

# Supplement to “Detection of Multiple Structural Breaks in Large Covariance Matrices”

Yu-Ning Li\*

Degui Li†

Piotr Fryzlewicz‡

This version: April 19, 2021

In this supplement, we give the technical assumptions, the detailed proofs of the main theoretical results as well as additional simulation studies. Appendix A lists the assumptions needed to establish the main asymptotic properties. 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 on selection of the factor number. Throughout the supplemental document, we let  $M$  be a generic positive constant whose value may change from line to line.

## Appendix A: Technical assumptions

The following assumptions are needed to establish the asymptotic theorems in Sections 3.1 and 3.2. Some of the conditions (such as the moment conditions) may be not the weakest possible but can be relaxed at the cost of more lengthy proofs.

ASSUMPTION 1. (i) For each  $k = 1, \dots, K_1 + 1$ , the process  $\{(\mathbf{F}_{t,k}, \boldsymbol{\epsilon}_t)\}$  is  $\alpha$ -mixing dependent with the mixing coefficient  $\alpha_k(\cdot)$  satisfying

$$\max_{1 \leq k \leq K_1 + 1} \alpha_k(s) \asymp \rho^s, \quad 0 < \rho < 1.$$

(ii) For  $k = 1, \dots, K_1 + 1$  and  $j = 1, \dots, d$ ,  $\mathbf{E}[\mathbf{F}_{t,k}] = \mathbf{0}$ ,  $\mathbf{E}[\boldsymbol{\epsilon}_t] = \mathbf{0}$  and  $\mathbf{E}[\boldsymbol{\epsilon}_{tj} \mathbf{F}_{t,k}] = \mathbf{0}$ . In

---

\*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.

addition,

$$\max_{1 \leq k \leq K_1+1} \max_{\eta_{k-1}^c+1 \leq t \leq \eta_k^c} \mathbb{E}[\|\mathbf{F}_{t,k}\|_2^{\delta_F}] < \infty, \quad \max_{1 \leq t \leq n} \max_{1 \leq j \leq d} \mathbb{E}[|\epsilon_{tj}|^{\delta_\epsilon}] < \infty,$$

where  $\delta_F > 2$  and  $\delta_\epsilon > 2$ .

ASSUMPTION 2. (i) For  $k = 1, \dots, K_1 + 1$ , the covariance matrix  $\text{Cov}(\mathbf{F}_{t,k}) = \Sigma_{F,k}$  of the latent factors when  $\eta_{k-1}^c + 1 \leq t \leq \eta_k^c$ , is  $r_k \times r_k$  positive definite. The positive integers  $K_1$  and  $r_k$ ,  $k = 1, \dots, K_1 + 1$ , are all bounded.

(ii) Letting  $\lambda_{k,j}^0$  be the  $j$ -th factor loading vector of  $\Lambda_k^0$  defined in (2.2),  $\|\lambda_{k,j}^0\|_2$  is bounded from above uniformly over  $j = 1, \dots, d$  and  $k = 1, \dots, K_1 + 1$ . In addition, letting  $\bar{\Lambda} = [\Lambda_1^0, \dots, \Lambda_{K_1+1}^0]$  which is a  $d \times \sum_{k=1}^{K_1+1} r_k$  matrix, the column rank of  $\bar{\Lambda}$  is  $q_0$  and there exist positive constants  $\underline{\mu}$  and  $\bar{\mu}$  such that

$$0 < \underline{\mu} \leq \mu_{q_0} \left( \frac{1}{d} \bar{\Lambda}^\top \bar{\Lambda} \right) \leq \mu_1 \left( \frac{1}{d} \bar{\Lambda}^\top \bar{\Lambda} \right) \leq \bar{\mu} < \infty. \quad (\text{A.1})$$

ASSUMPTION 3. (i) There exists a positive constant  $\iota_0$  such that

$$\max_{1 \leq k \leq K_1+1} \max_{1 \leq t \leq n} \mathbb{E} \left[ \left\| \sum_{j=1}^d \lambda_{k,j}^0 \epsilon_{tj} \right\|_2^{\delta_\epsilon} \right] \leq \iota_0 \cdot d^{\delta_\epsilon/2} \quad (\text{A.2})$$

and

$$\max_{1 \leq s, t \leq n} \mathbb{E} \left[ \left| \sum_{j=1}^d \{\epsilon_{sj} \epsilon_{tj} - \mathbb{E}[\epsilon_{sj} \epsilon_{tj}]\} \right|^{\delta_\epsilon} \right] \leq \iota_0 \cdot d^{\delta_\epsilon/2}. \quad (\text{A.3})$$

(ii) There exists a positive constant  $\iota_1 > 0$  such that

$$\max_{1 \leq j \leq d} \max_{1 \leq t \leq n} \mathbb{E} [\exp \{ \iota_1 \epsilon_{tj}^2 \}] < \infty, \quad \max_{1 \leq k \leq K_1+1} \max_{\eta_{k-1}^c+1 \leq t \leq \eta_k^c} \mathbb{E} [\exp \{ \iota_1 \|\mathbf{F}_{t,k}\|_2^2 \}] < \infty. \quad (\text{A.4})$$

ASSUMPTION 4. (i) There exist positive constants  $\iota_2$ ,  $\iota_3$  and  $\vartheta$  such that

$$\iota_2 \log^{5+\vartheta}(d) \leq n \leq \iota_3 d^{\delta/(\delta+4)}, \quad \delta = \delta_F \wedge \delta_\epsilon.$$

(ii) Letting  $\kappa_n^c = \min_{1 \leq k \leq K_1+1} (\eta_k^c - \eta_{k-1}^c)$ ,

$$\underline{\omega}_n^c = \frac{1}{d^2} \cdot \min_{1 \leq k \leq K_1} \|\Sigma_{k+1}^0(\Lambda, \mathbf{F}) - \Sigma_k^0(\Lambda, \mathbf{F})\|_F^2$$

and

$$\bar{\omega}_n^c = \frac{1}{d^2} \cdot \max_{1 \leq k \leq K_1} \|\Sigma_{k+1}^0(\mathbf{A}, \mathbf{F}) - \Sigma_k^0(\mathbf{A}, \mathbf{F})\|_F^2,$$

we have

$$\kappa_n^c \asymp n, \quad \frac{\omega_n^c}{\bar{\omega}_n^c} \cdot \frac{\kappa_n^c \omega_n^c}{\log^4 n} \rightarrow \infty, \quad M_n^c \rightarrow \infty,$$

where  $M_n^c$  is the number of random intervals drawn in the WBS-Cov algorithm.

(iii) Letting  $\kappa_n^e = \min_{1 \leq k \leq K_2+1} (\eta_k^e - \eta_{k-1}^e)$ ,

$$\underline{\omega}_n^e = \min_{1 \leq k \leq K_2} \min_{(i,j) \in \mathcal{J}_k} \left| \sigma_{k+1|i,j}^e - \sigma_{k|i,j}^e \right|^2, \quad \bar{\omega}_n^e = \max_{1 \leq k \leq K_2} \max_{(i,j) \in \mathcal{J}_k} \left| \sigma_{k+1|i,j}^e - \sigma_{k|i,j}^e \right|^2,$$

where  $\mathcal{J}_k = \left\{ (i,j) : \sigma_{k+1|i,j}^e \neq \sigma_{k|i,j}^e, 1 \leq i,j \leq d \right\}$  and  $\sigma_{k|i,j}^e$  is the  $(i,j)$ -entry of  $\Sigma_k^0(\boldsymbol{\epsilon})$ , we have

$$\bar{\omega}_n^e \asymp \underline{\omega}_n^e, \quad \frac{\kappa_n^e \underline{\omega}_n^e}{\log^4(nd)} \cdot \left( \frac{\kappa_n^e}{n} \right)^4 \rightarrow \infty, \quad \left[ 1 - (\kappa_n^e/(3n))^2 \right]^{M_n^e} \rightarrow 0,$$

where  $M_n^e$  is the number of random intervals drawn in the WSBS-Cov algorithm.

ASSUMPTION 5. There exist  $0 < \underline{\sigma} < \bar{\sigma} < \infty$  such that

$$\underline{\sigma} \leq \min_{(l,u): 1 \leq l < u \leq n} \min_{1 \leq i,j \leq d} \hat{\sigma}_{l,u}(i,j) \leq \max_{(l,u): 1 \leq l < u \leq n} \max_{1 \leq i,j \leq d} \hat{\sigma}_{l,u}(i,j) \leq \bar{\sigma}$$

with probability approaching one, where  $\hat{\sigma}_{l,u}(i,j)$  is defined in (2.15). In addition,  $\hat{\sigma}_{l,u}(i,j) = \hat{\sigma}_{l,u}(j,i)$ .

The assumption of mixing dependence on the latent common factor and idiosyncratic error process in Assumption 1(i) is very mild and covers some commonly-used vector time series models. Note that the  $\alpha$ -mixing dependence condition allows the underlying process to be nonstationary with time-varying covariance structure (e.g., the piece-wise stationary time series process with breaks in the covariance matrix). The moment conditions in Assumptions 1(ii) and 3 are crucial to consistently estimate the factors and factor loadings (with rotation) in the transformed factor model.

Assumption 2 contains some fundamental conditions for the approximate factor model (2.2), which can be seen as a generalised version of those for the stable factor model without breaks (e.g., Bai and Ng, 2002; Fan, Liao and Mincheva, 2013). These conditions, together with  $\kappa_n^c \asymp n$  in Assumption 4(ii), indicate that there exists a nonsingular limiting matrix for  $\frac{1}{n} \sum_{t=1}^n \mathbf{F}_t^* \mathbf{F}_t^{*\top}$  and  $\mathbf{A}^*$  is of full column rank (see Proposition 3.1), which are essential to derive the convergence result for the PCA estimated factors and factor loadings. In addition, Assumption 2 also guarantees that  $\bar{\omega}_n^c = O(1)$ .

The moment conditions (A.2) and (A.3) in Assumption 3(i) imply that the idiosyncratic errors  $\epsilon_{tj}$  are allowed to be weakly dependent over  $j$ . Similar high-level moment conditions can be found in Assumption 4 in Bai and Ng (2002) and Assumption 3.4 in Fan, Liao and Mincheva (2013). Assumption 3(ii) is similar to the exponential-tail condition commonly used in high-dimensional data analysis and is needed to allow the dimension  $d$  to be divergent at an exponential rate of  $n$ .

Assumption 4(i) indicates that both  $d$  and  $n$  diverge to infinity jointly, which is a typical large panel data setting necessary for consistent estimates of the rotated factors and factor loadings. A very similar assumption is also used in Theorem 1 of Fan, Liao and Mincheva (2013). In particular, Assumption 4(i) shows that  $d$  can be divergent to infinity at an exponential rate of  $n$ , further relaxing the corresponding condition assumed in Barigozzi, Cho and Fryzlewicz (2018) and Wang, Yu and Rinaldo (2018). Although the restriction  $n \leq \iota_3 d^{\delta/(\delta+4)}$  indicates that  $n$  is of smaller asymptotic order than  $d$ , our methodology is still applicable if  $n$  is much larger than  $d$  by switching the role of  $n$  and  $d$  and the role of the factor loadings and common factors in the PCA estimation (see Section 3 in Bai and Li, 2012). Assumption 4(ii) is mainly used for the asymptotics of WBS-Cov in the common components, and would be automatically satisfied when  $\iota_4 \leq \underline{\omega}_n^c \leq \bar{\omega}_n^c \leq \iota_5$  and  $M_n^c = \lfloor \log n \rfloor$ , where  $\iota_4 < \iota_5$  are two positive constants. Assumption 4(iii) is mainly used to derive the WSBS-Cov asymptotic theory in the idiosyncratic error components. The restriction  $\bar{\omega}_n^e \asymp \underline{\omega}_n^e$  is only imposed to facilitate the proofs. The restrictions on  $\underline{\omega}_n^c$  and  $\underline{\omega}_n^e$  are consistent with the argument in Andrews (1993) and Bai (1997) that the break points can be identified only if the size of breaks, i.e.,  $(\underline{\omega}_n^c)^{1/2}$  and  $(\underline{\omega}_n^e)^{1/2}$  in our paper, is asymptotically larger than  $n^{-1/2}$ . Assumption 4(iii) also indicates that there might be more frequent breaks in the idiosyncratic error components than in the common components, and the restriction  $\frac{\kappa_n^e \omega_n^e}{\log^4(nd)} \cdot \left(\frac{\kappa_n^e}{n}\right)^4 \rightarrow \infty$  would be automatically satisfied when  $\kappa_n^e \asymp n^\Theta$  and  $(\underline{\omega}_n^e n^{5\Theta-4}) / \log^4(nd) \rightarrow \infty$  with  $\Theta \in (4/5, 1]$ . In contrast, the minimum length  $\kappa_n^c$  of the sub-intervals separated by break points in the common components is of order  $n$ . Assumption 5 requires the random normalisers in WSBS-Cov to be uniformly bounded away from zero and infinity with probability approaching one.

## 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  $K_1 = 2$ :

$$\mathbf{X}_t = \boldsymbol{\Lambda}_{k+1}^0 \mathbf{F}_t + \boldsymbol{\epsilon}_t, \quad \eta_k^c + 1 \leq t \leq \eta_{k+1}^c,$$

where  $k = 0, 1, 2$ ,  $\eta_0^c = 0$  and  $\eta_3^c = n$ . 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  $(\boldsymbol{\Lambda}_1^0, \boldsymbol{\Lambda}_2^0)$  is  $r$ , indicating that there exists an  $r \times r$  matrix  $\mathbf{T}$  such that  $\boldsymbol{\Lambda}_2^0 = \boldsymbol{\Lambda}_1^0 \mathbf{T}$ ; and the column 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”<sup>1</sup>. The transformed factor loadings and factors can be defined as  $\boldsymbol{\Lambda}^* = (\boldsymbol{\Lambda}_1^0, \boldsymbol{\Lambda}_3^0)$  and

$$\mathbf{F}_t^* = \begin{cases} (\mathbf{F}_t^\top, \mathbf{0}^\top)^\top, & 1 \leq t \leq \eta_1^c, \\ (\mathbf{F}_t^\top \mathbf{T}^\top, \mathbf{0}^\top)^\top, & \eta_1^c + 1 \leq t \leq \eta_2^c, \\ (\mathbf{0}^\top, \mathbf{F}_t^\top)^\top, & \eta_2^c + 1 \leq t \leq n, \end{cases}$$

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

$$\mathbf{X}_t = \boldsymbol{\Lambda}^* \mathbf{F}_t^* + \boldsymbol{\epsilon}_t, \quad t = 1, \dots, n, \quad (\text{B.1})$$

the same as (2.6) in Proposition 2.1. Note that the number of latent common factors has increased from  $r$  to  $2r$  in model (B.1). Letting  $\boldsymbol{\Sigma}(\mathbf{F}) = \text{Cov}(\mathbf{F}_t)$ ,  $\boldsymbol{\Sigma}_t(\boldsymbol{\Lambda}, \mathbf{F})$  in (1.3) can be re-formulated as

$$\boldsymbol{\Sigma}_t(\boldsymbol{\Lambda}, \mathbf{F}) = \begin{cases} \boldsymbol{\Lambda}^* \text{diag}\{\boldsymbol{\Sigma}(\mathbf{F}), \mathbf{O}\} (\boldsymbol{\Lambda}^*)^\top, & 1 \leq t \leq \eta_1^c, \\ \boldsymbol{\Lambda}^* \text{diag}\{\mathbf{T} \boldsymbol{\Sigma}(\mathbf{F}) \mathbf{T}^\top, \mathbf{O}\} (\boldsymbol{\Lambda}^*)^\top, & \eta_1^c + 1 \leq t \leq \eta_2^c, \\ \boldsymbol{\Lambda}^* \text{diag}\{\mathbf{O}, \boldsymbol{\Sigma}(\mathbf{F})\} (\boldsymbol{\Lambda}^*)^\top, & \eta_2^c + 1 \leq t \leq n. \end{cases} \quad (\text{B.2})$$

where  $\text{diag}\{\mathbf{A}, \mathbf{B}\}$  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  $\boldsymbol{\Lambda}^*$  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  $\mathbf{F}_t^*$ .

<sup>1</sup>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.

