Supplementary material for 'On the non-standard distribution of empirical likelihood estimators with spatial data'

Matthew Van Hala[a], Soutir Bandyopadhyay[b,*], Soumendra N. Lahiri[c], Daniel J. Nordman[a]

[a]*Department of Statistics, Iowa State University, Ames, IA 50011, U.S.A*
[b]*Department of Mathematics, Lehigh University, Bethlehem, PA 18015, U.S.A.*
[c]*Department of Statistics, North Carolina State University, Raleigh, NC 27695, U.S.A.*

Abstract

This supplement provides regularity conditions and a proof of the main distributional result for the spatial empirical likelihood (EL) point estimator. The supplement is organized as follows. Section A outlines the notation used, followed by a description of the regularity conditions in Section B. Section C briefly outlines some preliminary, independent results. Section D then provides proofs for characterizing the distributional properties of spatial EL estimator in the frequency domain (i.e., Theorem 1 of the main manuscript). Any citations mentioned here will be provided in a reference section of this supplement.

Appendix A. Notation

Recall the spatial periodogram $I_n(\boldsymbol{\omega})$ of $\{Z(\boldsymbol{s}_1),\ldots,Z(\boldsymbol{s}_n)\}$ at a frequency $\boldsymbol{\omega}\in\mathbb{R}^d$ as $I_n(\boldsymbol{\omega})=|d_n(\boldsymbol{\omega})|^2$, where $\imath=\sqrt{-1}$ and $d_n(\boldsymbol{\omega})=\lambda_n^{d/2}n^{-1}\sum_{j=1}^n Z(\boldsymbol{s}_j)\exp(\imath\boldsymbol{\omega}'\boldsymbol{s}_j)$. Let $c_n=n/\lambda_n^d$ (defined slightly *different* here than $c_n=\lambda_n^d/n$ in Section 3 of the main manuscript) and $b_n^2=Nc_n^{-2}+\lambda_n^{\kappa d}$, where $N=|\mathcal{N}|$ denotes the number of frequencies in the grid $\mathcal{N}=\left\{\boldsymbol{j}\lambda_n^{-\kappa}:\boldsymbol{j}\in\mathbb{Z}^d,\boldsymbol{j}\in[-C\lambda_n^\eta,-C\lambda_n^\eta]^d\right\}=\{\boldsymbol{\omega}_{jn}: j=1,\ldots,N\}$. We suppose the indexing is done so that $\boldsymbol{\omega}_{1n}=\mathbf{0}\in\mathbb{R}^d$. The

*Corresponding author
Email addresses: mvanhala@iastate.edu (Matthew Van Hala), sob210@lehigh.edu (Soutir Bandyopadhyay), snlahiri@ncsu.edu (Soumendra N. Lahiri), dnordman@iastate.edu (Daniel J. Nordman)

bias corrected periodogram is $\widetilde{I}_n(\boldsymbol{\omega}) = I_n(\boldsymbol{\omega}) - c_n^{-1}\widehat{\sigma}_n(\mathbf{0})$ for the sample variance $\widehat{\sigma}_n(\mathbf{0}) = n^{-1}\sum_{i=1}^n (Z(\boldsymbol{s}_i) - \bar{Z}_n)^2$, with $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z(\boldsymbol{s}_i)$ denoting the sample mean. Let $I_n^*(\boldsymbol{\omega}) = I_n(\boldsymbol{\omega}) - c_n^{-1}\sigma(\mathbf{0})$, and $A_n(\boldsymbol{\omega}) = c_n^{-1}\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega}), \boldsymbol{\omega} \in \mathbb{R}^d$, where $K = (2\pi)^d \int f^2$. Set $\Sigma_n = 2\sum_{k=1}^N G_{\theta_0}(\boldsymbol{\omega}_{kn})G_{\theta_0}(\boldsymbol{\omega}_{kn})' A_n(\boldsymbol{\omega}_{kn})^2$ at the true parameter θ_0. Let $\widehat{f}(\boldsymbol{\omega}) = \int e^{\imath \boldsymbol{x}'\boldsymbol{\omega}} f(\boldsymbol{x})d\boldsymbol{x}$ and $\widehat{f^2}(\boldsymbol{\omega}) = \int e^{\imath \boldsymbol{x}'\boldsymbol{\omega}} f^2(\boldsymbol{x})d\boldsymbol{x}$ for $\boldsymbol{\omega} \in \mathbb{R}^d$.

In the following, for a random quantity Y depending on both $\boldsymbol{X} \equiv \{\boldsymbol{X}_i\}_{i\geq 1}$ and $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, we let $\mathrm{E}Y \equiv \mathrm{E}_{\cdot|\boldsymbol{X}}$ denote expectation conditional on $\boldsymbol{X}$ and likewise let $P(\cdot) = P_{\cdot|\boldsymbol{X}}(\cdot)$ denote probability conditional with respect to $\boldsymbol{X}$. The notation $\xrightarrow{p}$, $\xrightarrow{d}$, $O_p(\cdot)$ and $o_p(\cdot)$ will represent convergence in probability and distribution as well as probabilistic orders in terms of this conditional probability. Also, we let $P_{\boldsymbol{X}}$ and $\mathrm{E}_{\boldsymbol{X}}$ denote probability and expectation under the joint distribution of $\{\boldsymbol{X}_i\}_{i\geq 1}$. Let C or $C(\cdot)$ denote generic constants that depend on their arguments (if any), but do not depend on n or $\{\boldsymbol{X}_i\}_{i\geq 1}$.

Appendix B. Regularity Conditions

We require some mild assumptions on the dependence of the process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, formulated in terms of mixing/moment conditions, as well as some assumptions on the estimating functions $G_\theta(\cdot)$. For $\boldsymbol{x} = (x_1, ..., x_k)' \in \mathbb{R}^k$, let $\|\boldsymbol{x}\| \equiv (\sum_{i=1}^k |x_i|^2)^{1/2}$, $\|\boldsymbol{x}\|_1 \equiv \sum_{i=1}^k |x_i|$ and, for $E_1, E_2 \subset \mathbb{R}^k$, let $d_1(E_1, E_2) = \inf\{\|\boldsymbol{x} - \boldsymbol{s}\|_1 : \boldsymbol{x} \in E_1, \boldsymbol{s} \in E_2\}$. For $a, b \in (0, \infty)$, define the strong mixing coefficient of $Z(\cdot)$ as

$$\alpha(a,b) = \sup_{A_i \in \mathcal{F}_Z(E_i),\ i=1,2} \left\{|P(A_1 \cap A_2) - P(A_1)P(A_2)| : E_i \in \mathbb{C}_b, d_1(E_1, E_2) \geq a\right\},$$

where $\mathcal{F}_Z(E)$ is the σ-field generated by $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in E\}$ and $\mathbb{C}_b$ is the collection of d-dimensional rectangles with volume b or less. We shall suppose that

$$\alpha(a,b) \leq \gamma_1(a)\gamma_2(b), \qquad a, b \in (0, \infty)$$

for some left continuous, non-increasing function $\gamma_1 : (0,\infty) \to [0,\infty)$ and some right continuous, non-decreasing function $\gamma_2 : (0,\infty) \to (0,\infty)$ (Lahiri , 2003).

Note that we allow the function $\gamma_2(\cdot)$ in the above formulation to grow to infinity to ensure validity of the results for bonafide strongly mixing random fields in $d \geq 2$ (Bradley , 1989, 1993; Lahiri , 2003). We again assume that $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ is second order stationary with mean zero and spectral density $\phi(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \mathbb{R}^d$ and $\sigma(\mathbf{0}) = \mathrm{Var}[Z(\mathbf{0})]$. We write $c_n = n/\lambda_n^d$, recalling that $c_* = \lim_{n\to\infty} c_n$ determines the spatial asymptotic structure (Section 2.1) with $c_* \in (0,\infty)$ under pure increasing domain and $c_* = \infty$ under mixed increasing domain. Recall $b_n^2 = Nc_n^{-2} + \lambda_n^{\kappa d}$. We next list regularity conditions for establishing the empirical likelihood results. Recall $\theta_0 \in \mathbb{R}^p$ denotes the true parameter satisfying (1).

Conditions

(C.1): There exists $\delta \in (0,1]$ such that

$$\zeta_{8+\delta} \equiv \sup\{(\mathrm{E}|Z(\boldsymbol{s})|^{8+\delta})^{\frac{1}{8+\delta}} : \boldsymbol{s} \in \mathbb{R}^d\} \text{ and } \sum_{k=1}^{\infty} k^{7d}[\gamma_1(k)]^{\frac{\delta}{8+\delta}} < \infty.$$

(C.2): (i) The spatial sampling density $f(\cdot)$ is everywhere positive on $\mathcal{D}_0$ and satisfies a Lipschitz condition: for some $C_0 \in (0,\infty)$, $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq C_0\|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}_0$.

