ when n is sufficiently large. Then, using the definitions of ν_i, ν_h, ν_j and $\mathbf{V}^c_{\mathbf{G}}$, we readily have that

$$
\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_l)\right\|_2 = \|\boldsymbol{\beta}+\nu_l \mathbf{V}^c_{\mathbf{G}}\|_2 \cdot \left[\frac{\nu_i+\nu_j+\nu_h}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2},
$$

where $\boldsymbol{\beta}$ is defined as in case (i). Define

$$
D(\nu_l) = \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 - \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_l)\right\|_2
$$

and

$$
D_1 = \left[\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 - \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_h)\right\|_2\right] \cdot \frac{\nu_l}{\nu_h} \cdot \left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}.
$$

Note that

$$
\begin{aligned}
D(\nu_l) - D_1 = \ & \left\{1 - \frac{\nu_l}{\nu_h}\cdot\left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}\right\} \cdot \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \\
& - \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_l)\right\|_2 + \frac{\nu_l}{\nu_h}\cdot\left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2} \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0},u_{m_0}}(\eta^c_{k_\diamond}+\nu_h)\right\|_2
\end{aligned}
$$

$$
\begin{aligned}
&= \left\{1 - \frac{\nu_l}{\nu_h}\cdot\left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}\right\}\cdot\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \\
&\quad - \left[\frac{\nu_i+\nu_j+\nu_h}{\nu_h^2(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}(\|\nu_h\boldsymbol{\beta}+\nu_h\nu_l\mathbf{V}^c_{\mathbf{G}}\|_2 - \|\nu_l\boldsymbol{\beta}+\nu_h\nu_l\mathbf{V}^c_{\mathbf{G}}\|_2) \\
&\geqslant \left\{1 - \frac{\nu_l}{\nu_h}\cdot\left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}\right\}\cdot\left[\frac{\nu_i+\nu_j+\nu_h}{\nu_i(\nu_j+\nu_h)}\right]^{1/2}\cdot\|\boldsymbol{\beta}\|_2 \\
&\quad - \left[\frac{(\nu_h-\nu_l)^2(\nu_i+\nu_j+\nu_h)}{\nu_h^2(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}\cdot\|\boldsymbol{\beta}\|_2 \\
&= D_2\times(1+D_3), \qquad \text{(D.53)}
\end{aligned}
$$

where

$$
D_2 = \frac{\|\boldsymbol{\beta}\|_2\nu_l(\nu_h-\nu_l)\sqrt{\nu_i+\nu_j+\nu_h}}{\sqrt{\nu_i(\nu_j+\nu_h)}\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}(\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{\nu_i(\nu_j+\nu_h)})},
$$

and

$$
D_3 = \frac{(\nu_j-\nu_i)(\nu_j-\nu_i-\nu_l)}{(\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{(\nu_i+\nu_h)\nu_j})(\sqrt{\nu_i(\nu_j+\nu_h)}+\sqrt{(\nu_i+\nu_h)\nu_j})}.
$$

As ν_l is smaller than $\nu_h/2$ for n large enough, we have

$$
\begin{aligned}
D_2 &= \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \frac{\nu_l(\nu_h-\nu_l)}{\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\left[\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{\nu_i(\nu_j+\nu_h)}\right]} \\
&\geqslant \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 \frac{\nu_l\nu_h}{2\sqrt{2\nu_i(\nu_j+\nu_h)}\left[\sqrt{2\nu_i(\nu_j+\nu_h)}+\sqrt{\nu_i(\nu_j+\nu_h)}\right]} \\
&\geqslant 2\left(c_{13}\nu_l\kappa^c_n\right)\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2/(u_{m_0^c}-l_{m_0^c}+1)^2. \qquad \text{(D.54)}
\end{aligned}
$$

On the other hand, since $(\nu_j-\nu_i)(\nu_j-\nu_i-\nu_l)$ reaches its minimum at $\nu_j-\nu_i=\nu_l/2$, $\nu_i,\nu_j,\nu_h\geqslant(c_{12}\wedge 1)\kappa^c_n/3$ and $\nu_l=o(\kappa^c_n)$ for $0<\nu_l<c_{14}\gamma^c_n$, we have

$$
\begin{aligned}
D_3 &\geqslant \frac{-\nu_l^2}{4\left[\sqrt{\nu_i(\nu_j+\nu_h)/2}+\sqrt{(\nu_i+\nu_h)\nu_j}\right]\left[\sqrt{\nu_i(\nu_j+\nu_h)}+\sqrt{(\nu_i+\nu_h)\nu_j}\right]} \\
&\geqslant \frac{-\nu_l^2}{4(1+\sqrt{2})(\sqrt{2}+\sqrt{2})[(c_{12}\wedge 1)\kappa^c_n/3]^2}\to 0 \qquad \text{(D.55)}
\end{aligned}
$$

when n large enough. Noting that $\kappa^c_n\asymp n$ in Assumption 4(ii),

$$
\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2 - \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond}+\nu_h)\right\|_2 > -3c_6\log^2 n,
$$

and

$$\frac{\log^2 n}{\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\diamond})\right\|_2} \to 0$$

as $(\kappa_n^c \underline{\omega}_n^c)/\log^4 n \to \infty$, D_1 is dominated by D_2 when n is large enough. This, together with (D.53)–(D.55), indicates that the lower bound of $D(\nu_l)$ is dominated by D_2 when n is large enough. We have finally completed the proof of (D.41) for case (ii). □

LEMMA D.7. *Suppose that the assumptions of Lemma D.4 and (D.21)–(D.23) are satisfied. There exists* $k_0 + 1 \leqslant k_\bullet \leqslant k_0 + k_1$ *such that*

$$\left|s_0^c - \eta^c_{k_\bullet}\right| \leqslant c_{14}\gamma_n^c \tag{D.56}$$

with probability approaching one, as $n \to \infty$, *where* $\gamma_n^c = (\kappa_n^c/\underline{\omega}_n^c)^{1/2}\log^2 n$ *and* c_{14} *is a positive constant defined as in Lemma D.6.*

PROOF. By the definitions of m_0^c and s_0^c, Proposition 3.2 and Lemma D.3, we readily have for any $\eta_k^c \in [l_{m_0^c}, u_{m_0^c}]$,

$$\begin{aligned}
\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_k)\right\|_2 &\leqslant \left\|\mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_0^c},u_{m_0^c}}(\eta^c_k)\right\|_2 + (1+\tau/2)c_6\log^2 n \\
&\leqslant \left\|\mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_0^c},u_{m_0^c}}(s^c_0)\right\|_2 + (1+\tau/2)c_6\log^2 n \\
&\leqslant \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(s^c_0)\right\|_2 + (2+\tau)c_6\log^2 n
\end{aligned} \tag{D.57}$$

with probability approaching one, where τ is a very small positive constant and c_6 is defined in Lemma D.3. Without loss of generality, assume that $s_0^c \in [\eta^c_{\bar{k}}, \eta^c_{\bar{k}+1})$ with $k_0 + 1 \leqslant \bar{k} \leqslant k_0 + k_1$. We next show the consequence when (D.56) fails and consider two cases.

Case (i): only one of $\eta^c_{\bar{k}}$ and $\eta^c_{\bar{k}+1}$ locates within the interval $[l_{m_0^c}, u_{m_0^c})$. Without loss of generality, assume that $\eta^c_{\bar{k}}$ is in the interval $[l_{m_0^c}, u_{m_0^c})$. Let $\eta^c_{k_\bullet} = \eta^c_{\bar{k}}$. From Lemma D.5, without loss of generality, we may consider that $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ (treated as a function of s) locally decreases in the interval $[\eta^c_{k_\bullet}, u_{m_0^c})$ which includes the point $s = s_0^c$. From (D.57),

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\bullet})\right\|_2 \geqslant \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(s^c_0)\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\star})\right\|_2 - 3c_6\log^2 n,$$

where $k_\star$ is defined as in the proof of Lemma D.6. This indicates that (D.39) is satisfied with $k_\diamond = k_\bullet$ and $s_\star^c = \eta^c_{k_\star}$. By (D.41) in Lemma D.6, letting $c_{14} > 0$ be sufficiently large and noting that $\kappa_n^c = O(|u_{m_0^c} - l_{m_0^c}|)$, we may show that there exists $s_1 \in (\eta^c_{k_\bullet}, \eta^c_{k_\bullet} + c_{14}\gamma_n^c]$ such that

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\bullet})\right\|_2 \geqslant \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}^\star}_{l_{m_0^c},u_{m_0^c}}(s_1)\right\|_2 + (2+\tau)c_6\log^2 n$$

with probability approaching one. If (D.56) fails, as $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ is locally decreasing, we must have

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\bullet})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_1)\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c)\right\|_2,$$

and thus

$$\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k_\bullet})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c)\right\|_2 + (2+\tau)c_6\log^2 n,$$

leading to a contradiction with (D.57).

Case (ii): both η^c_k, and η^c_{k+1} locate in the interval $[l_{m_0^c}, u_{m_0^c})$. By Lemma D.5 again, we may show that $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ (treated as a function of s) is either monotonic or first decreasing and then increasing on the interval $\left[\eta^c_k, \eta^c_{k+1}\right]$, and consequently

$$\left\{\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_k)\right\|_2 \vee \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta^c_{k+1})\right\|_2\right\} \geqslant \left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c)\right\|_2.$$

We further consider two scenarios: (ii.1) $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ locally decreases at the point $s = s_0^c$; and (ii.2) $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ locally increases at the point $s = s_0^c$. To save the space, we only give the proof for scenario (ii.1) as that for (ii.2) is similar (by letting $\eta^c_{k_\bullet} = \eta^c_{k+1}$). When $\left\|\mathbf{C}^{\mathbf{G},\mathbf{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s)\right\|_2$ locally decreases at the point $s = s_0^c$, we let $\eta^c_{k_\bullet} = \eta^c_k$. If (D.56) fails, following the arguments as in case (i), there would be a contradiction with (D.57).

Combining cases (i) and (ii) above, the proof of the lemma has been completed. □

We next introduce some additional notation. For $k = 1, \cdots, q_0(q_0+1)/2$, let

$$\begin{aligned}
\mathbf{Z}^{\mathbf{HF}\star}_{\bullet,k} &= \left(Z^{\mathbf{HF}\star}_{l_{m_0^c},k}, \cdots Z^{\mathbf{HF}\star}_{u_{m_0^c},k}\right)^{\intercal},\\
\mathbf{G}^{\mathbf{HF}\star}_{\bullet,k} &= \left(G^{\mathbf{HF}\star}_{l_{m_0^c},k}, \cdots, G^{\mathbf{HF}\star}_{u_{m_0^c},k}\right)^{\intercal},\\
\mathbf{z}^{\mathbf{HF}\star}_{\bullet,k} &= \left(z^{\mathbf{HF}\star}_{l_{m_0^c},k}, \cdots, z^{\mathbf{HF}\star}_{u_{m_0^c},k}\right)^{\intercal},
\end{aligned}$$

where $Z^{\mathbf{HF}\star}_{t,k}$, $G^{\mathbf{HF}\star}_{t,k}$ and $z^{\mathbf{HF}\star}_{t,k}$ are the k-th element in the vectors $\mathbf{Z}^{\mathbf{HF}\star}_t$, $\mathbf{G}^{\mathbf{HF}\star}_t$ and $\mathbf{z}^{\mathbf{HF}\star}_t$, respectively. The following lemma further improves the convergence rate of the estimated break points given in Lemma D.7 above.

LEMMA D.8. *Suppose that the conditions of Lemma D.7 are satisfied. With probability approaching one, we have*

$$\left|s_0^c - \eta^c_{k_\bullet}\right| \leqslant c_{16}\varphi^c_n \tag{D.58}$$

as $n \to \infty$, *where* c_{16} *is a positive constant and* $\varphi^c_n = \log^4 n / \underline{\omega}^c_n$.

PROOF. Let $\langle\cdot,\cdot\rangle$ denote the inner product between two vectors and $\boldsymbol{\psi}_{l,u}^{s} = (\psi_l^s, \cdots, \psi_u^s)^\intercal$ be a vector of constants such that ψ_t^s is positive for $t = l, \cdots, s$ and negative for $t = s+1, \cdots, u$, $\sum_{t=l}^{u} \psi_t^s = 0$ and $\sum_{t=l}^{u} (\psi_t^s)^2 = 1$. Note that, for any vector $\boldsymbol{\nu} = (\nu_l, \cdots, \nu_u)^\intercal$, we have

$$\langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}^s, \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}^s \rangle = \langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \rangle - \langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\psi}_{l,u}^s \rangle^2, \tag{D.59}$$

where $\bar{\boldsymbol{\nu}}^s = \bar{\boldsymbol{\nu}} + \langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\psi}_{l,u}^s \rangle \boldsymbol{\psi}_{l,u}^s$, $\bar{\boldsymbol{\nu}} = \left[\left(\frac{1}{u-l+1} \right) \sum_{t=l}^{u} \nu_t \right] \mathbf{1}_{u-l+1}$, and $\mathbf{1}_q$ is a q-dimensional column vector with all the elements being ones. From (D.59), we readily have

$$\left| \langle \boldsymbol{\nu}, \boldsymbol{\psi}_{l,u}^s \rangle \right|^2 = \left| \langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\psi}_{l,u}^s \rangle \right|^2 = - \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}^s \|_2^2 + \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \|_2^2. \tag{D.60}$$

From (D.60), we can derive the following useful inequality: for $l \leqslant s \leqslant u$ and any vector $\boldsymbol{\omega} = (\omega_l, \cdots, \omega_u)^\intercal$,

$$\| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}^s \|_2^2 \leqslant \| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}}^s \|_2^2, \tag{D.61}$$

where $\bar{\boldsymbol{\omega}}^s$ is defined similarly to $\bar{\boldsymbol{\nu}}^s$ with $\boldsymbol{\nu}$ replaced by $\boldsymbol{\omega}$. In fact, (D.61) can be easily proved by noting that

$$\begin{aligned}
& \| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}}^s \|_2^2 - \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}^s \|_2^2 \\
= \; & \left\| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}} + \langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle \boldsymbol{\psi}_{l,u}^s \right\|_2^2 - \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \|_2^2 + \left\langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\psi}_{l,u}^s \right\rangle^2 \\
= \; & \| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}} \|_2^2 + \left\langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \right\rangle^2 + 2 \left\langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \right\rangle \left\langle \boldsymbol{\nu} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \right\rangle - \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \|_2^2 + \left\langle \boldsymbol{\nu} - \bar{\boldsymbol{\nu}}, \boldsymbol{\psi}_{l,u}^s \right\rangle^2 \\
= \; & \| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}} \|_2^2 - \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \|_2^2 + \left\langle \boldsymbol{\nu} + \boldsymbol{\omega} - \bar{\boldsymbol{\nu}} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \right\rangle^2 \geqslant 0
\end{aligned}$$

since $\| \boldsymbol{\nu} - \bar{\boldsymbol{\omega}} \|_2^2 \geqslant \| \boldsymbol{\nu} - \bar{\boldsymbol{\nu}} \|_2^2$.