## Appendix C: Proofs of Propositions 2.1 and 3.1

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

$$\max_{1 \leq k \leq K_1+1} \underline{r}_k \leq q_0 \leq \sum_{k=1}^{K_1+1} \underline{r}_k, \quad (\text{C.1})$$

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

$$\Lambda_k^0 = \Lambda^* \mathbf{T}_k, \quad k = 1, \dots, K_1 + 1. \quad (\text{C.2})$$

Then the transformed factors can be defined as

$$\mathbf{F}_t^* = \begin{cases} \mathbf{T}_1 \mathbf{F}_{t,1}, & 1 \leq t \leq \eta_1^c, \\ \mathbf{T}_2 \mathbf{F}_{t,2}, & \eta_1^c + 1 \leq t \leq \eta_2^c, \\ \vdots & \vdots \\ \mathbf{T}_{K_1+1} \mathbf{F}_{t,K_1+1}, & \eta_{K_1}^c + 1 \leq t \leq n. \end{cases} \quad (\text{C.3})$$

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

$$\mathbf{X}_t = \Lambda_k^0 \mathbf{F}_{t,k} + \boldsymbol{\epsilon}_t = \Lambda^* \mathbf{T}_k \mathbf{F}_{t,k} + \boldsymbol{\epsilon}_t = \Lambda^* \mathbf{F}_t^* + \boldsymbol{\epsilon}_t. \quad (\text{C.4})$$

The inequalities in (2.7) can be proved by combining (C.1) and the fact of  $\underline{r}_k \leq r_k$ .  $\square$

PROOF OF PROPOSITION 3.1. Letting  $\mathcal{L}(\Lambda)$  be defined as in the proof of Proposition 2.1, we may obtain a group of basis vectors for  $\mathcal{L}(\Lambda)$  directly from the column vectors of  $\Lambda_k^0$ , for  $k = 1, \dots, K_1 + 1$ . Specifically, define  $\Lambda^* = [\Lambda_1^0, \dots, \Lambda_{K_1+1}^0] \mathbf{S}$ , 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,  $\Lambda^*$  is of full column rank and the smallest eigenvalue of  $\frac{1}{d} \Lambda^{*\top} \Lambda^*$  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

$$\text{tr} \left( \frac{1}{d} \Lambda_k^{0\top} \Lambda_k^0 \right) = \text{tr} \left( \frac{1}{d} \mathbf{T}_k^\top \Lambda^{*\top} \Lambda^* \mathbf{T}_k \right) = \text{tr} \left( \frac{1}{d} \Lambda^{*\top} \Lambda^* \mathbf{T}_k \mathbf{T}_k^\top \right) \geq \sum_{j=1}^{q_0} \mu_j \left( \frac{1}{d} \Lambda^{*\top} \Lambda^* \right) \mu_{q_0-j+1} (\mathbf{T}_k \mathbf{T}_k^\top),$$

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

$$\mu_1(\mathbf{T}_k \mathbf{T}_k^\top) \leq \text{tr} \left( \frac{1}{d} \boldsymbol{\Lambda}_k^{0\top} \boldsymbol{\Lambda}_k^0 \right) / \mu_{q_0} \left( \frac{1}{d} \boldsymbol{\Lambda}^{*\top} \boldsymbol{\Lambda}^* \right),$$

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

$$\max_{1 \leq k \leq K_1 + 1} \|\mathbf{T}_k\|_F^2 = \max_{1 \leq k \leq K_1 + 1} \text{tr}(\mathbf{T}_k \mathbf{T}_k^\top) \leq \left( \max_{1 \leq k \leq K_1 + 1} r_k \right) \cdot \max_{1 \leq k \leq K_1 + 1} \mu_1(\mathbf{T}_k \mathbf{T}_k^\top) \leq M, \quad (\text{C.5})$$

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

$$\begin{aligned} \left\| \frac{1}{n} \sum_{t=1}^n \mathbf{F}_t^* \mathbf{F}_t^{*\top} \right\|_F &\leq \sum_{k=1}^{K_1+1} \left\| \frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t: \eta_{k-1}^c + 1 \leq t \leq \eta_k^c} \mathbf{T}_k \mathbf{F}_{t,k} \mathbf{F}_{t,k}^\top \mathbf{T}_k^\top \right\|_F \\ &\leq \sum_{k=1}^{K_1+1} \mu_1 \left( \frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t: \eta_{k-1}^c + 1 \leq t \leq \eta_k^c} \mathbf{F}_{t,k} \mathbf{F}_{t,k}^\top \right) \cdot \|\mathbf{T}_k\|_F^2. \end{aligned} \quad (\text{C.6})$$

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

$$\frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t: \eta_{k-1}^c + 1 \leq t \leq \eta_k^c} \mathbf{F}_{t,k} \mathbf{F}_{t,k}^\top \xrightarrow{P} \boldsymbol{\Sigma}_{F,k}, \quad k = 1, \dots, K_1 + 1. \quad (\text{C.7})$$

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

From (C.7), we readily have that

$$\frac{1}{n} \sum_{t=1}^n \mathbf{F}_t^* \mathbf{F}_t^{*\top} = \sum_{k=1}^{K_1+1} \frac{\eta_k^c - \eta_{k-1}^c}{n} \cdot \frac{1}{\eta_k^c - \eta_{k-1}^c} \sum_{t: \eta_{k-1}^c + 1 \leq t \leq \eta_k^c} \mathbf{T}_k \mathbf{F}_{t,k} \mathbf{F}_{t,k}^\top \mathbf{T}_k^\top \xrightarrow{P} \boldsymbol{\Sigma}_F, \quad (\text{C.8})$$

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

## 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)*

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

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

$$\max_{1 \leq j \leq d} \left\| \widehat{\lambda}_j - (\mathbf{H}^{-1})^\top \lambda_j^* \right\|_2 = O_P \left( \left( \frac{\log d}{n} \right)^{1/2} + \frac{n^{2/\delta}}{d^{1/2}} \right), \quad (\text{D.2})$$

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

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

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

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

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

$$\frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{F}}_t \widehat{\mathbf{F}}_t^\top = \mathbf{I}_{q_0}, \quad \frac{1}{n} \sum_{t=1}^n \mathbf{F}_t^* \mathbf{F}_t^{*\top} = O_P(1),$$

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

$$\left\| \sum_{s=1}^n \widehat{\mathbf{F}}_s \mathbf{F}_s^{*\top} \right\|_F = O_P(n). \quad (\text{D.4})$$



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

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

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  $\epsilon > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq t \leq n} \left\| \sum_{j=1}^d \lambda_j^* \epsilon_{tj} \right\|_2 > c_1 n^{1/\delta_\epsilon} d^{1/2} \right) &\leq \sum_{t=1}^n \mathbb{P} \left( \left\| \sum_{j=1}^d \lambda_j^* \epsilon_{tj} \right\|_2 > c_1 n^{1/\delta_\epsilon} d^{1/2} \right) \\ &\leq \max_{1 \leq t \leq n} \mathbb{E} \left[ \left\| \sum_{j=1}^d \lambda_j^* \epsilon_{tj} \right\|_2^{\delta_\epsilon} \right] / (c_1^{\delta_\epsilon} d^{\delta_\epsilon/2}) \leq \frac{c_0 \iota_0 (K_1 + 1)}{c_1^{\delta_\epsilon}} < \epsilon \end{aligned} \quad (\text{D.5})$$

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

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

By (C.5) and

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

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

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

By a basic inequality on the covariance bound for the  $\alpha$ -mixing sequence (e.g., Lemma 1.2.4 in [Lin and Lu, 1996](#)), we have

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

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

$$\begin{aligned}
\max_{1 \leq t \leq n} \|\mathbf{V}_{nt,3}\|_2 &= \frac{1}{nd} \cdot \max_{1 \leq t \leq n} \left\| \sum_{s=1}^n \sum_{j=1}^d \widehat{\mathbf{F}}_s \mathbf{E}[\epsilon_{sj} \epsilon_{tj}] \right\|_2 \\
&\leq \frac{1}{nd} \cdot \max_{1 \leq t \leq n} \left( \sum_{s=1}^n \|\widehat{\mathbf{F}}_s\|_2^2 \right)^{1/2} \left( \sum_{s=1}^n \left( \sum_{j=1}^d \mathbf{E}[\epsilon_{sj} \epsilon_{tj}] \right)^2 \right)^{1/2} \\
&= \frac{1}{nd} \cdot O_P(n^{1/2}) \cdot O \left( d \cdot \left[ \sum_{k=1}^n [\alpha(k)]^{2(1-2/\delta_\epsilon)} \right]^{1/2} \right) \\
&= O_P(n^{-1/2})
\end{aligned} \tag{D.8}$$

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  $\epsilon > 0$ ,

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

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

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

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  $\frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{F}}_t \widehat{\mathbf{F}}_t^\top = \mathbf{I}_{q_0}$ , we have

$$\begin{aligned}
\widehat{\lambda}_j &= \frac{1}{n} \sum_{t=1}^n X_{tj} \widehat{\mathbf{F}}_t = \frac{1}{n} \sum_{t=1}^n (\lambda_j^{*\top} \mathbf{F}_t^* + \epsilon_{tj}) \widehat{\mathbf{F}}_t \\
&= \frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{F}}_t \mathbf{F}_t^{*\top} \lambda_j^* + \frac{1}{n} \sum_{t=1}^n \epsilon_{tj} \widehat{\mathbf{F}}_t \\
&= (\mathbf{H}^{-1})^\top \lambda_j^* + \frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{F}}_t (\mathbf{F}_t^* - \mathbf{H}^{-1} \widehat{\mathbf{F}}_t)^\top \lambda_j^* \\
&\quad + \mathbf{H} \cdot \frac{1}{n} \sum_{t=1}^n \epsilon_{tj} \mathbf{F}_t^* + \frac{1}{n} \sum_{t=1}^n \epsilon_{tj} (\widehat{\mathbf{F}}_t - \mathbf{H} \mathbf{F}_t^*). \tag{D.10}
\end{aligned}$$

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

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

and

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

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

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

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

$$\zeta_{tj} = \zeta_{tj} - \mathbf{E}[\zeta_{tj}] = \bar{\zeta}_{tj} - \mathbf{E}[\bar{\zeta}_{tj}] + \tilde{\zeta}_{tj} - \mathbf{E}[\tilde{\zeta}_{tj}],$$

where

$$\bar{\zeta}_{tj} = \zeta_{tj} \cdot \mathcal{J}(\|\zeta_{tj}\|_2 \leq c_3 \log(dn)), \quad \tilde{\zeta}_{tj} = \zeta_{tj} \cdot \mathcal{J}(\|\zeta_{tj}\|_2 > c_3 \log(dn)),$$

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

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

and

$$\max_{1 \leq j \leq d} \left\| \frac{1}{n} \sum_{t=1}^n (\tilde{\zeta}_{tj} - \mathbf{E} [\tilde{\zeta}_{tj}]) \right\|_2 = O_P \left( \sqrt{(\log d)/n} \right). \quad (\text{D.15})$$

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  $\iota_1$ ) such that

$$\max_{1 \leq j \leq d} \max_{1 \leq t \leq n} \mathbf{E} [\exp \{ \iota_1^\diamond \|\epsilon_{tj} \mathbf{F}_t^*\|_2 \}] < \infty.$$

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

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

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

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

leading to (D.15).

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

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

where  $c_M > 0$  is a sufficiently large constant when  $M$  is large enough, completing the

proof of (D.14). □

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

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

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

$$\begin{aligned}
&\max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \sqrt{\frac{u-s}{(u-l+1)(s-l+1)}} \left\| \sum_{t=l}^s \text{vech} \left[ \left( \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right) \left( \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right)^\top \right] \right\|_2 \\
&= \max_{(l,u): 1 \leq l < u \leq n} (u-l+1)^{-1/2} \cdot O_p(n^{-1}) \max_{s: l \leq s < u} \sqrt{(u-s)(s-l+1)} \\
&= \max_{(l,u): 1 \leq l < u \leq n} (u-l+1)^{1/2} \cdot O_p(n^{-1}) = O_p(n^{-1/2}), \tag{D.16}
\end{aligned}$$

and similarly

$$\begin{aligned}
&\max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \sqrt{\frac{s-l+1}{(u-l+1)(u-s)}} \left\| \sum_{t=s+1}^u \text{vech} \left[ \left( \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right) \left( \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right)^\top \right] \right\|_2 \\
&= \max_{(l,u): 1 \leq l < u \leq n} (u-l+1)^{1/2} \cdot O_p(n^{-1}) = O_p(n^{-1/2}). \tag{D.17}
\end{aligned}$$

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

$$\begin{aligned}
&\left\| \sum_{t=l}^s \text{vech} \left[ \left( \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right) \mathbf{F}_t^{*\top} \mathbf{H}^\top \right] \right\|_2 \\
&\leq \left( \sum_{t=l}^s \left\| \widehat{\mathbf{F}}_t - \mathbf{HF}_t^* \right\|_2^2 \right)^{1/2} \left( \sum_{t=l}^s \left\| \mathbf{F}_t^* \right\|_2^2 \right)^{1/2} \cdot O_p(1) \\
&= O_p((s-l+1)/n^{1/2}),
\end{aligned}$$

and similarly

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

Consequently, we have

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

and

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

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

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

$$\begin{aligned} \mathbf{Z}_t^{\mathbf{F}^*} &= \text{vech} \left( \mathbf{F}_t^* \mathbf{F}_t^{*\top} \right) = \left( z_{t,1}^{\mathbf{F}^*}, \dots, z_{t,q_0(q_0+1)/2}^{\mathbf{F}^*} \right)^\top, \\ \mathbf{G}_t^{\mathbf{F}^*} &= \mathbf{E} \left[ \text{vech} \left( \mathbf{F}_t^* \mathbf{F}_t^{*\top} \right) \right] = \left( g_{t,1}^{\mathbf{F}^*}, \dots, g_{t,q_0(q_0+1)/2}^{\mathbf{F}^*} \right)^\top, \\ \mathbf{z}_t^{\mathbf{F}^*} &= \mathbf{Z}_t^{\mathbf{F}^*} - \mathbf{G}_t^{\mathbf{F}^*} = \left( z_{t,1}^{\mathbf{F}^*}, \dots, z_{t,q_0(q_0+1)/2}^{\mathbf{F}^*} \right)^\top. \end{aligned}$$

Define

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

Then

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

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

$$\eta_{k_0}^c \leq l < \eta_{k_0+1}^c < \cdots < \eta_{k_0+k_1}^c < u \leq \eta_{k_0+k_1+1}^c, \quad (\text{D.21})$$

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

$$l < \eta_{k_0+k}^c - c_4 \kappa_n^c < \eta_{k_0+k}^c + c_4 \kappa_n^c < u \quad (\text{D.22})$$

and

$$\{(l - \eta_{k_0}^c) \wedge (\eta_{k_0+1}^c - l)\} \vee \{(u - \eta_{k_0+k_1}^c) \wedge (\eta_{k_0+k_1+1}^c - u)\} \leq c_5 \varphi_n^c, \quad (\text{D.23})$$

where  $c_4$  and  $c_5$  are two positive constants,  $\kappa_n^c$  is defined in Assumption 4(ii), and  $\varphi_n^c$  is defined in Theorem 3.1. Define the intervals

$$\mathcal{J}_k^c = [\eta_{k-1}^c + (\eta_k^c - \eta_{k-1}^c)/3, \eta_{k-1}^c + 2(\eta_k^c - \eta_{k-1}^c)/3], \quad k = 1, \dots, K_1 + 1,$$

and the event

$$\mathcal{D}_n^c = \{\forall k = 1, \dots, K_1, \exists m = 1, \dots, M_n^c \text{ such that } l_m \in \mathcal{J}_k^c \text{ and } u_m \in \mathcal{J}_{k+1}^c\},$$

where  $M_n^c$  is defined in Section 2.3.

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

$$\mathbf{P}(\overline{\mathcal{D}}_n^c) \leq K_1 \left[1 - (\kappa_n^c/(3n))^2\right]^{M_n^c}, \quad (\text{D.24})$$

where  $\kappa_n^c$  is defined in Assumption 4(ii).

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

$$\begin{aligned} \mathbf{P}(\overline{\mathcal{D}}_n^c) &\leq \sum_{k=1}^{K_1} \prod_{m=1}^{M_n^c} [1 - \mathbf{P}(l_m \in \mathcal{J}_k^c \text{ and } u_m \in \mathcal{J}_{k+1}^c)] \\ &\leq K_1 \prod_{m=1}^{M_n^c} \left(1 - \frac{\eta_k^c - \eta_{k-1}^c}{3n} \cdot \frac{\eta_{k+1}^c - \eta_k^c}{3n}\right) \\ &\leq K_1 \left[1 - (\kappa_n^c/(3n))^2\right]^{M_n^c}, \end{aligned} \quad (\text{D.25})$$

completing the proof of Lemma D.2.  $\square$

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

$$\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left\| \mathbf{C}_{l,u}^{z,F^*}(s) \right\|_2 > c_6 \cdot \log^2 n \right) \rightarrow 0, \quad (\text{D.26})$$

as  $n \rightarrow \infty$ .

