Statistica Sinica **18**(2008): Supplement, S75∼S90

A BLOCKWISE EMPIRICAL LIKELIHOOD FOR SPATIAL LATTICE DATA

Daniel J. Nordman

Iowa State University

Supplementary Material

Appendix

Section A.1 details the spatial mixing and moment conditions used to establish the main results of the manuscript. Section A.2 provides some technical lemmas to facilitate the proofs of the main results, which are presented in Section A.3. In Section A.4, we describe a further result on EL inference under parameter constraints. Section A.5 describes the spatial bootstrap method used to implement the spatial EL Bartlett correction from Section 4 of the manuscript.

A.1. Assumptions

To establish the main results on the spatial EL, we require assumptions on the spatial process and the potential vector G_θ of estimating functions. Recall that we may collect observations from the real-valued, strictly stationary spatial process $\{Z_{\mathbf{s}} : \mathbf{s} \in \mathbb{Z}^d\}$ into m-dimensional vectors $Y_{\mathbf{s}} = (Z_{\mathbf{s}+\mathbf{h}_1}, \ldots, Z_{\mathbf{s}+\mathbf{h}_m})'$, $\mathbf{s} \in \mathbb{Z}^d$, using fixed lag vectors $\mathbf{h}_1, \ldots, \mathbf{h}_m \in \mathbb{Z}^d$ for a positive integer $m \geq 1$. Recall $\mathcal{R}_n = \lambda_n \mathcal{R}_0 \subset \mathbb{R}^d$ denotes the sampling region for the process $\{Z_{\mathbf{s}} : \mathbf{s} \in \mathbb{Z}^d\}$ and $\mathcal{R}_{m,n}$ is the sampling region of the observed $Y_{\mathbf{s}}, \mathbf{s} \in \mathbb{Z}^d$. We first outline some notation.

For $A \subset \mathbb{R}^d$, denote the Lesbegue volume of an uncountable set A as vol(A) and the cardinality of a uncountable set A as $|A|$. Limits in order symbols are taken letting $n \to \infty$ and, for two positive sequences, we write $s_n \sim t_n$ if $s_n/t_n \to 1$. For a vector $\mathbf{x} = (x_1, \ldots, x_d)' \in \mathbb{R}^d$, let $\|\mathbf{x}\|$ and $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq d} |x_i|$ denote the Euclidean and l^∞ norms of $\mathbf{x}$, respectively. Define the distance between two sets E_1, $E_2 \subset \mathbb{R}^d$ as: $\mathrm{dis}(E_1, E_2) = \inf\{\|\mathbf{x} - \mathbf{y}\|_\infty : \mathbf{x} \in E_1, \mathbf{y} \in E_2\}$.

Let $\mathcal{F}_Y(T)$ denote the σ-field generated by the random vectors $\{Y_{\mathbf{s}} : \mathbf{s} \in T\}$, $T \subset \mathbb{Z}^d$, and define the strong mixing coefficient for the strictly stationary random field $\{Y_{\mathbf{s}} : \mathbf{s} \in \mathbb{Z}^d\}$ as

$$\alpha_Y(v, w) = \sup\{\tilde{\alpha}(T_1, T_2) : T_i \subset \mathbb{Z}^d,\ |T_i| \leq w,\ i = 1, 2;\ \mathrm{dis}(T_1, T_2) \geq v\},\ \ v, w > 0, \tag{8}$$

where $\tilde{\alpha}_Y(T_1, T_2) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_Y(T_1), B \in \mathcal{F}_Y(T_2)\}$. In the following assumptions, let θ_0 denote the unique parameter value which satisfies (1).

Throughout the sequel, we use C to denote a generic positive constant that does not depend on n or any $\mathbb{Z}^d$ points and may vary from instance to instance.

Assumptions

1. As $n \to \infty$, $b_n^{-1} + b_n^2/\lambda_n = o(1)$ and, for any positive real sequence $a_n \to 0$, the number of cubes of $a_n\mathbb{Z}^d$ which intersect the closures $\overline{\mathcal{R}_0}$ and $\overline{\mathbb{R}^d \setminus \mathcal{R}_0}$ is $O(a_n^{-(d-1)})$.

2. There exist nonnegative functions $\alpha_1(\cdot)$ and $q(\cdot)$ such that $\alpha_1(v) \to 0$ as $v \to \infty$ and $\alpha_Y(v, w) \le \alpha_1(v)q(w)$, $v, w > 0$. The non-decreasing function $q(\cdot)$ is bounded for the time series case $d = 1$, but may be unbounded $q(w) \to \infty$ as $w \to \infty$ for $d \ge 2$.

3. For some $0 < \delta \le 1$, $0 < \kappa < (5d-1)(6+\delta)/(d\delta)$ and $C > 0$, it holds that $\mathrm{E}\{\|G_{\theta_0}(Y_{\mathbf{s}})\|^{6+\delta}\} < \infty$, $\sum_{v=1}^{\infty} v^{5d-1}\alpha_1(v)^{\delta/(6+\delta)} < \infty$, $q(w) \le Cw^{\kappa}$, $w \ge 1$.

4. The $r \times r$ matrix $\Sigma_{\theta_0} = \sum_{\mathbf{h}\in\mathbb{Z}^d} \mathrm{Cov}\{G_{\theta_0}(Y_{\mathbf{s}}), G_{\theta_0}(Y_{\mathbf{s}+\mathbf{h}})\}$ is positive definite.

The growth rate of the spatial block factor b_n in Assumption 1 represents a spatial extension of scaling conditions used for the blockwise EL for time series $d = 1$ in Kitamura (1997); this entails the block condition (3). Additionally, to avoid pathological sampling regions, a mild boundary condition on $\mathcal{R}_0$ implies that the number of $\mathbb{Z}^d$ lattice points near the boundary of $\mathcal{R}_n = \lambda_n\mathcal{R}_0$ is of smaller order $O(\lambda_n^{d-1})$ than the volume of the sampling region $\mathcal{R}_n$. As a consequence, the number n of $Z_{\mathbf{s}}$-sampling sites (i.e., $\mathbb{Z}^d$ points) contained in $\mathcal{R}_n$ is asymptotically equivalent to the volume of $\mathcal{R}_n$:

$$n = |\mathcal{R}_n \cap \mathbb{Z}^d| \sim \mathrm{vol}(\mathcal{R}_n) = \lambda_n^d \mathrm{vol}(\mathcal{R}_0).$$

Additionally, the boundary condition on $\mathcal{R}_0$ allows the number of blocks to be quantified under different EL blocking schemes; see Lemma 2(i) of the following Section A.3 for illustration.