(ii) There exist $C_1 \in (0,\infty)$ and $a_0 \in (d/2, d]$ such that for all $\|\boldsymbol{\omega}\| > C_1$,

$$\left|\int e^{\imath\boldsymbol{\omega}'\boldsymbol{x}} f(\boldsymbol{x}) d\boldsymbol{x}\right| + \left|\int e^{\imath\boldsymbol{\omega}'\boldsymbol{x}} f^2(\boldsymbol{x}) d\boldsymbol{x}\right| \leq C_1\|\boldsymbol{\omega}\|^{-a_0}.$$

(C.3): (i) $G_{\theta_0}(\cdot)$ is bounded, symmetric, and almost everywhere continuous on $\mathbb{R}^d$ (with respect to the Lebesgue measure) with $\int G_{\theta_0}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \mathbf{0}_r$;

(ii) There exist $C_2 \in (0,\infty)$ and a non-increasing function $h : [0,\infty) \to [0,\infty)$ such that $|\phi(\boldsymbol{\omega})| \leq h(\|\boldsymbol{\omega}\|)$ for all $\|\boldsymbol{\omega}\| > C_2$;

(iii) $\liminf_{n\to\infty} \det\left(N^{-1}\sum_{k=1}^N G_{\theta_0}(\boldsymbol{\omega}_{kn})G'_{\theta_0}(\boldsymbol{\omega}_{kn})\right) > 0$;

(iv) $\int G_{\theta_0}(\boldsymbol{\omega}_{in})G'_{\theta_0}(\boldsymbol{\omega}_{in})\phi(\boldsymbol{\omega})^2\, d\boldsymbol{\omega}$ is nonsingular.

(C.4): (i) $0 < \kappa < \eta < 1$ and (ii) $\Sigma_n^{-1/2}\sum_{k=1}^N G_{\theta_0}(\boldsymbol{\omega}_{kn})[I_n(\boldsymbol{\omega}_{kn}) - c_n^{-1}\sigma(\mathbf{0})] \xrightarrow{d} N(\mathbf{0}_r, I_{r\times r})$ for $\Sigma_n = 2\sum_{k=1}^N G_{\theta_0}(\boldsymbol{\omega}_{kn})G_{\theta_0}(\boldsymbol{\omega}_{kn})'[c_n^{-1}\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega}_{kn})]^2$, $K = (2\pi)^d\int f^2$.

(C.5): $\lambda_n^{-\kappa d} b_n^{3/4} N^{3/8}\log\lambda_n + N(n^{-1/2+\varepsilon} + \lambda_n^{-d/2})\lambda^{-\kappa d} = o(1)$ as $n \to \infty$ for some $\varepsilon > 0$. (C.6): For each $n \geq 1$, there exists a function $M_n(\cdot)$ where

$$\left\| \sum_{k=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{kn}) G_{\theta_0}(\boldsymbol{\omega}_{kn})' \exp(\imath \boldsymbol{t}' \boldsymbol{\omega}_{kn}) \right\| \leq M_n(\boldsymbol{t}) \text{ for all } \boldsymbol{t} \in \mathbb{R}^d,$$

such that, with $dv(\boldsymbol{t}, \boldsymbol{x}) = \|\boldsymbol{t}\| f(\boldsymbol{x}) \left[\gamma_1(\|\boldsymbol{t}\|)\right]^{\delta/(8+\delta)} d\boldsymbol{t}\, d\boldsymbol{x}$ and $\delta > 0$ from Condition (C.1) and for any $a_1, a_2, a_3 \in \{0, 1\}$,

$$\int \int M_n \left(\boldsymbol{t} + a_1 \left[\boldsymbol{s} + 2\lambda_n a_2 \boldsymbol{x} + 2\lambda_n a_3 \boldsymbol{y}\right]\right) d\nu(\boldsymbol{t}, \boldsymbol{x}) d\nu(\boldsymbol{s}, \boldsymbol{y}) = o\left(\lambda_n^{1+\kappa d}\right).$$

(C.7): In a neighborhood of θ_0 and for each $\boldsymbol{\omega} \in \mathbb{R}^d$, $h_\theta(\boldsymbol{\omega}) \equiv \partial G_\theta(\boldsymbol{\omega})/\partial\theta$ is continuous in θ and $\|\partial^2 G_\theta(\boldsymbol{\omega})/\partial\theta\partial\theta'\| \leq C$ for some $C > 0$; and $h_{\theta_0}(\cdot)$ is continuous almost everywhere.

(C.8): For $n \geq 1$, there exist functions $\widetilde{M}_n^{(1)}(\cdot)$ and $\widetilde{M}_n^{(2)}(\cdot)$ where, for all $\boldsymbol{t} \in \mathbb{R}^d$,

$$\begin{aligned} \left\| \sum_{j=1}^{N} h_{\theta_0}(\boldsymbol{\omega}_{jn}) \exp(\imath \boldsymbol{t}' \boldsymbol{\omega}_{jn}) \right\| &\leq \widetilde{M}_n^{(1)}(\boldsymbol{t}), \\ \left\| \sum_{j=1}^{N} h_{\theta_0}(\boldsymbol{\omega}_{jn}) h_{\theta_0}(\boldsymbol{\omega}_{jn})' \exp(\imath \boldsymbol{t}' \boldsymbol{\omega}_{jn}) \right\| &\leq \widetilde{M}_n^{(2)}(\boldsymbol{t}), \end{aligned}$$

such that with $dv(\boldsymbol{t}, \boldsymbol{x}) = \|\boldsymbol{t}\| f(\boldsymbol{x}) \left[\gamma_1(\|\boldsymbol{t}\|)\right]^{\delta/(8+\delta)} d\boldsymbol{t}\, d\boldsymbol{x}$ and $\delta > 0$ from Condition (C.1), $\int \widetilde{M}_n^{(1)}(\boldsymbol{t})\, dv(\boldsymbol{t}, \boldsymbol{x}) = o(\lambda_n^{1+\kappa d})$ and, for any $a_1, a_2, a_3 \in \{0, 1\}$,

$$\int \int \widetilde{M}_n^{(2)} \left(\boldsymbol{t} + a_1 \left[\boldsymbol{s} + 2\lambda_n a_2 \boldsymbol{x} + 2\lambda_n a_3 \boldsymbol{y}\right]\right) dv(\boldsymbol{t}, \boldsymbol{x})\, dv(\boldsymbol{s}, \boldsymbol{y}) = o\left(\lambda_n^{1+2\kappa d}\right).$$

Regarding the assumptions, Conditions (C.1)-(C.6) are essentially those of Bandyopadhyay et al. (2015a), with a slight strengthening of the number of finite moments. Condition (C.1) is a standard moment/mixing condition ensuring the periodogram has a fourth moment. Condition (C.2) are smoothness conditions on the location density f and the Fourier transforms of f and f^2; uniform and many smooth non-uniform densities f satisfy (C.2)(ii) with a decay rate $O(\|\omega\|^{-d})$. Condition (C.3) provides regularity conditions on the spectral estimating function G_{θ_0} at the true parameter θ_0. These conditions also ensure that certain Riemann sums over the frequency grid $\{\boldsymbol{\omega}_{kn}\}_{k=1}^{N}$ approximate a variance integral $\int G_{\theta_0}(\boldsymbol{\omega}) G'_{\theta_0}(\boldsymbol{\omega}) d\boldsymbol{\omega}$ asymptotically and that the $r \times r$ matrix Σ_n has a nonsingular limit along a sub-sequence. Considering Condition (C.4), as explained in Section 3, the choice $0 < \kappa < \eta < 1$ ensures that periodogram values $\{I_n(\boldsymbol{\omega}_{kn})\}_{k=1}^N$ will be asymptotically independent on frequency gird $\{\boldsymbol{\omega}_{kn}\}_{k=1}^N$ (consisting of asymptotically distance frequencies, Section 2.2). Consequently, the sum $\sum_{k=1}^N G_{\theta_0}(\boldsymbol{\omega}_{kn})[I_n(\boldsymbol{\omega}_{kn}) - c_n^{-1}\sigma(\mathbf{0})]$, involving a biased-corrected raw periodogram (cf. Section 2.1), can be expected to have a normal limit with mean zero under the spectral moment condition (1) for $G_{\theta_0}(\cdot)$. The central limit theorem statement in Condition (C.4) is a primitive one, and further sufficient conditions for such central limit theorems can be found in Bandyopadhyay et al. (2015b). Moving to Condition (C.5), the rate bounds involved depend on a technical quantity $b_n^2 = Nc_n^{-2} + \lambda_n^{\kappa d}$ whose order can change depending on the asymptotic frameworks and, as explained in the next section, different asymptotic regimes induce varying behavior in empirical likelihood statistics related to b_n^2. This condition can hold trivially in many cases, but we include this to minimize technicalities. The differentiability conditions on estimating functions in Condition (C.7) are standard in empirical likelihood frameworks (Qin and Lawless , 1994; Kitamura , 1997). However, the bounds in conditions (C.6)-(C.8) are also technical and related to certain Fourier transforms involving estimating functions, or their partial derivatives, over the discrete frequency grid $\{\boldsymbol{\omega}_{kn}\}_{k=1}^N$; these bounds ensure that certain remainders in the frequency domain arising from Taylor expansions are negligible. It can be verified that the estimating functions considered in Section 4 of the main manuscript, for example, satisfy (C.6)-(C.8).

Appendix C. Preliminary Technical Results

Here we collect some independent, technical results (Lemmas 1-3 to follow) regarding expansions and convergence properties of the spatial periodogram; proofs of these can be found in Van Hala et al. (2015).