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

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

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

$$\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left| C_{l,u,k}^{z,F^*}(s) \right| > \frac{2c_6}{q_0(q_0+1)} \cdot \log^2 n \right) \rightarrow 0 \quad (\text{D.27})$$

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

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

and

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

it suffices to prove that

$$\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left| C_{l,u,k}^{z,F^*}(s; j) \right| > \bar{c}(q_0) \cdot \log^2 n \right) \rightarrow 0 \quad (\text{D.28})$$

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

$$\bar{z}_{t,k}^{F^*} = z_{t,k}^{F^*} \cdot \mathcal{J}(|z_{t,k}^{F^*}| \leq c_7 \log n), \quad \tilde{z}_{t,k}^{F^*} = z_{t,k}^{F^*} \cdot \mathcal{J}(|z_{t,k}^{F^*}| > c_7 \log n),$$

where  $c_7 > 0$  is a sufficiently large constant. Letting  $\bar{C}_{l,u,k}^{z,F^*}(s; 1)$  and  $\tilde{C}_{l,u,k}^{z,F^*}(s; 1)$  be defined similarly to  $C_{l,u,k}^{z,F^*}(s; 1)$  but with  $z_{t,k}^{F^*}$  replaced by  $\bar{z}_{t,k}^{F^*} - \mathbf{E}[\bar{z}_{t,k}^{F^*}]$  and  $\tilde{z}_{t,k}^{F^*} - \mathbf{E}[\tilde{z}_{t,k}^{F^*}]$ , respectively. From Assumption 3(ii) and Proposition 3.1, there exists a positive constant  $\iota_6 > 0$



(which may be different from  $\iota_1$ ) such that

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

Consequently, we can show that

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

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

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

We next prove

$$\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left| \bar{\mathbf{C}}_{l,u,k}^{z,F^*}(s;1) \right| > \frac{\bar{c}(q_0)}{2} \cdot \log^2 n \right) \rightarrow 0. \quad (\text{D.30})$$

Consider the following two scenarios: (i)  $s-l+1 \leq c_8 \log^2 n$ , 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

$$\begin{aligned} \left| \bar{\mathbf{C}}_{l,u,k}^{z,F^*}(s;1) \right| &\leq \sqrt{\frac{u-s}{u-l+1}} \cdot \frac{1}{\sqrt{s-l+1}} \cdot \sum_{t=l}^s (|z_{t,k}^{F^*}| + \mathbf{E} [|\tilde{z}_{t,k}^{F^*}|]) \\ &\leq \sqrt{s-l+1} \cdot (2c_7 \log n) \leq (2c_7 c_8) \cdot \log^2 n. \end{aligned}$$

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

$$\begin{aligned} &\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left| \bar{\mathbf{C}}_{l,u,k}^{z,F^*}(s;1) \right| > \frac{\bar{c}(q_0)}{2} \cdot \log^2 n \right) \\ &\leq \mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l+c_8 \log^2 n - 1 \leq s < u} \left| \bar{\mathbf{C}}_{l,u,k}^{z,F^*}(s;1) \right| > \left[ \frac{\bar{c}(q_0)}{2} - 2c_7 c_8 \right] \cdot \log^2 n \right) \\ &\leq O(n^3 \exp\{-M \log n\} + n^{3+3/4} \rho^{\sqrt{c_8} \log n}) = o(1), \end{aligned}$$

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.  $\square$

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*

$$\mathbf{P} \left( \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\hat{\mathbf{F}}}(\mathbf{s}_0^c) \right\|_2 \geq c_9 \cdot (\kappa_n^c \underline{\omega}_n^c)^{1/2} \right) \rightarrow 1 \quad (\text{D.31})$$

as  $n \rightarrow \infty$ , 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  $\mathbf{F}_t^*$  given in (C.3), we readily have the following time-varying covariance structure for  $\mathbf{F}_t^*$ :

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

Hence, we have

$$\boldsymbol{\Sigma}_{k+1}^0(\boldsymbol{\Lambda}, \mathbf{F}) - \boldsymbol{\Sigma}_k^0(\boldsymbol{\Lambda}, \mathbf{F}) = \boldsymbol{\Lambda}^* [\boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^*) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^*)] \boldsymbol{\Lambda}^{*\top},$$

indicating that

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

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

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

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

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

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  $\eta_k^c$  be one of these break points within  $[l, u]$  satisfying  $l + c_5 \varphi_n^c < \eta_k^c < u - c_5 \varphi_n^c$ . On the set  $\mathcal{D}_n^c$ , there exists  $1 \leq m_k \leq M_n^c$  such that  $l_{m_k} \in \mathcal{J}_k^c$  and  $u_{m_k} \in \mathcal{J}_{k+1}^c$ , indicating that both  $\eta_k^c - l_{m_k}$  and  $u_{m_k} - \eta_k^c$  are larger than  $\kappa_n^c / 3$ . Define

$$\boldsymbol{\omega}_k^{\mathbf{F}^{\star}} = \text{vech} \left( \boldsymbol{\Sigma}_{k+1}^0(\mathbf{F}^{\star}) - \boldsymbol{\Sigma}_k^0(\mathbf{F}^{\star}) \right) =: \left( \omega_{k,1}^{\mathbf{F}^{\star}}, \dots, \omega_{k, q_0(q_0+1)/2}^{\mathbf{F}^{\star}} \right)^{\top}. \quad (\text{D.33})$$

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

$$\left| \mathbf{C}_{l_{m_k}, u_{m_k}, i}^{\mathbf{G}, \mathbf{F}^{\star}}(\eta_k^c) \right| = \sqrt{\frac{(\eta_k^c - l_{m_k} + 1)(u_{m_k} - \eta_k^c)}{u_{m_k} - l_{m_k} + 1}} \left| \omega_{k,i}^{\mathbf{F}^{\star}} \right| \geq \left( \frac{\kappa_n^c}{6} \right)^{1/2} \left| \omega_{k,i}^{\mathbf{F}^{\star}} \right|, \quad (\text{D.34})$$

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

$$\left\| \mathbf{C}_{l_{m_k}, u_{m_k}}^{\mathbf{G}, \mathbf{F}^{\star}}(\eta_k^c) \right\|_2 \geq c_{11} (\kappa_n^c)^{1/2} \left\| \boldsymbol{\omega}_k^{\mathbf{F}^{\star}} \right\|_2, \quad (\text{D.35})$$

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

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

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

$$\left\| \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \cdot \mathbf{C}_{l_{m_k}, u_{m_k}}^{\mathbf{G}, \mathbf{F}^{\star}}(\eta_k^c) \right\|_2 \geq 2c_9 (\kappa_n^c \underline{\omega}_n^c)^{1/2}. \quad (\text{D.36})$$

By the definitions of  $m_0^c$  and  $s_0^c$  in Algorithm 2 and using Proposition 3.2 and Lemma D.3, we may show that conditional on that  $\mathbf{H}$  is non-singular,

$$\begin{aligned} \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\hat{\mathbf{F}}} (s_0^c) \right\|_2 &\geq \left\| \mathbf{C}_{l_{m_k}, u_{m_k}}^{\hat{\mathbf{F}}} (\eta_k^c) \right\|_2 \\ &= \left\| \mathbf{C}_{l_{m_k}, u_{m_k}}^{\mathbf{H}\mathbf{F}^{\star}} (\eta_k^c) \right\|_2 + O_P(1) \\ &= \left\| \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \cdot \mathbf{C}_{l_{m_k}, u_{m_k}}^{\mathbf{G}, \mathbf{F}^{\star}} (\eta_k^c) \right\|_2 + O_P(\log^2 n) \end{aligned}$$

$$\geq 2c_9 (\kappa_n^c \underline{\omega}_n^c)^{1/2} + O_p(\log^2 n).$$

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

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

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

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*

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

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

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

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

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

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

As both  $G^*(x)$  and  $-\alpha[x(1-x)]^\gamma$  are convex, we readily prove that  $\mathcal{S}_\alpha$  is a convex set. Choosing  $G^*(x) = \sum_{i=1}^m |a_i x + b_i|^p$  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  $(\sum_{i=1}^m |a_i x + b_i|^p)^{1/p} / [x(1-x)]^{\gamma/p}$  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  $\mathbf{Z}_t^{\mathbf{F}^*}$ ,  $\mathbf{G}_t^{\mathbf{F}^*}$  and  $\mathbf{z}_t^{\mathbf{F}^*}$ , we define

$$\begin{aligned}\mathbf{Z}_t^{\mathbf{HF}^*} &= \text{vech}(\mathbf{HF}_t^{\mathbf{F}^*} \mathbf{F}_t^{\mathbf{F}^* \top} \mathbf{H}) = \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \text{vech}(\mathbf{F}_t^{\mathbf{F}^*} \mathbf{F}_t^{\mathbf{F}^* \top}) = \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \mathbf{Z}_t^{\mathbf{F}^*}, \\ \mathbf{G}_t^{\mathbf{HF}^*} &= \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \mathbf{E}[\mathbf{Z}_t^{\mathbf{F}^*}] = \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \mathbf{G}_t^{\mathbf{F}^*}, \\ \mathbf{z}_t^{\mathbf{HF}^*} &= \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} (\mathbf{Z}_t^{\mathbf{F}^*} - \mathbf{G}_t^{\mathbf{F}^*}) = \mathbf{L}_{q_0} (\mathbf{H} \otimes \mathbf{H}) \mathbf{D}_{q_0} \mathbf{z}_t^{\mathbf{F}^*},\end{aligned}$$

and then

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

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  $v$  with appropriate subscript to highlight the difference and similarity between Lemma 2.6 in Venkatraman (1992) and our lemma. For example,  $v_h, v_i, v_j$  and  $v_l$  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  $s_*^c \in [l_{m_\xi^c}, u_{m_\xi^c}]$  be the point of maximising  $\|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(s)\|_2$  with respect to  $s$ , i.e.,*

$$s_*^c = \arg \max_{l_{m_\xi^c} \leq s < u_{m_\xi^c}} \|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(s)\|_2, \quad (\text{D.38})$$

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

$$\|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(\eta_{k_\circ}^c)\|_2 > \|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(s_*^c)\|_2 - 3c_6 \log^2 n, \quad (\text{D.39})$$

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

$$(\eta_{k_\circ}^c - l_{m_\xi^c} + 1) \wedge (u_{m_\xi^c} - \eta_{k_\circ}^c) \geq c_{12} \kappa_n^c, \quad (\text{D.40})$$

and we further have

$$\|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(\eta_{k_\circ}^c)\|_2 > \|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(\eta_{k_\circ}^c + v_l)\|_2 + (c_{13} v_l \kappa_n^c) \cdot \frac{\|\mathbf{C}_{l_{m_\xi^c}, u_{m_\xi^c}}^{\mathbf{G},\mathbf{HF}^*}(\eta_{k_\circ}^c)\|_2}{(u_{m_\xi^c} - l_{m_\xi^c} + 1)^2}, \quad (\text{D.41})$$

where  $0 < v_l < c_{14} \gamma_n^c$  with  $\gamma_n^c = (\kappa_n^c / \underline{\omega}_n^c)^{1/2} \log^2 n$ , 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

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

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

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

where

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

If  $(\eta_{k_*}^c - l + 1) \wedge (u - \eta_{k_*}^c) \leq c_5 \varphi_n^c$  holds, we have  $(\eta_{k_*}^c - l_{m_0^c} + 1) \wedge (u_{m_0^c} - \eta_{k_*}^c) \leq c_5 \varphi_n^c$  as  $[l_{m_0^c}, u_{m_0^c}]$  is a random sub-interval of  $[l, u]$ . By Assumption 4(ii), we have

$$b_{l_{m_0^c}, u_{m_0^c}} \leq c_{15} \left( \overline{\omega}_{l_{m_0^c}, u_{m_0^c}}^c \right)^{1/2} \leq c_{15} \left( \overline{\omega}_{l,u}^c \right)^{1/2} \leq c_{15} \left( \overline{\omega}_n^c \right)^{1/2}, \quad (\text{D.43})$$

where  $c_{15}$  is a positive constant and

$$\overline{\omega}_{l,u}^c = \frac{1}{d^2} \cdot \max_{k: l+c_5 \varphi_n^c \leq \eta_k^c \leq u-c_5 \varphi_n^c} \left\| \Sigma_{k+1}^0(\boldsymbol{\Lambda}, \mathbf{F}) - \Sigma_k^0(\boldsymbol{\Lambda}, \mathbf{F}) \right\|_F^2. \quad (\text{D.44})$$

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

$$\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G},\text{HF}^*}(\eta_{k_*}^c) \right\|_2 \leq 2b_{l_{m_0^c}, u_{m_0^c}} (c_5 \varphi_n^c)^{1/2} \leq 2c_{15} (c_5 \overline{\omega}_n^c \varphi_n^c)^{1/2}. \quad (\text{D.45})$$

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

$$(\eta_{k_*}^c - l + 1) \wedge (u - \eta_{k_*}^c) \geq \kappa_n^c - c_5 \varphi_n^c, \quad (\text{D.46})$$

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

We next strengthen (D.46) to

$$(\eta_{k_*}^c - l_{m_*^c} + 1) \wedge (u_{m_*^c} - \eta_{k_*}^c) \geq c_{12} \kappa_n^c. \quad (\text{D.47})$$

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

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

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

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

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  $l_m \in \mathcal{J}_{k_0+k_1}^c$  and  $u_m \in \mathcal{J}_{k_0+k_1+1}^c$ .

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

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

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

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

Letting  $c_*$  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  $\eta_{k_0+k_1}^c \leq u_{m_0^c} \leq u \leq \eta_{k_0+k_1}^c + c_5 \varphi_n^c$ , by the triangle inequality, we have

$$\begin{aligned} & \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_*}^c) \right\|_2 \\ &= \sqrt{\frac{(\eta_{k_*}^c - l_{m_0^c} + 1)(u_{m_0^c} - \eta_{k_*}^c)}{u_{m_0^c} - l_{m_0^c} + 1}} \left\| \frac{1}{\eta_{k_*}^c - l_{m_0^c} + 1} \sum_{t=l_{m_0^c}}^{\eta_{k_*}^c} \mathbf{G}_t^{\text{HF}^*} - \frac{1}{u_{m_0^c} - \eta_{k_*}^c} \sum_{t=\eta_{k_*}^c+1}^{u_{m_0^c}} \mathbf{G}_t^{\text{HF}^*} \right\|_2 \\ &\leq \sqrt{\frac{(\eta_{k_*}^c - l_{m_0^c} + 1)(u_{m_0^c} - \eta_{k_*}^c)}{u_{m_0^c} - l_{m_0^c} + 1}} \cdot \left\| \frac{1}{\eta_{k_*}^c - l_{m_0^c} + 1} \sum_{t=l_{m_0^c}}^{\eta_{k_*}^c} \mathbf{G}_t^{\text{HF}^*} - \frac{1}{u_{m_0^c} - \eta_{k_*}^c} \sum_{t=\eta_{k_*}^c+1}^{u_{m_0^c}} \mathbf{G}_{t \wedge (u_{m_0^c} - c_5 \varphi_n^c)}^{\text{HF}^*} \right\|_2 \\ &\quad + \sqrt{\frac{(\eta_{k_*}^c - l_{m_0^c} + 1)(u_{m_0^c} - \eta_{k_*}^c)}{u_{m_0^c} - l_{m_0^c} + 1}} \cdot \frac{c_5 \varphi_n^c b_{u_{m_0^c} - c_5 \varphi_n^c, u_{m_0^c}}}{u_{m_0^c} - \eta_{k_*}^c} \\ &\leq \left( 2b_{l_{m_0^c}, u_{m_0^c} - c_5 \varphi_n^c} + c_5 \varphi_n^c b_{u_{m_0^c} - c_5 \varphi_n^c, u_{m_0^c}} / \kappa_n^c \right) \sqrt{(\eta_{k_*}^c - l_{m_0^c} + 1) \wedge (u_{m_0^c} - \eta_{k_*}^c)}. \end{aligned} \quad (\text{D.51})$$