Let $C_{l,u,k}^{\mathrm{HF}\star}(s)$ be the k-th element in the vector $\mathbf{C}_{l,u}^{\mathrm{HF}\star}(s)$. Using the notion of inner product, we may write $C_{l_{m_0^c}, u_{m_0^c}, k}^{\mathrm{HF}\star}(s)$ as $\langle \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star}, \boldsymbol{\psi}_{l_{m_0^c}, u_{m_0^c}}^s \rangle$. For $l_{m_0^c} \leqslant s < u_{m_0^c}$, define $Q_k^{\mathrm{HF}\star}(s;1) = |\langle \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star}, \boldsymbol{\psi}_{l_{m_0^c}, u_{m_0^c}}^s \rangle|^2$, and let $\bar{\mathbf{Z}}_{\bullet,k}^{\mathrm{HF}\star s}$ and $\bar{\mathbf{G}}_{\bullet,k}^{\mathrm{HF}\star s}$ be defined similarly to $\bar{\boldsymbol{\nu}}^s$ but with $\boldsymbol{\psi}_{l,u}^s$ replaced by $\boldsymbol{\psi}_{l_{m_0^c}, u_{m_0^c}}^s$, and $\boldsymbol{\nu}$ replaced by $\mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star}$ and $\mathbf{G}_{\bullet,k}^{\mathrm{HF}\star}$, respectively. By (D.60), we readily have

$$Q_k^{\mathrm{HF}\star}(s;1) = - \left\| \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathrm{HF}\star s} \right\|_2^2 + \left\| \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathrm{HF}\star} \right\|_2^2,$$

where $\bar{\mathbf{Z}}_{\bullet,k}^{\mathrm{HF}\star}$ is defined as $\bar{\boldsymbol{\nu}}$ but with $\boldsymbol{\nu}$ replaced by $\mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star}$. For $l_{m_0^c} \leqslant s < u_{m_0^c}$, define

$$Q_k^{\mathrm{HF}\star}(s;2) = - \left\| \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathrm{HF}\star s} \right\|_2^2 + \left\| \mathbf{Z}_{\bullet,k}^{\mathrm{HF}\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathrm{HF}\star} \right\|_2^2.$$

By (D.61), we may show that

$$Q_k^{\mathbf{HF}^\star}(s;1) \geqslant Q_k^{\mathbf{HF}^\star}(s;2), \quad k = 1, \cdots, q_0(q_0+1)/2. \tag{D.62}$$

Since $\mathbf{Z}_{\bullet,k}^{\mathbf{HF}^\star} = \mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} + \boldsymbol{z}_{\bullet,k}^{\mathbf{HF}^\star}$, we readily have

$$Q_k^{\mathbf{HF}^\star}(s;1) = -\left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\rangle,$$

and

$$Q_k^{\mathbf{HF}^\star}(s;2) = -\left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\rangle.$$

Letting

$$Q_k^{\mathbf{HF}^\star}(s;3) = -\left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^\star}\right\rangle,$$

by (D.61), we have

$$Q_k^{\mathbf{HF}^\star}(s;3) \geqslant Q_k^{\mathbf{HF}^\star}(s;1) \geqslant 0. \tag{D.63}$$

Next we prove the following result: there exists a sufficiently large constant $c_{17} > 0$,

$$\sum_{k=1}^{q_0(q_0+1)/2} \left[Q_k^{\mathbf{HF}^\star}(s_0^c;3) - Q_k^{\mathbf{HF}^\star}(\eta_{k_\bullet}^c;2)\right] \geqslant -c_{17} \tag{D.64}$$

holds with probability approaching one. Let $Q_k^{\widehat{\mathbf{F}}}(s;1)$ be defined similarly to $Q_k^{\mathbf{HF}^\star}(s;1)$ but with $\mathbf{HF}_t^\star$ replaced by $\widehat{\mathbf{F}}_t$. By (D.62), (D.63), Proposition 3.2 and the definition of s_0^c, we have

$$\begin{aligned}
\sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(s_0^c;3) &\geqslant \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(s_0^c;1) = \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\widehat{\mathbf{F}}}(s_0^c;1) + O_P(1) \\
&\geqslant \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\widehat{\mathbf{F}}}(\eta_{k_\bullet}^c;1) + O_P(1) = \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(\eta_{k_\bullet}^c;1) + O_P(1) \\
&\geqslant \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(\eta_{k_\bullet}^c;2) + O_P(1),
\end{aligned}$$

proving (D.64).

Letting $c_{16} > 0$ be sufficiently large, we next show that the assertion of $\left|s_0^c - \eta_{k_\bullet}^c\right| > c_{16}\varphi_n^c$

would lead to a contradiction with (D.64), which consequently proves (D.58). Defining

$$Q_k^{\mathbf{HF}^\star}(s;4) = \left|\langle \mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star}, \psi^s_{l_{m_0^c}, u_{m_0^c}}\rangle\right|^2 = -\left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\|^2 + \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^\star}\right\|^2,$$

we have

$$\begin{aligned}
& Q_k^{\mathbf{HF}^\star}(s;3) - Q_k^{\mathbf{HF}^\star}(\eta^c_{k_\bullet};2) \\
= & \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\|^2 - \left\|\mathbf{G}_{\bullet,k}^{\mathbf{HF}^\star} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\|^2 + 2\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\rangle \\
= & 2\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\rangle - \left[Q_k^{\mathbf{HF}^\star}(\eta^c_{k_\bullet};4) - Q_k^{\mathbf{HF}^\star}(s;4)\right]. \qquad \text{(D.65)}
\end{aligned}$$

We next show that with probability approaching one,

$$\sum_{k=1}^{q_0(q_0+1)/2}\left|\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\rangle\right| \leqslant c_{18}(\log^2 n)\max\left\{\frac{\left|s_0^c - \eta^c_{k_\star}\right|\cdot(\varpi^c_n)^{1/2}}{(\kappa^c_n)^{1/2}}, \left|s_0^c - \eta^c_{k_\bullet}\right|^{1/2}(\varpi^c_n)^{1/2}, \log^2 n\right\} \tag{D.66}$$

and

$$\sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(\eta^c_{k_\bullet};4) - \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\mathbf{HF}^\star}(s_0^c;4) \geqslant c_{19}\left|s_0^c - \eta^c_{k_\bullet}\right|\varpi^c_n, \tag{D.67}$$

where c_{18} and c_{19} are two positive constants.

Without loss of generality, we assume that $s_0^c \geqslant \eta^c_{k_\bullet}$. Note that the left hand side of (D.66) can be decomposed as follows:

$$\begin{aligned}
\sum_{k=1}^{q_0(q_0+1)/2}\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\rangle = & \sum_{k=1}^{q_0(q_0+1)/2}\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}}\right\rangle \\
& + \sum_{k=1}^{q_0(q_0+1)/2}\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star\eta^c_{k_\bullet}}}\right\rangle. \qquad \text{(D.68)}
\end{aligned}$$

Following standard calculations, we have

$$\begin{aligned}
\left\langle \mathbf{z}_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}\right\rangle = & \left(\sum_{t=l_{m_0^c}}^{s} + \sum_{t=s+1}^{u_{m_0^c}}\right) z_{t,k}^{\mathbf{HF}^\star}\left(\bar{Z}_{t,k}^{\mathbf{HF}^{\star s}} - \bar{G}_{t,k}^{\mathbf{HF}^{\star s}}\right) \\
= & \frac{1}{s - l_{m_0^c} + 1}\left(\sum_{t=l_{m_0^c}}^{s} z_{t,k}^{\mathbf{HF}^\star}\right)^2 + \frac{1}{u_{m_0^c} - s}\left(\sum_{t=s+1}^{u_{m_0^c}} z_{t,k}^{\mathbf{HF}^\star}\right)^2 \qquad \text{(D.69)}
\end{aligned}$$

for any s, where $Z_{t,k}^{\mathbf{HF}^{\star s}}$ and $\bar{G}_{t,k}^{\mathbf{HF}^{\star s}}$ are the $(t-l_{m_0^c}+1)$-th element in $\bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s}}$ and $\bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s}}$, respectively. By the definition of $z_{t,k}^{\mathbf{HF}^\star}$, the Cauchy-Schwarz inequality, using Lemma D.3 and noting that $\|\mathbf{L}_{q_0}(\mathbf{H}\otimes\mathbf{H})\mathbf{D}_{q_0}\|_F^2 < \infty$ with probability approaching one, we have, uniformly over s

$$\frac{1}{s-l_{m_0^c}+1}\left(\sum_{t=l_{m_0^c}}^{s} z_{t,k}^{\mathbf{HF}^\star}\right)^2 = O_P\left(\log^4 n\right), \quad \frac{1}{u_{m_0^c}-s}\left(\sum_{t=s+1}^{u_{m_0^c}} z_{t,k}^{\mathbf{HF}^\star}\right)^2 = O_P\left(\log^4 n\right),$$

which indicates that

$$\left\langle z_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{Z}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}}\right\rangle = O_P\left(\log^4 n\right) \tag{D.70}$$

for $k=1,\cdots,q_0(q_0+1)/2$. On the other hand,

$$\begin{aligned}\left\langle z_{\bullet,k}^{\mathbf{HF}^\star}, \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{\mathbf{G}}_{\bullet,k}^{\mathbf{HF}^{\star \eta_{k_\bullet}^c}}\right\rangle &= \left(\sum_{t=l_{m_0^c}}^{\eta_{k_\bullet}^c} + \sum_{t=\eta_{k_\bullet}^c+1}^{s_0^c} + \sum_{t=s_0^c+1}^{u_{m_0^c}}\right) z_{t,k}^{\mathbf{HF}^\star}\left(\bar{G}_{t,k}^{\mathbf{HF}^{\star s_0^c}} - \bar{G}_{t,k}^{\mathbf{HF}^{\star \eta_{k_\bullet}^c}}\right) \\ &=: \Pi_1+\Pi_2+\Pi_3. \end{aligned} \tag{D.71}$$

Recall that $b_{l,u} = \sup_{l\leqslant t\leqslant u}\left\|\mathbf{G}_t^{\mathbf{HF}^\star} - \frac{1}{u-l+1}\sum_{t=l}^{u}\mathbf{G}_t^{\mathbf{HF}^\star}\right\|_2$. As in (D.43),

$$b_{l+c_5\varphi_n^c, u-c_5\varphi_n^c} \leqslant \|\mathbf{L}_{q_0}(\mathbf{H}\otimes\mathbf{H})\mathbf{D}_{q_0}\|_F \cdot \sqrt{q_0(q_0+1)/2}\cdot(\overline{\omega}_n^c)^{1/2} = O_P\left((\overline{\omega}_n^c)^{1/2}\right),$$

which, together with Cauchy-Schwarz inequality and (D.70), indicates that

$$\begin{aligned}\sum_{k=1}^{q_0(q_0+1)/2}|\Pi_1| &\leqslant \sum_{k=1}^{q_0(q_0+1)/2}\left|\sum_{t=l_{m_0^c}}^{\eta_{k_\bullet}^c} z_{t,k}^{\mathbf{HF}^\star}\right|\cdot\left|\frac{1}{s_0^c-l_{m_0^c}+1}\sum_{t=l_{m_0^c}}^{s_0^c} G_{t,k}^{\mathbf{HF}^\star} - \frac{1}{\eta_{k_\bullet}^c-l_{m_0^c}+1}\sum_{t=l_{m_0^c}}^{\eta_{k_\bullet}^c} G_{t,k}^{\mathbf{HF}^\star}\right| \\ &\leqslant \left\|\sum_{t=l_{m_0^c}}^{\eta_{k_\bullet}^c} z_t^{\mathbf{HF}^\star}\right\|_2\cdot\left\|\frac{1}{s_0^c-l_{m_0^c}+1}\sum_{t=l_{m_0^c}}^{s_0^c}\mathbf{G}_t^{\mathbf{HF}^\star} - \frac{1}{\eta_{k_\bullet}^c-l_{m_0^c}+1}\sum_{t=l_{m_0^c}}^{\eta_{k_\bullet}^c}\mathbf{G}_t^{\mathbf{HF}^\star}\right\|_2 \\ &\leqslant O_P\left((\eta_{k_\bullet}^c-l_{m_0^c}+1)^{1/2}\log^2 n\right)\cdot\frac{|s_0^c-\eta_{k_\bullet}^c|\, b_{l+c_5\varphi_n^c, u-c_5\varphi_n^c}}{s_0^c-l_{m_0^c}+1} \\ &\leqslant O_P\left(\log^2 n\,|s_0^c-\eta_{k_\bullet}^c|\cdot(\overline{\omega}_n^c/\kappa_n^c)^{1/2}\right). \end{aligned} \tag{D.72}$$

This is also the asymptotic order for $\sum_{k=1}^{q_0(q_0+1)/2}\Pi_3$. Similarly, we may show that

$$\sum_{k=1}^{q_0(q_0+1)/2}\Pi_2 = O_P\left(\log^2 n\,|s_0^c-\eta_{k_\bullet}^c|^{1/2}(\overline{\omega}_n^c)^{1/2}\right). \tag{D.73}$$

With (D.68) and (D.70)–(D.73), we can complete the proof of (D.66).

We next turn to the proof of (D.67). By (D.41), we have

$$
\begin{aligned}
& \sum_{k=1}^{q_0(q_0+1)/2} \left[Q_k^{\mathrm{HF}\star}(\eta_{k_\bullet}^c;4) - Q_k^{\mathrm{HF}\star}(s_0^c;4)\right] \\
= & \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta_{k_\bullet}^c) \right\|_2^2 - \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c) \right\|_2^2 \\
= & \left[\left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta_{k_\bullet}^c) \right\|_2 - \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c) \right\|_2 \right] \cdot \left[\left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta_{k_\bullet}^c) \right\|_2 + \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c) \right\|_2 \right] \\
\geqslant & \left[\left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta_{k_\bullet}^c) \right\|_2 - \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(s_0^c) \right\|_2 \right] \cdot \left\| C^{G,\mathrm{HF}\star}_{l_{m_0^c},u_{m_0^c}}(\eta_{k_\bullet}^c) \right\|_2 \\
\geqslant & c_{19} \left| s_0^c - \eta_{k_\bullet}^c \right| (\varpi_n^c/\kappa_n^c)^{1/2} \cdot (\kappa_n^c \varpi_n^c)^{1/2} \\
= & c_{19} \left| s_0^c - \eta_{k_\bullet}^c \right| \varpi_n^c
\end{aligned}
\tag{D.74}
$$

with probability approaching one. This completes the proof of (D.67).

Suppose that (D.58) fails, i.e.,$\left| s_0^c - \eta_{k_\bullet}^c \right| > c_{16}\varphi_n^c$. By (D.65)–(D.67), Lemma D.7 and letting $c_{16} > 0$ be sufficiently large, we have

$$
\begin{aligned}
& \sum_{k=1}^{q_0(q_0+1)/2} \left[Q_k^{\mathrm{HF}\star}(s_0^c;3) - Q_k^{\mathrm{HF}\star}(\eta_{k_\bullet}^c;2)\right] \\
\leqslant & c_{18} \log^2 n \max\left\{ \frac{\left| s_0^c - \eta_{k_\bullet}^c \right| \cdot (\varpi_n^c)^{1/2}}{(\kappa_n^c)^{1/2}}, \left| s_0^c - \eta_{k_\bullet}^c \right|^{1/2} (\varpi_n^c)^{1/2}, \log^2 n \right\} - c_{19} \left| s_0^c - \eta_{k_\bullet}^c \right| \varpi_n^c \\
\leqslant & -c_{17} \log^4 n < -c_{17},
\end{aligned}
\tag{D.75}
$$

which contradicts with (D.64). We have finally proved (D.58), which completes the proof of Lemma D.8. □

Proof of Theorem 3.1. According to the WBS-Cov algorithm, we have $l = 1$ and $u = n$ at the start of the algorithm and (D.21)–(D.23) are automatically satisfied. Then, by (3.5), Lemmas D.4 and D.8, we can estimate a change point s_0^c which satisfies (D.58) with probability approaching one. Furthermore, (D.40) in Lemma D.6 shows that s_0^c is not close to l or u, thus it is a newly detected change point. By (D.58), we may show that (D.21)–(D.23) still hold within each segment until all of the change points in the common component are detected, and consequently, the estimated change points satisfy the convergence result (D.58) with probability approaching one. Once all of the change points are detected, the bounds of each segment l and u must fall into one of the following three scenarios: (i) there exists $1 \leqslant k \leqslant K_1$ such that $\eta_k^c < l < u \leqslant \eta_{k+1}^c$; (ii) there exists $1 \leqslant k \leqslant K_1$ such that $l \leqslant \eta_k^c \leqslant u$ and $(\eta_k^c - l + 1) \wedge (u - \eta_k^c) \leqslant c_{16}\varphi_n^c$; (iii) there exists $1 \leqslant k \leqslant K_1$

such that $l \leqslant \eta_k^c < \eta_{k+1}^c \leqslant u$ and $(\eta_k^c - l + 1) \vee (u - \eta_{k+1}^c) \leqslant c_{16}\varphi_n^c$, where c_{16} is defined in Lemma D.8. For l and u satisfy either of scenarios (i)–(iii), we may show that there exists a constant $c_{20} > 0$ such that

$$\mathsf{P}\left(\max_{l_{m_0^c} \leqslant s < u_{m_0^c}} \left\| \mathbf{C}^{\widehat{\mathbf{F}}}_{l_{m_0^c}, u_{m_0^c}}(s) \right\|_2 \leqslant c_{20} \cdot \log^2 n\right) \to 1 \tag{D.76}$$

as $n \to \infty$. By (3.5), Lemmas D.4 and D.6, no further change point would be detected. Letting $\iota^c = c_{16}$, the proof of Theorem 3.1 is completed. □

Appendix E: Proofs of the WSBS-Cov theory for the idiosyncratic components

We next give the detailed proofs of the asymptotic theory in Section 3.2.