Lemma 1 provides probabilistic bounds and expansions for sums of bias corrected periodogram values. To state the result, $A_n(\boldsymbol{\omega}) = c_n^{-1}\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega})$, where $K = (2\pi)^d \int f^2$, for $\boldsymbol{\omega} \in \mathbb{R}^d$.

Lemma 1. *Under Conditions (C.1), (C.2), (C.3), (C.5) and (C.6),*

(i) $\frac{1}{b_n^2}\sum_{j=1}^{N} \tilde{I}_n^2(\boldsymbol{\omega}_{jn}) = O_p(1)$ *a.s.* $(P_{\mathbf{X}})$.

(ii) $\frac{1}{b_n^2}\sum_{j=1}^{N} \tilde{I}_n^4(\boldsymbol{\omega}_{jn}) = O_p(1)$ *a.s.* $(P_{\mathbf{X}})$.

(iii) $\sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jn})G'_{\theta_0}(\boldsymbol{\omega}_{jn})\big[\tilde{I}_n^2(\boldsymbol{\omega}_{jn}) - \big(A_n(\boldsymbol{\omega}_{jn})^2 + K^2\phi(\boldsymbol{\omega}_{jn})^2\big)\big] = o_p(b_n^2)$ *a.s.* $(P_{\mathbf{X}})$,

$\sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jn})G'_{\theta_0}(\boldsymbol{\omega}_{jn})\big[I_n^2(\boldsymbol{\omega}_{jn}) - 2A_n(\boldsymbol{\omega}_{jn})^2\big] = o_p(b_n^2)$ *a.s.* $(P_{\mathbf{X}})$.

Lemma 2 next shows the convergence of Riemann sums of partial derivatives of estimating functions, with proper scaling. For $\theta \in \Theta$, define

$$D_{n,\theta} = \lambda_n^{-\kappa d} K \sum_{j=1}^{N} h_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn}) \tag{A.1}$$

for $h_\theta(\boldsymbol{\omega}) \equiv \partial G_\theta(\boldsymbol{\omega})/\partial\theta$, $\boldsymbol{\omega} \in \mathbb{R}^d$ and $K \equiv (2\pi)^d \int f^2$, where the partials exist in a neighborhood of θ_0.

Lemma 2. *Assume Conditions (C.1), (C.2), (C.3), (C.5), (C.7), and (C.8). Let* $D_{\theta_0} \equiv \int h_{\theta_0}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})d\boldsymbol{\omega}$.

(i) Then, $D_{n,\theta_0} \xrightarrow{p} D_{\theta_0}$ *as* $n \to \infty$ *a.s.* $(P_{\mathbf{X}})$.

(ii) For $B_n \equiv \{\theta \in \Theta : \|\theta - \theta_0\| \le \lambda_n^{-\kappa d} b_n \log \lambda_n\}$,

$$\sup_{\theta \in B_n} \|D_{n,\theta} - D_{\theta_0}\| \xrightarrow{p} 0 \quad a.s.\ (P_{\mathbf{X}}).$$

Finally, Lemma 3 concerns the positivity of EL function $\mathcal{R}_n(\theta)$, $\theta \in \Theta \subset \mathbb{R}^p$, (cf. Section 2.2) in a neighborhood around θ_0 defined by

$$\Theta_n \equiv \left\{\theta \in \Theta : \|\theta - \theta_0\| \le b_n \lambda_n^{-\kappa d} \log \lambda_n\right\}.$$

Recall that $b_n \lambda_n^{-\kappa d} \log \lambda_n \to 0$ as $n \to \infty$ under condition (C.5).

Lemma 3. *Under the assumptions of Theorem 1,* $P(\mathcal{R}_n(\theta) > 0$ *for* $\theta \in \Theta_n) \to 1$ *as* $n \to \infty$ *(a.s.* $(P_{\mathbf{X}})$*).*

Appendix D. Proof of Main Result (Theorem 1)

Section D.1 first provides some supporting technical results. Section D.2 then shows the existence of the the spatial EL point estimator $\widehat{\theta}_n$ (i.e., the maximizer of the EL function $\mathcal{R}_n(\theta)$) and establishes its limit distribution along different subsequence. The limit law will depend on the subsequence as well as the spatial sampling regime. Recall PID or MID refer to pure increasing domain or mixed increasing domain spatial asymptotics; see Section 3.2 of the manuscript for "slow rate" and "fast rate" MID definitions.

Appendix D.1. Background Results

We first require a distributional result which refines the CLT result in Condition (C.4) $\Sigma_n^{-1/2}\sum_{j=1}^N G_{\theta_0}(\boldsymbol{\omega}_{jn})I_n^*(\boldsymbol{\omega}_{jn}) \xrightarrow{d} N(\mathbf{0}_r, \mathbb{I}_r)$ a.s. ($P_{\mathbf{X}}$), for $I_n^*(\boldsymbol{\omega}) = I_n(\boldsymbol{\omega}_{jn}) - c_n^{-1}\sigma(\mathbf{0})$, $\boldsymbol{\omega} \in \mathbb{R}^d$, and $\Sigma_n = 2\sum_{j=1}^N G_{\theta_0}(\boldsymbol{\omega}_{jn})G'_{\theta_0}(\boldsymbol{\omega}_{jn})A_n(\boldsymbol{\omega}_{jn})^2$.

With $b_n^2 = Nc_n^{-2} + \lambda_n^{\kappa d}$, define

$$J_{n,\theta_0} = \frac{1}{b_n^2}\sum_{j=1}^N G_{\theta_0}(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn}), \quad W_{n,\theta_0} = \frac{1}{b_n^2}\sum_{j=1}^N G_{\theta_0}(\boldsymbol{\omega}_{jn})G'_{\theta_0}(\boldsymbol{\omega}_{jn})\tilde{I}_n^2(\boldsymbol{\omega}_{jn}). \tag{A.1}$$

Lemma 4. *Under Conditions (C.1), (C.2), (C.3), (C.4) and (C.6). Given any subsequence $\{n_j\} \subset \{n\}$, extract a further subsequence $\{k \geq n_k\} \subset \{n_j\}$ such that*

$$\frac{1}{N_k}\sum_{j=1}^{N_k} G_{\theta_0}(\boldsymbol{\omega}_{jk})G'_{\theta_0}(\boldsymbol{\omega}_{jk}) \to \Gamma^* \equiv \Gamma^*(n_k) \tag{A.2}$$

for a nonsingular $r \times r$ Γ^*. Then, as $k \to \infty$, it holds that

$$W_{k,\theta_0} \xrightarrow{p} V, \qquad b_k J_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, aV) \quad a.s.\ (P_{\mathbf{X}}) \tag{A.3}$$

for a positive definite matrix V and a constant $a \in \{1, 2\}$ as defined according to the following cases:

PID $b_n^2 \sim Nc_n^{-2}$, where $c_n^{-1} = n/\lambda^d \to c^* \in (0,\infty)$: $V = \sigma^2(\mathbf{0})\Gamma^*$, $a = 2$;
"slow infill" MID $b_n^2 \sim Nc_n^{-2}$, where $c_n^{-1} = n/\lambda^d \to 0$ and $\lambda^{\kappa d} \ll Nc_n^{-2}$: $V = \sigma^2(\mathbf{0})\Gamma^*$, $a = 2$;
"fast infill" MID $b_n^2 \sim \lambda_n^{-\kappa d}$, where $c_n^{-1} = n/\lambda^d \to 0$ and $Nc_n^{-2} \ll \lambda^{\kappa d}$: $V = 2\Gamma$, $a = 1$ *for*

$$\Gamma \equiv \int_{\mathbb{R}^d} G_{\theta_0}(\boldsymbol{\omega}_{jn}) G'_{\theta_0}(\boldsymbol{\omega}_{jn}) K^2 \phi^2(\boldsymbol{\omega})\, d\boldsymbol{\omega}.$$

Remark: A subsequence for which (A.2) always exists under (C.3)(iii) though the matrix $\Gamma^* \equiv \Gamma^*(\{n_k\})$ may change with the subsequence $k = n_k$.

Proof. Note that

$$\lambda_n^{-\kappa d} \sum_{j=1}^{N} \phi(\boldsymbol{\omega}_{jn}) \to \int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\, d\boldsymbol{\omega}$$

by the dominated convergence theorem from $\int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} < \infty$ and (C.3)(ii), so that

$$\sum_{j=1}^{N} \phi(\boldsymbol{\omega}_{jn}) = O(\lambda_n^{\kappa d}) = O(b_n^2), \qquad \sum_{j=1}^{N} \phi^2(\boldsymbol{\omega}_{jn}) = O(\lambda_n^{\kappa d}) = O(b_n^2), \qquad \text{(A.4)}$$

where the latter follows from $\sup_{1 \le j \le N} |\phi(\boldsymbol{\omega}_{jn})| \le C$ under (C.3)(ii).