Noting that

$$\varphi_n^c b_{u_{m_0^c} - c_5 \varphi_n^c, u_{m_0^c}} / \kappa_n^c = O\left(\overline{\omega}_n^c\right)^{1/2} \log^4 n / (\underline{\omega}_n^c \kappa_n^c) \quad \text{and} \quad 2b_{l_{m_0^c}, u_{m_0^c} - c_5 \varphi_n^c} \geq (\underline{\omega}_n^c)^{1/2},$$

as  $\left(\frac{\omega_n^c}{\bar{\omega}_n^c}\right)^{1/2} \cdot \frac{\kappa_n^c \omega_n^c}{\log^4 n} \rightarrow \infty$  from assumption 4(ii), we have

$$\varphi_n^c \mathbf{b}_{\mathbf{u}_{m_0^c} - c_5 \varphi_n^c, \mathbf{u}_{m_0^c}} / \kappa_n^c = o\left(\mathbf{b}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}\right),$$

which, together with (D.51), indicates that (D.50) holds as well. However, by letting  $c_*$  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)  $\mathbf{u}_{m_0^c} \leq \eta_{k_0+1}^c$  and (ii)  $\eta_{k_0+1}^c < \mathbf{u}_{m_0^c}$ . We start with case (i) of  $\mathbf{u}_{m_0^c} \leq \eta_{k_0+1}^c$ . For notational simplicity, we let  $v_i = \eta_{k_0}^c - \mathbf{l}_{m_0^c} + 1$  and  $v_h = \mathbf{u}_{m_0^c} - \eta_{k_0}^c$ , and define  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{q_0(q_0+1)/2})^\top$  with

$$\beta_k = \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}, k}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \left( \frac{v_i v_h}{v_i + v_h} \right)^{1/2},$$

where  $\mathbf{C}_{\mathbf{l}, \mathbf{u}, k}^{\mathbf{G}, \mathbf{HF}^*}(\cdot)$  is the  $k$ -th element of  $\mathbf{C}_{\mathbf{l}, \mathbf{u}, k}^{\mathbf{G}, \mathbf{HF}^*}(\cdot)$ . As  $\mathbf{u}_{m_0^c} \leq \eta_{k_0+1}^c$  in this case, it is easy to verify that

$$\mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}, k}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) = \beta_k \left( \frac{v_i v_h}{v_i + v_h} \right)^{-1/2}, \quad \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}, k}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c + v_l) = \beta_k \cdot \frac{v_h - v_l}{v_h} \cdot \left[ \frac{v_i + v_h}{(v_i + v_l)(v_h - v_l)} \right]^{1/2}$$

and

$$\left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \right\|_2 = \|\boldsymbol{\beta}\|_2 \left( \frac{v_i v_h}{v_i + v_h} \right)^{-1/2}.$$

As  $(\kappa_n^c \omega_n^c) / \log^4 n \rightarrow \infty$  by Assumption 4(ii), we have  $\gamma_n^c = o(\kappa_n^c)$  and consequently  $v_l < v_i$  when  $n$  is large enough. Hence, we have

$$\begin{aligned} & \left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \right\|_2 - \left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c + v_l) \right\|_2 \\ &= \|\boldsymbol{\beta}\|_2 \cdot \frac{\sqrt{v_i + v_h}}{v_h} \left( \sqrt{\frac{v_h}{v_i}} - \sqrt{\frac{v_h - v_l}{v_i + v_l}} \right) \\ &= v_l \left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \right\|_2 \cdot \frac{v_i + v_h}{\sqrt{v_h} \sqrt{v_i + v_l} (\sqrt{v_h} \sqrt{v_i + v_l} + \sqrt{v_h - v_l} \sqrt{v_i})} \\ &\geq v_l \left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \right\|_2 \cdot \frac{v_i + v_h}{\sqrt{v_h} \sqrt{2v_i} (\sqrt{v_h} \sqrt{2v_i} + \sqrt{v_h} \sqrt{v_i})} \\ &\geq v_l \left\| \mathbf{C}_{\mathbf{l}_{m_0^c}, \mathbf{u}_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_0}^c) \right\|_2 / [2(\mathbf{u}_{m_0^c} - \mathbf{l}_{m_0^c} + 1)], \end{aligned} \tag{D.52}$$

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

We next consider case (ii). Let  $v_i = \eta_{k_0}^c - \mathbf{l}_{m_0^c} + 1$ ,  $v_h = (c_{12} \wedge 1) \kappa_n^c / 3$ ,  $v_j = \mathbf{u}_{m_0^c} - \eta_{k_0}^c - v_h$



and

$$\mathbf{V}_G^c = \mathbf{G}_{\eta_{k_\circ}^c + 1}^{\text{HF}^*} - \frac{1}{\mathbf{u}_{m_\circ^c} - \mathbf{l}_{m_\circ^c} + 1} \sum_{t=\mathbf{l}_{m_\circ^c}}^{\mathbf{u}_{m_\circ^c}} \mathbf{G}_t^{\text{HF}^*}.$$

From the condition  $(\kappa_n^c \underline{\omega}_n^c) / \log^4 n \rightarrow \infty$ , we may show that  $0 \leq v_l \leq v_h$  when  $n$  is sufficiently large. Then, using the definitions of  $v_i, v_h, v_j$  and  $\mathbf{V}_G^c$ , we readily have that

$$\left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c + v_l) \right\|_2 = \|\boldsymbol{\beta} + v_l \mathbf{V}_G^c\|_2 \cdot \left[ \frac{v_i + v_j + v_h}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2},$$

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

$$D(v_l) = \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c) \right\|_2 - \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c + v_l) \right\|_2$$

and

$$D_1 = \left[ \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c) \right\|_2 - \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c + v_h) \right\|_2 \right] \cdot \frac{v_l}{v_h} \cdot \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2}.$$

Note that

$$\begin{aligned} D(v_l) - D_1 &= \left\{ 1 - \frac{v_l}{v_h} \cdot \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} \right\} \cdot \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c) \right\|_2 \\ &\quad - \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c + v_l) \right\|_2 + \frac{v_l}{v_h} \cdot \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c + v_h) \right\|_2 \\ &= \left\{ 1 - \frac{v_l}{v_h} \cdot \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} \right\} \cdot \left\| \mathbf{C}_{\mathbf{l}_{m_\circ^c}, \mathbf{u}_{m_\circ^c}}^{\text{G, HF}^*}(\eta_{k_\circ}^c) \right\|_2 \\ &\quad - \left[ \frac{v_i + v_j + v_h}{v_h^2 (v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} (\|v_h \boldsymbol{\beta} + v_h v_l \mathbf{V}_G^c\|_2 - \|v_l \boldsymbol{\beta} + v_h v_l \mathbf{V}_G^c\|_2) \\ &\geq \left\{ 1 - \frac{v_l}{v_h} \cdot \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} \right\} \cdot \left[ \frac{v_i + v_j + v_h}{v_i (v_j + v_h)} \right]^{1/2} \cdot \|\boldsymbol{\beta}\|_2 \\ &\quad - \left[ \frac{(v_h - v_l)^2 (v_i + v_j + v_h)}{v_h^2 (v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2} \cdot \|\boldsymbol{\beta}\|_2 \\ &= D_2 \times (1 + D_3), \end{aligned} \tag{D.53}$$

where

$$D_2 = \frac{\|\boldsymbol{\beta}\|_2 v_l (v_h - v_l) \sqrt{v_i + v_j + v_h}}{\sqrt{v_i (v_j + v_h)} \sqrt{(v_i + v_l)(v_j + v_h - v_l)} (\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{v_i (v_j + v_h)})},$$

and

$$D_3 = \frac{(v_j - v_i)(v_j - v_i - v_l)}{(\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{(v_i + v_h)v_j})(\sqrt{v_i(v_j + v_h)} + \sqrt{(v_i + v_h)v_j})}.$$

As  $v_l$  is smaller than  $v_h/2$  for  $n$  large enough, we have

$$\begin{aligned} D_2 &= \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c) \right\|_2 \frac{v_l(v_h - v_l)}{\sqrt{(v_i + v_l)(v_j + v_h - v_l)} [\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{v_i(v_j + v_h)}]} \\ &\geq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c) \right\|_2 \frac{v_l v_h}{2\sqrt{2v_i(v_j + v_h)} [\sqrt{2v_i(v_j + v_h)} + \sqrt{v_i(v_j + v_h)}]} \\ &\geq 2(c_{13}v_l\kappa_n^c) \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c) \right\|_2 / (u_{m_0^c} - l_{m_0^c} + 1)^2. \end{aligned} \quad (\text{D.54})$$

On the other hand, since  $(v_j - v_i)(v_j - v_i - v_l)$  reaches its minimum at  $v_j - v_i = v_l/2$ ,  $v_i, v_j, v_h \geq (c_{12} \wedge 1)\kappa_n^c/3$  and  $v_l = o(\kappa_n^c)$  for  $0 < v_l < c_{14}\gamma_n^c$ , we have

$$\begin{aligned} D_3 &\geq \frac{-v_l^2}{4 [\sqrt{v_i(v_j + v_h)}/2 + \sqrt{(v_i + v_h)v_j}] [\sqrt{v_i(v_j + v_h)} + \sqrt{(v_i + v_h)v_j}]} \\ &\geq \frac{-v_l^2}{4(1 + \sqrt{2})(\sqrt{2} + \sqrt{2})[(c_{12} \wedge 1)\kappa_n^c/3]^2} \rightarrow 0 \end{aligned} \quad (\text{D.55})$$

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

$$\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c + v_h) \right\|_2 > -3c_6 \log^2 n,$$

and

$$\frac{\log^2 n}{\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \text{HF}^*}(\eta_{k_\diamond}^c) \right\|_2} \rightarrow 0$$

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

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

$$|s_0^c - \eta_{k_\bullet}^c| \leq c_{14}\gamma_n^c \quad (\text{D.56})$$

with probability approaching one, as  $n \rightarrow \infty$ , 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  $\eta_{\bar{k}}^c \in [l_{m_0^c}, u_{m_0^c}]$ ,

$$\begin{aligned} \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{\bar{k}}^c) \right\|_2 &\leq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\hat{\mathbf{F}}}(\eta_{\bar{k}}^c) \right\|_2 + (1 + \tau/2)c_6 \log^2 n \\ &\leq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\hat{\mathbf{F}}}(s_0^c) \right\|_2 + (1 + \tau/2)c_6 \log^2 n \\ &\leq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s_0^c) \right\|_2 + (2 + \tau)c_6 \log^2 n \end{aligned} \quad (\text{D.57})$$

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

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

$$\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_\bullet}^c) \right\|_2 \geq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s_0^c) \right\|_2 > \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{k_\star}^c) \right\|_2 - 3c_6 \log^2 n,$$

where  $k_\star$  is defined as in the proof of Lemma D.6. This indicates that (D.39) is satisfied with  $k_\circ = k_\bullet$  and  $s_\star^c = \eta_{k_\star}^c$ . 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  $s_1 \in (\eta_{k_\bullet}^c, \eta_{k_\bullet}^c + c_{14}\gamma_n^c]$  such that

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

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

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

and thus

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

leading to a contradiction with (D.57).

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

$$\left\{ \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{\bar{k}}^c) \right\|_2 \vee \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(\eta_{\bar{k}+1}^c) \right\|_2 \right\} \geq \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s_0^c) \right\|_2.$$

We further consider two scenarios: (ii.1)  $\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s) \right\|_2$  locally decreases at the point  $s = s_0^c$ ; and (ii.2)  $\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s) \right\|_2$  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  $\eta_{k_\bullet}^c = \eta_{\bar{k}+1}^c$ ). When  $\left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\mathbf{G}, \mathbf{HF}^*}(s) \right\|_2$  locally decreases at the point  $s = s_0^c$ , we let  $\eta_{k_\bullet}^c = \eta_{\bar{k}}^c$ . 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.  $\square$

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

$$\begin{aligned} \mathbf{Z}_{\bullet, k}^{\mathbf{HF}^*} &= \left( Z_{l_{m_0^c}, k}^{\mathbf{HF}^*}, \dots, Z_{u_{m_0^c}, k}^{\mathbf{HF}^*} \right)^\top, \\ \mathbf{G}_{\bullet, k}^{\mathbf{HF}^*} &= \left( G_{l_{m_0^c}, k}^{\mathbf{HF}^*}, \dots, G_{u_{m_0^c}, k}^{\mathbf{HF}^*} \right)^\top, \\ \mathbf{z}_{\bullet, k}^{\mathbf{HF}^*} &= \left( z_{l_{m_0^c}, k}^{\mathbf{HF}^*}, \dots, z_{u_{m_0^c}, k}^{\mathbf{HF}^*} \right)^\top, \end{aligned}$$

where  $Z_{t, k}^{\mathbf{HF}^*}$ ,  $G_{t, k}^{\mathbf{HF}^*}$  and  $z_{t, k}^{\mathbf{HF}^*}$  are the  $k$ -th element in the vectors  $\mathbf{Z}_t^{\mathbf{HF}^*}$ ,  $\mathbf{G}_t^{\mathbf{HF}^*}$  and  $\mathbf{z}_t^{\mathbf{HF}^*}$ , 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*

$$|s_0^c - \eta_{k_\bullet}^c| \leq c_{16} \varphi_n^c \quad (\text{D.58})$$

as  $n \rightarrow \infty$ , where  $c_{16}$  is a positive constant and  $\varphi_n^c = \log^4 n / \underline{\omega}_n^c$ .

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

$$\langle \mathbf{v} - \bar{\mathbf{v}}^s, \mathbf{v} - \bar{\mathbf{v}}^s \rangle = \langle \mathbf{v} - \bar{\mathbf{v}}, \mathbf{v} - \bar{\mathbf{v}} \rangle - \langle \mathbf{v} - \bar{\mathbf{v}}, \boldsymbol{\psi}_{l, u}^s \rangle^2, \quad (\text{D.59})$$

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

$$|\langle \mathbf{v}, \boldsymbol{\psi}_{l, u}^s \rangle|^2 = |\langle \mathbf{v} - \bar{\mathbf{v}}, \boldsymbol{\psi}_{l, u}^s \rangle|^2 = -\|\mathbf{v} - \bar{\mathbf{v}}^s\|_2^2 + \|\mathbf{v} - \bar{\mathbf{v}}\|_2^2. \quad (\text{D.60})$$

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

$$\|\mathbf{v} - \bar{\mathbf{v}}^s\|_2^2 \leq \|\mathbf{v} - \bar{\boldsymbol{\omega}}^s\|_2^2, \quad (\text{D.61})$$

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

$$\|\mathbf{v} - \bar{\boldsymbol{\omega}}^s\|_2^2 - \|\mathbf{v} - \bar{\mathbf{v}}^s\|_2^2$$

$$\begin{aligned}
&= \|\mathbf{v} - \bar{\boldsymbol{\omega}} + \langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle \boldsymbol{\psi}_{l,u}^s\|_2^2 - \|\mathbf{v} - \bar{\mathbf{v}}\|_2^2 + \langle \mathbf{v} - \bar{\mathbf{v}}, \boldsymbol{\psi}_{l,u}^s \rangle^2 \\
&= \|\mathbf{v} - \bar{\boldsymbol{\omega}}\|_2^2 + \langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle^2 + 2 \langle \boldsymbol{\omega} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle \langle \mathbf{v} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle - \|\mathbf{v} - \bar{\mathbf{v}}\|_2^2 + \langle \mathbf{v} - \bar{\mathbf{v}}, \boldsymbol{\psi}_{l,u}^s \rangle^2 \\
&= \|\mathbf{v} - \bar{\boldsymbol{\omega}}\|_2^2 - \|\mathbf{v} - \bar{\mathbf{v}}\|_2^2 + \langle \mathbf{v} + \boldsymbol{\omega} - \bar{\mathbf{v}} - \bar{\boldsymbol{\omega}}, \boldsymbol{\psi}_{l,u}^s \rangle^2 \geq 0
\end{aligned}$$

since  $\|\mathbf{v} - \bar{\boldsymbol{\omega}}\|_2^2 \geq \|\mathbf{v} - \bar{\mathbf{v}}\|_2^2$ .

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

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

where  $\bar{\mathbf{Z}}_{\bullet,k}^{\text{HF}^*}$  is defined as  $\bar{\mathbf{v}}$  but with  $\mathbf{v}$  replaced by  $\mathbf{Z}_{\bullet,k}^{\text{HF}^*}$ . For  $l_m^c \leq s < u_m^c$ , define

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

By (D.61), we may show that

$$Q_k^{\text{HF}^*}(s; 1) \geq Q_k^{\text{HF}^*}(s; 2), \quad k = 1, \dots, q_0(q_0 + 1)/2. \quad (\text{D.62})$$

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

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

and

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

Letting

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

by (D.61), we have

$$Q_k^{\text{HF}^*}(s; 3) \geq Q_k^{\text{HF}^*}(s; 1) \geq 0. \quad (\text{D.63})$$

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

$$\sum_{k=1}^{q_0(q_0+1)/2} [Q_k^{\text{HF}^*}(s_0^c; 3) - Q_k^{\text{HF}^*}(\eta_{\mathbf{k}}^c; 2)] \geq -c_{17} \quad (\text{D.64})$$

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

$$\begin{aligned} \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(s_0^c;3) &\geq \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(s_0^c;1) = \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\widehat{F}_t}(s_0^c;1) + O_P(1) \\ &\geq \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\widehat{F}_t}(\eta_{k,\bullet}^c;1) + O_P(1) = \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(\eta_{k,\bullet}^c;1) + O_P(1) \\ &\geq \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(\eta_{k,\bullet}^c;2) + O_P(1), \end{aligned}$$

proving (D.64).