Assumption 2 describes a mild bound on the mixing coefficient from (8) with growth rates set in Assumption 3. These mixing assumptions permit moment bounds and a central limit theorem to be applied to sample means of the form $\bar{G}_n = \sum_{\mathbf{s}\in\mathcal{R}_{m,n}\cap\mathbb{Z}^2} G_{\theta_0}(Y_{\mathbf{s}})/n_{m,n}$ (Lahiri, 2003b); Lemma 2 in Section A.3 illustrates such moment bounds. The conditions on the mixing coefficient (8) in Assumptions 2-3 apply to many weakly dependent random fields including certain linear fields with a moving average representation, Gaussian fields with analytic spectral densities, Markov random fields as well as various time series; see Doukhan (1994). For $d > 1$, we allow (8) to become unbounded in w, which is important in the spatial case to avoid a more restrictive form of mixing; see Lahiri (2003a, p. 295). Assumption 4 implies that the limiting variance $\Sigma_{\theta_0} = \lim_{n\to\infty} n_{m,n}\mathrm{Var}\,(\bar{G}_n)$ is positive definite.

A.2. Preliminary results for main proofs

Lemma 2 gives moment bounds based on Doukhan (1994, p.9, 26) while Lemma 2 provides some important distributional results for proving the main EL results. In particular, parts (ii) and (iii) of Lemma 2 entail that, at the true parameter value θ_0, spatial block sample means $M_{\theta_0,\mathbf{i}}$, $\mathbf{i} \in \mathcal{I}_n$, from the EL construction (4) can be combined to produce normally distributed averages or consistent variance estimators. Parts (iv)−(vi) of this lemma are used to prove that, in a neighborhood of θ_0, the EL ratio $R_n(\theta)$ from (4) can be finitely computed and also that a sequence $\hat{\theta}_n$ of maximizers of $R_n(\theta)$ (i.e., the maximal EL estimator) must exist in probability. Lemma 3 establishes the distribution of the spatial log-EL ratio at the true parameter value θ_0. Proofs of Lemmas 2 and 3 appear subsequently.

Lemma 1. (i) *Suppose a random variable X_i is measurable with respect to $\mathcal{F}_Y(T_i)$ for bounded $T_i \subset \mathbb{Z}^d$, $i = 1, 2$ and let $s, t > 0$, $1/s+1/t < 1$. If $\mathrm{dis}(T_1, T_2) > 0$ and expectations are finite, then* $|\,\mathrm{Cov}\,(X_1, X_2)| \le 8\{\mathrm{E}\,(|X_1|^s)\}^{1/s}\{\mathrm{E}\,(|X_2|^t)\}^{1/t}\alpha_Y\big(\mathrm{dis}(T_1, T_2); \max_{i=1,2}|T_i|\big)^{1-1/s-1/t}$.

(ii) *Under Assumptions 2−3, for any real $1 \le k \le 6$ and $T \subset \mathbb{Z}^d$ it holds that* $\mathrm{E}\,\{\|\sum_{\mathbf{s}\in T}\tilde{G}_{\theta_0}(Y_{\mathbf{s}})\|^k\} \le C|T|^{k/2}$, *where* $\tilde{G}_{\theta_0}(Y_{\mathbf{s}}) = G_{\theta_0}(Y_{\mathbf{s}}) - \mathrm{E}\,\{G_{\theta_0}(Y_{\mathbf{s}})\}$.

Lemma 2. *Let $\mathcal{I}_n = \mathcal{I}_{b_n}^{OL}$ or $\mathcal{I}_{b_n}^{NOL}$ and $N_{\mathcal{I}} = |\mathcal{I}_n|$. Under Assumptions 1−4.*

(i) $|\mathcal{I}_{b_n}^{OL}| \sim \mathrm{vol}(\mathcal{R}_{m,n})$, $n_{m,n} \sim \mathrm{vol}(\mathcal{R}_{m,n})$, $|\mathcal{I}_{b_n}^{NOL}| \sim \mathrm{vol}(\mathcal{R}_{m,n})/b_n^d$ *and* $\mathrm{vol}(\mathcal{R}_{m,n}) \sim \mathrm{vol}(\mathcal{R}_n) = \lambda_n^d\mathrm{vol}(\mathcal{R}_0)$;

(ii) $n_{m,n}^{1/2}\bar{M}_{\theta_0} \xrightarrow{d} \mathcal{N}(0_r, \Sigma_{\theta_0})$, *where* $\bar{M}_{\theta_0} \equiv \sum_{\mathbf{i}\in\mathcal{I}_n} M_{\theta_0,\mathbf{i}}/N_{\mathcal{I}}$;

(iii) $\widehat{\Sigma}_{\theta_0} \equiv b_n^d\sum_{\mathbf{i}\in\mathcal{I}_n} M_{\theta_0,\mathbf{i}}M'_{\theta_0,\mathbf{i}}/N_{\mathcal{I}} \xrightarrow{p} \Sigma_{\theta_0}$, *with* Σ_{θ_0} *from Assumption 4*;

(iv) $P(R_n(\theta_0) > 0) \to 1$;

(v) $\max_{\mathbf{i}\in\mathcal{I}_n}\|M_{\theta_0,\mathbf{i}}\| = O_p\big(b_n^{-d}n_{m,n}^{5/12}\big)$;

(vi) $P(\inf_{v\in\mathbb{R}^r,\|v\|=1} N_{\mathcal{I}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_n} b_n^{d/2}v'M_{\theta_0,\mathbf{i}}\mathbb{I}(v'M_{\theta_0,\mathbf{i}} > 0) > C) \to 1$ *for some* $C > 0$, *letting* $\mathbb{I}(\cdot)$ *denote the indicator function.*

Lemma 3. *Under Assumptions 1−4 and $\mathcal{I}_n = \mathcal{I}_{b_n}^{OL}$ or $\mathcal{I}_{b_n}^{NOL}$, it holds in (6) that* $\ell_n(\theta_0) \xrightarrow{d} \chi_r^2$.

Proof of Lemma 2. Assumption 1 yields part(i) of the lemma. We shall sketch the proof for $\mathrm{vol}(\mathcal{R}_{m,n})$ and the number $|\mathcal{I}_{b_n}^{OL}|$ of OL blocks; the remaining cases follow similarly and more details on counting results can be found in Nordman and Lahiri (2004). For a positive integer j, define

$$J_n(j) = \{\mathbf{i} \in \mathbb{Z}^d : \ (\mathbf{i} + j[-1,1]^d) \cap \mathcal{R}_n \neq \varnothing, \ (\mathbf{i} + j[-1,1]^d) \cap (\mathbb{R}^d \setminus \mathcal{R}_n) \neq \varnothing\},$$

where again $\mathcal{R}_n = \lambda_n \mathcal{R}_0$, and note that for $a_n = j/\lambda_n$

$$\begin{aligned} |J_n(j)| &\leq (2j+1)^d \big|\{\mathbf{i} \in a_n \mathbb{Z}^d : \text{cube } \mathbf{i} + a_n[-1,1]^d \text{ intersects both } \overline{\mathcal{R}_0} \text{ and } \overline{\mathbb{R}^d \setminus \mathcal{R}_0}\}\big| \\ &= (2j+1)^d O(a_n^{-(d-1)}) = O(j\lambda_n^{d-1}) \end{aligned} \tag{9}$$

by the $\mathcal{R}_0$-boundary condition in Assumption 1. The bound in (9) also holds if we replace a fixed integer j by the sequence of block factors b_n (i.e., replace j, $J_n(j)$ with b_n, $J_n(b_n)$).