In the PID case, we have $b_n^2 \sim N c_*^{-2}$ and

$$\left\| \Sigma_n - 2c_*^{-2} \sigma^2(\mathbf{0}) \sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jk}) G'_{\theta_0}(\boldsymbol{\omega}_{jk}) \right\| = o(N)$$

by (C.3), (A.4) and $\|G_{\theta_0}(\boldsymbol{\omega})\| \le C$, $\boldsymbol{\omega} \in \mathbb{R}^d$. Then, under the subsequence (A.2), we have for $V = \sigma^2(\mathbf{0})\Gamma^*$ that $\|(2b_k^2)^{-1}\Sigma_k - V\| \to 0$ so that

$$\|W_{k,\theta_0} - V\| \le \|W_{k,\theta_0} - (2b_k^2)^{-1}\Sigma_k\| + \|(2b_k^2)^{-1}\Sigma_k - V\| = o_p(1)$$

using $\|W_{k,\theta_0} - (2b_k^2)^{-1}\Sigma_k\| = o_p(1)$ by Lemma 1(iii) and (A.4). From $\Sigma_k / b_k^2 \to 2V$ and $T^*_{k,\theta_0} \equiv \Sigma_k^{-1/2} \sum_{j=1}^{N_k} G_{\theta_0}(\boldsymbol{\omega}_{jk}) I_k^*(\boldsymbol{\omega}_{jk}) \xrightarrow{d} N(\mathbf{0}_r, \mathbb{I}_r)$ by (C.4), we have

$$b_k J_{k,\theta_0} = \frac{1}{b_k} \Sigma_k^{1/2} T^*_{k,\theta_0} + R_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, 2V)$$

since, by $|\hat{\sigma}_k(\mathbf{0}) - \sigma(\mathbf{0})| = O_p(\lambda_k^{-d/2})$,

$$R_{k,\theta_0} \equiv \|b_k J_{k,\theta_0} - b_k^{-1} \Sigma_k^{1/2} T^*_{k,\theta_0}\| \le b_k^{-1} N_k c_k^{-1} O_p(\lambda_k^{-d/2}) = o_p(1) \qquad \text{(A.5)}$$

as $N_k c_k^{-1} / b_k = O(N_k^{1/2}) = O(\lambda_k^{\eta d/2})$ here with $\eta < 1$.

In the first MID case, we have $b_n^2 \sim Nc_n^{-2}$ and

$$\left\| \Sigma_n - 2c_n^{-2}\sigma^2(\mathbf{0}) \sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jk}) G'_{\theta_0}(\boldsymbol{\omega}_{jk}) \right\| = o(b_n^2)$$

again by (C.3) and (A.4). Under the subsequence (A.2) and for $V = \sigma^2(\mathbf{0})\Gamma^*$ that $\|(2b_k^2)^{-1}\Sigma_k - V\| \to 0$ so that $\|W_{k,\theta_0} - V\| = o_p(1)$ by $\|W_{k,\theta_0} - (2b_k^2)^{-1}\Sigma_k\| = o_p(1)$ from Lemma 1(iii) and (A.4). As $\Sigma_k / b_k^2 \to 2V$ and $T^*_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, \mathbb{I}_r)$ by (C.4), we have $b_k J_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, 2V)$ where, analogous to (A.5), $R_{k,\theta_0} \equiv \|b_k J_{k,\theta_0} - b_k^{-1}\Sigma_k^{1/2} T^*_{k,\theta_0}\| \leq b_k^{-1} N_k c_k^{-1} O_p(\lambda_k^{-d/2}) = o_p(1)$ as $N_k c_k^{-1}/b_k = O(N_k^{1/2}) = O(\lambda_k^{\eta d/2})$ with $\eta < 1$.

In the second MID case, $b_n^2 \sim \lambda_n^{-\kappa d}$, $Nc_n^{-2} \ll \lambda_n^{\kappa d}$, holds and, by (C.3) and the Dominated Convergence Theorem, it follows that

$$\|\Sigma_n - 2\lambda_n^{\kappa d}\Gamma\| = o(b_n^2)$$

for $\Gamma \equiv \int_{\mathbb{R}^d} G_{\theta_0}(\boldsymbol{\omega}_{jn}) G'_{\theta_0}(\boldsymbol{\omega}_{jn}) K^2 \phi^2(\boldsymbol{\omega})\, d\boldsymbol{\omega}$. Under the subsequence (A.2) and for $V = 2\Gamma$ that $\|b_k^{-2}\Sigma_k - V\| \to 0$ so that $\|W_{k,\theta_0} - V\| = o_p(1)$ by $\|W_{k,\theta_0} - b_k^{-2}\Sigma_k\| = o_p(1)$ from Lemma 1(iii), $N_k c_k^{-2} \ll \lambda_k^{\kappa d}$ and the boundedness of $G_{\theta_0}(\cdot)$. Then, from $\Sigma_k / b_k^2 \to V$ (not $2V$ as in previous cases) and $T^*_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, \mathbb{I}_r)$ by (C.4), it follows that $b_k J_{k,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, V)$ where, again analogous to (A.5), $R_{k,\theta_0} \equiv \|b_k J_{k,\theta_0} - b_k^{-1}\Sigma_k^{1/2} T^*_{k,\theta_0}\| \leq b_k^{-1} N_k c_k^{-1} O_p(\lambda_k^{-d/2}) = o_p(1)$ from $N_k c_k^{-1}/b_k = o(N_k^{1/2}) = o(\lambda_k^{\eta d/2})$ by $N_k c_k^{-2} \ll \lambda_k^{\kappa d} \sim b_k^2$ with $\eta < 1$. $\square$

The next result re-states the main distributional finding of Bandyopadhyay et al. (2015a) for spatial EL, which shows the chi-square limit of the log-EL function at the true value of θ_0 (without consideration of point estimation); see Bandyopadhyay et al. (2015a) for its proof.

Lemma 5. *Under Conditions (C.1)-(C.8), as $n \to \infty$*

$$-a \log \mathcal{R}_n(\theta_0) \xrightarrow{d} \chi_r^2 \quad a.s.\ (P_{\mathbf{X}})$$

where $a = 1$ if $b_n^2 \sim Nc_n^{-2}$ (i.e., $\lambda_n^{-\kappa d} \ll Nc_n^{-2}$) and $a = 2$ if $b_n^2 \sim \lambda_n^{-\kappa d}$ (i.e., $Nc_n^{-2} \ll \lambda_n^{-\kappa d}$).

The last result of this section collects some useful distributional results, regarding extensions of the statistics J_{n,θ_0} and W_{n,θ_0} from (A.1) and their behaviors on shrinking neighborhoods of the true parameter θ_0. With $b_n^2 = Nc_n^{-2} + \lambda_n^{\kappa d}$ again, define

$$J_{n,\theta} \equiv \frac{1}{b_n^2} \sum_{j=1}^{N} G_\theta(\boldsymbol{\omega}_{jn}) \tilde{I}_n(\boldsymbol{\omega}_{jn}), \quad W_{n,\theta} \equiv \frac{1}{b_n^2} \sum_{j=1}^{N} G_\theta(\boldsymbol{\omega}_{jn}) G'_\theta(\boldsymbol{\omega}_{jn}) \tilde{I}_n^2(\boldsymbol{\omega}_{jn}), \tag{A.6}$$

and let $Z_{n,\theta} \equiv \max_{1 \le j \le N} \left\| G_\theta(\boldsymbol{\omega}_{jn}) \tilde{I}_n(\boldsymbol{\omega}_{jn}) \right\|$ for $\theta \in \Theta$. Let

$$\Theta_n \equiv \left\{ \theta \in \Theta : \|\theta - \theta_0\| \le b_n \lambda_n^{-\kappa d} \log \lambda_n \right\},$$

where $b_n \lambda_n^{-\kappa d} \log \lambda_n \to 0$ as $n \to \infty$ under condition (C.5). Define $\tau_{n,\theta} \equiv \max\{\lambda_n^{-\kappa d} b_n, \|\theta - \theta_0\|\}$ for $\theta \in \Theta$.

Lemma 6. *Under Conditions (C.1)-(C.8), as $n \to \infty$,*
(i) $\sup_{\theta \in \Theta_n} b_n \|J_{n,\theta}\| = O_p(\log \lambda_n)$ *(a.s. ($P_{\mathbf{X}}$)).*
(ii) $\sup_{\theta \in \Theta_n} \tau_{n,\theta}^{-1} \|J_{n,\theta}\| = O_p(1)$ *(a.s. ($P_{\mathbf{X}}$)).*
(iii) $\sup_{\theta \in \Theta_n} \|W_{n,\theta} - W_{n,\theta_0}\| = o_p(1)$ *(a.s. ($P_{\mathbf{X}}$)).*
(iv) $\sup_{\theta \in \Theta_n} \tau_{n,\theta} Z_{n,\theta} = o_p(1)$ *(a.s. ($P_{\mathbf{X}}$)).*

Proof. Consider part (i). By Lemma 4, any subsequence of $\{n\}$ has a further subsequence $\{n_k\}$ along which $b_{n_k} J_{n_k,\theta_0}$ has a normal limit. Hence, $b_n J_{n,\theta_0} = O_p(1)$ is tight. Then, for a given $\theta \in \Theta_n$, by Taylor expansion of $J_{n,\theta}$ around θ_0, for $\theta \in \Theta_n$ and $D_{n,\theta} \equiv \lambda_n^{-\kappa d} \sum_{j=1}^{N} (\partial G_\theta(\boldsymbol{\omega}_{jn}) / \partial \theta) I_n(\boldsymbol{\omega}_{jn})$, we have

$$J_{n,\theta} = J_{n,\theta_0} + \frac{\lambda_n^{\kappa d}}{b_n^2} D_{n,\theta_0 + c(\theta - \theta_0)} (\theta - \theta_0) \tag{A.7}$$

for some $c \in [0, 1]$ (depending on θ). Then,

$$\sup_{\theta \in \Theta_n} \|J_{n,\theta}\| \le \|J_{n,\theta_0}\| + \frac{\lambda_n^{\kappa d}}{b_n^2} \sup_{\theta \in \Theta_n} \|D_{n,\theta}\| \sup_{\theta \in \Theta_n} \|\theta - \theta_0\|$$