Letting  $c_{16} > 0$  be sufficiently large, we next show that the assertion of  $|s_0^c - \eta_{k,\bullet}^c| > c_{16}\varphi_n^c$  would lead to a contradiction with (D.64), which consequently proves (D.58). Defining

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

we have

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

We next show that with probability approaching one,

$$\sum_{k=1}^{q_0(q_0+1)/2} \left| \left\langle \mathbf{z}_{\bullet,k}^{\text{HF}^*}, \bar{\mathbf{z}}_{\bullet,k}^{\text{HF}^*s_0^c} - \bar{\mathbf{G}}_{\bullet,k}^{\text{HF}^*\eta_{k,\bullet}^c} \right\rangle \right| \leq c_{18}(\log^2 n) \max \left\{ \frac{|s_0^c - \eta_{k,\bullet}^c| \cdot (\bar{\omega}_n^c)^{1/2}}{(\kappa_n^c)^{1/2}}, |s_0^c - \eta_{k,\bullet}^c|^{1/2} (\bar{\omega}_n^c)^{1/2}, \log^2 n \right\} \quad (\text{D.66})$$

and

$$\sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(\eta_{k,\bullet}^c;4) - \sum_{k=1}^{q_0(q_0+1)/2} Q_k^{\text{HF}^*}(s_0^c;4) \geq c_{19} |s_0^c - \eta_{k,\bullet}^c| \bar{\omega}_n^c, \quad (\text{D.67})$$

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

Without loss of generality, we assume that  $s_0^c \geq \eta_{k,\bullet}^c$ . Note that the left hand side of

(D.66) can be decomposed as follows:

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

Following standard calculations, we have

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

for any  $s$ , where  $\bar{z}_{t,k}^{\text{HF}^*s}$  and  $\bar{G}_{t,k}^{\text{HF}^*s}$  are the  $(t - l_{m_0^c} + 1)$ -th element in  $\bar{\mathbf{z}}_{\bullet,k}^{\text{HF}^*s}$  and  $\bar{\mathbf{G}}_{\bullet,k}^{\text{HF}^*s}$ , respectively. By the definition of  $z_{t,k}^{\text{HF}^*}$ , 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  $s$

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

which indicates that

$$\left\langle \mathbf{z}_{\bullet,k}^{\text{HF}^*}, \bar{\mathbf{z}}_{\bullet,k}^{\text{HF}^*s_0^c} - \bar{\mathbf{G}}_{\bullet,k}^{\text{HF}^*s_0^c} \right\rangle = O_P(\log^4 n) \quad (\text{D.70})$$

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

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

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
&\sum_{k=1}^{q_0(q_0+1)/2} [Q_k^{\text{HF}^*}(s_0^c; 3) - Q_k^{\text{HF}^*}(\eta_{k_\bullet}^c; 2)] \\
&\leq c_{18} \log^2 n \max \left\{ \frac{|s_0^c - \eta_{k_\bullet}^c| \cdot (\bar{\omega}_n^c)^{1/2}}{(\kappa_n^c)^{1/2}}, |s_0^c - \eta_{k_\bullet}^c|^{1/2} (\bar{\omega}_n^c)^{1/2}, \log^2 n \right\} - c_{19} |s_0^c - \eta_{k_\bullet}^c| \bar{\omega}_n^c
\end{aligned}$$



$$\leq -c_{17} \log^4 n < -c_{17}, \quad (\text{D.75})$$

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

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  $1 \leq k \leq K_1$  such that  $\eta_k^c < l < u \leq \eta_{k+1}^c$ ; (ii) there exists  $1 \leq k \leq K_1$  such that  $l \leq \eta_k^c \leq u$  and  $(\eta_k^c - l + 1) \wedge (u - \eta_k^c) \leq c_{16} \varphi_n^c$ ; (iii) there exists  $1 \leq k \leq K_1$  such that  $l \leq \eta_k^c < \eta_{k+1}^c \leq u$  and  $(\eta_k^c - l + 1) \vee (u - \eta_{k+1}^c) \leq c_{16} \varphi_n^c$ , 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

$$\mathbb{P} \left( \max_{l_{m_0^c} \leq s < u_{m_0^c}} \left\| \mathbf{C}_{l_{m_0^c}, u_{m_0^c}}^{\hat{\mathbf{F}}}(s) \right\|_2 \leq c_{20} \cdot \log^2 n \right) \rightarrow 1 \quad (\text{D.76})$$

as  $n \rightarrow \infty$ . 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.  $\square$

## 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

$$\max_{1 \leq t \leq n} \|\mathbf{F}_t^*\|_2 = O_P \left( \sqrt{\log n} \right). \quad (\text{E.1})$$

Then, by the definition (2.9), (D.1), (D.2), (E.1), Proposition 3.1 and Assumption 4(i), we

readily have

$$\max_{1 \leq t \leq n} \max_{1 \leq j \leq d} |\hat{\epsilon}_{tj} - \epsilon_{tj}| = \max_{1 \leq t \leq n} \max_{1 \leq j \leq d} \left| \hat{\lambda}_j^\top \hat{\mathbf{F}}_t - ((\mathbf{H}^{-1})^\top \lambda_j^*)^\top \mathbf{H} \mathbf{F}_t^* \right| = O_P \left( \left[ \frac{(\log d)(\log n)}{n} \right]^{1/2} \right). \quad (\text{E.2})$$

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

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

$$\eta_{k_0}^e \leq l < \eta_{k_0+1}^e < \cdots < \eta_{k_0+k_1}^e < u \leq \eta_{k_0+k_1+1}^e, \quad (\text{E.3})$$

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

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

and

$$\{(l - \eta_{k_0}^e) \wedge (\eta_{k_0+1}^e - l)\} \vee \{(u - \eta_{k_0+k_1}^e) \wedge (\eta_{k_0+k_1+1}^e - u)\} \leq c_{22} \varphi_{n,d}^e, \quad (\text{E.5})$$

where  $c_{21}$  and  $c_{22}$  are two positive constants,  $\kappa_n^e$  and  $\varphi_{n,d}^e$  are defined in Theorem 3.2. Define the intervals

$$\mathcal{J}_k^e = [\eta_{k-1}^e + (\eta_k^e - \eta_{k-1}^e)/3, \eta_{k-1}^e + 2(\eta_k^e - \eta_{k-1}^e)/3], \quad k = 1, \dots, K_2 + 1,$$

and the event

$$\mathcal{D}_n^e = \left\{ \forall k = 1, \dots, K_2, \exists m = 1, \dots, M_n^e \text{ such that } l_m \in \mathcal{J}_k^e \text{ and } u_m \in \mathcal{J}_{k+1}^e \right\},$$

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

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

$$\mathbf{P} \left( \overline{\mathcal{D}}_n^e \right) \leq K_2 \left[ 1 - (\kappa_n^e / (3n))^2 \right]^{M_n^e}, \quad (\text{E.6})$$

where  $\kappa_n^e$  is defined in Theorem 3.2.

PROOF. The proof is the same as Lemma D.2. Details are omitted here.  $\square$

Note that

$$\epsilon_{ti}\epsilon_{tj} = \mathbf{E}[\epsilon_{ti}\epsilon_{tj}] + (\epsilon_{ti}\epsilon_{tj} - \mathbf{E}[\epsilon_{ti}\epsilon_{tj}]) =: \mathbf{G}_{t,ij}^\epsilon + z_{t,ij}^\epsilon$$

and from (3.7)

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

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

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

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

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

as  $n, d \rightarrow \infty$ .

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

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

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

By Assumption 5, we readily have

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

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

$$\mathbf{P} \left( \max_{(l,u): 1 \leq l < u \leq n} \max_{(i,j): 1 \leq i, j \leq d} \max_{s: l \leq s < u} |c_{l,u}^{z,\epsilon}(s; i, j, k)| > \frac{c_{23}\sigma}{2} \log^2(nd) \right) \rightarrow 0, \quad (\text{E.11})$$

for  $k = 1$  and  $2$ .

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

$$\bar{z}_{t,ij}^\epsilon = z_{t,ij}^\epsilon \cdot \mathcal{J}(|z_{t,ij}^\epsilon| \leq c_{24} \log(nd)), \quad \tilde{z}_{t,ij}^\epsilon = z_{t,ij}^\epsilon \cdot \mathcal{J}(|z_{t,ij}^\epsilon| > c_{24} \log(nd)),$$

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

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

$$\max_{(i,j): 1 \leq i, j \leq d} \max_{1 \leq t \leq n} \mathbf{E}[\exp\{\iota_1 |z_{t,ij}^\epsilon|\}] < \infty.$$

Consequently, we can show that

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

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

$$\begin{aligned} &\mathbf{P} \left( \max_{(i,j): 1 \leq i, j \leq d} \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} |\bar{c}_{l,u}^{z,\epsilon}(s; i, j, 1)| > \frac{c_{23}\sigma}{4} \cdot \log^2(nd) \right) \\ &\leq \mathbf{P} \left( \max_{(i,j): 1 \leq i, j \leq d} \max_{(l,u): 1 \leq l < u \leq n} \max_{s: l \leq s < u} \left| \sqrt{\frac{u-s}{u-l+1}} \cdot \frac{1}{\sqrt{s-l+1}} \cdot \sum_{t=l}^s \tilde{z}_{t,ij}^\epsilon \right| > \frac{c_{23}\sigma}{5} \cdot \log^2(nd) \right) \\ &\leq \mathbf{P} \left( \max_{(i,j): 1 \leq i, j \leq d} \max_{1 \leq t \leq n} |z_{t,ij}^\epsilon| > c_{24} \log(nd) \right) \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{i=1}^d \sum_{j=i}^d \sum_{t=1}^n \frac{\mathbb{E}[\exp\{t_1 |z_{t,ij}^\epsilon|\}]}{\exp\{t_1 c_{24} \log(nd)\}} \\
&= O(d^{2-t_1 c_{24}} n^{1-t_1 c_{24}}) = o(1).
\end{aligned} \tag{E.12}$$

We next prove

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

As in the proof of (D.30), we consider the following two scenarios: (i)  $s - l + 1 \leq c_{25} \log^2(nd)$ , 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

$$\begin{aligned}
|\bar{c}_{l,u}^{z,\epsilon}(s; i, j, 1)| &\leq \sqrt{\frac{u-s}{u-l+1}} \cdot \frac{1}{\sqrt{s-l+1}} \cdot \sum_{t=1}^s (|\bar{z}_{t,ij}^\epsilon| + \mathbb{E}[|\bar{z}_{t,ij}^\epsilon|]) \\
&\leq \sqrt{s-l+1} \cdot (2c_{24} \log(nd)) \leq (2c_{24}\sqrt{c_{25}}) \cdot \log^2(nd).
\end{aligned}$$

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

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

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.  $\square$

For notational simplicity, we let

$$c_{l_m, u_m}^{\hat{\epsilon}, \hat{\sigma}, \mathcal{J}}(s; i, j) = c_{l_m, u_m}^{\hat{\epsilon}, \hat{\sigma}}(s; i, j) \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{\hat{\epsilon}, \hat{\sigma}}(t; i, j)| > \xi_n^e \right)$$

and

$$c_{l_m, u_m}^{\mathbb{G}, \hat{\epsilon}, \hat{\sigma}, \mathcal{J}}(s; i, j) = c_{l_m, u_m}^{\mathbb{G}, \hat{\epsilon}, \hat{\sigma}}(s; i, j) \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{\mathbb{G}, \hat{\epsilon}, \hat{\sigma}}(t; i, j)| > \xi_n^e \right)$$

for  $m \in \mathcal{M}_{l, u}^e$  such that  $[l_m, u_m]$  is a random sub-interval of  $[l, u]$ , where  $c_{l, u}^{\hat{\epsilon}, \hat{\sigma}}(s; i, j)$  is defined in (2.15) and  $c_{l, u}^{\mathbb{G}, \hat{\epsilon}, \hat{\sigma}}(s; i, j)$  is defined in (E.7). Define  $\mathbf{C}_{l_m, u_m}^{\hat{\epsilon}, \mathcal{J}}(s)$  and  $\mathbf{C}_{l_m, u_m}^{\mathbb{G}, \mathcal{J}}(s)$  as half-vectorisation of the two symmetric  $d \times d$  matrices with the  $(i, j)$ -entry being

$c_{l_m, u_m}^{\hat{\epsilon}, \hat{\sigma}, \hat{J}}(s; i, j)$  and  $c_{l_m, u_m}^{G, \epsilon, \hat{\sigma}, J}(s; i, j)$ , respectively. By (2.16) in Section 2.4, we readily have that

$$C_{l_m, u_m}^{\hat{\epsilon}}(s) = \left\| \mathbf{C}_{l_m, u_m}^{\hat{\epsilon}, \hat{J}}(s) \right\|_2^2.$$

Let

$$\mathcal{J}_{l, u}^e = \bigcup_{k: l + c_{22} \varphi_{n, d}^e \leq \eta_k^e \leq u - c_{22} \varphi_{n, d}^e} \mathcal{J}_k \quad (\text{E.14})$$

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

$$\tilde{\mathcal{J}}_{l, u}^e = \mathcal{J} \left\{ (i, j) : \max_{t: l \leq t < u} \left| c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j) \right| > \xi_n^e, 1 \leq i, j \leq d \right\} \quad (\text{E.15})$$

and

$$\hat{\mathcal{J}}_{l, u}^e = \mathcal{J} \left\{ (i, j) : \max_{t: l \leq t < u} \left| c_{l, u}^{\hat{\epsilon}, \hat{\sigma}}(t; i, j) \right| > \xi_n^e, 1 \leq i, j \leq d \right\}, \quad (\text{E.16})$$

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

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

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

The following lemma derives the asymptotic property of  $\tilde{\mathcal{J}}_{l, u}^e$  and  $\hat{\mathcal{J}}_{l, u}^e$  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*

$$\mathbf{P} \left( \mathcal{J}_{l, u}^e = \tilde{\mathcal{J}}_{l, u}^e \right) \rightarrow 1, \quad \mathbf{P} \left( \tilde{\mathcal{J}}_{l, u}^e = \hat{\mathcal{J}}_{l, u}^e \right) \rightarrow 1 \quad (\text{E.18})$$

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

$$(|\mathcal{J}_{l, u}^e|/K_2) \cdot \underline{\omega}_n^e \leq |\mathcal{J}_k| \cdot \underline{\omega}_n^e \leq \omega_k^e \leq \bar{\omega}_{l, u}^e \leq |\mathcal{J}_{l, u}^e| \cdot \bar{\omega}_n^e. \quad (\text{E.19})$$

Furthermore,

$$\mathbf{P} \left( \left\| \mathbf{C}_{l_m^e, u_m^e}^{\hat{\epsilon}, \hat{J}}(s_0^e) \right\|_2 \geq c_{26} \left( |\mathcal{J}_{l, u}^e| \kappa_n^e \underline{\omega}_n^e \right)^{1/2} \right) \rightarrow 1 \quad (\text{E.20})$$

as  $n, d \rightarrow \infty$ , 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),

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

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  $(i, j) \notin \mathcal{T}_{l,u}^\epsilon$ ; (iii)  $(i, j) \in \mathcal{T}_{l,u}^\epsilon$ . For case (i), it is obvious that  $|c_{l,u}^{G,\epsilon}(s; i, j)| = 0 \forall s \in [l, u]$ . 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),  $|c_{l,u}^{G,\epsilon}(s; i, j)|$  takes the maximum at one of the change points, which, together with the first equality in (E.21), leads to

$$\max_{s:l \leq s < u} |c_{l,u}^{G,\epsilon}(s; i, j)| \leq (c_{22} \varphi_{n,d}^\epsilon)^{1/2} \cdot 2(\bar{\omega}_n^\epsilon)^{1/2} \leq 2\sqrt{c_{22}} \log^2(nd). \tag{E.22}$$

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

$$\left| G_{\eta_{k_0+k}^\epsilon+1, ij}^\epsilon - G_{\eta_{k_0+k}^\epsilon, ij}^\epsilon \right| \geq (\underline{\omega}_n^\epsilon)^{1/2},$$

we readily have  $\left| G_{\eta_{k_0+k}^\epsilon, ij}^\epsilon \right| \vee \left| G_{\eta_{k_0+k}^\epsilon+1, ij}^\epsilon \right| \geq (\underline{\omega}_n^\epsilon)^{1/2}/2$ , implying that

$$\left| \sum_{t=\eta_{k_0+k}^\epsilon - c_{21} \kappa_n^\epsilon}^{\eta_{k_0+k}^\epsilon} G_{t,ij}^\epsilon \right| \vee \left| \sum_{t=\eta_{k_0+k}^\epsilon+1}^{\eta_{k_0+k}^\epsilon + c_{21} \kappa_n^\epsilon} G_{t,ij}^\epsilon \right| \geq c_{21} \kappa_n^\epsilon (\underline{\omega}_n^\epsilon)^{1/2}/2,$$

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

$$\left| \sum_{t=\eta_{k_0+k}^\epsilon - c_{21} \kappa_n^\epsilon}^{\eta_{k_0+k}^\epsilon} G_{t,ij}^\epsilon \right| \geq c_{21} \kappa_n^\epsilon (\underline{\omega}_n^\epsilon)^{1/2}/2. \tag{E.23}$$

By the triangle inequality, we have that

$$\max_{s:l \leq s < u} \left| \sum_{t=l}^s G_{t,ij}^\epsilon \right| \geq \left| \sum_{t=l}^{\eta_{k_0+k}^\epsilon} G_{t,ij}^\epsilon \right| = \left| \sum_{t=l}^{\eta_{k_0+k}^\epsilon - c_{21} \kappa_n^\epsilon - 1} G_{t,ij}^\epsilon + \sum_{t=\eta_{k_0+k}^\epsilon - c_{21} \kappa_n^\epsilon}^{\eta_{k_0+k}^\epsilon} G_{t,ij}^\epsilon \right|$$

$$\begin{aligned}
&\geq \left| \sum_{t=\eta_{k_0+k}^e - c_{21}\kappa_n^e}^{\eta_{k_0+k}^e} G_{t,ij}^e \right| - \left| \sum_{t=l}^{\eta_{k_0+k}^e - c_{21}\kappa_n^e - 1} G_{t,ij}^e \right| \\
&\geq \left| \sum_{t=\eta_{k_0+k}^e - c_{21}\kappa_n^e}^{\eta_{k_0+k}^e} G_{t,ij}^e \right| - \max_{s:l \leq s < u} \left| \sum_{t=l}^s G_{t,ij}^e \right|,
\end{aligned}$$

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

$$\max_{s:l \leq s < u} \left| \sum_{t=l}^s G_{t,ij}^e \right| \geq c_{21}\kappa_n^e (\underline{\omega}_n^e)^{1/2}/4. \quad (\text{E.24})$$

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

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

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

$$\max_{s:l \leq s < u} |c_{l,u}^{G,e}(s; i, j)| \geq c_{21}\kappa_n^e (\underline{\omega}_n^e/n)^{1/2}/2. \quad (\text{E.25})$$

Combining the above three cases and using (3.10), we can prove  $\mathbf{P}(\mathcal{J}_{l,u}^e = \tilde{\mathcal{J}}_{l,u}^e) \rightarrow 1$ .