Recall that $\mathcal{R}_{m,n} = \{\mathbf{s} \in \mathcal{R}_n : \mathbf{s} + \mathbf{h}_1, \ldots, \mathbf{s} + \mathbf{h}_m \in \mathcal{R}_n\} \subset \mathbb{R}^d$ is defined with respect to m fixed lags $\{\mathbf{h}_i\}_{i=1}^m \subset \mathbb{Z}^d$. Let $h = \max_{1 \leq i \leq m} \|\mathbf{h}_i\|_\infty$ and note that

$$\text{vol}(\mathcal{R}_n) - \text{vol}(\mathcal{R}_n \setminus \mathcal{R}^*_{m,n}) = \text{vol}(\mathcal{R}^*_{m,n}) \leq \text{vol}(\mathcal{R}_{m,n}) \leq \text{vol}(\mathcal{R}_n)$$

where $\mathcal{R}^*_{m,n} = \{\mathbf{s} \in \mathcal{R}_n : \mathbf{s} + h[-1,1]^d \subset \mathcal{R}_n\}$. Then, for fixed h by (9), we find $\text{vol}(\mathcal{R}_n \setminus \mathcal{R}^*_{m,n}) \leq (2h)^d |J_n(h)| = O(\lambda_n^{d-1})$ so that $\text{vol}(\mathcal{R}_{m,n}) \sim \text{vol}(\mathcal{R}_n) = \lambda_n^d \text{vol}(\mathcal{R}_0)$ follows. Likewise, $n = |\mathbb{Z}^d \cap \mathcal{R}_n| \sim \text{vol}(\mathcal{R}_n)$ holds from $\big|n - \text{vol}(\mathcal{R}_n)\big| \leq 2^d |J_n(1)|$ and then $|\mathcal{I}^{OL}_{b_n}| \sim \text{vol}(\mathcal{R}_n)$ follows from $n - |J_n(b_n)| \leq |\mathcal{I}^{OL}_{b_n}| \leq n$ and $|J_n(b_n)| = O(b_n \lambda^{d-1}) = o(\text{vol}(\mathcal{R}_n))$.

To prove parts of Lemma 2(ii) and (iii), we treat only the OL block case $\mathcal{I}_n = \mathcal{I}^{OL}_{b_n}$; the NOL case follows similarly and we shall describe the modifications required for handling NOL blocks. Defining the overall sample mean $\bar{G}_n \equiv n_{m,n}^{-1} \sum_{\mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d} G_{\theta_0}(Y_\mathbf{s})$, it holds that $n_{m,n}^{1/2} \bar{G}_n \xrightarrow{d} \mathcal{N}(0_r, \Sigma_{\theta_0})$ under Assumptions 1-3 by applying the spatial central limit theorem result in Theorem 4.2 of Lahiri (2003b). Now define a scaled difference between $\bar{G}_n$ and the average of block sample means $\bar{M}_{\theta_0}$ as

$$A_n \ \equiv \ \bar{G}_n - n_{m,n}^{-1} N_\mathcal{I} \bar{M}_{\theta_0} \ = \ n_{m,n}^{-1} \sum_{\mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d} w_\mathbf{s} G_{\theta_0}(Y_\mathbf{s}),$$

where the last representation uses weights $w_\mathbf{s} \in [0,1]$ for each $\mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d$ where

$$w_\mathbf{s} = 1 - b_n^{-d} \times \text{“\# of OL blocks among } \{\mathcal{B}_{b_n}(\mathbf{i}) \equiv \mathbf{i} + b_n(-\tfrac{1}{2}, \tfrac{1}{2}]^d : \mathbf{i} \in \mathcal{I}^{OL}_{b_n}\} \text{ containing } \mathbf{s}\text{”}.$$

Because $w_\mathbf{s} = 0$ if $\mathbf{s} + b_n[-1,1]^d \subset \mathcal{R}_{m,n}$, it holds that $|\{\mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d : w_\mathbf{s} \neq 0\}| \leq |J_n(b_n)| \leq C b_n \lambda_n^{d-1}$ from (9) and $\mathcal{R}_{m,n} \subset \mathcal{R}_n$. Consequently, letting $\mathbf{0} \in \mathbb{Z}^d$ denote the zero vector, we have

$$n_{m,n} \text{E}\,(A_n^2) \leq n_{m,n}^{-1} |\{\mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d : w_\mathbf{s} \neq 0\}| \sum_{\mathbf{h} \in \mathbb{Z}^d} \|\,\text{Cov}\,\{G_{\theta_0}(Y_\mathbf{0}), G_{\theta_0}(Y_\mathbf{h})\}\|$$

$$\leq Cb_n\lambda_n^{d-1}n_{m,n}^{-1} = O\Big(\frac{b_n}{\lambda_n}\Big) = o(1)$$

follows from Lemma 2(i) along with

$$\sum_{\mathbf{h}\in\mathbb{Z}^d,\mathbf{h}\neq\mathbf{0}} \|\operatorname{Cov}(G_{\theta_0}(Y_{\mathbf{0}}), G_{\theta_0}(Y_{\mathbf{h}}))\| \leq C\sum_{v=1}^{\infty}\alpha_Y(v;1)^{\frac{\delta}{6+\delta}}|\{\mathbf{h}\in\mathbb{Z}^d : \|\mathbf{h}\|_\infty = v\}| < \infty, \tag{10}$$

which holds by Lemma 1 with Assumptions 2-3 and $|\{\mathbf{h}\in\mathbb{Z}^d : \|\mathbf{h}\|_\infty = v\}| \leq 2(2v+1)^{d-1}$, $v \geq 1$. Hence, in the OL block case, $n_{m,n}^{1/2}A_n \xrightarrow{p} 0$ and part(ii) follows from the normal limit of $n_{m,n}^{1/2}\bar{G}_n$ along with Slutsky's theorem and $n_{m,n}^{-1}N_{\mathcal{I}} \to 1$ for OL blocks by Lemma 2(i). (In the NOL block case, we define a difference $A_n \equiv \bar{G}_n - n_{m,n}^{-1}b_n^dN_{\mathcal{I}}\bar{M}_{\theta_0} = n_{m,n}^{-1}\sum_{\mathbf{s}\in\mathcal{R}_{m,n}\cap\mathbb{Z}^d} w_{\mathbf{s}}G_{\theta_0}(Y_{\mathbf{s}})$, where weight $w_{\mathbf{s}} = 1$ if site $\mathbf{s}\in\mathcal{R}_{m,n}\cap\mathbb{Z}^d$ belongs to some NOL block in the collection $\{\mathcal{B}_{b_n}(\mathbf{i}) : \mathbf{i}\in\mathcal{I}_{b_n}^{NOL}\}$ and $w_{\mathbf{s}} = 0$ otherwise. Then, $n_{m,n}^{1/2}A_n \xrightarrow{p} 0$ holds for NOL blocks by the same argument and part(ii) then follows by Slutsky's theorem along with $n_{m,n}^{-1}b_n^dN_{\mathcal{I}} \to 1$ for NOL blocks by Lemma 2(i).)