As $\sup_{\theta \in \Theta_n} \|D_{n,\theta}\| = O_p(1)$ by Lemma 2(ii) and $\|J_{n,\theta_0}\| = O_p(b_n^{-1})$, we have

$$\sup_{\theta \in \Theta_n} b_n \|J_{n,\theta}\| \le O_p(1) + b_n \frac{\lambda_n^{\kappa d}}{b_n^2} O_p(1) b_n \lambda_n^{-\kappa d} \log \lambda_n = O_p(\log \lambda_n).$$

For part (ii), note that $\tau_{n,\theta}^{-1} \leq b_n^{-1}\lambda_n^{\kappa d}$. Using a similar Taylor expansion (A.7), we have

$$\begin{aligned} \sup_{\theta\in\Theta_n} \frac{\|J_{n,\theta}\|}{\tau_{n,\theta}} &\leq b_n^{-1}\lambda^{\kappa d}\|J_{n,\theta_0}\| + \frac{\lambda_n^{\kappa d}}{b_n^2} \sup_{\theta\in\Theta_n} \|D_{n,\theta}\| \sup_{\theta\in\Theta_n} \frac{\|\theta-\theta_0\|}{\tau_{n,\theta}} \\ &\leq \lambda_n^{\kappa d} b_n^{-1} O_p\left(b_n^{-1}\right) + \frac{\lambda_n^{\kappa d}}{b_n^2} O_p(1) = O_p(1) \end{aligned}$$

using that $\|J_{n,\theta_0}\| = O_p(b_n^{-1})$, $\sup_{\theta\in\Theta_n}\|D_{n,\theta}\| = O_p(1)$ and $\lambda^{\kappa d}/b_n^2 = O(1)$.

To show part (iii), note that by Taylor expansion of $G_\theta(\boldsymbol{\omega}_{jn})$ around θ_0, we have

$$G_\theta(\boldsymbol{\omega}_{jn}) = G_{\theta_0}(\boldsymbol{\omega}_{jn}) + (\theta-\theta_0)\frac{\partial G_\theta(\boldsymbol{\omega}_{jn})}{\partial\theta} + S_n(\boldsymbol{\omega}_{jn}),$$

where $\|S_n(\boldsymbol{\omega}_{jn})\| \leq C_0\|\theta-\theta_0\|^2$ for some C_0 bounding $\|G_\theta(\cdot)\|, \|\partial G_\theta(\cdot)/\partial\theta\|$, and $\left\|\partial^2 G_{\theta_0}(\cdot)/\partial\theta\partial\theta'\right\|$ over Θ_n. Hence, it holds that

$$\begin{aligned} &\sup_{\theta\in\Theta_n} \|W_{n,\theta} - W_{n,\theta_0}\| \\ &\leq \sup_{\theta\in\Theta_n} b_n^{-2} \sum_{k=1}^{N} \left(\|G_\theta(\boldsymbol{\omega}_{jn})\| + \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|\right) \|G_\theta(\boldsymbol{\omega}_{jn}) - G_{\theta_0}(\boldsymbol{\omega}_{jn})\| \tilde{I}_n^2(\boldsymbol{\omega}_{jn}) \\ &\leq \frac{4C_0^2}{b_n^2} \sup_{\theta\in\Theta_n} (1+\|\theta-\theta_0\|)\|\theta-\theta_0\| \sum_{j=1}^{N} \tilde{I}_n^2(\boldsymbol{\omega}_{jn}) \\ &= O(b_n\lambda_n^{-\kappa d}\log\lambda_n)O_p(1) = o_p(1), \end{aligned}$$

by Lemma 1(i).

Considering part (iv), note first that, by Lemma 1(ii) and the bounded $G_{\theta_0}(\cdot)$, $\mathrm{E}Z_{n,\theta_0} \leq \left(\mathrm{E}\sum_{j=1}^N \left\|G_{\theta_0}(\boldsymbol{\omega}_{jn})^4 \tilde{I}_n^4(\boldsymbol{\omega}_{jn})\right\|\right)^{1/4} = O(b_n^{1/2})$, so that

$$Z_{n,\theta_0} = O_p(b_n^{1/2}) \qquad \text{a.s. } (P_{\mathbf{X}}). \tag{A.8}$$

Using a Taylor expansion of $G_\theta(\cdot)$ as in the proof of part(iii) and $\tau_{n,\theta} \leq$

$\lambda_n^{-\kappa d} b_n \log \lambda_n$,

$$\begin{aligned}
&\sup_{\theta\in\Theta_n} \tau_{n,\theta} Z_{n,\theta} \\
&\quad \leq \lambda_n^{-\kappa d} b_n (\log \lambda_n) \sup_{\theta\in\Theta_n} Z_{n,\theta} \\
&\quad \leq \lambda_n^{-\kappa d} b_n (\log \lambda_n) \left(Z_{n,\theta_0} + \sup_{\theta\in\Theta_n} C_0 \|\theta - \theta_0\| \max_{1\leq j\leq N} \tilde{I}_n(\boldsymbol{\omega}_{jn}) \right) \\
&\quad \leq \lambda_n^{-\kappa d} b_n^{3/2} (\log \lambda_n) \frac{Z_{n,\theta_0}}{b_n^{1/2}} + C_0 \lambda_n^{-\kappa d} b_n (\log \lambda_n) \sup_{\theta\in\Theta_n} \|\theta - \theta_0\| \left(\sum_{j=1}^{N} \tilde{I}_n^2(\boldsymbol{\omega}_{jn}) \right)^{1/2} \\
&\quad = o(1) O_p(1) + O([\lambda_n^{-\kappa d} b_n \log \lambda_n]^2) O_p(b_n) = o(1)
\end{aligned}$$

(a.s. ($P_{\mathbf{X}}$)) from Lemma 1(i), (A.8) and $\lambda_n^{-\kappa d} b_n^{3/2} \log \lambda_n = o(1)$ by (C.5). □

Appendix D.2. Distributional Properties of the Spatial Maximum EL Estimator

We next show that a non-trivial maximizer of the EL function $\mathcal{R}_n(\cdot)$ is guaranteed to exist. For $\Theta_n \equiv \left\{\theta \in \Theta : \|\theta - \theta_0\| \leq b_n \lambda_n^{-\kappa d} \log \lambda_n \right\}$, define its interior $\Theta_n^\circ = \left\{\theta \in \Theta : \|\theta - \theta_0\| < b_n \lambda_n^{-\kappa d} \log \lambda_n \right\}$ and its boundary $\partial\Theta_n = \left\{\theta \in \Theta : \|\theta - \theta_0\| = b_n \lambda_n^{-\kappa d} \log \lambda_n \right\}$. For $\theta \in \Theta_n$ and $\boldsymbol{t} \in \mathbb{R}$, define

$$Q_n(\theta, \boldsymbol{t}) \equiv \frac{1}{b_n^2} \sum_{j=1}^{N} \frac{\tilde{I}_n(\boldsymbol{\omega}_{jn}) G_\theta(\boldsymbol{\omega}_{jn})}{1 + \boldsymbol{t}' \tilde{I}_n(\boldsymbol{\omega}_{jn}) G_\theta(\boldsymbol{\omega}_{jn})}; \quad \tilde{Q}_n(\theta, \boldsymbol{t}) \equiv \sum_{j=1}^{N} \frac{[\partial G_\theta(\boldsymbol{\omega}_{jn})/\partial\theta]' \tilde{I}_n(\boldsymbol{\omega}_{jn}) \boldsymbol{t}}{1 + \boldsymbol{t}' G_\theta(\boldsymbol{\omega}_{jn}) \tilde{I}_n(\boldsymbol{\omega}_{jn})}.$$

Lemma 7. *Under the assumptions of Theorem 1, a maximizer* $\hat{\theta}_n = \mathrm{argmax}_{\theta\in\Theta_n} \mathcal{R}_n(\theta)$ *exists on* Θ_n *satisfying* $\hat{\theta}_n \in \Theta_n^\circ$, $Q_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n}) = \mathbf{0}_r$ *and* $\tilde{Q}_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n}) = \mathbf{0}_p$ *with* P*-probability converging to 1 as* $n \to \infty$ *a.s.* ($P_{\mathbf{X}}$).