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

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

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

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

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

$$\mathcal{J} \left( \max_{t:l \leq t < u} |c_{l,u}^{\hat{\epsilon},\hat{\sigma}}(t; i, j)| > \xi_n^e \right) = \mathcal{J} \left( \max_{t:l \leq t < u} |c_{l,u}^{G,\epsilon,\hat{\sigma}}(t; i, j)| > \xi_n^e \right), \quad (\text{E.27})$$



with probability approaching one, i.e.,  $\mathbf{P} \left( \tilde{\mathcal{J}}_{l,u}^e = \hat{\mathcal{J}}_{l,u}^e \right) \rightarrow 1$ . We have completed the proof of the two equalities in (E.18).

The proof of

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

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

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  $1 \leq m_k \leq M_n^e$  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  $1 \leq i \leq j \leq d$  and  $k$  such that  $l + c_{22}\varphi_{n,d}^e < \eta_k^e < u - c_{22}\varphi_{n,d}^e$ , we have

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

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

$$\left| c_{l_{m_k}, u_{m_k}}^{G,\epsilon,\hat{\sigma}}(\eta_k^e; i, j) \right| \geq c_{28} (\kappa_n^e)^{1/2} |\omega_{k,ij}^e| / \bar{\sigma}. \quad (\text{E.29})$$

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

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

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

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

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

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

with probability approaching one. Choosing  $c_{26} = c_{28}/(2K_2^{1/2}\bar{\sigma})$ , we can complete the

proof of (E.20). The proof of Lemma E.3 has been completed.  $\square$

LEMMA E.4. *The function  $\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2$  (as a function of  $s$ ) is either monotonic or first decreasing and then increasing on the interval  $[\eta_{k^e}^e, \eta_{k+1}^e]$  if  $s_0^e \in [\eta_{k^e}^e, \eta_{k+1}^e] \subseteq [l_{m_0^e}, u_{m_0^e}]$ . Furthermore,*

$$\left\{ \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k^e}^e) \right\|_2 \vee \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k+1}^e) \right\|_2 \right\} \geq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s_0^e) \right\|_2. \quad (\text{E.32})$$

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.  $\square$

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  $s_\star^e \in [l_{m_0^e}, u_{m_0^e}]$  be the point of maximising  $\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2$  with respect to  $s$ , i.e.,*

$$s_\star^e = \arg \max_{l_{m_0^e} \leq s < u_{m_0^e}} \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2 \quad (\text{E.33})$$

and define  $\eta_{k_\diamond}^e$  as a change point that satisfies

$$\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\diamond}^e) \right\|_2 > \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s_\star^e) \right\|_2 - 3c_{23} |\mathcal{T}_{l, u}^e|^{1/2} \log^2(nd), \quad (\text{E.34})$$

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

$$(\eta_{k_\diamond}^e - l_{m_0^e} + 1) \wedge (u_{m_0^e} - \eta_{k_\diamond}^e) \geq c_{29} \kappa_n^e \quad (\text{E.35})$$

when  $n$  is sufficiently large, and furthermore,

$$\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\diamond}^e) \right\|_2 > \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\diamond}^e + v_l) \right\|_2 + c_{30} v_l \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\diamond}^e) \right\|_2 \kappa_n^e / (u_{m_0^e} - l_{m_0^e} + 1)^2, \quad (\text{E.36})$$

where  $0 < v_l < c_{31} \gamma_n^e$  and  $\gamma_n^e = (\kappa_n^e / \underline{\omega}_n^e)^{1/2} \log^2(nd)$ ,  $c_{30}$  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  $s_\star^e$  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  $s_\star^e = \eta_{k_\star}^e$ . First we prove that

$$(\eta_{k_\star}^e - l + 1) \wedge (u - \eta_{k_\star}^e) \geq \kappa_n^e - c_{22} \varphi_{n, d}^e, \quad (\text{E.37})$$

where  $c_{22}$  is the same as that in (E.5). By (E.4) and (E.5), we have  $(\eta_{k_\star}^e - l + 1) \wedge (u - \eta_{k_\star}^e)$

is either smaller than  $c_{22}\varphi_{n,d}^e$  or larger than  $\kappa_n^e - c_{22}\varphi_{n,d}^e$ . Let

$$\mathbf{G}_s^{\epsilon,J} = \left[ G_{s,11}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} |c_{l,u}^{G,\epsilon,\hat{\sigma}}(t; 1, 1)| > \xi_n^e \right), \dots, G_{s,dd}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} |c_{l,u}^{G,\epsilon,\hat{\sigma}}(t; d, d)| > \xi_n^e \right) \right]^\top,$$

a  $d(d+1)/2$  column vector which denotes half-vectorisation of a  $d \times d$  symmetric matrix with the  $(i, j)$ -entry being  $G_{s,ij}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} |c_{l,u}^{G,\epsilon,\hat{\sigma}}(t; i, j)| > \xi_n^e \right)$ . Note that

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

where

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

If  $(\eta_{k_*}^e - l + 1) \wedge (u - \eta_{k_*}^e) \leq c_{22}\varphi_{n,d}^e$ , we must have  $(\eta_{k_*}^e - l_{m_0^e} + 1) \wedge (u_{m_0^e} - \eta_{k_*}^e) \leq c_{22}\varphi_{n,d}^e$ , implying that

$$c_{28} (|\mathcal{J}_{l,u}^e| \kappa_n^e \underline{\omega}_n^e)^{1/2} / \bar{\sigma} \leq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{G,\epsilon,J}(\eta_{k_*}^e) \right\|_2 \leq 2b_{l_{m_0^e}, u_{m_0^e}}^{\epsilon,J} (c_{22}\varphi_{n,d}^e)^{1/2} / \underline{\sigma}, \quad (\text{E.39})$$

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

$$b_{l_{m_0^e}, u_{m_0^e}}^{\epsilon,J} \leq K_2 (|\mathcal{J}_{l,u}^e| \bar{\omega}_n^e)^{1/2}, \quad (\text{E.40})$$

the inequalities in (E.39) would lead to a contradiction with the condition  $\kappa_n^e \underline{\omega}_n^e / \log^4(nd) \rightarrow \infty$  in Assumption 4(iii). Hence (E.37) has been proved, which indicates that there exists  $m_*^e \in \mathcal{M}_{l,u}^e$  such that  $l_{m_*^e} \in \mathcal{J}_{k_*}^e$  and  $u_{m_*^e} \in \mathcal{J}_{k_*+1}^e$ .

We next prove that for  $n$  large enough,

$$(\eta_{k_*}^e - l_{m_0^e} + 1) \wedge (u_{m_0^e} - \eta_{k_*}^e) \geq c_{29}\kappa_n^e. \quad (\text{E.41})$$

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

$$(\eta_{k_*}^e - l_{m_0^e} + 1) \wedge (u_{m_0^e} - \eta_{k_*}^e) < c_* \kappa_n^e. \quad (\text{E.42})$$

As in the proof of (D.48), without loss of generality, we let  $\eta_{k_*}^e - l_{m_0^e} + 1 < c_* \kappa_n^e$ , and consider the following two cases of  $u_{m_0^e}$ :

(i)  $\eta_{k_*}^e \leq u_{m_0^e} < \eta_{k_0+k_1}^e$ , or  $\eta_{k_0+k_1+1}^e - c_{22}\varphi_{n,d}^e \leq u_{m_0^e} \leq \eta_{k_0+k_1+1}^e$  and  $\eta_{k_*}^e < \eta_{k_0+k_1}^e < u_{m_0^e} \leq u$ ;

(ii)  $\eta_{k_0+k_1}^e \leq \mathbf{u} \leq \eta_{k_0+k_1}^e + c_{22}\varphi_{n,d}^e$  and  $\eta_{k_*}^e < \eta_{k_0+k_1}^e < \mathbf{u}_{m_0^e} \leq \mathbf{u}$ .

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

$$c_{28} (|\mathcal{J}_{l,u}^e| \kappa_n^e \underline{\omega}_n^e)^{1/2} / \bar{\sigma} \leq \left\| \mathbf{C}_{l_{m_0^e}, \mathbf{u}_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}} (\eta_{k_*}^e) \right\|_2 \leq 2K_2 (c_* |\mathcal{J}_{l,u}^e| \kappa_n^e \bar{\omega}_n^e)^{1/2} / \underline{\sigma}, \quad (\text{E.43})$$

which would result in a contradiction if we choose a sufficiently small  $c_* > 0$ . We next consider case (ii). Since  $\eta_{k_0+k_1}^e \leq \mathbf{u}_{m_0^e} \leq \mathbf{u} \leq \eta_{k_0+k_1}^e + c_{22}\varphi_{n,d}^e$  in this case, we may show that

$$\begin{aligned} & c_{28} (|\mathcal{J}_{l,u}^e| \kappa_n^e \underline{\omega}_n^e)^{1/2} / \bar{\sigma} \leq \left\| \mathbf{C}_{l_{m_0^e}, \mathbf{u}_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}} (\eta_{k_*}^e) \right\|_2 \\ & \leq \frac{1}{\underline{\sigma}} \cdot \sqrt{\frac{(\eta_{k_*}^e - l_{m_0^e} + 1)(\mathbf{u}_{m_0^e} - \eta_{k_*}^e)}{\mathbf{u}_{m_0^e} - l_{m_0^e} + 1}} \left\| \frac{1}{\eta_{k_*}^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta_{k_*}^e} \mathbf{G}_t^{\epsilon, \mathcal{J}} - \frac{1}{\mathbf{u}_{m_0^e} - \eta_{k_*}^e} \sum_{t=\eta_{k_*}^e+1}^{\mathbf{u}_{m_0^e}} \mathbf{G}_t^{\epsilon, \mathcal{J}} \right\|_2 \\ & \leq \frac{1}{\underline{\sigma}} \cdot \sqrt{\frac{(\eta_{k_*}^e - l_{m_0^e} + 1)(\mathbf{u}_{m_0^e} - \eta_{k_*}^e)}{\mathbf{u}_{m_0^e} - l_{m_0^e} + 1}} \cdot \left\| \frac{1}{\eta_{k_*}^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta_{k_*}^e} \mathbf{G}_t^{\epsilon, \mathcal{J}} - \frac{1}{\mathbf{u}_{m_0^e} - \eta_{k_*}^e} \sum_{t=\eta_{k_*}^e+1}^{\mathbf{u}_{m_0^e}} \mathbf{G}_{t \wedge (\mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e)}^{\epsilon, \mathcal{J}} \right\|_2 \\ & \quad + \frac{1}{\underline{\sigma}} \cdot \sqrt{\frac{(\eta_{k_*}^e - l_{m_0^e} + 1)(\mathbf{u}_{m_0^e} - \eta_{k_*}^e)}{\mathbf{u}_{m_0^e} - l_{m_0^e} + 1}} \cdot \frac{c_{22}\varphi_{n,d}^e \mathbf{b}_{\mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e, \mathbf{u}_{m_0^e}}^{\epsilon, \mathcal{J}}}{\mathbf{u}_{m_0^e} - \eta_{k_*}^e} \\ & \leq \left[ \left( 2\mathbf{b}_{l_{m_0^e}, \mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e}^{\epsilon, \mathcal{J}} + c_{22}\varphi_{n,d}^e \mathbf{b}_{\mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e, \mathbf{u}_{m_0^e}}^{\epsilon, \mathcal{J}} / \kappa_n^e \right) / \underline{\sigma} \right] \cdot \sqrt{(\eta_{k_*}^e - l_{m_0^e} + 1) \wedge (\mathbf{u}_{m_0^e} - \eta_{k_*}^e)}. \end{aligned}$$

As  $\frac{\kappa_n^e \underline{\omega}_n^e}{\log^4(nd)} \rightarrow \infty$  in Assumption 4(iii), we have

$$\frac{\varphi_{n,d}^e \mathbf{b}_{\mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e, \mathbf{u}_{m_0^e}}^{\epsilon, \mathcal{J}}}{\kappa_n^e} = \frac{1}{\kappa_n^e} \mathcal{O} \left( (|\mathcal{J}_{l,u}^e| \cdot \bar{\omega}_n^e)^{1/2} \cdot \log^4(nd) / \underline{\omega}_n^e \right) = o \left( (|\mathcal{J}_{l,u}^e| \cdot \bar{\omega}_n^e)^{1/2} \right),$$

which, together with the fact that  $2\mathbf{b}_{l_{m_0^e}, \mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e}^{\epsilon, \mathcal{J}} \geq \bar{\omega}_{l,u} \geq (|\mathcal{J}_{l,u}^e| \underline{\omega}_n^e / K_2)^{1/2}$ , leads to

$$\varphi_{n,d}^e \mathbf{b}_{\mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e, \mathbf{u}_{m_0^e}}^{\epsilon, \mathcal{J}} / \kappa_n^e = o \left( \mathbf{b}_{l_{m_0^e}, \mathbf{u}_{m_0^e} - c_{22}\varphi_{n,d}^e}^{\epsilon, \mathcal{J}} \right).$$

Hence, we have

$$c_{28} (|\mathcal{J}_{l,u}^e| \kappa_n^e \underline{\omega}_n^e)^{1/2} / \bar{\sigma} \leq \left\| \mathbf{C}_{l_{m_0^e}, \mathbf{u}_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}} (\eta_{k_*}^e) \right\|_2 \leq 2K_2 (c_* |\mathcal{J}_{l,u}^e| \kappa_n^e \bar{\omega}_n^e)^{1/2} / \underline{\sigma}$$

which would lead a contradiction as  $\bar{\omega}_n^e \asymp \underline{\omega}_n^e$  in Assumption 4(iii) when  $c_* > 0$  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)  $\mathbf{u}_{m_0^e} \leq \eta_{k_\diamond+1}^e$  and (ii)  $\eta_{k_\diamond+1}^e < \mathbf{u}_{m_0^e}$ . For case (i), we define  $v_i = \eta_{k_\diamond}^e - l_{m_0^e} + 1$  and  $v_h = \mathbf{u}_{m_0^e} - \eta_{k_\diamond}^e$ . Let

$\boldsymbol{\beta} = (\beta_1, \dots, \beta_{d(d+1)/2})^\top$  with

$$\beta_k = c_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \hat{\sigma}}(\eta_{k_\circ}^e; i, j) \left( \frac{v_i v_h}{v_i + v_h} \right)^{1/2} \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{\mathbf{G}, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^e \right)$$