We next establish Lemma 2(iii) for OL blocks $\mathcal{I}_n = \mathcal{I}_{b_n}^{OL}$. Writing $\mathbf{h} = (h_1,\ldots,h_d)' \in \mathbb{Z}^d$, note that by the Dominated Convergence Theorem and (10) we have that

$$\begin{aligned}\mathrm{E}(\widehat{\Sigma}_{\theta_0}) &= b_n^d\mathrm{E}(M_{\theta_0,\mathbf{0}}M_{\theta_0,\mathbf{0}}') = b_n^{-d}\operatorname{Var}\Bigg(\sum_{\mathbf{s}\in\mathcal{B}_{b_n}(\mathbf{0})\cap\mathbb{Z}^d} G_{\theta_0}(Y_{\mathbf{s}})\Bigg)\\ &= b_n^{-d}\sum_{\|\mathbf{h}\|_\infty\leq b_n}\operatorname{Cov}(G_{\theta_0}(Y_{\mathbf{0}}), G_{\theta_0}(Y_{\mathbf{h}}))\prod_{i=1}^d(b_n-|h_i|) \to \Sigma_{\theta_0},\end{aligned}$$

for expectation over the cube $\mathcal{B}_{b_n}(\mathbf{0}) = b_n(-1/2,1/2]^d$. Hence, for part(iii) it suffices to show $\operatorname{Var}(v_1'\widehat{\Sigma}_{\theta_0}v_2) = o(1)$ for any $v_i\in\mathbb{R}^r, \|v_i\| = 1$, $i = 1,2$. Fix v_1, v_2 and expand the variance

$$\begin{aligned}&\operatorname{Var}(v_1'\widehat{\Sigma}_{\theta_0}v_2)\\ &= N_{\mathcal{I}}^{-2}b_n^{2d}\sum_{\mathbf{h}\in\mathbb{Z}^d}|\{\mathbf{i}\in\mathcal{I}_n : \mathbf{i}+\mathbf{h}\in\mathcal{I}_n\}|\operatorname{Cov}\{(v_1'M_{\theta_0,\mathbf{0}}M_{\theta_0,\mathbf{0}}'v_2), (v_1'M_{\theta_0,\mathbf{h}}M_{\theta_0,\mathbf{h}}'v_2)\}\\ &\equiv A_{1n} + A_{2n}\end{aligned}$$

by considering two sums of covariances at displacements $\mathbf{h}\in\mathbb{Z}^d$ with $\|\mathbf{h}\|_\infty \leq b_n$ (i.e., A_{1n}) or $\|\mathbf{h}\|_\infty > b_n$ (i.e., A_{2n}). Then, applying the Cauchy-Schwartz inequality with Lemma 1(ii) and Assumption 3, we have for $\mathbf{h}\in\mathbb{Z}^d$

$$|\operatorname{Cov}\{(v_1'M_{\theta_0,\mathbf{0}}M_{\theta_0,\mathbf{0}}'v_2), (v_1'M_{\theta_0,\mathbf{h}}M_{\theta_0,\mathbf{h}}'v_2)\}|$$

$$\leq \mathrm{Var}\,(v_1' M_{\theta_0,\mathbf{0}} M_{\theta_0,\mathbf{0}}' v_2) \leq \mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^4) \leq C b_n^{-2d},$$

so that $|A_{1n}| \leq C N_{\mathcal{I}}^{-1} |\{\mathbf{h} \in \mathbb{Z}^d : \|\mathbf{h}\|_\infty \leq b_n\}| = O(b_n^d/\lambda_n^d) = o(1)$ by Lemma 2(i) for OL blocks. For $\mathbf{h} \in \mathbb{Z}^d$ with $\|\mathbf{h}\|_\infty > b_n$, it holds that $\mathrm{dis}[\mathcal{B}_{b_n}(\mathbf{0}), \mathcal{B}_{b_n}(\mathbf{h})] \geq 1$ so that by Assumption 3 and Lemma 1(i) (i.e., taking $s = t = 3/(6+\delta)$ there for δ in Assumption 3), we may bound the covariance $|\,\mathrm{Cov}\,\{(v_1' M_{\theta_0,\mathbf{0}} M_{\theta_0,\mathbf{0}}' v_2), (v_1' M_{\theta_0,\mathbf{h}} M_{\theta_0,\mathbf{h}}' v_2)\}|$ by the quantity $C\{\mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^{(12+2\delta)/3})\}^{6/(6+\delta)} \alpha_Y(\mathrm{dis}[\mathcal{B}_{b_n}(\mathbf{0}), \mathcal{B}_{b_n}(\mathbf{h})], b_n^d)^{\delta/(6+\delta)}$ where the moment satisfies $\{\mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^{(12+2\delta)/3})\}^{6/(6+\delta)} \leq C b_n^{-2d}$ by Lemma 1(ii). By Lemma 2(i) and Assumptions 2-3, we then bound

$$\begin{aligned}
|A_{2n}| &\leq \frac{b_n^{2d}}{N_{\mathcal{I}}} \sum_{\mathbf{h} \in \mathbb{Z}^d, \|\mathbf{h}\|_\infty > b_n} \{\mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^{\frac{12+2\delta}{3}})\}^{\frac{6}{6+\delta}} \alpha_Y\left(\mathrm{dis}[\mathcal{B}_{b_n}(\mathbf{0}), \mathcal{B}_{b_n}(\mathbf{h})], b_n^d\right)^{\frac{\delta}{6+\delta}} \\
&\leq \frac{C}{N_{\mathcal{I}}} \sum_{k=1}^{\infty} k(k+b_n)^{d-1} \alpha_Y(k, b_n^d)^{\frac{\delta}{6+\delta}} \\
&\leq \frac{C}{N_{\mathcal{I}}} \sum_{k=1}^{b_n} k(k+b_n)^{d-1} + \frac{C b_n^{\frac{d\kappa\delta}{6+\delta}}}{N_{\mathcal{I}}} \sum_{k=b_n+1}^{\infty} \left(\frac{k}{b_n}\right)^{4d-1} k^d \alpha_1(k)^{\frac{\delta}{6+\delta}} \\
&\leq C\lambda_n^{-d} b_n^{d+1} + C\lambda_n^{-d} b_n^d \sum_{k=b_n+1}^{\infty} k^{5d-1} \alpha_1(k)^{\frac{\delta}{6+\delta}} = o(1),
\end{aligned}$$

using $|\{\mathbf{h} \in \mathbb{Z}^d : \mathrm{dis}[\mathcal{B}_{b_n}(\mathbf{0}), \mathcal{B}_{b_n}(\mathbf{h})] = k\}| \leq Ck(k+b_n)^{d-1}$, $k \geq 1$, in the second inequality and substituting $(k/b_n)^{4d-1} \geq 1$ in the second sum of the third inequality. So part(iii) follows for OL blocks. (We note that, in the case of NOL blocks, the above argument that $\mathrm{Var}\,(v_1' \widehat{\Sigma}_{\theta_0} v_2) = o(1)$ must be slightly modified. When $\mathcal{I}_n = \mathcal{I}_{b_n}^{NOL}$ and $N_{\mathcal{I}} = |\mathcal{I}_{b_n}^{NOL}|$, then