Proof. It suffices to show that, for any subsequence $\{n_j\}$, there exists a further subsequence $\{n_k\} \subset \{n_j\}$ such that the P-probability of the event in Lemma 7 converges to one along $\{n_k\}$ (a.s. ($P_{\mathbf{X}}$)). From a given subsequence $\{n_j\}$, one can extract $\{n_k\}$ such that (A.2) holds by Condition (C.3)(iii), i.e., $N_k^{-1} \sum_{j=1}^{N_k} G_{\theta_0}(\boldsymbol{\omega}_{jn_k}) G'_{\theta_0}(\boldsymbol{\omega}_{jn_k}) \to \Gamma^*$ for a nonsingular $r \times r$ Γ^*. We shall assume (A.2) setting $n_k = n$ in the following to ease the notation throughout the remainder of the proof. For simplicity, we will also suppress (a.s. ($P_{\mathbf{X}}$)) notation so that all probability statements to follow are to be understood as holding (a.s. ($P_{\mathbf{X}}$)).

For $\ell(\theta) = -\log \mathcal{R}_n(\theta)$, we will first show that, with arbitrarily high P-probability as $n \to \infty$, $\ell_n(\theta)$ exists and is continuously differentiable on the neighborhood Θ_n of θ_0. This, in turn, implies that $\ell(\theta)$ has a minimizer $\hat{\theta}_n$ (or equivalently $\mathcal{R}_n(\theta)$ has a maximizer) on the compact set Θ_n.

Now by Lemma 3, the event that $\mathcal{R}_n(\theta) > 0$ for all $\theta \in \Theta_n$ has arbitrarily high P-probability as $n \to \infty$. When $\mathcal{R}_n(\theta) > 0$, the spatial EL function admits an expansion

$$\ell(\theta) = \sum_{i=1}^{N} \log\left(1 + \boldsymbol{t}'\tilde{I}_n(\boldsymbol{\omega}_{jn})G_\theta(\boldsymbol{\omega}_{jn})\right) \tag{A.9}$$

for a Lagrange multiplier $\boldsymbol{t}_\theta$ satisfying $Q_n(\theta, \boldsymbol{t}_\theta) = \mathbf{0}_r$; see Owen (1990). By Lemma 4, $W_{n,\theta_0} \xrightarrow{p} V$ holds for a positive definite V, so that by Lemma 6(iii) (i.e., $\sup_{\theta\in\Theta_n} \|W_{n,\theta} - W_{n,\theta_0}\| = o_p(1)$) it follows that $W_{n,\theta}$ is positive definite for all $\theta \in \Theta_n$ with P-probability approaching 1 as $n \to \infty$. When $W_{n,\theta}$ is positive definite and $\mathcal{R}_n(\theta) > 0$, $\partial Q_n(\theta, \boldsymbol{t})/\partial \boldsymbol{t}$ is negative definite for all $\theta \in \Theta_n$. By this fact combined with the implicit function theorem and $Q_n(\theta, \boldsymbol{t}_\theta) = \mathbf{0}_r$ (cf. Qin and Lawless, 1994), we have that $\boldsymbol{t}_\theta$ is a continuously differentiable function of θ on Θ_n. Hence, with P-probability approaching 1 as $n \to \infty$, $\ell_n(\theta)$ is then continuously differentiable on Θ_n and consequently has a minimizer $\hat{\theta}_n$ on Θ_n.

We next show that the minimizer $\hat{\theta}_n$ cannot be on $\partial\Theta_n$, the boundary of Θ_n, and then must lie in the interior Θ_n°. Write $\boldsymbol{t}_\theta = \|\boldsymbol{t}_\theta\|\boldsymbol{u}_\theta$, $\|\boldsymbol{u}_\theta\| = 1$, $\theta \in \Theta_n$. Using a standard argument to expand $Q_n(\theta, \boldsymbol{t}_\theta) = \mathbf{0}_r$, we get

$$\mathbf{0}_r = Q_n(\theta, \boldsymbol{t}_\theta) = -b_n^{-2} \sum_{j=1}^{N} \frac{G_\theta(\boldsymbol{\omega}_{jn})G'_\theta(\boldsymbol{\omega}_{jn})'\boldsymbol{t}_\theta \tilde{I}_n^2(\boldsymbol{\omega}_{jn})}{1 + \boldsymbol{t}'_\theta G_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})} + J_{n,\theta}.$$

Multiplying both sides above by $\boldsymbol{u}'_\theta$, adding $\boldsymbol{u}'_\theta J_{n,\theta}$, and taking norms gives

$$\|J_{n,\theta}\| \geq \|\boldsymbol{t}_\theta\| \frac{\boldsymbol{u}'_\theta W_{n,\theta}\boldsymbol{u}_\theta}{1 + \|\boldsymbol{t}_\theta\| Z_{n,\theta}},$$

implying

$$\sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|J_{n,\theta}\| \geq \frac{\sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\| \sup_{\theta\in\Theta_n} \boldsymbol{u}'_\theta W_{n,\theta}\boldsymbol{u}_\theta}{1 + \left(\sup_{\theta\in\Theta_n} \tau_{n,\theta} Z_{n,\theta}\right)\left(\sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\|\right)}. \tag{A.10}$$

As $W_{n,\theta_0} \xrightarrow{p} V$ holds for a positive definite V by construction under Lemma 4 and $\sup_{\theta\in\Theta_n} \|W_{n,\theta} - W_{n,\theta_0}\| = o_p(1)$ by Lemma 6(iii), we have in (A.10) that

$\sup_{\theta\in\Theta_n} \boldsymbol{u}_\theta' W_{n,\theta} \boldsymbol{u}_\theta \geq \sigma_1(1+o_p(1))$ where $\sigma_1 > 0$ is the smallest eigenvalue of V. As $\sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|J_{n,\theta}\| = O_p(1)$ and $\sup_{\theta\in\Theta_n} \tau_{n,\theta}\|Z_{n,\theta}\| = o_p(1)$ by Lemma 6, we have from (A.10) that $O_p(1) \geq \sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\|(\sigma_1 + o_p(1))$ or

$$\sup_{\theta\in\Theta_n} \tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\| = O_p(1). \tag{A.11}$$

Then again expanding $Q_n(\theta, \boldsymbol{t}_\theta) = \mathbf{0}_r$, we may write

$$\mathbf{0}_r = J_{n,\theta} - W_{n,\theta}\boldsymbol{t}_\theta + \frac{1}{b_n^2}\sum_{j=1}^{N} \frac{G_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\left(G_\theta'(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_\theta\right)^2}{1 + \boldsymbol{t}_\theta' G_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})}.$$

Using $\sup_{\theta\in\Theta_n} \|W_{n,\theta} - V\| \xrightarrow{p} 0$ and consequently, $\sup_{\theta\in\Theta_n} \|W_{n,\theta}^{-1} - V^{-1}\| \xrightarrow{p} 0$, we obtain

$$\boldsymbol{t}_\theta = W_{n,\theta}^{-1}J_{n,\theta} + R_\theta, \tag{A.12}$$

where

$$\begin{aligned}
\sup_{\theta\in\Theta_n} \|R_\theta\| &\leq \sup_{\theta\in\Theta_n} \left\| \frac{W_{n,\theta}^{-1}}{b_n^2} \sum_{j=1}^{N} \frac{G_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\left(G_\theta'(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_\theta\right)^2}{1 + \boldsymbol{t}_\theta' G_\theta(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})} \right\| \\
&\leq \sup_{\theta\in\Theta_n} \frac{\left\|W_{n,\theta}^{-1}\right\| \|W_{n,\theta}\|\, Z_{n,\theta}\, \|\boldsymbol{t}_\theta\|^2}{1 + Z_{n,\theta}\|\boldsymbol{t}_\theta\|} \\
&= \sup_{\theta\in\Theta_n} \frac{\left\|W_{n,\theta}^{-1}\right\| \|W_{n,\theta}\| \left(\tau_{n,\theta}Z_{n,\theta}\right) \left(\tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\|\right)\|\boldsymbol{t}_\theta\|}{1 + \left(\tau_{n,\theta}Z_{n,\theta}\right)\left(\tau_{n,\theta}^{-1}\|\boldsymbol{t}_\theta\|\right)} \\
&= o_p\left(\sup_{\theta\in\Theta_n} \|\boldsymbol{t}_\theta\|\right),
\end{aligned} \tag{A.13}$$

using Lemma 6(iv) (i.e., $\sup_{\theta\in\Theta_n} \tau_{n,\theta}Z_{n,\theta} = o_p(1)$), (A.11), $\sup_{\theta\in\Theta_n} \|W_{n,\theta}^{-1}\| = O_p(1)$, $\sup_{\theta\in\Theta_n} \|W_{n,\theta}\| = O_p(1)$ and $\sup_{\theta\in\Theta_n} Z_{n,\theta}\|\boldsymbol{t}_\theta\| = o_p(1)$.