PROOF OF PROPOSITION 3.3. By (A.4) in Assumption 3(ii) and Proposition 3.1, the Bonferroni and Markov inequalities, we may show that

$$\max_{1 \leqslant t \leqslant n} \|\mathbf{F}_t^\star\|_2 = O_P\left(\sqrt{\log n}\right). \tag{E.1}$$

Then, by the definition (2.7), (D.1), (D.2), (E.1), Proposition 3.1 and Assumption 4(i), we readily have

$$\max_{1 \leqslant t \leqslant n} \max_{1 \leqslant j \leqslant d} |\widehat{\epsilon}_{tj} - \epsilon_{tj}| = \max_{1 \leqslant t \leqslant n} \max_{1 \leqslant j \leqslant d} \left| \widehat{\lambda}_j^\intercal \widehat{\mathbf{F}}_t - ((\mathbf{H}^{-1})^\intercal \lambda_j^\star)^\intercal \mathbf{H} \mathbf{F}_t^\star \right| = O_P\left(\left[\frac{(\log d)(\log n)}{n}\right]^{1/2}\right). \tag{E.2}$$

Following the proof of Proposition 3.2 and using Assumption 5, we may complete the proof of Proposition 3.3. □

We next turn to proof of Theorem 3.2. As in Appendix D, we let the two positive integers l and u denote the "lower" and "upper" bounds of a segment, and assume that

$$\eta_{k_0}^e \leqslant l < \eta_{k_0+1}^e < \cdots < \eta_{k_0+k_1}^e < u \leqslant \eta_{k_0+k_1+1}^e, \tag{E.3}$$

where $k_0 \in \{0, \cdots, K_2 - k_1\}$ and $k_1 \in \{1, \cdots, K_2 - k_0\}$. Like in the proofs of the lemmas in Appendix D, the following two conditions are key to the WSBS-Cov asymptotic analysis: for some $1 \leqslant k \leqslant k_1$,

$$l < \eta_{k_0+k}^e - c_{21}\kappa_n^e < \eta_{k_0+k}^e + c_{21}\kappa_n^e < u \tag{E.4}$$

and

$$\left\{(l-\eta^e_{k_0})\wedge(\eta^e_{k_0+1}-l)\right\}\vee\left\{(u-\eta^e_{k_0+k_1})\wedge(\eta^e_{k_0+k_1+1}-u)\right\}\leqslant c_{22}\varphi^e_{n,d}, \tag{E.5}$$

where c_{21} and c_{22} are two positive constants, κ^e_n and $\varphi^e_{n,d}$ are defined in Theorem 3.2. Define the intervals

$$\mathcal{I}^e_k=\left[\eta^e_{k-1}+(\eta^e_k-\eta^e_{k-1})/3,\ \eta^e_{k-1}+2(\eta^e_k-\eta^e_{k-1})/3\right],\ \ k=1,\cdots,K_2+1,$$

and the event

$$\mathcal{D}^e_n=\left\{\forall\, k=1,\cdots,K_2,\ \exists\ m=1,\cdots,M^e_n \text{ such that } l_m\in\mathcal{I}^e_k \text{ and } u_m\in\mathcal{I}^e_{k+1}\right\},$$

where M^e_n is defined in Section 2.4. The following lemma is an extension of Lemma D.2 to WSBS-Cov.

LEMMA E.1. *Letting $\overline{\mathcal{D}}^e_n$ be the complement of $\mathcal{D}^e_n$, we have*

$$\mathsf{P}\left(\overline{\mathcal{D}}^e_n\right)\leqslant K_2\left[1-(\kappa^e_n/(3n))^2\right]^{M^e_n}, \tag{E.6}$$

where κ^e_n is defined in Theorem 3.2.

PROOF. The proof is the same as Lemma D.2. Details are omitted here. □

Note that

$$\epsilon_{ti}\epsilon_{tj}=\mathsf{E}\left[\epsilon_{ti}\epsilon_{tj}\right]+\left(\epsilon_{ti}\epsilon_{tj}-\mathsf{E}\left[\epsilon_{ti}\epsilon_{tj}\right]\right)=:G^\epsilon_{t,ij}+z^\epsilon_{t,ij}$$

and from (3.7)

$$\begin{aligned}
c^{\epsilon,\widehat{\sigma}}_{l,u}(s;i,j) &= \frac{1}{\widehat{\sigma}_{l,u}(i,j)}\sqrt{\frac{(s-l+1)(u-s)}{(u-l+1)}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}\epsilon_{ti}\epsilon_{tj}-\frac{1}{u-s}\sum_{t=s+1}^{u}\epsilon_{ti}\epsilon_{tj}\right)\\
&= \frac{1}{\widehat{\sigma}_{l,u}(i,j)}\sqrt{\frac{(s-l+1)(u-s)}{(u-l+1)}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}G^\epsilon_{t,ij}-\frac{1}{u-s}\sum_{t=s+1}^{u}G^\epsilon_{t,ij}\right)\\
&\quad+\frac{1}{\widehat{\sigma}_{l,u}(i,j)}\sqrt{\frac{(s-l+1)(u-s)}{(u-l+1)}}\left(\frac{1}{s-l+1}\sum_{t=l}^{s}z^\epsilon_{t,ij}-\frac{1}{u-s}\sum_{t=s+1}^{u}z^\epsilon_{t,ij}\right)\\
&=: \frac{1}{\widehat{\sigma}_{l,u}(i,j)}c^{G,\epsilon}_{l,u}(s;i,j)+\frac{1}{\widehat{\sigma}_{l,u}(i,j)}c^{z,\epsilon}_{l,u}(s;i,j)\\
&=: c^{G,\epsilon,\widehat{\sigma}}_{l,u}(s;i,j)+c^{z,\epsilon,\widehat{\sigma}}_{l,u}(s;i,j).
\end{aligned} \tag{E.7}$$

Let $\mathbf{C}^\epsilon_{l,u}(s)$ denote half-vectorisation of a symmetric $d\times d$ matrix with the (i,j)-entry being $c^{\epsilon,\widehat{\sigma}}_{l,u}(s;i,j)$. The definitions of $\mathbf{C}^{\mathbf{G},\epsilon}_{l,u}(s)$ and $\mathbf{C}^{z,\epsilon}_{l,u}(s)$ are similar to $\mathbf{C}^\epsilon_{l,u}(s)$ but with $c^{\epsilon,\widehat{\sigma}}_{l,u}(s;i,j)$ replaced by $c_{l,u}^{G,\epsilon,\widehat{\sigma}}(s;i,j)$ and $c_{l,u}^{z,\epsilon,\widehat{\sigma}}(s;i,j)$, respectively. Note that

$$\mathbf{C}_{l,u}^{\epsilon}(s) = \mathbf{C}_{l,u}^{G,\epsilon}(s) + \mathbf{C}_{l,u}^{z,\epsilon}(s), \quad l \leqslant s < u. \tag{E.8}$$

The following lemma derives an asymptotic order for $\|\mathbf{C}_{l,u}^{z,\epsilon}(s)\|_\infty$ uniformly over l, u and s, where $\|\cdot\|_\infty$ denotes the l_∞-norm.

LEMMA E.2. *Suppose that Assumptions 1, 3(ii) and 5 in Appendix A are satisfied. There exists a positive constant c_{23} such that*

$$\mathsf{P}\left(\max_{(l,u):\ 1\leqslant l<u\leqslant n}\ \max_{s:\ l\leqslant s<u}\left\|\mathbf{C}_{l,u}^{z,\epsilon}(s)\right\|_\infty > c_{23}\cdot \log^2(nd)\right) \to 0 \tag{E.9}$$

as $n, d \to \infty$.

PROOF. From the definition of the l_∞-norm, we only need to show that

$$\mathsf{P}\left(\max_{(l,u):\ 1\leqslant l<u\leqslant n}\ \max_{(i,j):\ 1\leqslant i,j\leqslant d}\ \max_{s:\ l\leqslant s<u}\left|c_{l,u}^{z,\epsilon,\widehat{\sigma}}(s;i,j)\right| > c_{23}\log^2(nd)\right) \to 0, \tag{E.10}$$

where $c_{l,u}^{z,\epsilon,\widehat{\sigma}}(s;i,j)$ is defined in (E.7).

By Assumption 5, we readily have

$$\max_{(l,u):\ 1\leqslant l<u\leqslant n}\ \max_{(i,j):\ 1\leqslant i,j\leqslant d}\frac{1}{\widehat{\sigma}_{l,u}(i,j)} \leqslant \frac{1}{\underline{\sigma}}.$$

Letting

$$c_{l,u}^{z,\epsilon}(s;i,j,1) = \sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s} z_{t,ij}^{\epsilon}$$

and

$$c_{l,u}^{z,\epsilon}(s;i,j,2) = \sqrt{\frac{s-l+1}{u-l+1}}\cdot\frac{1}{\sqrt{u-s}}\cdot\sum_{t=s+1}^{u} z_{t,ij}^{\epsilon},$$

it suffices to prove that

$$\mathsf{P}\left(\max_{(l,u):\ 1\leqslant l<u\leqslant n}\ \max_{(i,j):\ 1\leqslant i,j\leqslant d}\ \max_{s:\ l\leqslant s<u}\left|c_{l,u}^{z,\epsilon}(s;i,j,k)\right| > \frac{c_{23}\underline{\sigma}}{2}\log^2(nd)\right) \to 0, \tag{E.11}$$

for $k = 1$ and 2.

The proof of (E.11) is similar to the proof of (D.28) in Lemma D.3. Define

$$\overline{z}_{t,ij}^{\epsilon} = z_{t,ij}^{\epsilon}\cdot \mathcal{I}\left(\left|z_{t,ij}^{\epsilon}\right| \leqslant c_{24}\log(nd)\right),\ \ \widetilde{z}_{t,ij}^{\epsilon} = z_{t,ij}^{\epsilon}\cdot \mathcal{I}\left(\left|z_{t,ij}^{\epsilon}\right| > c_{24}\log(nd)\right),$$

where $c_{24} > 0$ is a sufficiently large constant to be determined later. Let $\bar{c}^{z,\epsilon}_{l,u}(s;i,j,1)$ and $\widetilde{c}^{z,\epsilon}_{l,u}(s;i,j,1)$ be defined similarly to $c^{z,\epsilon}_{l,u}(s;i,j,1)$ but with $z^{\epsilon}_{t,ij}$ replaced by $\bar{z}^{\epsilon}_{t,ij} - \mathsf{E}\left[\bar{z}^{\epsilon}_{t,ij}\right]$ and $\widetilde{z}^{\epsilon}_{t,ij} - \mathsf{E}\left[\widetilde{z}^{\epsilon}_{t,ij}\right]$, respectively.

From (A.4) in Assumption 3(ii), there exists a positive constant $\iota_1 > 0$ such that

$$\max_{(i,j):\,1\leqslant i,j\leqslant d}\max_{1\leqslant t\leqslant n}\mathsf{E}\left[\exp\left\{\iota_1\left|z^{\epsilon}_{t,ij}\right|\right\}\right] < \infty.$$

Consequently, we can show that

$$\begin{aligned}
\mathsf{E}\left[\left|\widetilde{z}^{\epsilon}_{t,ij}\right|\right] &\leqslant \left\{\mathsf{E}\left[\left|z^{\epsilon}_{t,ij}\right|^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\left|z^{\epsilon}_{t,k}\right| > c_{24}\log(nd)\right)\right\}^{1/2}\\
&= \left\{\mathsf{E}\left[\left|z^{\epsilon}_{t,ij}\right|^2\right]\right\}^{1/2}\left\{\mathsf{P}\left(\exp\left\{\iota_1\left|z^{\epsilon}_{t,ij}\right|\right\} > \exp\{\iota_1 c_{24}\log(nd)\}\right)\right\}^{1/2}\\
&\leqslant O\left((nd)^{-\iota_1 c_{24}/2}\right) = o\left(n^{-1/2}\right)
\end{aligned}$$

uniformly over i, j and t, where the constant c_{24} is chosen so that $c_{24}\iota_1 > 2$. Therefore, we can prove that

$$\begin{aligned}
&\mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\widetilde{c}^{z,\epsilon}_{l,u}(s;i,j,1)\right| > \frac{c_{23}\sigma}{4}\cdot\log^2(nd)\right)\\
\leqslant\ &\mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s}\widetilde{z}^{\epsilon}_{t,ij}\right| > \frac{c_{23}\sigma}{5}\cdot\log^2(nd)\right)\\
\leqslant\ &\mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\max_{1\leqslant t\leqslant n}\left|z^{\epsilon}_{t,ij}\right| > c_{24}\log(nd)\right)\\
\leqslant\ &\sum_{i=1}^{d}\sum_{j=i}^{d}\sum_{t=1}^{n}\frac{\mathsf{E}[\exp\{\iota_1\left|z^{\epsilon}_{t,ij}\right|\}]}{\exp\{\iota_1 c_{24}\log(nd)\}}\\
=\ &O\left(d^{2-\iota_1 c_{24}}n^{1-\iota_1 c_{24}}\right) = o(1).
\end{aligned}\tag{E.12}$$

We next prove

$$\mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\max_{(l,u):\,1\leqslant l<u\leqslant n}\max_{s:\,l\leqslant s<u}\left|\bar{c}^{z,\epsilon}_{l,u}(s;i,j,1)\right| > \frac{c_{23}\sigma}{4}\cdot\log^2(nd)\right)\to 0.\tag{E.13}$$

As in the proof of (D.30), we consider the following two scenarios: (i) $s-l+1 \leqslant c_{25}\log^2(nd)$, and (ii) $s-l+1 > c_{25}\log^2(nd)$, where c_{25} is a sufficiently large positive constant. For scenario (i), it is easy to see that

$$\left|\bar{c}^{z,\epsilon}_{l,u}(s;i,j,1)\right| \leqslant \sqrt{\frac{u-s}{u-l+1}}\cdot\frac{1}{\sqrt{s-l+1}}\cdot\sum_{t=l}^{s}\left(\left|\bar{z}^{\epsilon}_{t,ij}\right| + \mathsf{E}\left[\left|\bar{z}^{\epsilon}_{t,ij}\right|\right]\right)$$

$$\leqslant \sqrt{s-l+1}\cdot(2c_{24}\log(nd)) \leqslant (2c_{24}\sqrt{c_{25}})\cdot\log^2(nd).$$

For scenario (ii), by Theorem 1.3(2) in Bosq (1998), we have

$$\begin{aligned}
& \mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\ \max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l\leqslant s<u}\left|\bar{c}^{z,\epsilon}_{l,u}(s;i,j,1)\right| > \frac{c_{23}\sigma}{4}\cdot\log^2(nd)\right)\\
\leqslant\ & \mathsf{P}\left(\max_{(i,j):\,1\leqslant i,j\leqslant d}\ \max_{(l,u):\,1\leqslant l<u\leqslant n}\ \max_{s:\,l+c_{25}\log^2(nd)-1\leqslant s<u}\left|\bar{c}^{z,\epsilon}_{l,u}(s;i,j,1)\right| > \left[\frac{c_{23}\sigma}{4}-2c_{24}\sqrt{c_{25}}\right]\cdot\log^2(nd)\right)\\
\leqslant\ & O\left(d^2n^3\exp\{-M\log(nd)\}+d^2n^{3+3/4}\rho^{\sqrt{c_{25}}\log(nd)}\right)=o(1),
\end{aligned}$$

where the constant c_{23} is chosen to be sufficiently large such that $\frac{c_{23}\sigma}{4}-2c_{24}\sqrt{c_{25}}$ is strictly larger than zero and $M>3$, and the constant c_{25} is chosen to be larger than $(-15/(4\log\rho))^2$. This proves (E.13).

With (E.12) and (E.13), we can show (E.11), completing the proof of the lemma. □

For notational simplicity, we let

$$c^{\widehat{\epsilon},\widehat{\sigma},\mathcal{I}}_{l_m,u_m}(s;i,j) = c^{\widehat{\epsilon},\widehat{\sigma}}_{l_m,u_m}(s;i,j)\cdot\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{\widehat{\epsilon},\widehat{\sigma}}_{l,u}(t;i,j)\right|>\xi^e_n\right)$$

and

$$c^{G,\epsilon,\widehat{\sigma},\mathcal{I}}_{l_m,u_m}(s;i,j) = c^{G,\epsilon,\widehat{\sigma}}_{l_m,u_m}(s;i,j)\cdot\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right|>\xi^e_n\right)$$

for $m\in\mathcal{M}^e_{l,u}$ such that $[l_m,u_m]$ is a random sub-interval of $[l,u]$, where $c^{\widehat{\epsilon},\widehat{\sigma}}_{l,u}(s;i,j)$ is defined in (2.12) and $c^{G,\epsilon,\widehat{\sigma}}_{l,u}(s;i,j)$ is defined in (E.7). Define $\mathbf{C}^{\widehat{\epsilon},\mathcal{I}}_{l_m,u_m}(s)$ and $\mathbf{C}^{G,\epsilon,\mathcal{I}}_{l_m,u_m}(s)$ as half-vectorisation of the two symmetric $d\times d$ matrices with the (i,j)-entry being $c^{\widehat{\epsilon},\widehat{\sigma},\mathcal{I}}_{l_m,u_m}(s;i,j)$ and $c^{G,\epsilon,\widehat{\sigma},\mathcal{I}}_{l_m,u_m}(s;i,j)$, respectively. By (2.13) in Section 2.4, we readily have that

$$C^{\widehat{\epsilon}}_{l_m,u_m}(s) = \left\|\mathbf{C}^{\widehat{\epsilon},\mathcal{I}}_{l_m,u_m}(s)\right\|_2^2.$$

Let

$$\mathcal{T}^e_{l,u} = \bigcup_{k:l+c_{22}\varphi^e_{n,d}\leqslant\eta^e_k\leqslant u-c_{22}\varphi^e_{n,d}}\mathcal{I}_k \tag{E.14}$$

be a set of index pairs which have breaks between $l+c_{22}\varphi^e_{n,d}$ and $u-c_{22}\varphi^e_{n,d}$, where $\mathcal{I}_k$ is defined in Assumption 4(iii). Define

$$\widetilde{\mathcal{T}}^e_{l,u} = \mathcal{I}\left\{(i,j):\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right|>\xi^e_n,\ 1\leqslant i,j\leqslant d\right\} \tag{E.15}$$

and

$$\widehat{\mathcal{T}}_{l,u}^{e} = \mathcal{I}\left\{(i,j) : \max_{t:l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right| > \xi_n^e,\ 1\leqslant i,j\leqslant d\right\}, \tag{E.16}$$

which can be regarded as the infeasible and feasible estimates of $\mathcal{T}_{l,u}^e$, respectively. Let

$$\overline{\omega}_{l,u}^e = \max_{k:l+c_{22}\varphi_{n,d}^e\leqslant\eta_k^e\leqslant u-c_{22}\varphi_{n,d}^e}\omega_k^e\ \text{ with }\ \omega_k^e = \sum_{(i,j)\in\mathcal{I}_k}\left|\sigma_{k+1|i,j}^e-\sigma_{k|i,j}^e\right|^2, \tag{E.17}$$

where $\sigma_{k|i,j}^e$ is defined in Assumption 4(iii).