$$\begin{aligned}
&\mathrm{Var}\,(v' \widehat{\Sigma}_{\theta_0} v) \\
&= \frac{b_n^{2d}}{N_{\mathcal{I}}} \sum_{\mathbf{h} \in \mathbb{Z}^d} |\{\mathbf{i} \in \mathcal{I}_n : \mathbf{i} + b_n \mathbf{h} \in \mathcal{I}_n\}|\, \mathrm{Cov}\,\{(v_1' M_{\theta_0,\mathbf{0}} M_{\theta_0,\mathbf{0}}' v_2), (v_1' M_{\theta_0,b_n\mathbf{h}} M_{\theta_0,b_n\mathbf{h}}' v_2)\} \\
&\equiv A_{1n} + A_{2n}
\end{aligned}$$

where $A_{1n} = N_{\mathcal{I}}^{-1} b_n^{2d} \mathrm{Var}\,(v_1' M_{\theta_0,b_n\mathbf{h}} M_{\theta_0,b_n\mathbf{h}}' v_2) = O(N_{\mathcal{I}}^{-1}) = o(1)$ corresponds to the covariance sum at lag $\mathbf{h} = \mathbf{0}$ and $A_{2n} = o(1)$ represents the sum of covariance terms over non-zero lags $\|\mathbf{h}\| > 0$.)

In proving the remaining parts of Lemma 2, we need not make a distinction between OL or NOL blocks. To show part(iv) of Lemma 2, we will assume part(vi) holds. We argue that a contradiction arises by supposing that the event in probability statement of part(vi) holds and the zero vector $0_r \in \mathbb{R}^r$ is *not* interior to the convex hull of $\{M_{\theta_0,\mathbf{i}} : \mathbf{i} \in \mathcal{I}_n\}$. If 0_r is not interior, then by supporting/separating hyperplane theorem there exists some $v \in \mathbb{R}^r, \|v\| = 1$ where $v'M_{\theta_0,\mathbf{i}} \leq v'0_r = 0$ holds for all $\mathbf{i} \in \mathcal{I}_n$; however, this contradicts the event in the probability statement of part(vi), which implies that $v'M_{\theta_0,\mathbf{i}} > 0$ holds for some $\mathbf{i} \in \mathcal{I}_n$. Therefore, whenever the event in part(vi) holds, then 0_r must be interior to the convex hull of $\{M_{\theta_0,\mathbf{i}} : \mathbf{i} \in \mathcal{I}_n\}$, which implies $R_n(\theta_0) > 0$ by (5). Hence, part(vi) implies part(iv) of the lemma.

To show part(v), note

$$\mathrm{E}\,(\max_{\mathbf{i}\in\mathcal{I}_n} \|M_{\theta_0,\mathbf{i}}\|) \leq \mathrm{E}\left\{\Big(\sum_{\mathbf{i}\in\mathcal{I}_n} \|M_{\theta_0,\mathbf{i}}\|^6\Big)^{\frac{1}{6}}\right\} = \left\{\sum_{\mathbf{i}\in\mathcal{I}_n} \mathrm{E}\,(\|M_{\theta_0,\mathbf{i}}\|^6)\right\}^{\frac{1}{6}} \leq Cb_n^{-\frac{d}{2}} N_{\mathcal{I}}^{\frac{1}{6}}$$

by Lemma 1(ii) so that $n_{m,n}^{-5/12} b_n^d \max_{\mathbf{i}\in\mathcal{I}_n} \|M_{\theta_0,\mathbf{i}}\| = O_p(n_{m,n}^{-1/4} b_n^{d/2}) = O_p(\lambda_n^{-d/4} b_n^{d/2}) = o_p(1)$ by Assumption 1, Lemma 2(i) and $N_{\mathcal{I}} \leq n_{m,n}$.

Finally, to establish part(vi), we employ an empirical distribution of block means $\hat{F}(v) = N_{\mathcal{I}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_n} \mathbb{I}(b_n^{d/2} M_{\theta_0,\mathbf{i}} \leq v)$, $v \in \mathbb{R}^r$. For fixed $v \in \mathbb{R}^d$, it holds that $|\hat{F}_n(v) - P(Z \leq v)| = o_p(1)$ where Z denotes a normal $\mathcal{N}(0_r, \Sigma_{\theta_0})$ random vector. This can be shown using $\mathrm{E}\,\{\hat{F}_n(v)\} = P(b_n^{d/2} M_{\theta_0,\mathbf{0}} \leq v) \to P(Z \leq v)$ under Assumptions 1-3 by applying a central limit theorem for the block sample mean $b_n^{d/2} M_{\theta_0,\mathbf{0}}$ (Theorem 4.2, Lahiri, 2003b) and verifying $\mathrm{Var}\,\{\hat{F}_n(v)\} = o(1)$ similar to the proof of Lemma 2(iii). Consequently, $\sup_{v\in\mathbb{R}^r} |\hat{F}_n(v) - P(Z \leq v)| = o_p(1)$ holds by Polya's theorem and, from this and part(iii), one can prove convergence of the following absolute "half-space" moments of $\hat{F}_n(\cdot)$

$$\sup_{v\in\mathbb{R}^r, \|v\|=1} \big|N_{\mathcal{I}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_n} b_n^{d/2} |v'M_{\theta_0,\mathbf{i}}| \,-\, \mathrm{E}\,|v'Z|\big| = o_p(1).$$

Using this along with $b_n^{1/2} \bar{M}_{\theta_0} \xrightarrow{p} 0_r$ by part(ii), where $\bar{M}_{\theta_0} = N_{\mathcal{I}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_n} M_{\theta_0,\mathbf{i}}$, we have

$$\sup_{v\in\mathbb{R}^r, \|v\|=1} \big|N_{\mathcal{I}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_n} b_n^{d/2} v'M_{\theta_0,\mathbf{i}} \mathbb{I}(v'M_{\theta_0,\mathbf{i}} > 0) - 2^{-1}\mathrm{E}\,|v'Z|\big| = o_p(1)$$

because $v'M_{\theta_0,\mathbf{i}}\mathbb{I}(v'M_{\theta_0,\mathbf{i}} > 0) = (|v'M_{\theta_0,\mathbf{i}}| + v'M_{\theta_0,\mathbf{i}})/2$ for $\mathbf{i} \in \mathcal{I}_n$, $v \in \mathbb{R}^r$. Now part(vi) follows using the fact that $\inf_{v\in\mathbb{R}^r, \|v\|=1} \mathrm{E}\,|v'Z| \geq C$ holds for some $C > 0$ since $\mathrm{Var}\,(Z) = \Sigma_{\theta_0}$ is positive definite by Assumption 4.