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

$$c_{l_m, u_m}^{\mathbf{G}, \epsilon, \hat{\sigma}, \mathcal{J}}(s; i, j) = c_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \hat{\sigma}}(\eta_{k_\circ}^e; i, j) \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{\mathbf{G}, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^e \right) = \beta_k \left( \frac{v_i + v_h}{v_i v_h} \right)^{1/2}$$

and similarly

$$c_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \hat{\sigma}, \mathcal{J}}(\eta_{k_\circ}^e + v_l; i, j) = \beta_k \left( \frac{v_h - v_l}{v_h} \right) \cdot \left[ \frac{v_i + v_h}{(v_i + v_l)(v_h - v_l)} \right]^{1/2},$$

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

$$\begin{aligned} \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e + v_l) \right\|_2 &= \|\boldsymbol{\beta}\|_2 \cdot \frac{\sqrt{v_i + v_h}}{v_h} \left( \sqrt{\frac{v_h}{v_i}} - \sqrt{\frac{v_h - v_l}{v_i + v_l}} \right) \\ &\geq \frac{v_l \cdot \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e) \right\|_2}{2(u_{m_0^e} - l_{m_0^e} + 1)}. \end{aligned} \quad (\text{E.44})$$

For case (ii), we let  $v_i = \eta_{k_\circ}^e - l_{m_0^e} + 1$ ,  $v_h = (c_{29} \wedge 1) \kappa_n^e / 3$ ,  $v_j = u_{m_0^e} - \eta_{k_\circ}^e - v_h$ , and

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

Then, for  $0 \leq v_l \leq v_h$ , we readily have that

$$\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e + v_l) \right\|_2 = \|\boldsymbol{\beta} + v_l \mathbf{V}_{\mathbf{G}}^e\|_2 \left[ \frac{v_i + v_j + v_h}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2},$$

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

$$\mathbb{E}(v_l) = \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e + v_l) \right\|_2$$

and

$$\mathbb{E}_1 = \left[ \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_\circ}^e + v_h) \right\|_2 \right] \cdot \frac{v_l}{v_h} \left[ \frac{(v_i + v_h)v_j}{(v_i + v_l)(v_j + v_h - v_l)} \right]^{1/2}.$$

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

$$E(v_l) - E_1 \geq E_2 \times (1 + E_3), \quad (\text{E.45})$$

where

$$E_2 = \frac{\|\boldsymbol{\beta}\|_2 v_l (v_h - v_l) \sqrt{v_i + v_j + v_h}}{\sqrt{v_i(v_j + v_h)} \sqrt{(v_i + v_l)(v_j + v_h - v_l)} (\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{v_i(v_j + v_h)})},$$

and

$$E_3 = \frac{(v_j - v_i)(v_j - v_i - v_l)}{(\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{(v_i + v_h)v_j})(\sqrt{v_i(v_j + v_h)} + \sqrt{(v_i + v_h)v_j})}.$$

Noting that  $v_l$  is smaller than  $v_h/2$  and  $v_i$  for large  $n$ , we have

$$\begin{aligned} E_2 &= \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_o}^e) \right\|_2 \frac{v_l (v_h - v_l)}{\sqrt{(v_i + v_l)(v_j + v_h - v_l)} [\sqrt{(v_i + v_l)(v_j + v_h - v_l)} + \sqrt{v_i(v_j + v_h)}]} \\ &\geq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_o}^e) \right\|_2 \frac{(v_l v_h / 2)}{\sqrt{2v_i(v_j + v_h)} [\sqrt{2v_i(v_j + v_h)} + \sqrt{v_i(v_j + v_h)}]} \\ &\geq (2c_{30} v_l \kappa_n^e) \cdot \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_o}^e) \right\|_2 / (u_{m_0^e} - l_{m_0^e} + 1)^2. \end{aligned} \quad (\text{E.46})$$

Meanwhile, as  $(v_j - v_i)(v_j - v_i - v_l)$  reaches its minimum at  $v_j - v_i = v_l/2$ ,  $v_i, v_j, v_h \geq (c_{29} \wedge 1) \kappa_n^e / 3$  by (E.35) and  $v_l = o(\kappa_n^e)$ , following the proof of (D.55), we have

$$E_3 \geq \frac{-v_l^2}{4(1 + \sqrt{2})(\sqrt{2} + \sqrt{2})[(c_{29} \wedge 1) \kappa_n^e / 3]^2} \rightarrow 0. \quad (\text{E.47})$$

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  $E(v_l)$  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).  $\square$

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  $k_0 + 1 \leq k_o \leq k_0 + k_1$  such that*

$$|s_0^e - \eta_{k_o}^e| \leq c_{31} \gamma_{n,d}^e \quad (\text{E.48})$$

with probability approaching one, as  $n \rightarrow \infty$ , where  $\gamma_{n,d}^e = (\kappa_n^e / \underline{\omega}_n^e)^{1/2} \log^2(nd)$  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  $s_0^e \in [\eta_{\tilde{k}}^e, \eta_{\tilde{k}+1}^e)$  for  $k_0 \leq \tilde{k} \leq k_0 + k_1$ . We next show the consequence if (E.48) fails and consider two cases.

Case (i): only one of  $\eta_{\tilde{k}}^e$  and  $\eta_{\tilde{k}+1}^e$  locates in the interval  $[l_{m_0^e}, u_{m_0^e})$ . Without loss of generality, consider that  $\eta_{\tilde{k}}^e$  belongs to the interval  $[l_{m_0^e}, u_{m_0^e})$  and choose  $\eta_{k_0}^e = \eta_{\tilde{k}}^e$ . By the definitions of  $m_0^e$  and  $s_0^e$ , (E.18) and following the proof of Proposition 3.3, we readily have that

$$\begin{aligned} \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_0}^e) \right\|_2 &\leq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\hat{\epsilon}, \hat{\mathcal{J}}}(\eta_{k_0}^e) \right\|_2 + \left[ c_{27} \sqrt{(\log d)(\log n)} + c_{23} \log^2(nd) \right] |\mathcal{J}_{l, u}^e|^{1/2} \\ &\leq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\hat{\epsilon}, \hat{\mathcal{J}}}(s_0^e) \right\|_2 + \left[ c_{27} \sqrt{(\log d)(\log n)} + c_{23} \log^2(nd) \right] |\mathcal{J}_{l, u}^e|^{1/2} \\ &\leq \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s_0^e) \right\|_2 + 2 \left[ c_{27} \sqrt{(\log d)(\log n)} + c_{23} \log^2(nd) \right] |\mathcal{J}_{l, u}^e|^{1/2} \end{aligned} \quad (\text{E.49})$$

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  $\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2$  (treated as a function of  $s$ ) locally decreases at  $[\eta_{k_0}^e, u_{m_0^e})$  which includes the point of  $s = s_0^e$ . When (E.48) fails, we have

$$\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s_0^e) \right\|_2 < \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2 < \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_0}^e) \right\|_2. \quad (\text{E.50})$$

for any  $s \in (\eta_{k_0}^e, \eta_{k_0}^e + c_{31}\gamma_n^e]$ . 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

$$\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(\eta_{k_0}^e) \right\|_2 > \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s_0^e) \right\|_2 + 2 \left[ c_{27} \sqrt{(\log d)(\log n)} + c_{23} \log^2(nd) \right] |\mathcal{J}_{l, u}^e|^{1/2} \quad (\text{E.51})$$

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

Case (ii): both  $\eta_{\tilde{k}}^e$  and  $\eta_{\tilde{k}+1}^e$  are in the interval  $[l_{m_0^e}, u_{m_0^e}]$ . As in the proof of Lemma D.7, we consider two scenarios: (ii.1)  $\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2$  locally decreases at the point  $s = s_0^e$ ; and (ii.2)  $\left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, \mathcal{J}}(s) \right\|_2$  locally increases at the point  $s = s_0^e$ . For scenario (ii.1), we choose  $\eta_{k_0}^e = \eta_{\tilde{k}}^e$ , and for scenario (ii.2), we choose  $\eta_{k_0}^e = \eta_{\tilde{k}+1}^e$ . 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.  $\square$

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

$$Z_{t, ij}^e = \mathbf{E}[\epsilon_{ti} \epsilon_{tj}] + (\epsilon_{ti} \epsilon_{tj} - \mathbf{E}[\epsilon_{ti} \epsilon_{tj}]) =: G_{t, ij}^e + z_{t, ij}^e.$$

For  $(i, j)$  satisfying  $1 \leq i \leq j \leq d$ , 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

$$\begin{aligned} \mathbf{Z}_{\bullet, k}^{\epsilon, J} &= \left( Z_{l_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right), \dots, Z_{u_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right) \right)^\top, \\ \mathbf{G}_{\bullet, k}^{\epsilon, J} &= \left( G_{l_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right), \dots, G_{u_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right) \right)^\top, \\ \mathbf{z}_{\bullet, k}^{\epsilon, J} &= \left( z_{l_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right), \dots, z_{u_{m_0^\epsilon}, ij}^\epsilon \cdot \mathcal{J} \left( \max_{t: l \leq t < u} |c_{l, u}^{G, \epsilon, \hat{\sigma}}(t; i, j)| > \xi_n^\epsilon \right) \right)^\top. \end{aligned}$$

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*

$$|s_0^\epsilon - \eta_{k_0}^\epsilon| \leq c_{32} \varphi_{n, d}^\epsilon \quad (\text{E.52})$$

as  $n \rightarrow \infty$ , where  $c_{32}$  is a positive constant and  $\varphi_{n, d}^\epsilon$  is defined in Theorem 3.2.

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

By (D.60), we readily have

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

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

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

By (D.61), we may show that

$$Q_k^{\epsilon, J}(s; 1) \geq Q_k^{\epsilon, J}(s; 2), \quad k = 1, \dots, d(d+1)/2. \quad (\text{E.53})$$

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

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



and

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

Letting

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

by (D.61), we have

$$Q_k^{\varepsilon, J}(s; 3) \geq Q_k^{\varepsilon, J}(s; 1) \geq 0. \quad (\text{E.54})$$

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

$$\begin{aligned} & \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\varepsilon, J}(s_0^e; 3)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \geq \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\varepsilon, J}(s_0^e; 1)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} = \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\hat{\varepsilon}, \hat{J}}(s_0^e; 1)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} + O_P(|\mathcal{J}_{l, u}^e|(\log d)(\log n)) \\ & \geq \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\hat{\varepsilon}, \hat{J}}(\eta_{k_0^e}^e; 1)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} + O_P(|\mathcal{J}_{l, u}^e|(\log d)(\log n)) = \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\varepsilon, J}(\eta_{k_0^e}^e; 1)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} + O_P(|\mathcal{J}_{l, u}^e|(\log d)(\log n)) \\ & \geq \sum_{k=1}^{d(d+1)/2} \frac{Q_k^{\varepsilon, J}(\eta_{k_0^e}^e; 2)}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} + O_P(|\mathcal{J}_{l, u}^e|(\log d)(\log n)), \end{aligned}$$

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

$$\sum_{k=1}^{d(d+1)/2} \frac{1}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} [Q_k^{\varepsilon, J}(s_0^e; 3) - Q_k^{\varepsilon, J}(\eta_{k_0^e}^e; 2)] \geq -c_{33} |\mathcal{J}_{l, u}^e|(\log d)(\log n) \quad (\text{E.55})$$

holds with probability approaching one.

Letting  $c_{32} > 0$  be sufficiently large, we next show that the assertion of  $|s_0^e - \eta_{k_0^e}^e| > c_{32} \log^4(nd)/\underline{\omega}_n^e$  would lead to a contradiction with (E.55), which consequently proves (E.52). Defining

$$Q_k^{\varepsilon, J}(s; 4) = \left| \left\langle \mathbf{G}_{\bullet, k}^{\varepsilon, J}, \boldsymbol{\psi}_{l_{m_0^e}, u_{m_0^e}}^s \right\rangle \right|^2 = - \left\| \mathbf{G}_{\bullet, k}^{\varepsilon, J} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J^s} \right\|_2^2 + \left\| \mathbf{G}_{\bullet, k}^{\varepsilon, J} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J} \right\|_2^2,$$

we have

$$\begin{aligned} Q_k^{\varepsilon, J}(s; 3) - Q_k^{\varepsilon, J}(\eta_{k_0^e}^e; 2) &= \left\| \mathbf{G}_{\bullet, k}^{\varepsilon, J} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J^{\eta_{k_0^e}^e}} \right\|_2^2 - \left\| \mathbf{G}_{\bullet, k}^{\varepsilon, J} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J^s} \right\|_2^2 + 2 \left\langle \mathbf{z}_{\bullet, k}^{\varepsilon, J}, \bar{\mathbf{Z}}_{\bullet, k}^{\varepsilon, J^s} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J^{\eta_{k_0^e}^e}} \right\rangle \\ &= 2 \left\langle \mathbf{z}_{\bullet, k}^{\varepsilon, J}, \bar{\mathbf{Z}}_{\bullet, k}^{\varepsilon, J^s} - \bar{\mathbf{G}}_{\bullet, k}^{\varepsilon, J^{\eta_{k_0^e}^e}} \right\rangle - [Q_k^{\varepsilon, J}(\eta_{k_0^e}^e; 4) - Q_k^{\varepsilon, J}(s; 4)]. \quad (\text{E.56}) \end{aligned}$$

We next show that with probability approaching one,

$$\begin{aligned} & \sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \left| \left\langle \mathbf{z}_{\bullet, k}^{\epsilon, J}, \bar{\mathbf{z}}_{\bullet, k}^{\epsilon, J^{s_0^e}} - \bar{\mathbf{G}}_{\bullet, k}^{\epsilon, J^{\eta_{k_0^e}^e}} \right\rangle \right| \\ & \leq c_{34} |\mathcal{J}_{l, u}^e| \log^2(nd) \max \left\{ \frac{|s_0^e - \eta_{k_0^e}^e| \cdot (\bar{\omega}_n^e)^{1/2}}{(\kappa_n^e)^{1/2}}, |s_0^e - \eta_{k_0^e}^e|^{1/2} (\bar{\omega}_n^e)^{1/2}, \log^2(nd) \right\} \end{aligned} \quad (\text{E.57})$$

and

$$\sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} [Q_k^{\epsilon, J}(\eta_{k_0^e}^e; 4) - Q_k^{\epsilon, J}(s_0^e; 4)] \geq c_{35} |\mathcal{J}_{l, u}^e| |s_0^e - \eta_{k_0^e}^e| \underline{\omega}_n^e (\kappa_n^e/n)^2, \quad (\text{E.58})$$

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

Without loss of generality, we assume that  $s_0^e \geq \eta_{k_0^e}^e$ . Note that

$$\begin{aligned} \sum_{k=1}^{d(d+1)/2} \left\langle \mathbf{z}_{\bullet, k}^{\epsilon, J}, \bar{\mathbf{z}}_{\bullet, k}^{\epsilon, J^{s_0^e}} - \bar{\mathbf{G}}_{\bullet, k}^{\epsilon, J^{\eta_{k_0^e}^e}} \right\rangle &= \sum_{k=1}^{d(d+1)/2} \left\langle \mathbf{z}_{\bullet, k}^{\epsilon, J}, \bar{\mathbf{z}}_{\bullet, k}^{\epsilon, J^{s_0^e}} - \bar{\mathbf{G}}_{\bullet, k}^{\epsilon, J^{s_0^e}} \right\rangle \\ &+ \sum_{k=1}^{d(d+1)/2} \left\langle \mathbf{z}_{\bullet, k}^{\epsilon, J}, \bar{\mathbf{G}}_{\bullet, k}^{\epsilon, J^{s_0^e}} - \bar{\mathbf{G}}_{\bullet, k}^{\epsilon, J^{\eta_{k_0^e}^e}} \right\rangle. \end{aligned} \quad (\text{E.59})$$

Following standard calculations, we have

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

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

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

uniformly over  $s$  and  $k$ . This indicates that

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

with probability approaching one. On the other hand,

$$\begin{aligned} \left\langle \mathbf{z}_{\bullet,k}^{\epsilon,J}, \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,J^{s_0^e}} - \bar{\mathbf{G}}_{\bullet,k}^{\epsilon,J^{\eta_{k_0}^e}} \right\rangle &= \left( \sum_{t=l_{m_0^e}}^{\eta_{k_0}^e} + \sum_{t=\eta_{k_0}^e+1}^{s_0^e} + \sum_{t=s_0^e+1}^{u_{m_0^e}} \right) \mathbf{z}_{t,k}^{\epsilon,J} \left( \bar{\mathbf{G}}_{t,k}^{\epsilon,J^{s_0^e}} - \bar{\mathbf{G}}_{t,k}^{\epsilon,J^{\eta_{k_0}^e}} \right) \\ &=: \Xi_1 + \Xi_2 + \Xi_3. \end{aligned} \quad (\text{E.63})$$