The following lemma derives the asymptotic property of $\widetilde{\mathcal{T}}_{l,u}^e$ and $\widehat{\mathcal{T}}_{l,u}^e$ as well as a lower bound of the CUSUM statistic when there exists a change point which is an extension of Lemma D.4 to the WSBS-Cov method.

LEMMA E.3. *Suppose that the assumptions in Lemma D.3, Assumptions 4(iii) and 5 are satisfied, and let* l *and* u *satisfy the conditions (E.4) and (E.5). If the condition (3.10) in Theorem 3.2 is satisfied, we have*

$$\mathsf{P}\left(\mathcal{T}_{l,u}^e = \widetilde{\mathcal{T}}_{l,u}^e\right)\to 1,\ \ \mathsf{P}\left(\widetilde{\mathcal{T}}_{l,u}^e = \widehat{\mathcal{T}}_{l,u}^e\right)\to 1 \tag{E.18}$$

as $n,d\to\infty$. *There exists a positive integer* k *satisfying* $l+c_{22}\varphi_{n,d}^e\leqslant\eta_k^e\leqslant u-c_{22}\varphi_{n,d}^e$, *and*

$$(|\mathcal{T}_{l,u}^e|/K_2)\cdot\underline{\omega}_n^e\leqslant|\mathcal{I}_k|\cdot\underline{\omega}_n^e\leqslant\omega_k^e\leqslant\overline{\omega}_{l,u}^e\leqslant|\mathcal{T}_{l,u}^e|\cdot\overline{\omega}_n^e. \tag{E.19}$$

Furthermore,

$$\mathsf{P}\left(\left\|\mathbf{C}_{l_{m_0^e},u_{m_0^e}}^{\widehat{\epsilon},\widehat{\mathcal{I}}}(s_0^e)\right\|_2\geqslant c_{26}\left(|\mathcal{T}_{l,u}^e|\kappa_n^e\underline{\omega}_n^e\right)^{1/2}\right)\to 1 \tag{E.20}$$

as $n,d\to\infty$, *where* $|\cdot|$ *denotes the cardinality of a set and* c_{26} *is a positive constant.*

PROOF. We start with the proof of (E.18). The conditions (E.4) and (E.5) imply that l and u are sufficiently bounded away from the previously undetected break points. Note that from (E.7),

$$\begin{aligned}
\left|c_{l,u}^{G,\epsilon}(s;i,j)\right| &= \sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left|\frac{1}{s-l+1}\sum_{t=l}^{s}G_{t,ij}^{\epsilon}-\frac{1}{u-s}\sum_{t=s+1}^{u}G_{t,ij}^{\epsilon}\right| \\
&= \sqrt{\frac{u-l+1}{(s-l+1)(u-s)}}\left|\frac{u-s}{u-l+1}\sum_{t=l}^{s}G_{t,ij}^{\epsilon}-\frac{s-l+1}{u-l+1}\sum_{t=s+1}^{u}G_{t,ij}^{\epsilon}\right| \\
&= \sqrt{\frac{u-l+1}{(s-l+1)(u-s)}}\left|\frac{s-l+1}{u-l+1}\sum_{t=l}^{u}G_{t,ij}^{\epsilon}-\sum_{t=l}^{s}G_{t,ij}^{\epsilon}\right|.
\end{aligned} \tag{E.21}$$

Without loss of generality, we assume that $\sum_{t=l}^{u}G_{t,ij}^{\epsilon}=0$. For a given index pair (i,j), we consider the following three cases: (i) there is no change point within $[l,u)$; (ii) there are change points within $[l,u)$ but $(i,j) \notin \mathcal{T}^{e}_{l,u}$; (iii) $(i,j) \in \mathcal{T}^{e}_{l,u}$. For case (i), it is obvious that $\left|c^{G,\epsilon}_{l,u}(s;i,j)\right| = 0$ $\forall s \in [l,u)$. Case (ii) indicates that the change points may have been detected but are close to either l or u and there are at most two such change points. By Lemma 2.2 in Venkatraman (1992), $\left|c^{G,\epsilon}_{l,u}(s;i,j)\right|$ takes the maximum at one of the change points, which, together with the first equality in (E.21), leads to

$$\max_{s:l\leqslant s<u} \left|c^{G,\epsilon}_{l,u}(s;i,j)\right| \leqslant (c_{22}\varphi^{e}_{n,d})^{1/2} \cdot 2(\overline{\omega}^{e}_{n})^{1/2} \leqslant 2\sqrt{c_{22}}\log^2(nd). \tag{E.22}$$

Consider case (iii) and let k_0 and k be defined in (E.3) and (E.4). As

$$\left|G^{\epsilon}_{\eta^{e}_{k_0+k}+1,ij} - G^{\epsilon}_{\eta^{e}_{k_0+k},ij}\right| \geqslant (\underline{\omega}^{e}_{n})^{1/2},$$

we readily have $\left|G^{\epsilon}_{\eta^{e}_{k_0+k},ij}\right| \vee \left|G^{\epsilon}_{\eta^{e}_{k_0+k}+1,ij}\right| \geqslant (\underline{\omega}^{e}_{n})^{1/2}/2$, implying that

$$\left|\sum_{t=\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| \vee \left|\sum_{t=\eta^{e}_{k_0+k}+1}^{\eta^{e}_{k_0+k}+c_{21}\kappa^{e}_{n}} G^{\epsilon}_{t,ij}\right| \geqslant c_{21}\kappa^{e}_{n}(\underline{\omega}^{e}_{n})^{1/2}/2,$$

where c_{21} is defined in (E.4). Without loss of generality, we only consider that

$$\left|\sum_{t=\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| \geqslant c_{21}\kappa^{e}_{n}(\underline{\omega}^{e}_{n})^{1/2}/2. \tag{E.23}$$

By the triangle inequality, we have that

$$\begin{aligned}
\max_{s:l\leqslant s<u}\left|\sum_{t=l}^{s} G^{\epsilon}_{t,ij}\right| &\geqslant \left|\sum_{t=l}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| = \left|\sum_{t=l}^{\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}-1} G^{\epsilon}_{t,ij} + \sum_{t=\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| \\
&\geqslant \left|\sum_{t=\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| - \left|\sum_{t=l}^{\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}-1} G^{\epsilon}_{t,ij}\right| \\
&\geqslant \left|\sum_{t=\eta^{e}_{k_0+k}-c_{21}\kappa^{e}_{n}}^{\eta^{e}_{k_0+k}} G^{\epsilon}_{t,ij}\right| - \max_{s:l\leqslant s<u}\left|\sum_{t=l}^{s} G^{\epsilon}_{t,ij}\right|,
\end{aligned}$$

which, together with (E.23), leads to

$$\max_{s:l\leqslant s<u}\left|\sum_{t=l}^{s} G^{\epsilon}_{t,ij}\right| \geqslant c_{21}\kappa^{e}_{n}(\underline{\omega}^{e}_{n})^{1/2}/4. \tag{E.24}$$

Combining (E.21) and (E.24) and noting that

$$(s-l+1)(u-s)/(u-l+1) \leqslant (u-l+1)/4 \leqslant n/4$$

as $(s-l+1)(u-s)$ achieves the maximum when $s-l+1=u-s$, we readily have that

$$\max_{s:l\leqslant s<u} \left|c_{l,u}^{G,\epsilon}(s;i,j)\right| \geqslant c_{21}\kappa_n^e(\underline{\omega}_n^e/n)^{1/2}/2. \tag{E.25}$$

Combining the above three cases and using (3.10), we can prove $\mathsf{P}\left(\mathcal{T}_{l,u}^e=\widetilde{\mathcal{T}}_{l,u}^e\right)\to 1$.

By Proposition 3.3 and Lemma E.2, we readily have that, uniformly over $1\leqslant i\leqslant j\leqslant d$,

$$\begin{aligned}
&\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t;i,j)\right| > \xi_n^e + c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right)\\
\leqslant\;& \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right| > \xi_n^e\right)\\
\leqslant\;& \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t;i,j)\right| > \xi_n^e - c_{27}\sqrt{(\log d)(\log n)} - c_{23}\log^2(nd)\right)
\end{aligned} \tag{E.26}$$

with probability approaching one, where $c_{27}>0$ is a constant. Furthermore, following the proof of $\mathsf{P}\left(\mathcal{T}_{l,u}^e=\widetilde{\mathcal{T}}_{l,u}^e\right)\to 1$ and using (3.10) again, we may show that

$$\begin{aligned}
&\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t;i,j)\right| > \xi_n^e + c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right)\\
=\;& \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t;i,j)\right| > \xi_n^e - c_{27}\sqrt{(\log d)(\log n)} - c_{23}\log^2(nd)\right),
\end{aligned}$$

which, together with (E.26), indicates that, uniformly over $1\leqslant i\leqslant j\leqslant d$,

$$\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right| > \xi_n^e\right) = \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t;i,j)\right| > \xi_n^e\right), \tag{E.27}$$

with probability approaching one, i.e., $\mathsf{P}\left(\widetilde{\mathcal{T}}_{l,u}^e=\widehat{\mathcal{T}}_{l,u}^e\right)\to 1$. We have completed the proof of the two equalities in (E.18).

The proof of

$$|\mathcal{J}_k|\cdot\underline{\omega}_n^e \leqslant \omega_k^e \leqslant \overline{\omega}_{l,u}^e \leqslant |\mathcal{T}_{l,u}^e|\cdot\overline{\omega}_n^e$$

is straightforward. Then we can prove the inequalities in (E.19) by noting that $|\mathcal{T}_{l,u}^e|\leqslant K_2\cdot|\mathcal{J}_k|$ for at least one k satisfying $l+c_{22}\varphi_{n,d}^e\leqslant\eta_k^e\leqslant u-c_{22}\varphi_{n,d}^e$.

Finally, we turn to the proof of (E.20). As in the proof of Lemma D.4, on the set $\mathcal{D}_n^e$, there exists

$1 \leqslant m_k \leqslant M_n^e$ such that $l_{m_k} \in \mathcal{J}_k^e$ and $u_{m_k} \in \mathcal{J}_{k+1}^e$, indicating that both $\eta_k^e - l_{m_k}$ and $u_{m_k} - \eta_k^e$ are larger than $\kappa_n^e/3$. For $1 \leqslant i \leqslant j \leqslant d$ and k such that $l + c_{22}\varphi_{n,d}^e < \eta_k^e < u - c_{22}\varphi_{n,d}^e$, we have

$$\left|c^{G,\epsilon}_{l_{m_k},u_{m_k}}(\eta_k^e; i,j)\right| = \sqrt{\frac{(\eta_k^e - l_{m_k}+1)(u_{m_k}-\eta_k^e)}{u_{m_k}-l_{m_k}+1}}\left|\varpi_{k,ij}^e\right| \geqslant c_{28}\left(\kappa_n^e\right)^{1/2}\left|\varpi_{k,ij}^e\right|, \tag{E.28}$$

where $\varpi_{k,ij}^e = \sigma_{k+1|i,j}^e - \sigma_{k|i,j}^e$ and c_{28} is a positive constant. By (E.28) and Assumption 5, we have

$$\left|c^{G,\epsilon,\widehat{\sigma}}_{l_{m_k},u_{m_k}}(\eta_k^e; i,j))\right| \geqslant c_{28}\left(\kappa_n^e\right)^{1/2}\left|\varpi_{k,ij}^e\right|/\overline{\sigma}. \tag{E.29}$$

Following the proof of Proposition 3.3, and using Lemma E.2 and $\mathsf{P}\left(\widetilde{\mathcal{T}}_{l,u}^e = \widehat{\mathcal{T}}_{l,u}^e\right) \to 1$ from (E.18), we have, for k such that $l + c_{22}\varphi_{n,d}^e < \eta_k^e < u - c_{22}\varphi_{n,d}^e$,

$$\begin{aligned}
&\left|c^{\widehat{\epsilon},\widehat{\sigma}}_{l_{m_k},u_{m_k}}(\eta_k^e; i,j)\right| \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{\widehat{\epsilon},\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi_n^e\right) \\
\geqslant\ &\left(\left|c^{G,\epsilon,\widehat{\sigma}}_{l_{m_k},u_{m_k}}(\eta_k^e; i,j)\right| - c_{27}\sqrt{(\log d)(\log n)} - c_{23}\log^2(nd)\right)\cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi_n^e\right)
\end{aligned}$$

with probability approaching one uniformly over $1 \leqslant i \leqslant j \leqslant d$. This, together with (E.29), implies that

$$\begin{aligned}
\left\|\mathbf{C}^{\widehat{\epsilon},\widehat{\mathcal{J}}}_{l_{m_k},u_{m_k}}(\eta_k^e)\right\|_2 &\geqslant \left\|\mathbf{C}^{G,\epsilon,\mathcal{J}}_{l_{m_k},u_{m_k}}(\eta_k^e)\right\|_2 - \left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}_{l,u}^e|^{1/2} \\
&\geqslant (c_{28}/\overline{\sigma})\cdot(\kappa_n^e\omega_k^e)^{1/2} - \left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}_{l,u}^e|^{1/2}
\end{aligned} \tag{E.30}$$

with probability approaching one. Then, by the definitions of m_0^e and s_0^e, (E.19) and (E.30), and noting that $\kappa_n^e\underline{\omega}_n^e/\log^4(nd) \to \infty$ in Assumption 4(iii), we have

$$\begin{aligned}
\left\|\mathbf{C}^{\widehat{\epsilon},\widehat{\mathcal{J}}}_{l_{m_0^e},u_{m_0^e}}(s_0^e)\right\|_2 &\geqslant \max_{k:l+c_{22}\varphi_{n,d}^e<\eta_k^e<u-c_{22}\varphi_{n,d}^e}\left\|\mathbf{C}^{\widehat{\epsilon},\widehat{\mathcal{J}}}_{l_{m_k},u_{m_k}}(\eta_k^e)\right\|_2 \\
&\geqslant [c_{28}/(2\overline{\sigma})]\cdot\left(|\mathcal{T}_{l,u}^e|\kappa_n^e\underline{\omega}_n^e/K_2\right)^{1/2}
\end{aligned} \tag{E.31}$$

with probability approaching one. Choosing $c_{26} = c_{28}/(2K_2^{1/2}\overline{\sigma})$, we can complete the proof of (E.20). The proof of Lemma E.3 has been completed. □

LEMMA E.4. *The function* $\left\|\mathbf{C}^{G,\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s)\right\|_2$ *(as a function of s) is either monotonic or first decreasing and* *then increasing on the interval* $\left[\eta^e_{\tilde{k}}, \eta^e_{\tilde{k}+1}\right]$ *if* $s^e_0 \in \left[\eta^e_{\tilde{k}}, \eta^e_{\tilde{k}+1}\right] \subseteq [l_{m^e_0}, u_{m^e_0})$. *Furthermore,*

$$\left\{\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{\tilde{k}})\right\|_2 \vee \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{\tilde{k}+1})\right\|_2\right\} \geqslant \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s^e_0)\right\|_2. \tag{E.32}$$

PROOF. As the involvement of the indicator function (which does not depend on s) does not change the quasi-convexity of the function, the result directly follows from Lemma D.5. □

We next provide an extension of Lemma 2.6 in Venkatraman (1992) and Lemma D.6 in Appendix D to the WSBS-Cov method. Note that some notation used in Lemma E.5 below and its proof is similar to that in Lemma D.6.