Proof of Lemma 3. By Lemma 2(iv), a positive $R_n(\theta_0)$ exists in probability and can be written, from (5), as $R_n(\theta_0) = \prod_{\mathbf{i}\in\mathcal{I}_n} (1 + \gamma_{\theta_0,\mathbf{i}})^{-1}$ with $\gamma_{\theta_0,\mathbf{i}} = t'_{\theta_0} M_{\theta_0,\mathbf{i}} < 1$, where $t_{\theta_0} \in \mathbb{R}^r$ satisfies $Q_{1n}(\theta_0, t_{\theta_0}) = 0_r$ in (15). By Lemma 2, it holds that $Z_{\theta_0} \equiv \max_{\mathbf{i}\in\mathcal{I}_n} \|M_{\theta_0,\mathbf{i}}\| = o_p(b_n^{-d} n_{m,n}^{1/2})$. We now modify an argument from

Owen (1990, p. 101) by writing $t_{\theta_0} = \|t_{\theta_0}\| u_{\theta_0}$ with $u_{\theta_0} \in \mathbb{R}^r$, $\|u_{\theta_0}\| = 1$, and then expanding $Q_{1n}(\theta_0, t_{\theta_0}) = 0_r$ to find

$$0 = -n_{m,n}^{1/2} u'_{\theta_0} Q_{1n}(\theta_0, t_{\theta_0}) = \frac{n_{m,n}^{\frac{1}{2}} \|t_{\theta_0}\|}{N_{\mathcal{I}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \frac{u'_{\theta_0} M_{\theta_0,\mathbf{i}} M'_{\theta_0,\mathbf{i}} u_{\theta_0}}{1 + \gamma_{\theta_0,\mathbf{i}}} - n_{m,n}^{\frac{1}{2}} u'_{\theta_0} \bar{M}_{\theta_0}$$
$$\geq \frac{n_{m,n}^{\frac{1}{2}} b_n^{-d} \|t_{\theta_0}\| u'_{\theta_0} \widehat{\Sigma}_{\theta_0} u_{\theta_0}}{1 + (n_{m,n}^{-\frac{1}{2}} b_n^d Z_{\theta_0})(n_{m,n}^{\frac{1}{2}} b_n^{-d} \|t_{\theta_0}\|)} - n_{m,n}^{\frac{1}{2}} \|\bar{M}_{\theta_0}\| \quad (11)$$

where the inequality follows upon replacing each $\gamma_{\theta_0,\mathbf{i}}$ with $Z_{\theta_0}\|t_{\theta_0}\|$ and $u'_{\theta_0}\bar{M}_{\theta_0}$ with $\|\bar{M}_{\theta_0}\|$ and using the definitions of $\bar{M}_{\theta_0}, \widehat{\Sigma}_{\theta_0}$ from Lemma 2. Then combining the facts that $n_{m,n}^{-1/2} b_n^d Z_{\theta_0} = o_p(1)$, that $n_{m,n}^{1/2}\|\bar{M}_{\theta_0}\| = O_p(1)$ by Lemma 2(ii), and that $P(u'_{\theta_0}\widehat{\Sigma}_{\theta_0} u_{\theta_0} > C) \to 1$ for some $C > 0$ by Lemma 2(iii) and Assumption 4, we deduce $\|t_{\theta_0}\| = O_p(b_n^d n_{m,n}^{-1/2})$ from (11). From this, we also have $\max_{\mathbf{i}\in\mathcal{I}_n} |\gamma_{\theta_0,\mathbf{i}}| \leq \|t_{\theta_0}\| Z_{\theta_0} = o_p(1)$.

As $\widehat{\Sigma}_{\theta_0}$ is positive definite in probability, we may algebraically solve $Q_{1n}(\theta_0, t_{\theta_0}) = 0_r$ for $t_{\theta_0} = b_n^d \widehat{\Sigma}_{\theta_0}^{-1} \bar{M}_{\theta_0} + \phi_{\theta_0}$ where

$$\|\phi_{\theta_0}\| \leq \frac{Z_{\theta_0}\|t_{\theta_0}\|^2 \|\widehat{\Sigma}_{\theta_0}^{-1}\| \|\widehat{\Sigma}_{\theta_0}\|}{1 - \|t_{\theta_0}\| Z_{\theta_0}} = o_p(b_n^d n_{m,n}^{-\frac{1}{2}}). \quad (12)$$

Applying a Taylor expansion gives $\log(1 + \gamma_{\theta_0,\mathbf{i}}) = \gamma_{\theta_0,\mathbf{i}} - \gamma_{\theta_0,\mathbf{i}}^2/2 + \Delta_{\mathbf{i}}$ for each $\mathbf{i} \in \mathcal{I}_n$ so that

$$\ell_n(\theta_0) = 2B_n \sum_{\mathbf{i}\in\mathcal{I}_n} \log(1+\gamma_{\theta_0,\mathbf{i}}) = n_{m,n}(\bar{M}'_{\theta_0}\widehat{\Sigma}_{\theta_0}^{-1}\bar{M}_{\theta_0} - b_n^{-2d}\phi'_{\theta_0}\widehat{\Sigma}_{\theta_0}\phi_{\theta_0}) + 2B_n \sum_{\mathbf{i}\in\mathcal{I}_n} \Delta_{\mathbf{i}} \quad (13)$$

where $B_n = n_{m,n}/(b_n^d N_{\mathcal{I}})$. By Lemma 2(ii)$-$(iii), $n_{m,n}\bar{M}'_{\theta_0}\widehat{\Sigma}_{\theta_0}^{-1}\bar{M}_{\theta_0} \xrightarrow{d} \chi_r^2$ and it also holds that $b_n^{-2d} n_{m,n} \phi'_{\theta_0}\widehat{\Sigma}_{\theta_0}\phi_{\theta_0} = o_p(1)$ from (12). Finally, we may bound

$$2B_n \sum_{\mathbf{i}\in\mathcal{I}_n} |\Delta_{\mathbf{i}}| \leq \frac{b_n^{-2d} n_{m,n} 2 Z_{\theta_0} \|t_{\theta_0}\|^3 \|\widehat{\Sigma}_{\theta_0}\|}{(1 - Z_{\theta_0}\|t_{\theta_0}\|)^2} = o_p(1). \quad (14)$$

Lemma 3 then follows by Slutsky's Theorem.

A.3. Proofs of the main results

Proof of Theorem 1. In the case that $H(\theta) = \theta$ is the identity mapping, the result follows immediately from Lemma 3. From this, Theorem 1 follows for a general smooth $H(\cdot)$ as in the proof of Theorem 2.1 of Hall and La Scala (1990).