For $\theta \in \Theta_n$ and $j = 1, \ldots, N$, define $\gamma_{j,\theta} \equiv G_\theta'(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_\theta$, noting that

$$\sup_{\theta\in\Theta_n} \max_{1\leq j\leq n} |\gamma_{j,\theta}| \leq \sup_{\theta\in\Theta_n} Z_{n,\theta}\|\boldsymbol{t}_\theta\| = o_p(1). \tag{A.14}$$

When $\sup_{\theta\in\Theta_n}\max_{1\le j\le n}|\gamma_{j,\theta}|$ is small, we may use a Taylor expansion and (A.12) to express $\ell_n(\theta)$ in (A.9) as

$$\begin{aligned}
\ell_n(\theta) &= \sum_{j=1}^{N}\log(1+\gamma_{j,\theta}) = \sum_{j=1}^{N}\gamma_{j,\theta} - \frac{1}{2}\sum_{j=1}^{N}\gamma_{j,\theta}^2 + \tilde{R}_{n,\theta} \\
&= \sum_{j=1}^{N} G_\theta'(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_\theta - \frac{1}{2}\boldsymbol{t}_\theta'\sum_{j=1}^{N} G_\theta(\boldsymbol{\omega}_{jn})G_\theta'(\boldsymbol{\omega}_{jn})\tilde{I}_n^2(\boldsymbol{\omega}_{jn})\boldsymbol{t}_\theta + \tilde{R}_{n,\theta} \\
&= b_n^2 J_{n,\theta}'\left(W_{n,\theta}^{-1}J_{n,\theta}+R_\theta\right) - \frac{1}{2}\left(W_{n,\theta}^{-1}J_{n,\theta}+R_\theta\right)' b_n^2 W_{n,\theta}\left(W_{n,\theta}^{-1}J_{n,\theta}+R_\theta\right) + \tilde{R}_{n,\theta} \\
&= \frac{1}{2}b_n^2 J_{n,\theta}W_{n,\theta}^{-1}J_{n,\theta} - \frac{1}{2}b_n^2 R_\theta' W_{n,\theta}R_\theta + \tilde{R}_{n,\theta}. \qquad \text{(A.15)}
\end{aligned}$$

where R_θ is from (A.12) and $\tilde{R}_{n,\theta}$ is a remainder (from Taylor expansion) bounded by

$$\sup_{\theta\in\Theta_n}\left|\tilde{R}_{n,\theta}\right| \le \sup_{\theta\in\Theta_n}\frac{1}{3}\frac{\|W_{n,\theta}\|Z_{n,\theta}\|\boldsymbol{t}_\theta\|^3 b_n^2}{(1-\|\boldsymbol{t}_\theta\|Z_{n,\theta})^3} = o_p\left(\sup_{\theta\in\Theta_n}\|\boldsymbol{t}_\theta b_n\|^2\right). \qquad \text{(A.16)}$$

from (A.14). By (A.12)-(A.13) and Lemma 6(i),

$$\begin{aligned}
\sup_{\theta\in\Theta_n}\|b_n\boldsymbol{t}_\theta\| &\le \sup_{\theta\in\Theta_n}\left\|W_{n,\theta}^{-1}\right\|\sup_{\theta\in\Theta_n} b\,\|J_{n,\theta}\| + \sup_{\theta\in\Theta_n} b_n\,\|R_\theta\| \\
&= O_p(1)O_p(\log\lambda_n) + o_p\left(\sup_{\theta\in\Theta_n}\|\boldsymbol{t}_\theta b_n\|\right)
\end{aligned}$$

so that

$$\begin{aligned}
\sup_{\theta\in\Theta_n}\|b_n\boldsymbol{t}_\theta\| &= O_p(\log\lambda_n), && \text{(A.17)} \\
\sup_{\theta\in\Theta_n}\|b_n R_\theta\| &= o_p(\log\lambda_n), && \text{(A.18)} \\
\sup_{\theta\in\Theta_n}|\tilde{R}_{n,\theta}| &= o_p(\log^2\lambda_n), && \text{(A.19)}
\end{aligned}$$

where the last bound follows by modifying (A.16). Hence, combining (A.15), (A.18), and (A.19), we get

$$\sup_{\theta\in\partial\Theta_n}\left|\ell_n(\theta) - \frac{1}{2}b_n^2 J_{n,\theta}' W_{n,\theta}^{-1}J_{n,\theta}\right| = o_p(\log^2\lambda_n)$$

which can be further re-written as

$$\sup_{\theta\in\partial\Theta_n}\left|\ell_n(\theta) - \frac{1}{2}b_n^2 J_{n,\theta}' V^{-1}J_{n,\theta}\right| = o_p(\log^2\lambda_n)$$

from

$$\sup_{\theta\in\Theta_n}\left\|b_n^2 J'_{n,\theta}W_{n,\theta}^{-1}J_{n,\theta} - b_n^2 J'_{n,\theta}V^{-1}J_{n,\theta}\right\| \leq \sup_{\theta\in\Theta_n} b_n^2\|J_{n,\theta}\|^2 \sup_{\theta\in\Theta_n}\left\|W_{n,\theta}^{-1} - V^{-1}\right\|$$
$$= O_p(\log^2\lambda_n)o_p(1)$$

by $\sup_{\theta\in\Theta_n}\left\|W_{n,\theta}^{-1} - V^{-1}\right\| = o_p(1)$ and Lemma 6(i). For $\theta \in \partial\Theta_n$, it holds that $\theta = \theta_0 + \left(\lambda_n^{-\kappa d} b_n \log\lambda_n\right)\boldsymbol{v}_\theta$ for some $\boldsymbol{v}_\theta \in \mathbb{R}^p$ with $\|\boldsymbol{v}_\theta\| = 1$, so that by Lemma 2(ii)

$$b_n J_{n,\theta} = b_n J_{n,\theta_0} + \frac{\lambda_n^{\kappa d}}{b_n}\left(D_{\theta_0} + o_p(1)\right)\lambda_n^{-\kappa d} b_n \log\lambda_n \boldsymbol{v}_\theta = M_{n,\theta} + o_p(\log(\lambda_n)).$$

$M_{n,\theta} \equiv b_n J_{n,\theta_0} + D_{\theta_0}\boldsymbol{v}_\theta \log\lambda_n$, where the $o_p(\log(\lambda_n))$ term is uniform in $\theta \in \Theta_n$ so that we may re-express

$$\sup_{\theta\in\partial\Theta_n}\left|\ell_n(\theta) - \frac{1}{2}M'_{n,\theta}V^{-1}M_{n,\theta}\right| = o_p(\log^2\lambda_n).$$

Then we have, uniformly in $\theta \in \partial\Theta_n$, $\|\boldsymbol{v}_\theta\| = 1$,

$$\begin{aligned}
2\ell_n(\theta) &= M'_{n,\theta}V^{-1}M_{n,\theta} + o_p\left(\log^2\lambda_n\right)\\
&= b_n^2 J'_{n,\theta_0}V^{-1}J_{n,\theta_0} + b_n J'_{n,\theta_0}V^{-1}\log(\lambda_n)D_{\theta_0}\boldsymbol{v}_\theta\\
&\qquad + \log(\lambda_n)\boldsymbol{v}'_\theta D'_{\theta_0}V^{-1}J_{\theta_0}b_n + \log^2(\lambda_n)\boldsymbol{v}'_\theta D'_{\theta_0}V^{-1}D_{\theta_0}\boldsymbol{v}_\theta + o_p(\log^2\lambda_n)\\
&= \log^2(\lambda_n)\boldsymbol{v}'_\theta D'_{\theta_0}V^{-1}D_{\theta_0}\boldsymbol{v}_\theta + o_p(\log^2\lambda_n)\\
&\geq \frac{\sigma}{2}\log^2(\lambda_n)(1 + o_p(1)),
\end{aligned}$$

where $\sigma > 0$ is the smallest eigenvalue of positive definite $D'_{\theta_0}V^{-1}D_{\theta_0}$ (as D_{θ_0} has full column rank p).

Consequently, $\inf_{\theta\in\partial\Theta_n}\ell_n(\theta) \geq 2^{-1}\sigma\log^2(\lambda_n)(1+o_p(1))$, while by Lemma 5 $\ell_n(\theta_0) = O_p(1)$. Hence, the minimizer $\hat{\theta}_n$ of $\ell_n(\theta)$ on Θ_n must lie in Θ_n° and, as $\ell(\theta)$ is continuously differentiable on Θ, it follows that

$$\mathbf{0}_p = \left.\frac{\partial\ell_n(\theta)}{\partial\theta}\right|_{\hat{\theta}_n} = \tilde{Q}_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n}) + \left[\frac{\partial\boldsymbol{t}_{\hat{\theta}_n}}{\partial\theta}\right]' Q_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n})$$

using $Q_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n}) = \mathbf{0}_r$. This completes the proof. □

The final result is a distributional characterization of the spatial EL point estimator.