LEMMA E.5. *Suppose that the assumptions of Lemma E.3 and (E.3)–(E.5) are satisfied. Let* $s^e_\star \in [l_{m^e_0}, u_{m^e_0}]$ *be the point of maximising* $\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0},u_{m_0}}(s)\right\|_2$ *with respect to* s*, i.e.,*

$$s^e_\star = \arg\max_{l_{m^e_0} \leqslant s < u_{m^e_0}} \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s)\right\|_2 \tag{E.33}$$

and define $\eta^e_{k_\diamond}$ *as a change point that satisfies*

$$\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\diamond})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s^e_\star)\right\|_2 - 3c_{23}|\mathcal{T}^e_{l,u}|^{1/2}\log^2(nd), \tag{E.34}$$

where c_{23} *is defined in Lemma E.2. Then, there exists a positive constant* c_{29} *such that*

$$(\eta^e_{k_\diamond} - l_{m^e_0} + 1) \wedge (u_{m^e_0} - \eta^e_{k_\diamond}) \geqslant c_{29}\kappa^e_n \tag{E.35}$$

when n *is sufficiently large, and furthermore,*

$$\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\diamond})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\diamond}+\nu_l)\right\|_2 + c_{30}\nu_l\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\diamond})\right\|_2 \kappa^e_n/(u_{m^e_0} - l_{m^e_0}+1)^2, \tag{E.36}$$

where $0 < \nu_l < c_{31}\gamma^e_n$ *and* $\gamma^e_n = (\kappa^e_n/\underline{\omega}^e_n)^{1/2}\log^2(nd)$, c_{30} *and* c_{31} *are two positive constants.*

PROOF. The proof is similar to the proof of Lemma D.6 in Appendix D. From the definition of $s^e_\star$ in (E.33) and using Lemma E.4, there exists a positive integer $k_\star$ (whose value is often different from $k_\star$ used in the proof of Lemma D.6) such that $s^e_\star = \eta^e_{k_\star}$. First we prove that

$$(\eta^e_{k_\star} - l + 1) \wedge (u - \eta^e_{k_\star}) \geqslant \kappa^e_n - c_{22}\varphi^e_{n,d}, \tag{E.37}$$

where c_{22} is the same as that in (E.5). By (E.4) and (E.5), we have $(\eta^e_{k_\star} - l + 1) \wedge (u - \eta^e_{k_\star})$ is either smaller than $c_{22}\varphi^e_{n,d}$ or larger than $\kappa^e_n - c_{22}\varphi^e_{n,d}$. Let

$$\mathbf{G}^{\epsilon,\mathcal{I}}_s = \left[G^\epsilon_{s,11}\cdot\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;1,1)\right| > \xi^e_n\right),\cdots,G^\epsilon_{s,dd}\cdot\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;d,d)\right| > \xi^e_n\right)\right]^\intercal,$$

a $d(d+1)/2$ column vector which denotes half-vectorisation of a $d\times d$ symmetric matrix with the (i,j)-entry being $G^\epsilon_{s,ij}\cdot\mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^e_n\right)$. Note that

$$\begin{aligned}\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{I}}_{l,u}(s)\right\|_2 &\leqslant \frac{1}{\underline{\sigma}}\cdot\sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left\|\frac{1}{s-l+1}\sum_{t=l}^{s}\mathbf{G}^{\epsilon,\mathcal{I}}_t - \frac{1}{u-s}\sum_{t=s+1}^{u}\mathbf{G}^{\epsilon,\mathcal{I}}_t\right\|_2\\ &\leqslant 2b^{\epsilon,\mathcal{I}}_{l,u}\sqrt{(s-l+1)\wedge(u-s)}/\underline{\sigma},\end{aligned}\tag{E.38}$$

where

$$b^{\epsilon,\mathcal{I}}_{l,u} = \sup_{l\leqslant s\leqslant u}\left\|\mathbf{G}^{\epsilon,\mathcal{I}}_s - \frac{1}{u-l+1}\sum_{t=l}^{u}\mathbf{G}^{\epsilon,\mathcal{I}}_t\right\|_2.$$

If $(\eta^e_{k_\star}-l+1)\wedge(u-\eta^e_{k_\star})\leqslant c_{22}\varphi^e_{n,d}$, we must have $(\eta^e_{k_\star}-l_{m^e_0}+1)\wedge(u_{m^e_0}-\eta^e_{k_\star})\leqslant c_{22}\varphi^e_{n,d}$, implying that

$$c_{28}\left(|\mathcal{I}^e_{l,u}|\kappa^e_n\underline{\omega}^e_n\right)^{1/2}/\overline{\sigma}\leqslant\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{I}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\star})\right\|_2\leqslant 2b^{\epsilon,\mathcal{I}}_{l_{m^e_0},u_{m^e_0}}\left(c_{22}\varphi^e_{n,d}\right)^{1/2}/\underline{\sigma},\tag{E.39}$$

where the first inequality is proved by (E.19) and (E.29), and the second inequality is obtained using (E.38). Noting that

$$b^{\epsilon,\mathcal{I}}_{l_{m^e_0},u_{m^e_0}}\leqslant K_2(|\mathcal{I}^e_{l,u}|\overline{\omega}^e_n)^{1/2},\tag{E.40}$$

the inequalities in (E.39) would lead to a contradiction with the condition $\kappa^e_n\underline{\omega}^e_n/\log^4(nd)\to\infty$ in Assumption 4(iii). Hence (E.37) has been proved, which indicates that there exists $m^e_\star\in\mathcal{M}^e_{l,u}$ such that $l_{m^e_\star}\in\mathcal{I}^e_{k_\star}$ and $u_{m^e_\star}\in\mathcal{I}^e_{k_\star+1}$.

We next prove that for n large enough,

$$(\eta^e_{k_\star}-l_{m^e_0}+1)\wedge(u_{m^e_0}-\eta^e_{k_\star})\geqslant c_{29}\kappa^e_n.\tag{E.41}$$

Suppose that (E.41) fails, i.e., for any $c_\star$ and N, we have some $n>N$ such that

$$(\eta^e_{k_\star}-l_{m^e_0}+1)\wedge(u_{m^e_0}-\eta^e_{k_\star})<c_\star\kappa^e_n.\tag{E.42}$$

As in the proof of (D.48), without loss of generality, we let $\eta^e_{k_\star}-l_{m^e_0}+1<c_\star\kappa^e_n$, and consider the following two cases of $u_{m^e_0}$:

(i) $\eta^e_{k_\star}\leqslant u_{m^e_0}<\eta^e_{k_0+k_1}$, or $\eta^e_{k_0+k_1+1}-c_{22}\varphi^e_{n,d}\leqslant u\leqslant\eta^e_{k_0+k_1+1}$ and $\eta^e_{k_\star}<\eta^e_{k_0+k_1}<u_{m^e_0}\leqslant u$;

(ii) $\eta^e_{k_0+k_1} \leqslant u \leqslant \eta^e_{k_0+k_1} + c_{22}\varphi^e_{n,d}$ and $\eta^e_{k_\star} < \eta^e_{k_0+k_1} < u_{m^e_0} \leqslant u$.

The difference between cases (i) and (ii) is that in case (ii) we cannot find $m \in \mathcal{M}^e_{l,u}$ such that $l_m \in \mathcal{J}^e_{k_0+k_1}$ and $u_m \in \mathcal{J}^e_{k_0+k_1+1}$. Consider case (i) first. By (E.39) and (E.40), we readily have that

$$c_{28}\left(|\mathcal{T}^e_{l,u}|\kappa^e_n\underline{\omega}^e_n\right)^{1/2}/\overline{\sigma} \leqslant \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\star})\right\|_2 \leqslant 2K_2\left(c_\star|\mathcal{T}^e_{l,u}|\kappa^e_n\overline{\omega}^e_n\right)^{1/2}/\underline{\sigma}, \tag{E.43}$$

which would result in a contradiction if we choose a sufficiently small $c_\star > 0$. We next consider case (ii). Since $\eta^e_{k_0+k_1} \leqslant u_{m^e_0} \leqslant u \leqslant \eta^e_{k_0+k_1} + c_{22}\varphi^e_{n,d}$ in this case, we may show that

$$\begin{aligned}
& c_{28}\left(|\mathcal{T}^e_{l,u}|\kappa^e_n\underline{\omega}^e_n\right)^{1/2}/\overline{\sigma} \leqslant \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\star})\right\|_2 \\
\leqslant\ & \frac{1}{\underline{\sigma}}\cdot\sqrt{\frac{(\eta^e_{k_\star}-l_{m^e_0}+1)(u_{m^e_0}-\eta^e_{k_\star})}{u_{m^e_0}-l_{m^e_0}+1}}\left\|\frac{1}{\eta^e_{k_\star}-l_{m^e_0}+1}\sum_{t=l_{m^e_0}}^{\eta^e_{k_\star}}\mathbf{G}^{\boldsymbol{\epsilon},\mathcal{J}}_t - \frac{1}{u_{m^e_0}-\eta^e_{k_\star}}\sum_{t=\eta^e_{k_\star}+1}^{u_{m^e_0}}\mathbf{G}^{\boldsymbol{\epsilon},\mathcal{J}}_t\right\|_2 \\
\leqslant\ & \frac{1}{\underline{\sigma}}\cdot\sqrt{\frac{(\eta^e_{k_\star}-l_{m^e_0}+1)(u_{m^e_0}-\eta^e_{k_\star})}{u_{m^e_0}-l_{m^e_0}+1}}\cdot\left\|\frac{1}{\eta^e_{k_\star}-l_{m^e_0}+1}\sum_{t=l_{m^e_0}}^{\eta^e_{k_\star}}\mathbf{G}^{\boldsymbol{\epsilon},\mathcal{J}}_t - \frac{1}{u_{m^e_0}-\eta^e_{k_\star}}\sum_{t=\eta^e_{k_\star}+1}^{u_{m^e_0}}\mathbf{G}^{\boldsymbol{\epsilon},\mathcal{J}}_{t\wedge(u_{m^e_0}-c_{22}\varphi^e_{n,d})}\right\|_2 \\
& +\frac{1}{\underline{\sigma}}\cdot\sqrt{\frac{(\eta^e_{k_\star}-l_{m^e_0}+1)(u_{m^e_0}-\eta^e_{k_\star})}{u_{m^e_0}-l_{m^e_0}+1}}\cdot\frac{c_{22}\varphi^e_{n,d}b^{\boldsymbol{\epsilon},\mathcal{J}}_{u_{m^e_0}-c_{22}\varphi^e_{n,d},u_{m^e_0}}}{u_{m^e_0}-\eta^e_{k_\star}} \\
\leqslant\ & \left[\left(2b^{\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}-c_{22}\varphi^e_{n,d}} + c_{22}\varphi^e_{n,d}b^{\boldsymbol{\epsilon},\mathcal{J}}_{u_{m^e_0}-c_{22}\varphi^e_{n,d},u_{m^e_0}}/\kappa^e_n\right)/\underline{\sigma}\right]\cdot\sqrt{(\eta^e_{k_\star}-l_{m^e_0}+1)\wedge(u_{m^e_0}-\eta^e_{k_\star})}.
\end{aligned}$$

As $\frac{\kappa^e_n\underline{\omega}^e_n}{\log^4(nd)} \to \infty$ in Assumption 4(iii), we have

$$\frac{\varphi^e_{n,d}b^{\boldsymbol{\epsilon},\mathcal{J}}_{u_{m^e_0}-c_{22}\varphi^e_{n,d},u_{m^e_0}}}{\kappa^e_n} = \frac{1}{\kappa^e_n}O\left(\left(|\mathcal{T}^e_{l,u}|\cdot\overline{\omega}^e_n\right)^{1/2}\cdot\log^4(nd)/\underline{\omega}^e_n\right) = o\left(\left(|\mathcal{T}^e_{l,u}|\cdot\overline{\omega}^e_n\right)^{1/2}\right),$$

which, together with the fact that $2b^{\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}-c_{22}\varphi^e_{n,d}} \geqslant \overline{\omega}_{l,u} \geqslant \left(|\mathcal{T}^e_{l,u}|\underline{\omega}^e_n/K_2\right)^{1/2}$, leads to

$$\varphi^e_{n,d}b^{\boldsymbol{\epsilon},\mathcal{J}}_{u_{m^e_0}-c_{22}\varphi^e_{n,d},u_{m^e_0}}/\kappa^e_n = o\left(b^{\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}-c_{22}\varphi^e_{n,d}}\right).$$

Hence, we have

$$c_{28}\left(|\mathcal{T}^e_{l,u}|\kappa^e_n\underline{\omega}^e_n\right)^{1/2}/\overline{\sigma} \leqslant \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\star})\right\|_2 \leqslant 2K_2\left(c_\star|\mathcal{T}^e_{l,u}|\kappa^e_n\overline{\omega}^e_n\right)^{1/2}/\underline{\sigma}$$

which would lead a contradiction as $\overline{\omega}^e_n \asymp \underline{\omega}^e_n$ in Assumption 4(iii) when $c_\star > 0$ is chosen to be sufficiently small. Combining the above arguments, neither case (i) nor case (ii) holds, completing the proof of (E.41). Following the similar argument and using (E.34), we may prove (E.35).

We finally give the proof of (E.36). Consider two cases: (i) $u_{m_0^e} \leqslant \eta_{k_\diamond+1}^e$ and (ii) $\eta_{k_\diamond+1}^e < u_{m_0^e}$. For case (i), we define $\nu_i = \eta_{k_\diamond}^e - l_{m_0^e} + 1$ and $\nu_h = u_{m_0^e} - \eta_{k_\diamond}^e$. Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_{d(d+1)/2})^\intercal$ with

$$\beta_k = c_{l_{m_0^e}, u_{m_0^e}}^{G,\epsilon,\widehat{\sigma}}(\eta_{k_\diamond}^e; i, j) \left(\frac{\nu_i \nu_h}{\nu_i + \nu_h}\right)^{1/2} \cdot \mathcal{I}\left(\max_{t: l \leqslant t < u} \left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t; i, j)\right| > \xi_n^e\right)$$

and $k := k(i, j) = (i-1)d + j - (i-1)j/2$. Then we readily have that

$$c_{l_m, u_m}^{G,\epsilon,\widehat{\sigma},\mathcal{I}}(s; i, j) = c_{l_{m_0^e}, u_{m_0^e}}^{G,\epsilon,\widehat{\sigma}}(\eta_{k_\diamond}^e; i, j) \cdot \mathcal{I}\left(\max_{t: l \leqslant t < u} \left|c_{l,u}^{G,\epsilon,\widehat{\sigma}}(t; i, j)\right| > \xi_n^e\right) = \beta_k \left(\frac{\nu_i + \nu_h}{\nu_i \nu_h}\right)^{1/2}$$

and similarly

$$c_{l_{m_0^e}, u_{m_0^e}}^{G,\epsilon,\widehat{\sigma},\mathcal{I}}(\eta_{k_\diamond}^e + \nu_l; i, j) = \beta_k \left(\frac{\nu_h - \nu_l}{\nu_h}\right) \cdot \left[\frac{\nu_i + \nu_h}{(\nu_i + \nu_l)(\nu_h - \nu_l)}\right]^{1/2},$$

where the subscript $k = (i-1)d + j - (i-1)j/2$. Following the same arguments as in the proof of (D.52), we can show that

$$\begin{aligned} \left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e)\right\|_2 - \left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e + \nu_l)\right\|_2 &= \|\boldsymbol{\beta}\|_2 \cdot \frac{\sqrt{\nu_i + \nu_h}}{\nu_h}\left(\sqrt{\frac{\nu_h}{\nu_i}} - \sqrt{\frac{\nu_h - \nu_l}{\nu_i + \nu_l}}\right) \\ &\geqslant \frac{\nu_l \cdot \left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e)\right\|_2}{2(u_{m_0^e} - l_{m_0^e} + 1)}. \end{aligned} \tag{E.44}$$

For case (ii), we let $\nu_i = \eta_{k_\diamond}^e - l_{m_0^e} + 1$, $\nu_h = (c_{29} \wedge 1)\kappa_n^e/3$, $\nu_j = u_{m_0^e} - \eta_{k_\diamond}^e - \nu_h$, and

$$\mathbf{V}_{\mathbf{G}}^e = \mathbf{G}_{\eta_{k_\diamond}^e+1}^{\boldsymbol{\epsilon},\mathcal{I}} - \frac{1}{u - l + 1}\sum_{t=l}^{u} \mathbf{G}_t^{\boldsymbol{\epsilon},\mathcal{I}}.$$

Then, for $0 \leqslant \nu_l \leqslant \nu_h$, we readily have that

$$\left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e + \nu_l)\right\|_2 = \|\boldsymbol{\beta} + \nu_l \mathbf{V}_{\mathbf{G}}^e\|_2 \left[\frac{\nu_i + \nu_j + \nu_h}{(\nu_i + \nu_l)(\nu_j + \nu_h - \nu_l)}\right]^{1/2},$$

where $\boldsymbol{\beta}$ is defined as in case (i). Define

$$E(\nu_l) = \left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e)\right\|_2 - \left\|\mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{I}}(\eta_{k_\diamond}^e + \nu_l)\right\|_2$$

and

$$E_1 = \left[\left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta^e_{k_\diamond})\right\|_2 - \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta^e_{k_\diamond}+\nu_h)\right\|_2\right] \cdot \frac{\nu_l}{\nu_h}\left[\frac{(\nu_i+\nu_h)\nu_j}{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\right]^{1/2}.$$