Proof of Theorem 2. Set $\Theta_n = \{\theta \in \Theta : \|\theta - \theta_0\| \le n_{m,n}^{-5/12}\}$, $\partial\Theta_n = \{\theta \in \Theta : \|\theta - \theta_0\| = n_{m,n}^{-5/12}\}$ and define $\bar{M}_\theta = \sum_{\mathbf{i}\in\mathcal{I}_n} M_{\theta,\mathbf{i}}/N_{\mathcal{I}}$, $\widehat{\Sigma}_\theta = b_n^d \sum_{\mathbf{i}\in\mathcal{I}_n} M_{\theta,\mathbf{i}} M'_{\theta,\mathbf{i}}/N_{\mathcal{I}}$, $\theta \in \Theta_n$ and functions

$$Q_{1n}(\theta, t) = \frac{1}{N_{\mathcal{I}}} \sum_{\mathbf{i}\in\mathcal{I}_n} \frac{M_{\theta,\mathbf{i}}}{1 + t' M_{\theta,\mathbf{i}}}, \qquad Q_{2n}(\theta, t) = \frac{b_n^{-d}}{N_{\mathcal{I}}} \sum_{\mathbf{i}\in\mathcal{I}_n} \frac{\left(\frac{\partial M_{\theta,\mathbf{i}}}{\partial\theta}\right)' t}{1 + t' M_{\theta,\mathbf{i}}}, \tag{15}$$

on $\Theta \times \mathbb{R}^r$. For $i = 1, 3$, set $J_{n,i} = \sum_{\mathbf{s}\in\mathcal{R}_{m,n}\cap\mathbb{Z}^d} J^i(Y_{\mathbf{s}})/n_{m,n}$, noting $J_{i,n} = O_p(1)$ by $\mathrm{E}\, J^3(Y_{\mathbf{s}}) < \infty$; again $J(\cdot)$ is assumed to be nonnegative. To establish Theorem 2, we proceed in three steps to show, that with arbitrarily large probability as $n \to \infty$, the following hold: Step 1. the log EL ratio $\ell_n(\theta)$ exists finitely on Θ_n and is continuously differentiable and hence a sequence of minimums $\hat{\theta}_n$ exists of $\ell_n(\theta)$ on Θ_n (i.e., $\hat{\theta}_n$ is a maximizer of $R_n(\theta)$); Step 2. $\hat{\theta}_n \notin \partial\Theta_n$ and $\partial\ell_n(\theta)/\partial\theta = 0_p$ at $\theta = \hat{\theta}_n$; Step 3. $\hat{\theta}_n$ has the normal limit stated in Theorem 2.

Step 1. Note that

$$\begin{aligned}
&\sup_{\substack{v\in\mathbb{R}^r, \|v\|=1 \\ \theta\in\Theta_n}} \left| N_{\mathcal{I}}^{-1} \sum_{\mathbf{i}\in\mathcal{I}_n} \left(v' M_{\theta,\mathbf{i}} \mathbb{I}(v' M_{\theta,\mathbf{i}} > 0) - v' M_{\theta_0,\mathbf{i}} \mathbb{I}(v' M_{\theta_0,\mathbf{i}} > 0) \right) \right| \\
&\quad \le \sup_{\theta\in\Theta_n} \sum_{\mathbf{i}\in\mathcal{I}_n} \frac{\|M_{\theta,\mathbf{i}} - M_{\theta_0,\mathbf{i}}\|}{N_{\mathcal{I}}},
\end{aligned}$$

which is bounded by $C J_{n,1} \sup_{\theta\in\Theta_n} \|\theta - \theta_0\| = O_p(n_{m,n}^{-5/12}) = o_p(b_n^{-d/2})$. From this and Lemma 2(vi), it holds that, for some $C > 0$,

$$P\left(\inf_{\|v\|=1, \theta\in\Theta_n} \sum_{\mathbf{i}\in\mathcal{I}_n} b_n^{\frac{d}{2}} v' M_{\theta,\mathbf{i}} \frac{\mathbb{I}(v' M_{\theta,\mathbf{i}} > 0)}{N_{\mathcal{I}}} > C \right) \to 1$$

As proof of Lemma 2(iv), when the event in the above probably statement holds, then for any $\theta \in \Theta_n$, we may write $R_n(\theta) = \prod_{\mathbf{i}\in\mathcal{I}_n} (1 + \gamma_{\theta,\mathbf{i}}) > 0$ where $\gamma_{\theta,\mathbf{i}} = t'_\theta M_{\theta,\mathbf{i}}$ and $Q_{1n}(\theta, t_\theta) = 0_r$.

Let $\Omega_\theta = \max\{n_{m,n}^{-1/2}, \|\theta - \theta_0\|\}$, $\theta \in \Theta_n$. Expanding both $\bar{M}_\theta$ and $\widehat{\Sigma}_\theta$ around θ_0, we find

$$\begin{aligned}
\sup_{\theta\in\Theta_n} \frac{\|\bar{M}_\theta\|}{\Omega_\theta} &\le n_{m,n}^{\frac{1}{2}} \|\bar{M}_{\theta_0}\| + C J_{n,1} \sup_{\theta\in\Theta_n} \Omega_\theta^{-1} \|\theta - \theta_0\| = O_p(1), \\
\sup_{\theta\in\Theta_n} \|\widehat{\Sigma}_\theta - \Sigma_{\theta_0}\| &\le \sup_{\theta\in\Theta_n} \|\widehat{\Sigma}_\theta - \widehat{\Sigma}_{\theta_0}\| + \|\widehat{\Sigma}_{\theta_0} - \Sigma_{\theta_0}\| = o_p(1),
\end{aligned} \tag{16}$$

by applying Lemma 2(ii)$-$(iii) above along with $\Omega_\theta^{-1} \le n_{m,n}^{1/2}$ and

$$\sup_{\theta\in\Theta_n} \|\widehat{\Sigma}_\theta - \widehat{\Sigma}_{\theta_0}\| \le \sup_{\theta\in\Theta_n} \frac{b_n^d}{N_{\mathcal{I}}} \sum_{\mathbf{i}\in\mathcal{I}_n} \|M_{\theta_0,\mathbf{i}}\| \|M_{\theta_0,\mathbf{i}} - M_{\theta,\mathbf{i}}\| (1 + \|M_{\theta_0,\mathbf{i}} - M_{\theta,\mathbf{i}}\|) \equiv A_n$$

$$\begin{aligned}\mathrm{E}\,(A_n) &\le C n_{m,n}^{-\frac{5}{12}} b_n^d \{\mathrm{E}\,[J(Y_{\mathbf{0}})^3]\}^{\frac{2}{3}} \left\{\mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^3) + [\mathrm{E}\,(\|M_{\theta_0,\mathbf{0}}\|^3)]^2\right\}^{\frac{1}{3}} \\ &\le C n_{m,n}^{-\frac{5}{12}} b_n^{\frac{d}{2}} = o(1),\end{aligned}$$

which follows from Holder's inequality, $n_{m,n} \sim \mathrm{vol}(\mathcal{R}_0)\lambda_n^d$ by Lemma 2(i), and using Lemma 2(ii) in the last line. Hence, by the positive definiteness of Σ_{θ_0} in Assumption 4, $\widehat{\Sigma}_\theta^{-1}$ exists uniformly in $\theta \in \Theta_n$. Also, the positive definiteness of $\widehat{\Sigma}_n$ by (16) implies, for each fixed $\theta \in \Theta_n$, $\partial Q_{1n}(\theta, t)/\partial t$ is negative definitive for $t \in \{t \in \mathbb{R}^r : 1 + t' M_{\theta,\mathbf{i}} \ge 1/N_{\mathcal{I}}, \mathbf{i} \in \mathcal{I}_n\}$ so that, by implicit function theorem using $Q_{1n}(\theta, t_\theta) = 0_r$, t_θ is a continuously differentiable function of θ on Θ_n and the function $\ell_n(\theta) = -2B_n \log R_n(\theta)$ is as well (e.g., Qin and Lawless, 1994, p. 304-305). Hence, with large probability as $n \to \infty$, the minimizer of $\ell_n(\theta)$ exists on Θ_n.