Lemma 8. *Under the assumptions of Theorem 1, given any subsequence* $\{n_j\} \subset \{n\}$*, extract a further subsequence* $\{k \equiv n_k\} \subset \{n_j\}$ *such that (A.2) holds. Then, for the maximizer* $\hat{\theta}_k$ *and its Lagrange multiplier* $\boldsymbol{t}_{\theta_k}$ *along the subsequence, it holds that*

$$\begin{pmatrix} b_k \boldsymbol{t}_{\hat{\theta}_k} \\ (\hat{\theta}_k - \theta_0)\dfrac{\lambda_k^{\kappa d}}{b_k} \end{pmatrix} \xrightarrow{d} N\left(0, a\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}\right) \quad a.s.\ (P_{\mathbf{X}})$$

as $k \to \infty$*, where* $U_1 = V^{-1} - V^{-1}D_{\theta_0}U_2D'_{\theta_0}V^{-1}$ *and* $U_2 = (D'_{\theta_0}V^{-1}D_{\theta_0})^{-1}$ *are positive definite,* $D_{\theta_0} \equiv \int h_{\theta_0}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})d\boldsymbol{\omega}$*, and the positive definite matrix* V *and constant* $a \in \{1, 2\}$ *are defined in Lemma 4 (depending on both the subsequence* $\{k \equiv n_k\}$ *and the three PID/MID subcases).*

Proof. We shall use notation and preliminary results established in the proof of Lemma 7. From a given subsequence $\{n_j\}$, one can again extract $\{n_k\}$ such that (A.2) holds by Condition (C.3)(iii), i.e., $N_k^{-1}\sum_{j=1}^{N_k} G_{\theta_0}(\boldsymbol{\omega}_{jn_k})G'_{\theta_0}(\boldsymbol{\omega}_{jn_k}) \to \Gamma^*$ for a nonsingular $r \times r$ Γ^*. We shall henceforth assume (A.2) setting $n_k = n$ in the following for simplicity and also suppress (a.s. $(P_{\mathbf{X}})$) notation in P-probability statements. Recall that, by Lemma 4, (A.3) holds for the subsequence, i.e., $W_{n,\theta_0} \xrightarrow{p} V$ and $b_nJ_{n,\theta_0} \xrightarrow{d} N(\mathbf{0}_r, aV)$ for a positive definite V and constant $a \in 1, 2$ defined in Lemma 4.

With arbitrarily high P-probability as $n \to \infty$, the maximizer $\hat{\theta}_n$ exists with the properties stated in Lemma 7. From $\tilde{Q}_n(\hat{\theta}_n, \boldsymbol{t}_{\hat{\theta}_n}) = \mathbf{0}_p$, we have

$$\begin{aligned} \mathbf{0}_p &= \lambda_n^{-\kappa d}\sum_{j=1}^{N} \frac{[\partial G_{\hat{\theta}_n}(\boldsymbol{\omega}_{jn})/\partial\theta]'\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_{\hat{\theta}_n}b_n}{1 + \boldsymbol{t}'_{\hat{\theta}_n}G_{\hat{\theta}_n}(\boldsymbol{\omega}_{jn})\tilde{I}_n(\boldsymbol{\omega}_{jn})} \\ &= \lambda_n^{-\kappa d}\sum_{j=1}^{N}\left[\frac{\partial G_{\hat{\theta}_n}(\boldsymbol{\omega}_{jn})}{\partial\theta}\right]'\tilde{I}_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_{\hat{\theta}_n}b_n + S_n, \end{aligned}$$

where, by the boundedness of partial derivatives under Condition (C.7), Lemma 1(i) and $\sup_{\theta\in\Theta_n}\|\boldsymbol{t}_\theta\|Z_{n,\theta}=o_p(1)$, we have

$$\|S_n\| = \left\| \lambda_n^{-\kappa d}\sum_{j=1}^{N}\frac{\boldsymbol{t}'_{\hat\theta_n}G_{\hat\theta_n}(\boldsymbol{\omega}_{jn})[\partial G_{\hat\theta_n}(\boldsymbol{\omega}_{jn})/\partial\theta]'\tilde I_n^2(\boldsymbol{\omega}_{jn})\boldsymbol{t}_{\hat\theta_n}b_n}{1+\boldsymbol{t}'_{\hat\theta_n}G_{\hat\theta_n}(\boldsymbol{\omega}_{jn})\tilde I_n(\boldsymbol{\omega}_{jn})}\right\|$$

$$\le C_0\frac{\lambda_n^{-\kappa d}\|\boldsymbol{t}_{\hat\theta_n}\|^2 b_n O_p(b_n^2)}{1-\|\boldsymbol{t}_{\hat\theta_n}\|Z_{n,\hat\theta_n}} = O_p\left(\lambda_n^{-\kappa d}b_n^3\|\boldsymbol{t}_{\hat\theta_n}\|^2\right).$$

As $\|b_n\boldsymbol{t}_{\hat\theta_n}\|=O_p(\log\lambda_n)$ by (A.17) and $\lambda_n^{-\kappa d}b_n\log\lambda_n=o(1)$ by (C.5), it follows then that $\|S_n\|=o_p(b_n\|\boldsymbol{t}_{\hat\theta_n}\|)$. Then, by Lemma 2(ii), we obtain

$$\begin{aligned}\mathbf{0}_p &= \lambda_n^{-\kappa d}\sum_{j=1}^{N}\left[\frac{\partial G_{\hat\theta_n}(\boldsymbol{\omega}_{jn})}{\partial\theta}\right]'\tilde I_n(\boldsymbol{\omega}_{jn})\boldsymbol{t}_{\hat\theta_n}b_n+S_n\\ &= [D_{n,\hat\theta_n}]'\boldsymbol{t}_{\hat\theta_n}b_n+o_p(b_n\|\boldsymbol{t}_{\hat\theta_n}\|)\\ &= D'_{\theta_0}\boldsymbol{t}_{\hat\theta_n}b_n+o_p(b_n\|\boldsymbol{t}_{\hat\theta_n}\|).\end{aligned}$$

From (A.7), (A.12), (A.13) along with Lemma 2(ii) and $\sup_{\theta\in\Theta_n}\|W_{n,\theta}-V\|=o(1)$ by Lemma 6(iii), we also have

$$Vb_n\boldsymbol{t}_{\hat\theta_n}=b_nJ_{n,\theta_0}+D_{\theta_0}(\hat\theta_n-\theta_0)\lambda_n^{\kappa d}b_n^{-1}+o_p(\delta_n)$$

for $\delta_n=\|\hat\theta_n-\theta_0\|\lambda_n^{\kappa d}b_n^{-1}+b_n\|\boldsymbol{t}_{\hat\theta_n}\|$. The two previous expansions then may be combined to yield

$$\Sigma\begin{pmatrix} b_n\boldsymbol{t}_{\hat\theta_n}\\ (\hat\theta_n-\theta_0)\dfrac{\lambda_n^{\kappa d}}{b_n}\end{pmatrix}=\begin{pmatrix} b_nJ_{\theta_0}+o_p(\delta_n)\\ o_p(\delta_n)\end{pmatrix} \tag{A.20}$$

for

$$\Sigma=\begin{bmatrix} V & -D_{\theta_0}\\ D'_{\theta_0} & 0\end{bmatrix},\quad \Sigma^{-1}=\begin{bmatrix} U_1 & V^{-1}D_{\theta_0}U_2\\ -U_2D'_{\theta_0}V^{-1} & U_2\end{bmatrix},$$

where Σ_n, $U_1=V^{-1}-V^{-1}D_{\theta_0}U_2D'_{\theta_0}V^{-1}$, and $U_2=(D'_{\theta_0}V^{-1}D_{\theta_0})^{-1}$ are positive definite. Taking norms in (A.20) and recalling that $b_nJ_{n,\theta_0}\xrightarrow{d}N(0,aV)$, one can deduce $O_p(\delta_n)(1+o_p(1))=O_p(1)$ or $\delta_n=O_p(1)$ so that

$$\begin{pmatrix} b_n\boldsymbol{t}_{\hat\theta_n}\\ (\hat\theta_n-\theta_0)\dfrac{\lambda^{\kappa d}}{b_n}\end{pmatrix}=\Sigma^{-1}\begin{pmatrix} b_nJ_{\theta_0}+o_p(1)\\ o_p(1)\end{pmatrix}=\begin{pmatrix} U_1\\ -U_2D'_{\theta_0}V^{-1}\end{pmatrix}b_nJ_{n,\theta_0}+o_p(1). \tag{A.21}$$

Noting that $U_1 V U_1' = U_1$, $(U_2 D'_{\theta_0} V^{-1}) V (U_2 D'_{\theta_0} V^{-1})' = U_2$ and $(U_2 D'_{\theta_0} V^{-1}) V U_1' = \mathbf{0}_{p\times r}$, it now follows from $b_n J_{n,\theta_0} = O_p(1)$ and (A.21) that

$$\begin{pmatrix} b_n \boldsymbol{t}_{\hat{\theta}_n} \\ (\hat{\theta}_n - \theta_0)\dfrac{\lambda^{\kappa d}}{b_n} \end{pmatrix} \xrightarrow{d} N\left(0, a \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}\right).$$

This completes the proof. □

References

Bandyopadhyay, S., Lahiri, S. N., and Nordman, D. J. (2015). A central limit theorem for periodogram based statistics for irregularly spaced spatial data and Whittle estimation. *Preprint.*

Bandyopadhyay, S., Lahiri, S. N., and Nordman, D. J. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *The Annals of Statistics* **43**, 519-545.

Bradley, R. C. (1989). A caution on mixing conditions for random fields. *Statistics and Probability Letters* **8**, 489-491.

Bradley, R. C. (1993). Equivalent mixing conditions for random fields. *The Annals of Probability* **21**, 1921-1926.

Kitamura, Y. (1997). Empirical likelihood methods for weakly dependent processes. *The Annals of Statistics* **25**, 2084-2102.

Lahiri, S. N. (2003). Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhya, Series A* **65**, 356-388.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics* **18**, 90-120.

Qin, J., and Lawless, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics* **22**, 300-325.

Van Hala, M, Bandyopadhyay, S., Lahiri, S. N., and Nordman, D. J.(2015) A general frequency domain method for assessing spatial covariance structure. *Preprint*.