Following the same argument in the proof of (D.53), we have

$$E(\nu_l) - E_1 \geqslant E_2 \times (1+E_3), \tag{E.45}$$

where

$$E_2 = \frac{\|\boldsymbol{\beta}\|_2 \nu_l(\nu_h-\nu_l)\sqrt{\nu_i+\nu_j+\nu_h}}{\sqrt{\nu_i(\nu_j+\nu_h)}\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\left(\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{\nu_i(\nu_j+\nu_h)}\right)},$$

and

$$E_3 = \frac{(\nu_j-\nu_i)(\nu_j-\nu_i-\nu_l)}{\left(\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{(\nu_i+\nu_h)\nu_j}\right)\left(\sqrt{\nu_i(\nu_j+\nu_h)}+\sqrt{(\nu_i+\nu_h)\nu_j}\right)}.$$

Noting that ν_l is smaller than $\nu_h/2$ and ν_i for large n, we have

$$\begin{aligned}
E_2 &= \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta^e_{k_\diamond})\right\|_2 \frac{\nu_l(\nu_h-\nu_l)}{\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}\left[\sqrt{(\nu_i+\nu_l)(\nu_j+\nu_h-\nu_l)}+\sqrt{\nu_i(\nu_j+\nu_h)}\right]} \\
&\geqslant \left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta^e_{k_\diamond})\right\|_2 \frac{(\nu_l\nu_h/2)}{\sqrt{2\nu_i(\nu_j+\nu_h)}\left[\sqrt{2\nu_i(\nu_j+\nu_h)}+\sqrt{\nu_i(\nu_j+\nu_h)}\right]} \\
&\geqslant (2c_{30}\nu_l\kappa^e_n)\cdot\left\|\mathbf{C}^{\mathbf{G},\boldsymbol{\epsilon},\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta^e_{k_\diamond})\right\|_2 / (u_{m_0^e}-l_{m_0^e}+1)^2.
\end{aligned} \tag{E.46}$$

Meanwhile, as $(\nu_j-\nu_i)(\nu_j-\nu_i-\nu_l)$ reaches its minimum at $\nu_j-\nu_i=\nu_l/2$, $\nu_i,\nu_j,\nu_h \geqslant (c_{29}\wedge 1)\kappa^e_n/3$ by (E.35) and $\nu_l = o(\kappa^e_n)$, following the proof of (D.55), we have

$$E_3 \geqslant \frac{-\nu_l^2}{4(1+\sqrt{2})(\sqrt{2}+\sqrt{2})[(c_{29}\wedge 1)\kappa^e_n/3]^2} \to 0. \tag{E.47}$$

Following the same arguments as in the proof of Lemma D.6, E_1 is dominated by E_2 when n is sufficiently large, which, together with (E.45)–(E.47), indicates that the lower bound of $E(\nu_l)$ is dominated by E_2 when n is large enough. Combining the arguments for cases (i) and (ii), we may complete the proof of (E.36). □

The following lemma can be seen as an extension of Lemma D.7 from WBS-Cov to WSBS-Cov.

LEMMA E.6. *Suppose that (3.10), Assumptions 1–3, 4(i)(iii) and 5 in Appendix A, and (E.3)–(E.5) are satisfied. There exists* $k_0+1 \leqslant k_\circ \leqslant k_0+k_1$ *such that*

$$\left|s^e_0 - \eta^e_{k_\circ}\right| \leqslant c_{31}\gamma^e_{n,d} \tag{E.48}$$

with probability approaching one, as $n \to \infty$, where $\gamma^e_{n,d} = (\kappa^e_n/\underline{\omega}^e_n)^{1/2}\log^2(nd)$ and c_{31} is a positive constant defined as in Lemma E.5.

PROOF. The proof is similar to the proof of Lemma D.7 in Appendix D. Without loss of generality, assume that $s^e_0 \in \left[\eta^e_{\tilde{k}}, \eta^e_{\tilde{k}+1}\right)$ for $k_0 \leqslant \tilde{k} \leqslant k_0 + k_1$. We next show the consequence if (E.48) fails and consider two cases.

Case (i): only one of $\eta^e_{\tilde{k}}$ and $\eta^e_{\tilde{k}+1}$ locates in the interval $[l_{m^e_0}, u_{m^e_0})$. Without loss of generality, consider that $\eta^e_{\tilde{k}}$ belongs to the interval $[l_{m^e_0}, u_{m^e_0})$ and choose $\eta^e_{k_\circ} = \eta^e_{\tilde{k}}$. By the definitions of m^e_0 and s^e_0, (E.18) and following the proof of Proposition 3.3, we readily have that

$$
\begin{aligned}
\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\circ})\right\|_2 &\leqslant \left\|\mathbf{C}^{\widehat{\epsilon},\widehat{\mathcal{J}}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\circ})\right\|_2 + \left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}^e_{l,u}|^{1/2} \\
&\leqslant \left\|\mathbf{C}^{\widehat{\epsilon},\widehat{\mathcal{J}}}_{l_{m^e_0},u_{m^e_0}}(s^e_0)\right\|_2 + \left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}^e_{l,u}|^{1/2} \\
&\leqslant \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s^e_0)\right\|_2 + 2\left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}^e_{l,u}|^{1/2}
\end{aligned} \tag{E.49}
$$

with probability approaching one, where c_{23} is defined in Lemma E.2 and c_{27} is defined as in (E.26). On the other hand, by Lemma E.4, without loss of generality, we only consider that $\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s)\right\|_2$ (treated as a function of s) locally decreases at $[\eta^e_{k_\circ}, u_{m^e_0})$ which includes the point of $s = s^e_0$. When (E.48) fails, we have

$$
\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s^e_0)\right\|_2 < \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s)\right\|_2 < \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\circ})\right\|_2. \tag{E.50}
$$

for any $s \in (\eta^e_{k_\circ}, \eta^e_{k_\circ} + c_{31}\gamma^e_n]$. By (E.34) and (E.36) in Lemma E.5 and (E.39), following the same arguments as in case (i) in the proof of Lemma D.7, we have

$$
\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(\eta^e_{k_\circ})\right\|_2 > \left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s^e_0)\right\|_2 + 2\left[c_{27}\sqrt{(\log d)(\log n)} + c_{23}\log^2(nd)\right]|\mathcal{T}^e_{l,u}|^{1/2} \tag{E.51}
$$

by choosing c_{31} in Lemma E.5 to be sufficiently large. This leads to a contradiction with (E.49).

Case (ii): both $\eta^e_{\tilde{k}}$, and $\eta^e_{\tilde{k}+1}$ are in the interval $[l_{m^e_0}, u_{m^e_0}]$. As in the proof of Lemma D.7, we consider two scenarios: (ii.1) $\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s)\right\|_2$ locally decreases at the point $s = s^e_0$; and (ii.2) $\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m^e_0},u_{m^e_0}}(s)\right\|_2$ locally increases at the point $s = s^e_0$. For scenario (ii.1), we choose $\eta^e_{k_\circ} = \eta^e_{\tilde{k}}$, and for scenario (ii.2), we choose $\eta^e_{k_\circ} = \eta^e_{\tilde{k}+1}$. In either of the two scenarios, we can similarly prove (E.51) when (E.48) fails. This would lead to a contradiction with (E.49). The proof of the lemma has been completed. □

We next introduce some additional notation to be used in the subsequent proof. Let $Z^{\epsilon}_{t,ij} = \epsilon_{ti}\epsilon_{tj}$ and recall that

$$Z^{\epsilon}_{t,ij} = \mathsf{E}\left[\epsilon_{ti}\epsilon_{tj}\right] + \left(\epsilon_{ti}\epsilon_{tj} - \mathsf{E}\left[\epsilon_{ti}\epsilon_{tj}\right]\right) =: G^{\epsilon}_{t,ij} + z^{\epsilon}_{t,ij}.$$

For (i,j) satisfying $1 \leqslant i \leqslant j \leqslant d$, consider a one-to-one map: $k(i,j) = d(i-1) + j - j(i-1)/2$ and let $k := k(i,j)$ for notational simplicity. Define

$$\begin{aligned}
\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k} &= \left(Z^{\epsilon}_{l_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right), \cdots, Z^{\epsilon}_{u_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right)\right)^{\intercal}, \\
\mathbf{G}^{\epsilon,\mathcal{J}}_{\bullet,k} &= \left(G^{\epsilon}_{l_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right), \cdots, G^{\epsilon}_{u_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right)\right)^{\intercal}, \\
\mathbf{z}^{\epsilon,\mathcal{J}}_{\bullet,k} &= \left(z^{\epsilon}_{l_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right), \cdots, z^{\epsilon}_{u_{m_0^e},ij} \cdot \mathcal{I}\left(\max_{t:l\leqslant t<u}\left|c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j)\right| > \xi^{e}_{n}\right)\right)^{\intercal}.
\end{aligned}$$

The following lemma further improves the break point estimation rate obtained in Lemma E.6.

LEMMA E.7. *Suppose that the conditions of Lemma E.6 are satisfied. With probability approaching one, we have*

$$\left|s^{e}_{0} - \eta^{e}_{k_\circ}\right| \leqslant c_{32}\varphi^{e}_{n,d} \tag{E.52}$$

as $n \to \infty$, *where* c_{32} *is a positive constant and* $\varphi^{e}_{n,d}$ *is defined in Theorem 3.2.*

PROOF. For $1 \leqslant i \leqslant j \leqslant d$, we let $k := k(i,j) = d(i-1) + j - j(i-1)/2$ throughout the proof. Let $C^{\epsilon,\mathcal{J}}_{l,u}(s;k)$ be the k-th element of $\mathbf{C}^{\epsilon,\mathcal{J}}_{l,u}(s)$ and write $C^{\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s;k) = \langle \mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \boldsymbol{\psi}^{s}_{l_{m_0^e},u_{m_0^e}}\rangle / \widehat{\sigma}_{l,u}(i,j)$ using the notion of inner product, where $\boldsymbol{\psi}^{s}_{l,u}$ is defined as in the proof of Lemma D.8. For $l_{m_0^e} \leqslant s < u_{m_0^e}$, define $Q^{\epsilon,\mathcal{J}}_{k}(s;1) = \left|\langle \mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \boldsymbol{\psi}^{s}_{l_{m_0^e},u_{m_0^e}}\rangle\right|^2$, and let $\bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}$ and $\bar{\mathbf{G}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}$ be defined similarly to $\bar{\mathbf{v}}^{s}$ in the proof of Lemma D.8 with $\mathbf{v}$ replaced by $\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k}$ and $\mathbf{G}^{\epsilon,\mathcal{J}}_{\bullet,k}$, respectively.

By (D.60), we readily have

$$Q^{\epsilon,\mathcal{J}}_{k}(s;1) = -\left\|\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k} - \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}\right\|_2^2 + \left\|\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k} - \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}}_{\bullet,k}\right\|_2^2,$$

where $\bar{\mathbf{Z}}^{\epsilon,\mathcal{J}}_{\bullet,k}$ is defined as $\bar{\mathbf{v}}$ but with $\mathbf{v}$ replaced by $\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k}$. For $l_{m_0^e} \leqslant s < u_{m_0^e}$, define

$$Q^{\epsilon,\mathcal{J}}_{k}(s;2) = -\left\|\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}\right\|_2^2 + \left\|\mathbf{Z}^{\epsilon,\mathcal{J}}_{\bullet,k} - \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}}_{\bullet,k}\right\|_2^2.$$

By (D.61), we may show that

$$Q^{\epsilon,\mathcal{J}}_{k}(s;1) \geqslant Q^{\epsilon,\mathcal{J}}_{k}(s;2), \quad k = 1, \cdots, d(d+1)/2. \tag{E.53}$$

Since $\mathbf{Z}_{\bullet,k}^{\epsilon,\mathcal{J}} = \mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} + \boldsymbol{z}_{\bullet,k}^{\epsilon,\mathcal{J}}$, we have

$$Q_k^{\epsilon,\mathcal{J}}(s;1) = -\left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}^s}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\epsilon,\mathcal{J}}, \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}^s} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\rangle,$$

and

$$Q_k^{\epsilon,\mathcal{J}}(s;2) = -\left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^s}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\epsilon,\mathcal{J}}, \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^s} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\rangle.$$

Letting

$$Q_k^{\epsilon,\mathcal{J}}(s;3) = -\left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^s}\right\|_2^2 + \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\|_2^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\epsilon,\mathcal{J}}, \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}^s} - \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\rangle,$$

by (D.61), we have

$$Q_k^{\epsilon,\mathcal{J}}(s;3) \geqslant Q_k^{\epsilon,\mathcal{J}}(s;1) \geqslant 0. \tag{E.54}$$

By (E.18), (E.53), (E.54), Proposition 3.3 and the definition of s_0^e, we have

$$\begin{aligned}
&\sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\epsilon,\mathcal{J}}(s_0^e;3)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} \geqslant \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\epsilon,\mathcal{J}}(s_0^e;1)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} = \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\widehat{\epsilon},\widehat{\mathcal{J}}}(s_0^e;1)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} + O_P\left(|\mathcal{T}_{l,u}^e|(\log d)(\log n)\right) \\
\geqslant\ &\sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\widehat{\epsilon},\widehat{\mathcal{J}}}(\eta_{k_\circ}^e;1)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} + O_P\left(|\mathcal{T}_{l,u}^e|(\log d)(\log n)\right) = \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;1)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} + O_P(|\mathcal{T}_{l,u}^e|(\log d)(\log n)) \\
\geqslant\ &\sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;2)}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} + O_P(|\mathcal{T}_{l,u}^e|(\log d)(\log n)),
\end{aligned}$$

where $\widehat{\sigma}_{l,u}(k) := \widehat{\sigma}_{l,u}(k(i,j)) = \widehat{\sigma}_{l,u}(i,j)$, and $Q_k^{\widehat{\epsilon},\widehat{\mathcal{J}}}(s;1)$ is defined similarly to $Q_k^{\epsilon,\mathcal{J}}(s;1)$ but with $\mathbf{Z}_t^{\epsilon,\mathcal{J}}$ replaced by $\mathbf{Z}_t^{\widehat{\epsilon},\widehat{\mathcal{J}}}$. Hence, there exists a sufficiently large constant $c_{33} > 0$ such that

$$\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} \left[Q_k^{\epsilon,\mathcal{J}}(s_0^e;3) - Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;2)\right] \geqslant -c_{33}|\mathcal{T}_{l,u}^e|(\log d)(\log n) \tag{E.55}$$

holds with probability approaching one.

Letting $c_{32} > 0$ be sufficiently large, we next show that the assertion of $\left|s_0^e - \eta_{k_\circ}^e\right| > c_{32}\log^4(nd)/\underline{\omega}_n^e$ would lead to a contradiction with (E.55), which consequently proves (E.52). Defining

$$Q_k^{\epsilon,\mathcal{J}}(s;4) = \left|\langle \mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}}, \boldsymbol{\psi}^s_{l_{m_0^e},u_{m_0^e}}\rangle\right|^2 = -\left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^s}\right\|^2 + \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}}\right\|^2,$$

we have

$$Q_k^{\epsilon,\mathcal{J}}(s;3) - Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;2) \ = \ \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^{\eta_{k_\circ}^e}}\right\|^2 - \left\|\mathbf{G}_{\bullet,k}^{\epsilon,\mathcal{J}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^s}\right\|^2 + 2\left\langle \boldsymbol{z}_{\bullet,k}^{\epsilon,\mathcal{J}}, \bar{\mathbf{Z}}_{\bullet,k}^{\epsilon,\mathcal{J}^s} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,\mathcal{J}^{\eta_{k_\circ}^e}}\right\rangle$$

$$= 2\left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^s}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{\eta^e_{k_\circ}}}_{\bullet,k} \right\rangle - \left[Q^{\epsilon,\mathcal{J}}_k(\eta^e_{k_\circ};4) - Q^{\epsilon,\mathcal{J}}_k(s;4)\right]. \tag{E.56}$$

We next show that with probability approaching one,

$$\begin{aligned} &\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} \left|\left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{\eta^e_{k_\circ}}}_{\bullet,k} \right\rangle\right| \\ \leqslant\ & c_{34}|\mathcal{T}^e_{l,u}|\log^2(nd)\max\left\{\frac{\left|s^e_0-\eta^e_{k_\circ}\right|\cdot(\overline{\varpi}^e_n)^{1/2}}{(\kappa^e_n)^{1/2}}, \left|s^e_0-\eta^e_{k_\circ}\right|^{1/2}(\overline{\varpi}^e_n)^{1/2}, \log^2(nd)\right\} \end{aligned} \tag{E.57}$$

and

$$\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)} \left[Q^{\epsilon,\mathcal{J}}_k(\eta^e_{k_\circ};4) - Q^{\epsilon,\mathcal{J}}_k(s^e_0;4)\right] \geqslant c_{35}|\mathcal{T}^e_{l,u}|\left|s^e_0-\eta^e_{k_\circ}\right|\underline{\omega}^e_n(\kappa^e_n/n)^2, \tag{E.58}$$

where c_{34} and c_{35} are two positive constants.