For  $\Xi_1$ , we note that

$$|\Xi_1| \leq \sqrt{\eta_{k_0}^e - l_{m_0^e} + 1} \left| \frac{1}{\sqrt{\eta_{k_0}^e - l_{m_0^e} + 1}} \sum_{t=l_{m_0^e}}^{\eta_{k_0}^e} \mathbf{z}_{t,k}^{\epsilon,J} \cdot \left| \frac{1}{s_0^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{s_0^e} \mathbf{G}_{t,k}^{\epsilon,J} - \frac{1}{\eta_{k_0}^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta_{k_0}^e} \mathbf{G}_{t,k}^{\epsilon,J} \right| \right|,$$

and recall that

$$\mathbf{b}_{l,u}^{\epsilon,J} = \sup_{l \leq t \leq u} \left\| \mathbf{G}_t^{\epsilon,J} - \frac{1}{u-l+1} \sum_{t=l}^u \mathbf{G}_t^{\epsilon,J} \right\|_2 \leq K_2 (|\mathcal{J}_{l,u}^e| \bar{\omega}_n^e)^{1/2}.$$

Let

$$\mathbf{z}_s^{\epsilon,J} = \left[ \mathbf{z}_{s,11}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} \left| \mathbf{c}_{l,u}^{G,\epsilon,\hat{\sigma}}(t; 1, 1) \right| > \xi_n^e \right), \dots, \mathbf{z}_{s,dd}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} \left| \mathbf{c}_{l,u}^{G,\epsilon,\hat{\sigma}}(t; d, d) \right| > \xi_n^e \right) \right]^\top,$$

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  $\mathbf{z}_{s,ij}^e \cdot \mathcal{J} \left( \max_{t:l \leq t < u} \left| \mathbf{c}_{l,u}^{G,\epsilon,\hat{\sigma}}(t; i, j) \right| > \xi_n^e \right)$ . Then, by (E.61), Assumption 5 and the Cauchy-Schwarz inequality,

$$\begin{aligned} & \sum_{k=1}^{d(d+1)/2} \frac{1}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \cdot |\Xi_1| \\ & \leq \frac{\sqrt{\eta_{k_0}^e - l_{m_0^e} + 1}}{\underline{\sigma}^2} \left\| \frac{1}{\sqrt{\eta_{k_0}^e - l_{m_0^e} + 1}} \sum_{t=l_{m_0^e}}^{\eta_{k_0}^e} \mathbf{z}_t^{\epsilon,J} \right\|_2 \cdot \left\| \frac{1}{s_0^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{s_0^e} \mathbf{G}_t^{\epsilon,J} - \frac{1}{\eta_{k_0}^e - l_{m_0^e} + 1} \sum_{t=l_{m_0^e}}^{\eta_{k_0}^e} \mathbf{G}_t^{\epsilon,J} \right\|_2 \\ & \leq \frac{\sqrt{\eta_{k_0}^e - l_{m_0^e} + 1}}{\underline{\sigma}^2} \cdot O_P \left( |\mathcal{J}_{l,u}^e|^{1/2} \log^2(nd) \right) \cdot \frac{|s_0^e - \eta_{k_0}^e| 2\mathbf{b}_{l+c_{22}\varphi_{n,d}^e, u-c_{22}\varphi_{n,d}^e}^{\epsilon,J}}{s_0^e - l_{m_0^e} + 1} \\ & = O_P \left( |\mathcal{J}_{l,u}^e| \log^2(nd) |s_0^e - \eta_{k_0}^e| \cdot (\bar{\omega}_n^e / \kappa_n^e)^{1/2} \right), \end{aligned} \quad (\text{E.64})$$

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

$$\sum_{k=1}^{d(d+1)/2} \frac{\Xi_2}{\hat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \leq (c_{34}/4) \cdot |\mathcal{J}_{l,u}^e| \log^2(nd) |s_0^e - \eta_{k_0}^e|^{1/2} (\bar{\omega}_n^e)^{1/2} \quad (\text{E.65})$$

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

of (E.57).

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

$$\begin{aligned}
& \sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \cdot [Q_k^{\epsilon, J}(\eta_{k_0}^e; 4) - Q_k^{\epsilon, J}(s_0^e; 4)] \\
&= \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(\eta_{k_0}^e) \right\|_2^2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(s_0^e) \right\|_2^2 \\
&= \left[ \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(\eta_{k_0}^e) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(s_0^e) \right\|_2 \right] \cdot \left[ \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(\eta_{k_0}^e) \right\|_2 + \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(s_0^e) \right\|_2 \right] \\
&\geq \left[ \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(\eta_{k_0}^e) \right\|_2 - \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(s_0^e) \right\|_2 \right] \cdot \left\| \mathbf{C}_{l_{m_0^e}, u_{m_0^e}}^{\mathbf{G}, \epsilon, J}(\eta_{k_0}^e) \right\|_2 \\
&\geq c_{35} |\mathcal{J}_{l, u}^e| |s_0^e - \eta_{k_0}^e| \underline{\omega}_n^e (\kappa_n^e / n)^2, \tag{E.66}
\end{aligned}$$

completing the proof of (E.58).

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

$$\begin{aligned}
& \sum_{k=1}^{d(d+1)/2} \frac{1}{\widehat{\sigma}_{l_{m_0^e}, u_{m_0^e}}^2(k)} \cdot [Q_k^{\epsilon, J}(s_0^e; 3) - Q_k^{\epsilon, J}(\eta_{k_0}^e; 2)] \\
&\leq c_{34} |\mathcal{J}_{l, u}^e| \log^2(nd) \max \left\{ \frac{|s_0^e - \eta_{k_0}^e| \cdot (\overline{\omega}_n^e)^{1/2}}{(\kappa_n^e)^{1/2}}, |s_0^e - \eta_{k_0}^e|^{1/2} (\overline{\omega}_n^e)^{1/2}, \log^2(nd) \right\} \\
&\quad - c_{35} |\mathcal{J}_{l, u}^e| |s_0^e - \eta_{k_0}^e| \underline{\omega}_n^e (\kappa_n^e / n)^2, \\
&\leq -c_{33} |\mathcal{J}_{l, u}^e| \log^4(nd), \tag{E.67}
\end{aligned}$$

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.  $\square$

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  $1 \leq k \leq K_2$  such that  $\eta_k^e < l < u \leq \eta_{k+1}^e$ ; (ii) there exists  $1 \leq k \leq K_2$  such that  $l \leq \eta_k^e < u$  and  $(\eta_k^e - l + 1) \wedge (u - \eta_k^e) \leq c_{32} \varphi_{n, d}^e$ ; (iii) there exists  $1 \leq k \leq K_2$  such that  $l \leq \eta_k^e < \eta_{k+1}^e < u$  and  $(\eta_k^e - l + 1) \vee (u - \eta_{k+1}^e) \leq c_{32} \varphi_{n, d}^e$ , where  $c_{32}$  is defined in Lemma E.7. For  $l$  and  $u$

satisfy either of scenarios (i)–(iii), we may show that

$$\max_{1 \leq i, j \leq d} \max_{l_{m_0^e} \leq s < u_{m_0^e}} \left| c_{l_{m_0^e}, u_{m_0^e}}^{\hat{e}}(s_0^e; i, j) \right| = O_P \left( \log^2(nd) \right), \quad (\text{E.68})$$

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.  $\square$

## Appendix F: Additional simulation results

We next provide an additional simulated example to compare the finite-sample performance of some commonly-used criteria for factor number selection in the approximate factor model with structural breaks.

EXAMPLE F.1. As in Example 5.1, we consider the following approximate factor model to generate data:

$$X_{ti} = \sum_{j=1}^r \lambda_{ij,t} F_{tj} + \epsilon_{ti}, \quad i = 1, \dots, d, \quad t = 1, \dots, n, \quad (\text{F.1})$$

where  $r = 5$  is the number of common factors, each factor process is generated via an AR(1) model:

$$F_{tj} = \rho_j F_{t-1,j} + u_{tj}, \quad t = 1, \dots, n, \quad (\text{F.2})$$

with  $\rho_j = 0.4 - 0.05(j-1)$  for  $j = 1, \dots, r$ , and  $u_{tj}$  following a standard normal distribution independently over  $t$  and  $j$ . As in Example 5.1, there is one break point  $\eta_1^c = \lfloor n/3 \rfloor + 1 = 67$  in the factor loadings  $\lambda_{ij,t}$ : 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  $\eta_1^c$ ; and they are shifted by a random amount  $N(0, 4)$  after the break point  $\eta_1^c$  as in [Barigozzi, Cho and Fryzlewicz \(2018\)](#). The idiosyncratic errors  $\epsilon_t$  follow a multivariate normal distribution  $N_d(\mathbf{0}, \Sigma_\epsilon)$  independently over  $t$ , where  $\phi_j$ , the square root of the  $j$ -th diagonal element of  $\Sigma_\epsilon$ , is generated from independent uniform distribution  $U(0.5, 1.5)$ , and the  $(i, j)$ -entry of  $\Sigma_\epsilon$  is  $\phi_i \phi_j (-0.5)^{|i-j|}$ , for  $1 \leq i \neq j \leq d$ . For the covariance structure of  $\epsilon_t$ , there is one break point  $\eta_1^e = \lfloor 2n/3 \rfloor = 133$  after which we swap the orders of  $\lfloor \rho_1^e d/2 \rfloor$  randomly selected pairs of elements of  $\epsilon_t$  (c.f., [Cho and Fryzlewicz, 2015](#)) with  $\rho_1^e$  chosen as 0.1, 0.5 or 1. In this simulation, we set the sample size  $n = 100, 200, 500$ , the dimension  $d = 100, 200, 500$  and the replication number  $R = 100$ .

Note that, with existence of a “type 1 break” in the factor loadings, the number of transformed factors in model (2.6) is  $q_0 = 10$ . The performance of the IC are showed in Table S.1. [Bai and Ng \(2002\)](#) give 3 different information criteria (with different choices of

penalty), which are defined by:

$$\begin{aligned} \text{IC}_1(q) &= \ln [\mathbf{V}_n(q)] + q \left( \frac{n+d}{nd} \right) \ln \left( \frac{nd}{n+d} \right), \\ \text{IC}_2(q) &= \ln [\mathbf{V}_n(q)] + q \left( \frac{n+d}{nd} \right) \ln(n \wedge d), \\ \text{IC}_3(q) &= \ln [\mathbf{V}_n(q)] + q \left( \frac{1}{n \wedge d} \right) \ln(n \wedge d). \end{aligned}$$

We also consider using the ratio method (denoted by “Ratio”) to estimate the number of factors (c.f., [Lam and Yao, 2012](#)):

$$\hat{q} = \arg \max_{1 \leq q \leq \bar{Q}-1} \mu_q(\mathcal{X}_n \mathcal{X}_n^\top / (nd)) / \mu_{q+1}(\mathcal{X}_n \mathcal{X}_n^\top / (nd)), \quad (\text{F.3})$$

where  $\mu_q(\cdot)$  denotes the  $q$ -th largest eigenvalue. [Barigozzi, Cho and Fryzlewicz \(2018\)](#) suggest using the information criteria to obtain an initial estimate of the factor number, and then screening over a range of candidate factor numbers which are equal to or larger than the initial estimate, since over-estimation of the factor number does not affect consistency of break estimation. The candidate that gives the largest number of break points is used as the final estimate of the factor number. We denote such a method by “Screen” in Table S.1, where the first number is the number of factors estimated by the  $\text{IC}_2$  and the second number is the final estimate of the factor number.

Table S.1 presents the estimates of the factor number via different methods (averaged over 100 replications). The numbers in the parentheses are the corresponding standard errors. When the sample size and dimension are relatively small,  $(n, d) = (100, 100)$ , only “ $\text{IC}_1$ ” and “ $\text{IC}_2$ ” are able to accurately estimate the true number of transformed factors ( $q_0 = 10$ ). As the penalty term in the  $\text{IC}_3$  function is too small, the estimated factor number is equal to the upper bound  $\bar{Q} = 50$  used in the maximisation. In contrast, the “Ratio” method tends to under-estimate the transformed factor number with the estimated number close to the true number of factors in the original factor model ( $r = 5$ ). When the sample size or dimension increases, say  $(n, d) = (200, 500)$  or  $(n, d) = (500, 200)$ , most of the estimated factor numbers are closer to the true value 10. Note that the “ $\text{IC}_3$ ” method often performs poorly when the sample size is close to the dimension, say  $(n, d) = (100, 100)$  or  $(n, d) = (200, 200)$ . For the “Screen” method recommended by [Barigozzi, Cho and Fryzlewicz \(2018\)](#), the first estimated number is close to the true value, but its performance is slightly worse than the “ $\text{IC}_2$ ” method (as different R packages are used in the relevant eigen-decomposition). In addition, the final estimated number of factors in the “Screen” method is often close to 20. In the numerical studies reported in the main document, we use “ $\text{IC}_2$ ” to determine the number of factors.

## References

- ANDREWS, D.W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**, 821–856.
- BAI, J. (1997). Estimating multiple breaks on at a time. *Econometric Theory* **13**, 315–352.
- BAI, J. AND LI, K. (2012). Statistical analysis of factor models of high dimension. *Annals of Statistics* **40**, 436–465.
- 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.
- LAM, C. AND YAO, Q. (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *Annals of Statistics* **40**, 694–726.
- 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. (2018). Optimal covariance change point localization in high dimensions. Working paper available at <https://arxiv.org/pdf/1712.09912.pdf>.

Table S.1: The estimated number of factors for Example F.1

$\rho_f^c$	NoF	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen
1	10	10.00 (0.00)	9.99 (0.10)	50.00+ (0.00)	5.45 (1.44)	10.04/20.04 (0.20/0.49)	10.00 (0.00)	10.00 (0.00)	11.43 (1.28)	5.600 (1.63)	10.38/20.60 (0.57/1.57)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	6.10 (2.08)	10.79/21.52 (0.88/1.84)
0.5	10	10.00 (0.00)	10.00 (0.00)	50.00+ (0.00)	5.300 (1.19)	10.12/19.99 (0.33/1.66)	10.00 (0.00)	10.00 (0.00)	14.16 (2.34)	5.60 (1.63)	10.39/20.65 (0.63/1.67)	10.00 (0.00)	10.00 (0.00)	10.18 (0.41)	5.35 (1.28)	11.09/22.09 (1.13/2.40)
0.1	10	10.03 (0.17)	9.97 (0.17)	50.00+ (0.00)	5.30 (1.19)	10.26/20.48 (0.54/1.03)	10.03 (0.17)	10.00 (0.00)	48.55+ (5.80)	5.300 (1.19)	11.12/21.83 (1.12/3.00)	10.06 (0.24)	10.00 (0.00)	12.23 (1.98)	5.25 (1.10)	12.38/24.64 (1.44/3.40)
NoF		IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen
1	10	10.00 (0.00)	10.00 (0.00)	10.01 (0.10)	6.05 (2.05)	10.01/19.96 (0.10/0.40)	10.00 (0.00)	10.00 (0.00)	15.76 (2.58)	7.80 (2.49)	10.04/20.06 (0.20/0.45)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	9.85 (0.86)	10.22/20.32 (0.46/1.08)
0.5	10	10.00 (0.00)	10.00 (0.00)	10.04 (0.20)	6.00 (2.01)	10.01/19.88 (0.10/1.06)	10.00 (0.00)	10.00 (0.00)	29.75+ (11.03)	7.5 (2.51)	10.04/19.90 (0.20/1.18)	10.00 (0.00)	10.00 (0.00)	10.07 (0.26)	9.10 (1.93)	10.38/20.59 (0.65/1.65)
0.1	10	10.00 (0.00)	10.00 (0.00)	10.34 (0.50)	5.75 (1.80)	10.02/20.02 (0.14/0.35)	10.00 (0.00)	10.00 (0.00)	50.00+ (0.00)	7.05 (2.47)	10.05/20.03 (0.22/0.70)	10.00 (0.00)	10.00 (0.00)	10.93 (0.92)	8.30 (2.38)	10.55/20.88 (0.87/2.24)
NoF		IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen	IC <sub>1</sub>	IC <sub>2</sub>	IC <sub>3</sub>	Ratio	Screen
1	10	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	8.600 (2.26)	10.00/19.95 (0.00/0.50)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	9.95 (0.50)	10.00/19.83 (0.00/0.97)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	10.00/19.93 (0.00/0.54)
0.5	10	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	8.65 (2.23)	10.00/19.83 (0.00/1.13)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	9.90 (0.70)	10.00/19.73 (0.00/1.38)	10.00 (0.00)	10.00 (0.00)	10.78 (0.84)	10.00 (0.00)	10.01/19.72 (0.10/1.47)
0.1	10	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	8.40 (2.34)	10.00/19.76 (0.00/1.24)	10.00 (0.00)	10.00 (0.00)	10.00 (0.00)	9.80 (0.98)	10.00/19.81 (0.00/1.37)	10.00 (0.00)	10.00 (0.00)	16.43 (1.91)	10.00 (0.00)	10.01/19.81 (0.10/0.94)