Without loss of generality, we assume that $s^e_0 \geqslant \eta^e_{k_\circ}$. Note that

$$\begin{aligned} \sum_{k=1}^{d(d+1)/2} \left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{\eta^e_{k_\circ}}}_{\bullet,k} \right\rangle &= \sum_{k=1}^{d(d+1)/2} \left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{s_0^e}}_{\bullet,k} \right\rangle \\ &+ \sum_{k=1}^{d(d+1)/2} \left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^{\eta^e_{k_\circ}}}_{\bullet,k} \right\rangle. \end{aligned} \tag{E.59}$$

Following standard calculations, we have

$$\begin{aligned} \left\langle \boldsymbol{z}^{\epsilon,\mathcal{J}}_{\bullet,k}, \bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^s}_{\bullet,k} - \bar{\mathbf{G}}^{\epsilon,\mathcal{J}^s}_{\bullet,k} \right\rangle &= \left(\sum_{t=l_{m_0^e}}^{s} + \sum_{t=s+1}^{u_{m_0^e}}\right) z^{\epsilon,\mathcal{J}}_{t,k}\left(\bar{Z}^{\epsilon,\mathcal{J}^s}_{t,k} - \bar{G}^{\epsilon,\mathcal{J}^s}_{t,k}\right) \\ &= \frac{1}{s-l_{m_0^e}+1}\left(\sum_{t=l_{m_0^e}}^{s} z^{\epsilon,\mathcal{J}}_{t,k}\right)^2 + \frac{1}{u_{m_0^e}-s}\left(\sum_{t=s+1}^{u_{m_0^e}} z^{\epsilon,\mathcal{J}}_{t,k}\right)^2 \end{aligned} \tag{E.60}$$

for any s, where $Z^{\epsilon,\mathcal{J}^s}_{t,k}$ and $\bar{G}^{\epsilon,\mathcal{J}^s}_{t,k}$ are the $(t-l_{m_0^e}+1)$-th element in $\bar{\mathbf{Z}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}$ and $\bar{\mathbf{G}}^{\epsilon,\mathcal{J}^s}_{\bullet,k}$, respectively. By the definition of $z^{\epsilon,\mathcal{J}}_{t,k}$ and the Cauchy-Schwarz inequality, we have

$$\frac{1}{s-l_{m_0^e}+1}\left(\sum_{t=l_{m_0^e}}^{s} z^{\epsilon,\mathcal{J}}_{t,k}\right)^2 = O_P\left(\log^4(nd)\right),\ \ \frac{1}{u_{m_0^e}-s}\left(\sum_{t=s+1}^{u_{m_0^e}} z^{\epsilon,\mathcal{J}}_{t,k}\right)^2 = O_P\left(\log^4(nd)\right), \tag{E.61}$$

uniformly over s and k. This indicates that

$$\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}^2_{l_{m_0^e}, u_{m_0^e}}(k)} \left\langle \boldsymbol{z}^{\epsilon, \mathcal{J}}_{\bullet,k}, \bar{\boldsymbol{Z}}^{\epsilon, \mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\boldsymbol{G}}^{\epsilon, \mathcal{J}^{s_0^e}}_{\bullet,k} \right\rangle \leqslant (c_{34}/4) \cdot |\mathcal{T}^e_{l,u}| \log^4(nd) \tag{E.62}$$

with probability approaching one. On the other hand,

$$\begin{aligned} \left\langle \boldsymbol{z}^{\epsilon, \mathcal{J}}_{\bullet,k}, \bar{\boldsymbol{G}}^{\epsilon, \mathcal{J}^{s_0^e}}_{\bullet,k} - \bar{\boldsymbol{G}}^{\epsilon, \mathcal{J}^{\eta^e_{k_\circ}}}_{\bullet,k} \right\rangle &= \left(\sum_{t=l_{m_0^e}}^{\eta^e_{k_\circ}} + \sum_{t=\eta^e_{k_\circ}+1}^{s_0^e} + \sum_{t=s_0^e+1}^{u_{m_0^e}} \right) z^{\epsilon,\mathcal{J}}_{t,k} \left(\bar{G}^{\epsilon, \mathcal{J}^{s_0^e}}_{t,k} - \bar{G}^{\epsilon, \mathcal{J}^{\eta^e_{k_\circ}}}_{t,k} \right) \\ &=: \ \Xi_1 + \Xi_2 + \Xi_3. \end{aligned} \tag{E.63}$$

For Ξ_1, we note that

$$|\Xi_1| \leqslant \sqrt{\eta^e_{k_\circ} - l_{m_0^e} + 1} \left| \frac{1}{\sqrt{\eta^e_{k_\circ} - l_{m_0^e} + 1}} \sum_{t=l_{m_0^e}}^{\eta^e_{k_\circ}} z^{\epsilon,\mathcal{J}}_{t,k} \right| \cdot \left| \frac{1}{s_0^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{s_0^e} G^{\epsilon,\mathcal{J}}_{t,k} - \frac{1}{\eta^e_{k_\circ} - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta^e_{k_\circ}} G^{\epsilon,\mathcal{J}}_{t,k} \right|,$$

and recall that

$$b^{\epsilon,\mathcal{J}}_{l,u} = \sup_{l \leqslant t \leqslant u} \left\| \boldsymbol{G}^{\epsilon,\mathcal{J}}_t - \frac{1}{u-l+1} \sum_{t=l}^{u} \boldsymbol{G}^{\epsilon,\mathcal{J}}_t \right\|_2 \leqslant K_2 \left(|\mathcal{T}^e_{l,u}| \varpi^e_n \right)^{1/2}.$$

Let

$$\boldsymbol{z}^{\epsilon,\mathcal{J}}_s = \left[z^{\epsilon}_{s,11} \cdot \mathcal{I}\left(\max_{t: l \leqslant t < u} \left| c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;1,1) \right| > \xi^e_n \right), \cdots, z^{\epsilon}_{s,dd} \cdot \mathcal{I}\left(\max_{t: l \leqslant t < u} \left| c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;d,d) \right| > \xi^e_n \right) \right]^{\intercal},$$

which is a $d(d+1)/2$ column vector obtained via half-vectorisation of a $d \times d$ symmetric matrix with the (i,j)-entry being $z^{\epsilon}_{s,ij} \cdot \mathcal{I}\left(\max_{t: l \leqslant t < u} \left| c^{G,\epsilon,\widehat{\sigma}}_{l,u}(t;i,j) \right| > \xi^e_n \right)$. Then, by (E.61), Assumption 5 and the Cauchy-Schwarz inequality,

$$\begin{aligned} &\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}^2_{l_{m_0^e}, u_{m_0^e}}(k)} \cdot |\Xi_1| \\ \leqslant\ & \frac{\sqrt{\eta^e_{k_\circ} - l_{m_0^e} + 1}}{\underline{\sigma}^2} \left\| \frac{1}{\sqrt{\eta^e_{k_\circ} - l_{m_0^e} + 1}} \sum_{t=l_{m_0^e}}^{\eta^e_{k_\circ}} \boldsymbol{z}^{\epsilon,\mathcal{J}}_t \right\|_2 \cdot \left\| \frac{1}{s_0^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{s_0^e} \boldsymbol{G}^{\epsilon,\mathcal{J}}_t - \frac{1}{\eta^e_{k_\circ} - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta^e_{k_\circ}} \boldsymbol{G}^{\epsilon,\mathcal{J}}_t \right\|_2 \\ \leqslant\ & \frac{\sqrt{\eta^e_{k_\circ} - l_{m_0^e} + 1}}{\underline{\sigma}^2} \cdot O_P\left(|\mathcal{T}^e_{l,u}|^{1/2} \log^2(nd) \right) \cdot \frac{\left| s_0^e - \eta^e_{k_\circ} \right| 2 b^{\epsilon,\mathcal{J}}_{l + c_{22}\varphi^e_{n,d}, u - c_{22}\varphi^e_{n,d}}}{s_0^e - l_{m_0^e} + 1} \\ =\ & O_P\left(|\mathcal{T}^e_{l,u}| \log^2(nd) \left| s_0^e - \eta^e_{k_\circ} \right| \cdot (\varpi^e_n / \kappa^e_n)^{1/2} \right), \end{aligned} \tag{E.64}$$

where $\mathbf{G}_t^{\epsilon,\mathcal{J}}$ is defined in the proof of Lemma E.5. The asymptotic order for $\sum_{k=1}^{d(d+1)/2}\left[\Xi_3/\widehat{\sigma}_{l_{m_0^e},u_{m_0^e}}(k)\right]$ is the same as that for $\sum_{k=1}^{d(d+1)/2}\left[\Xi_1/\widehat{\sigma}_{l_{m_0^e},u_{m_0^e}}(k)\right]$. Similarly, we may show that

$$\sum_{k=1}^{d(d+1)/2}\frac{\Xi_2}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)}\leqslant (c_{34}/4)\cdot|\mathcal{T}_{l,u}^e|\log^2(nd)\left|s_0^e-\eta_{k_\circ}^e\right|^{1/2}(\overline{\omega}_n^e)^{1/2} \tag{E.65}$$

with probability approaching one. Using (E.59) and (E.62)–(E.65), we compete the proof of (E.57).

By Lemmas E.3 and E.5, we have

$$\begin{aligned}
&\sum_{k=1}^{d(d+1)/2}\frac{1}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)}\cdot\left[Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;4)-Q_k^{\epsilon,\mathcal{J}}(s_0^e;4)\right]\\
=\;&\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta_{k_\circ}^e)\right\|_2^2-\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s_0^e)\right\|_2^2\\
=\;&\left[\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta_{k_\circ}^e)\right\|_2-\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s_0^e)\right\|_2\right]\cdot\left[\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta_{k_\circ}^e)\right\|_2+\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s_0^e)\right\|_2\right]\\
\geqslant\;&\left[\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta_{k_\circ}^e)\right\|_2-\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(s_0^e)\right\|_2\right]\cdot\left\|\mathbf{C}^{\mathbf{G},\epsilon,\mathcal{J}}_{l_{m_0^e},u_{m_0^e}}(\eta_{k_\circ}^e)\right\|_2\\
\geqslant\;&c_{35}|\mathcal{T}_{l,u}^e|\left|s_0^e-\eta_{k_\circ}^e\right|\underline{\omega}_n^e(\kappa_n^e/n)^2,
\end{aligned}\tag{E.66}$$

completing the proof of (E.58).

Finally, by (E.57), (E.58) and Lemma E.6, we have

$$\begin{aligned}
&\sum_{k=1}^{d(d+1)/2}\frac{1}{\widehat{\sigma}^2_{l_{m_0^e},u_{m_0^e}}(k)}\cdot\left[Q_k^{\epsilon,\mathcal{J}}(s_0^e;3)-Q_k^{\epsilon,\mathcal{J}}(\eta_{k_\circ}^e;2)\right]\\
\leqslant\;&c_{34}|\mathcal{T}_{l,u}^e|\log^2(nd)\max\left\{\frac{\left|s_0^e-\eta_{k_\circ}^e\right|\cdot(\overline{\omega}_n^e)^{1/2}}{(\kappa_n^e)^{1/2}},\left|s_0^e-\eta_{k_\circ}^e\right|^{1/2}(\overline{\omega}_n^e)^{1/2},\log^2(nd)\right\}\\
&-c_{35}|\mathcal{T}_{l,u}^e|\left|s_0^e-\eta_{k_\circ}^e\right|\underline{\omega}_n^e(\kappa_n^e/n)^2,\\
\leqslant\;&-c_{33}|\mathcal{T}_{l,u}^e|\log^4(nd),
\end{aligned}\tag{E.67}$$

which would lead to a contradiction with (E.55) if we choose c_{32} to be sufficiently large. The proof of Lemma E.7 has been completed. □

PROOF OF THEOREM 3.2. When starting with the WSBS-Cov algorithm, we have $l=1$ and $u=n$ and we may show that (E.3)–(E.5) are satisfied. Then, by (3.10), Lemmas E.3 and E.6, the estimated change point s_0^e satisfies (E.52) with probability approaching one. In addition, Lemma E.5 shows that s_0^e is not close to l and u, so it is a newly detected change point. By (E.52), we may show that (E.3)–(E.5) still hold within each segment until all of the change points in the idiosyncratic error component are detected. By Lemma E.7, the estimated change points satisfy the convergence result (E.52) with probability approaching one. Once all of the change points are detected, the bounds of each segment l and u must fall into one of the following three scenarios: (i) there exists $1 \leqslant k \leqslant K_2$ such that $\eta_k^e < l < u \leqslant \eta_{k+1}^e$; (ii) there exists $1 \leqslant k \leqslant K_2$ such that $l \leqslant \eta_k^e < u$ and $(\eta_k^e - l + 1) \wedge (u - \eta_k^e) \leqslant c_{32}\varphi_{n,d}^e$; (iii) there exists $1 \leqslant k \leqslant K_2$ such that $l \leqslant \eta_k^e < \eta_{k+1}^e < u$ and $(\eta_k^e - l + 1) \vee (u - \eta_{k+1}^e) \leqslant c_{32}\varphi_{n,d}^e$, where c_{32} is defined in Lemma E.7. For l and u satisfy either of scenarios (i)–(iii), we may show that

$$\max_{1 \leqslant i,j \leqslant d} \max_{l_{m_0^e} \leqslant s < u_{m_0^e}} \left| c^{\widehat{\epsilon}}_{l_{m_0^e}, u_{m_0^e}}(s_0^e; i, j) \right| = O_P\left(\log^2(nd)\right), \tag{E.68}$$

which together with (3.10), Lemmas E.3 and E.5, indicates that no further change point could be detected. Letting $\iota^e = c_{32}$, the proof of Theorem 3.2 is completed. □

Appendix F: Additional simulation results

We next provide simulation studies to further compare the finite-sample performance between the proposed methods and various other competing methods. As in Section 5 of the main document, we consider the following factor model to generate data:

$$X_{ti} = \sum_{j=1}^{r} \lambda_{ij,t} F_{tj} + \sqrt{\theta}\epsilon_{ti}, \quad i = 1, \cdots, d, \quad t = 1, \cdots, n. \tag{F.1}$$

The replication number in each simulation cases is set to $R = 100$. For the 100 simulated samples, we report the estimated number of break(s) as well as the accuracy measure ACU_k for each break defined in (5.2). In Example F.1 below, we compare the numerical performance among the WBS-Cov and WSBS-Cov, BS-Cov and SBS-Cov algorithms, and examine the finite-sample influence of different norms used in aggregation of the CUSUM quantities and various transformation techniques used in construction of the CUSUM statistics.

Example F.1. Consider the factor model in (F.1) with $\theta = 1$. The sample size is $n = 200$, and the dimension is $d = 200$. In this example, we consider the scenario of a single break in both the common and idiosyncratic components: $\eta_1^c = \lfloor n/3 \rfloor + 1 = 67$ and $\eta_1^e = \lfloor 2n/3 \rfloor = 133$. The number of factors is set to be $r = 5$, and each factor process is generated via an AR(1) model:

$$F_{tj} = \rho_j F_{t-1,j} + u_{tj}, \quad t = 1, \cdots, n, \tag{F.2}$$

where u_{tj} follows a standard normal distribution independently over t and j, and $\rho_j = 0.4 -$

$0.05(j-1)$ for $j = 1, \cdots, 5$. The factor loadings $\lambda_{ij,t}$ are first generated from a standard normal distribution independently over i and j when t is from 1 to η_1^c; whereas after the break point η_1^c, the factor loadings $\lambda_{ij,t}$ are shifted by a random amount $\mathsf{N}(0,4)$ as in Barigozzi, Cho and Fryzlewicz (2018). The sudden change on the factor loadings leads to break in the second-order moment structure of the common components. The idiosyncratic errors $\boldsymbol{\epsilon}_t$ follow a multivariate normal distribution $\mathsf{N}_d(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$ independently over t, where ϕ_j, the square root of the j-th diagonal element of $\boldsymbol{\Sigma}_\epsilon$, is generated from an independent uniform distribution $\mathsf{U}(0.5, 1.5)$, and the (i,j)-entry of $\boldsymbol{\Sigma}_\epsilon$ is $\phi_i\phi_j(-0.5)^{|i-j|}$ for $1 \leqslant i \neq j \leqslant d$. After the break point η_1^e, we swap the orders of $\lfloor \rho_1^e d/2 \rfloor$ randomly selected pairs of elements of $\boldsymbol{\epsilon}_t$ (c.f., Cho and Fryzlewicz, 2015) with ρ_1^e chosen as 0.1, 0.5 or 1. Note that $\rho_1^e = 0.1$ indicates that the structural breaks are relatively sparse in the high-dimensional error components, whereas $\rho_1^e = 1$ indicates that the breaks are dense.

Table 1: Comparison of detection results using different BS-based methods

	Common components					Idiosyncratic error components				
	Methods	# break(%)			ACU_1(%)	Methods	# break(%)			ACU_1(%)
		<1	1	>1	$\eta_1^c = 67$		<1	1	>1	$\eta_1^e = 133$
$\rho_1^e = 1$	BS-Cov	0	99	1	100	BS-Cov	0	97	3	100
						SBS-Cov	0	97	3	100
	WBS-Cov	0	99	1	100	WBS-Cov	0	98	2	98
						WSBS-Cov	0	99	1	100
$\rho_1^e = 0.5$	BS-Cov	0	100	0	100	BS-Cov	0	96	4	99
						SBS-Cov	0	99	1	98
	WBS-Cov	0	100	0	100	WBS-Cov	0	94	6	95
						WSBS-Cov	0	100	0	98
$\rho_1^e = 0.1$	BS-Cov	0	99	1	100	BS-Cov	24	72	4	53
						SBS-Cov	20	80	0	61
	WBS-Cov	0	99	1	100	WBS-Cov	28	70	2	31
						WSBS-Cov	20	80	0	61

In Table 1, we compare the proposed WBS-Cov with the classical BS-Cov in detecting breaks in the common components, and compare the proposed WSBS-Cov with the BS-Cov, WBS-Cov and SBS-Cov in detecting breaks in the idiosyncratic components. For the break detection in the common component, the finite-sample performance of WBS-Cov and BS-Cov are the same. For the break detection in the idiosyncratic components, the four methods behave differently in finite samples. When the breaks are sparse in the high-dimensional error covariance matrix ($\rho_1^e = 0.1$), the sparsified detection techniques (WSBS-Cov and SBS-Cov) outperform the non-sparsified ones (BS-Cov and WBS-Cov) in both the break number and location estimation; when the breaks are dense ($\rho_1^e = 0.5$ and 1), the proposed WSBS-Cov has the best performance in estimating the break number whereas the BS-Cov performs better than the other three methods in estimating the break location.

In Table 2, we examine the finite-sample influence of different norms used in the aggregation

Table 2: Comparison of detection results using different norms in the CUSUM statistics

		Common components				Idiosyncratic error components			
		# break(%)			$\mathrm{ACU}_1(\%)$	# break(%)			$\mathrm{ACU}_1(\%)$
		<1	1	>1	$\eta_1^c = 67$	<1	1	>1	$\eta_1^e = 133$
Breaks in common components									
$\rho_1^e = 1$	l_1	0	99	1	99	0	99	1	98
	l_2	0	99	1	100	0	99	1	100
	l_∞	0	79	21	64	0	90	10	77
	op	0	100	0	99	0	89	11	89
$\rho_1^e = 0.5$	l_1	0	100	0	100	0	97	3	94
	l_2	0	100	0	100	0	100	0	98
	l_∞	0	79	21	69	0	94	6	69
	op	0	99	1	99	0	94	6	74
$\rho_1^e = 0.1$	l_1	0	99	1	100	23	72	5	50
	l_2	0	99	1	100	20	80	0	61
	l_∞	0	78	22	65	23	75	2	42
	op	0	100	0	98	32	66	2	39

of the CUSUM quantities. For the idiosyncratic components, as in (2.12), the CUSUM statistic aggregated with the l_1-norm is defined by

$$\sum_{i=1}^{d}\sum_{j=i}^{d}\left|c_{l_m,u_m}^{\widehat{\epsilon},\widehat{\sigma}}(s;i,j)\right|\mathcal{I}\left(\max_{l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right|>\xi_n^e\right)$$

and the CUSUM statistic aggregated with the l_∞-norm is defined by

$$\max_{1\leqslant i\leqslant j\leqslant d}\left\{\left|c_{l_m,u_m}^{\widehat{\epsilon},\widehat{\sigma}}(s;i,j)\right|\mathcal{I}\left(\max_{l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right|>\xi_n^e\right)\right\};$$

and the construction is similar for the common components. In addition, we also consider aggregating via the operator norm, as suggested in Wang, Yu and Rinaldo (2021). For the idiosyncratic components, let $C_{l_m,u_m}^{M,\widehat{\epsilon}}(s)$ be a $d\times d$ matrix with the (i,j)-th entry being

$$c_{l_m,u_m}^{\widehat{\epsilon},\widehat{\sigma}}(s;i,j)\mathcal{I}\left(\max_{l\leqslant t<u}\left|c_{l,u}^{\widehat{\epsilon},\widehat{\sigma}}(t;i,j)\right|>\xi_n^e\right)$$

and then obtain the CUSUM statistic by taking the operator norm of $C_{l_m,u_m}^{M,\widehat{\epsilon}}(s)$. For the common components, the CUSUM statistic is defined by taking the operator norm of the matrix:

$$\sqrt{\frac{(s-l+1)(u-s)}{u-l+1}}\left[\frac{1}{s-l+1}\sum_{t=l}^{s}\widehat{F}_t\widehat{F}_t^{\intercal}-\frac{1}{u-s}\sum_{t=s+1}^{u}\widehat{F}_t\widehat{F}_t^{\intercal}\right].$$

It is obvious from Table 2 that the l_2-based detection method has the best finite-sample performance with more accurate estimated break number and higher ACU. The operator norm based detection method performs well in break detection for the common components, but it performs poorly when breaks are sparse in the idiosyncratic components.

Table 3: Comparison of detection results using different transformations in break detection

	Common components					Idiosyncratic error components				
	Methods	# break(%)			$\mathsf{ACU}_1(\%)$	Methods	# break(%)			$\mathsf{ACU}_1(\%)$
		<1	1	>1	$\eta_1^c = 67$		<1	1	>1	$\eta_1^e = 133$
$\rho_1^e = 1$	BCF	0	95	5	100	BCF(D)	0	100	0	100
						BCF	0	100	0	100
	WBS-Cov	0	99	1	100	WSBS-Cov(D)	0	100	0	99
						WSBS-Cov	0	99	1	100
	WAVELET	0	92	8	100	WAVELET	0	93	7	100
	ADD-MNS	0	99	1	100	ADD-MNS	0	83	17	98
$\rho_1^e = 0.5$	BCF	0	94	6	100	BCF(D)	0	100	0	100
						BCF	0	100	0	100
	WBS-Cov	0	100	0	100	WSBS-Cov(D)	0	100	0	100
						WSBS-Cov	0	100	0	98
	WAVELET	0	96	4	100	WAVELET	0	100	0	98
	ADD-MNS	0	100	0	100	ADD-MNS	0	90	10	95
$\rho_1^e = 0.1$	BCF	0	96	4	100	BCF(D)	24	76	0	55
						BCF	50	50	0	44
	WBS-Cov	0	99	1	100	WSBS-Cov(D)	21	79	0	65
						WSBS-Cov	20	80	0	61
	WAVELET	0	92	8	100	WAVELET	7	77	16	64
	ADD-MNS	0	100	0	100	ADD-MNS	0	87	13	61

Table 3 reports the simulation result when different transformation techniques are used in construction of the CUSUM statistics. In the table, "BCF" denotes the method proposed by Barigozzi, Cho and Fryzlewicz (2018) which combines the wavelet-based transformation and the double-CUSUM method, "WBS-Cov" denotes the proposed method in Section 2.3, and "WSBS-Cov" denotes the proposed method in Section 2.4. For structural breaks in the covariance matrix of the error components, we may detect the breaks only for its diagonal elements (variance) rather than all the elements in the high-dimensional covariance matrix in order to save computational time. This is considered in our simulation with "BCF(D)" and "WSBS-Cov(D)" denoting the "BCF" and "WSBS-Cov" methods by only detecting breaks for the diagonal elements. Letting a_i and a_j be either the common factors or the idiosyncratic errors, "ADD-MNS" denotes a transformation of $(a_i + a_j)^2$ and $(a_i - a_j)^2$ (e.g., Cho and Fryzlewicz, 2015) in the construction of the CUSUM statistics (instead of $a_i a_j$ in our proposed method), whereas "WAVELET" denotes the wavelet transformation on a_i and a_j (e.g., Barigozzi, Cho and Fryzlewicz, 2018) in the construction of the CUSUM statistics. The algorithms introduced in Sections 2.3 and 2.4 are used after making the "WAVELET" and "ADD-MINS" transformations. The R package "factorcpt" is used to implement Barigozzi, Cho and Fryzlewicz (2018)'s method in the simulation.

From the table, the proposed WBS-Cov algorithm and the "ADD-MINS" method have the best finite-sample performance in estimating the break in the common components. In terms of the idiosyncratic components, the "WSBS" method has similar performance to the "BCF" method, and the best performance is from the "WSBS-Cov(D)" method. In terms of "WAVELET" method, we find that the thresholding parameter ξ_n^e selected in pre-estimation is too small, and thus use $\sqrt{2}\xi_n^e$ as the threshold. However, this method tends to over-estimate the break number. The performance of the "ADD-MINS" method in estimating the break location is not as good as the other methods, which might be caused by selection of the thresholding parameter ξ_n^e.

In the following example, we consider an alternative weak factor structure which is different from that in Example 5.2 of the main document. The factor loadings are not sparse but have small magnitude.

EXAMPLE F.2. We use model (F.1) to generate the data in simulation, where the number of factors is $r = 3$, the sample size is $n = 400$, the dimension is $d = 200$, and $\theta = 1$. The factor process F_t is generated from a multivariate normal distribution $N_3(\mathbf{0}, \mathbf{\Sigma}_F^*)$ independently over t, where $\mathbf{\Sigma}_F^*$ is the covariance matrix specified as follows: the square root of the j-th diagonal element of $\mathbf{\Sigma}_F^*$, is independently generated from a uniform distribution $U(0.5, 1.5)$, and the (i, j)-entry of $\mathbf{\Sigma}_F^*$ is defined as $\phi_i^F\phi_j^F(0.5)^{|i-j|}$ for $1 \leqslant i \neq j \leqslant 3$. For $1 \leqslant t \leqslant \eta_1^c = 100$, the factor loadings for the first factor, λ_{i1} are independently generated from a uniform distribution $U(-w, w)$, and the factor loadings for the second and third factors, λ_{i2} and λ_{i3}, are independently generated from a uniform distribution $U(-1, 1)$; for $\eta_1^c < t \leqslant \eta_2^c = 300$, the factor loadings λ_{i1} are regenerated from a uniform distribution $U(-w, w)$; whereas for $\eta_2^c < t \leqslant 400$, the factor loadings corresponding to the first two factors are regenerated by uniform distribution $U(-w, w)$ and $U(-1, 1)$, respectively. We consider five different cases by setting $w = n^{(a_i-1)/2}$ with $(a_1, \cdots, a_5) = (1, 0.85, 0.75, 2/3, 0.6)$.

The idiosyncratic errors $\boldsymbol{\epsilon}_t$ follow a multivariate normal distribution $N_d(\mathbf{0}, \mathbf{\Sigma}_\epsilon)$ independently over t, where ϕ_j, the square root of the j-th diagonal element of $\mathbf{\Sigma}_\epsilon$, is generated from an independent uniform distribution $U(0.5, 1.5)$, and the (i, j)-entry of $\mathbf{\Sigma}_\epsilon$ is $\phi_i\phi_j(-0.5)^{|i-j|}$ for $1 \leqslant i \neq j \leqslant d$. We set three breaks $\eta_1^e = \lfloor n/8 \rfloor = 50$, $\eta_2^e = \lfloor n/2 \rfloor = 200$ and $\eta_3^e = \lfloor 7n/8 \rfloor = 350$. At each of the three break points η_1^e and η_2^e, we swap the orders of $\lfloor 0.8d/2 \rfloor$ randomly selected pairs of elements of $\boldsymbol{\epsilon}_t$.

Table 4 shows that under-estimation of the factor number would negatively impact break detection. In this example, the number of factors for the transformed factor model (2.4) is 6 (3 original factors plus 3 factors due to factor transformation accommodating breaks). However, the mean value of $\widehat{q}$ is only 5.01 when $w = 1$ in case 1 and is even smaller in other cases when factors are weaker. The information criterion tends to under-estimate the number of factors in all cases. To see the impact of under-estimating the factor number, we set r to be 6 and 9, and

Table 4: Break detection results for the weak factor model with non-sparse factor loadings

	$\widehat{q}$	Common components						Idiosyncratic error components					
		# break(%)				ACU_1(%)	ACU_2(%)	# break(%)			ACU_1(%)	ACU_2(%)	ACU_3(%)
		0	1	2	>2	$\eta_1^c = 100$	$\eta_2^c = 300$	<3	3	>3	$\eta_1^e = 50$	$\eta_2^e = 200$	$\eta_3^e = 350$
Case 1	5.01	1	12	87	0	77	89	0	99	1	99	100	99
	$\widehat{q} = 9$ fixed	0	0	100	0	79	94	1	98	1	99	100	99
	$\widehat{q} = 6$ fixed	0	0	100	0	81	95	0	100	0	99	100	100
	$\widehat{q} = 3$ fixed	3	52	25	0	42	75	13	84	3	53	88	49
Case 2	4.11	3	33	64	0	54	78	0	99	1	98	100	99
	$\widehat{q} = 9$ fixed	0	3	97	0	77	92	0	99	1	99	100	100
	$\widehat{q} = 6$ fixed	0	7	93	0	75	93	0	100	0	99	100	100
	$\widehat{q} = 3$ fixed	3	68	29	0	26	68	2	93	5	84	98	81
Case 3	3.42	4	64	32	0	21	69	0	98	2	99	100	100
	$\widehat{q} = 9$ fixed	0	16	84	0	68	86	0	99	1	99	100	100
	$\widehat{q} = 6$ fixed	0	16	84	0	69	86	0	100	0	99	100	100
	$\widehat{q} = 3$ fixed	3	89	8	0	6	68	0	96	4	94	100	97
Case 4	3.01	10	74	16	0	12	65	0	98	2	98	100	100
	$\widehat{q} = 9$ fixed	0	34	66	0	52	83	1	99	0	99	100	99
	$\widehat{q} = 6$ fixed	0	36	64	0	54	83	0	100	0	99	100	100
	$\widehat{q} = 3$ fixed	1	98	1	0	2	73	0	98	2	98	100	100
Case 5	2.81	10	88	2	0	2	65	0	99	1	98	100	100
	$\widehat{q} = 9$ fixed	0	53	47	0	35	76	1	99	0	99	100	99
	$\widehat{q} = 6$ fixed	0	58	42	0	31	78	0	100	0	99	100	100
	$\widehat{q} = 3$ fixed	0	100	0	0	0	75	0	99	1	98	100	100

then detect the breaks again. We find that the performance of detection is improved significantly. On the contrary, if we set r to be 3, the proposed break detection method performs worse for the common components. Although under-estimation of the factor number also affects the detection of breaks in the idiosyncratic components, the impact is not as significant as that on the common components.

References

BAI, J. AND NG, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

BARIGOZZI, M., CHO, H. AND FRYZLEWICZ, P. (2018). Simultaneous multiple change-point and factor analysis for high-dimensional time series. *Journal of Econometrics* **206**, 187–225.

BOSQ, D. (1998). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction* (2nd Edition). Lecture Notes in Statistics 110, Springer-Verlag, Berlin.

CHEN, J., LI, D., LINTON, O. AND LU, Z. (2018). Semiparametric ultra-high dimensional model averaging of nonlinear dynamic time series. *Journal of the American Statistical Association* **113**, 919–932.

CHO, H. AND FRYZLEWICZ P. (2015). Multiple change-point detection for high-dimensional time series via Sparsified Binary Segmentation. *Journal of the Royal Statistical Society Series B* **77**, 475–507.

FAN, J., LIAO, Y. AND MINCHEVA, M. (2013). Large covariance estimation by thresholding principal orthogonal complements (with discussion). *Journal of the Royal Statistical Society, Series B* **75**, 603–680.

HAN X. AND INOUE, A. (2015). Tests for parameter instability in dynamic factor models. *Econometric Theory* **31**, 1117–1152.

LIN, Z. AND BAI, Z. (2010). *Probability Inequalities*. Springer Science & Business Media.

LIN, Z. AND LU, C. (1996). *Limit Theory for Mixing Dependent Random Variables*. Science Press / Kluwer Academic Publishers.

MARSHALL, A. W., OLKIN, I. AND ARNOLD, B. (2011). *Inequalities: Theory of Majorization and Its Applications* (2nd ed.). Springer, New York.

VENKATRAMAN, E. S. (1992). Consistency results in multiple change-point problems. *Technical Report No. 24, Department of Statistics, Stanford University*.

WANG, D., YU, Y. AND RINALDO, A. (2021). Optimal covariance change point localization in high dimensions. *Bernoulli* **27**, 554–575.