Step 2. Let $Z_\theta \equiv \max_{\mathbf{i}\in\mathcal{I}_n} \|M_{\mathbf{i},\theta}\|$, $\theta \in \Theta_n$. Using $b_n^2/\lambda_n = o(1)$ by Assumption 1, $\sup_{\theta\in\Theta_n} \Omega_\theta \le n_{m,n}^{-5/12}$, and Lemma 2 [parts (i) and (v)], we may expand the block means $M_{\theta,\mathbf{i}}, \mathbf{i} \in \mathcal{I}_n$ around θ_0 to find

$$\begin{aligned}\sup_{\theta\in\Theta_n} \Omega_\theta b_n^d Z_\theta &\le b_n^d n_{m,n}^{-\frac{5}{12}} \left(\max_{\mathbf{i}\in\mathcal{I}_n} \|M_{\mathbf{i},\theta_0}\| + \sup_{\theta\in\Theta_n} C\|\theta-\theta_0\| (n_{m,n} J_{n,3})^{\frac{1}{3}} \right) \\ &\le o_p(1) + O_p(b_n^d n_{m,n}^{-\frac{1}{2}}) = o_p(1). \qquad (17)\end{aligned}$$

Now using (16) and (17) and that $Q_{1n}(\theta, t_\theta) = 0_r$ for $\theta \in \Theta_n$, we can repeat the same essential argument in (11) (i.e., replace $\theta_0, n_{m,n}^{1/2}$ there with $\theta, \Omega_\theta^{-1}$) to find

$$0 \ge \frac{\Omega_\theta^{-1} b_n^{-d} \|t_\theta\| u_\theta' \widehat{\Sigma}_\theta u_\theta}{1 + (\Omega_\theta b_n^d Z_\theta)(\Omega_\theta^{-1} b_n^{-d} \|t_\theta\|)} - \Omega_\theta^{-1} \|\bar{M}_\theta\| \qquad \big(\text{with } t_\theta = \|t_\theta\| u_\theta,\ \|u_\theta\| = 1\big)$$

and then show $\sup_{\theta\in\Theta_n} \Omega_\theta^{-1} b_n^{-d} \|t_\theta\| = O_p(1)$. From this (and analogous to (12) from the proof of Lemma 3), we expand $Q_{1n}(\theta, t_\theta) = 0_r$ to yield $t_\theta = b_n^d \widehat{\Sigma}_\theta^{-1} \bar{M}_\theta + \phi_\theta$ for $\theta \in \Theta_n$ where $\sup_{\theta\in\Theta_n} \Omega_\theta^{-1} b_n^{-d} \|\phi_\theta\| = o_p(1)$. Using now these orders of $\|\phi_\theta\|, \|t_\theta\|$ and Z_θ with arguments as in (13) and (14), we may then expand $\ell_n(\theta)$ uniformly in $\theta \in \Theta_n$ as

$$\begin{aligned}&\sup_{\theta\in\Theta_n} n_{m,n}^{-1} \Omega_\theta^{-2} |\ell_n(\theta) - n_{m,n} \bar{M}_\theta' \widehat{\Sigma}_\theta^{-1} \bar{M}_\theta| \\ &\quad \le O_p\left(\Omega_\theta^{-2} b_n^{-2d} \sup_{\theta\in\Theta_n} \left[\phi_\theta' \widehat{\Sigma}_\theta \phi_\theta + \frac{2Z_\theta \|t_\theta\|^2 \|\widehat{\Sigma}_\theta\|}{(1 - Z_\theta\|t_\theta\|)^2} \right] \right) = o_p(1)\end{aligned}$$

and then using (16)

$$\sup_{\theta\in\Theta_n} n_{m,n}^{-1} \Omega_\theta^{-2} |\ell_n(\theta) - n_{m,n} \bar{M}_\theta' \widehat{\Sigma}_{\theta_0}^{-1} \bar{M}_\theta| = o_p(1)$$

follows. For each $\theta \in \Theta_n$, we may write $\bar{M}_\theta = \bar{M}_{\theta_0} + \bar{D}_{\theta_0}(\theta - \theta_0) + E_\theta$ for $\bar{D}_{\theta_0} = N_{\mathcal{I}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_n} \partial M_{\theta_0,\mathbf{i}}/\partial\theta$ and a remainder E_θ satisfying $\sup_{\theta\in\Theta_n}\|E_\theta\| \leq C\|\theta-\theta_0\|^2 J_{n,1}$. Note that $\bar{D}_{\theta_0} \xrightarrow{p} D_{\theta_0} \equiv \mathrm{E}\,\partial G_{\theta_0}(Y_{\mathbf{t}})/\partial\theta$ because $\mathrm{E}\,\bar{D}_{\theta_0} = D_{\theta_0}$ and, as in (10),

$$\mathrm{Var}\,(\bar{D}_{\theta_0}) \leq Cn_{m,n}^{-1}\sum_{\mathbf{h}\in\mathbb{Z}^d}\left\|\mathrm{Cov}\left\{\frac{\partial G_{\theta_0}(Y_{\mathbf{0}})}{\partial\theta}, \frac{\partial G_{\theta_0}(Y_{\mathbf{h}})}{\partial\theta}\right\}\right\| \leq Cn_{m,n}^{-1}$$

by Lemma 1 and Assumptions 2–3. Hence, we have

$$\sup_{\theta\in\Theta_n}|\bar{M}_\theta - [\bar{M}_{\theta_0} + D_{\theta_0}(\theta-\theta_0)]| = o_p(\Omega_\theta) \tag{18}$$

and so it now follows that

$$\sup_{\theta\in\Theta_n} n_{m,n}^{-1}\Omega_\theta^{-2}\left|\ell_n(\theta) - n_{m,n}\left[\bar{M}_{\theta_0} + D_{\theta_0}(\theta-\theta_0)\right]'\Sigma_{\theta_0}^{-1}\left[\bar{M}_{\theta_0} + D_{\theta_0}(\theta-\theta_0)\right]\right| = o_p(1). \tag{19}$$

For $\theta = v_\theta n_{m,n}^{-5/12} + \theta_0 \in \partial\Theta_n$, $\|v_\theta\| = 1$, we have $\Omega_\theta = n_{m,n}^{-5/12}$ so that from (19) we find that $\ell_n(\theta) \geq \sigma n_{m,n}^{1/6}/2$ holds uniformly in $\theta \in \partial\Theta_n$ when n is large, where σ denotes the smallest eigenvalue of $D'_{\theta_0}\Sigma_{\theta_0}^{-1}D_{\theta_0}$. At the same time, by Lemma 3, we have $\ell_n(\theta_0) = O_p(1)$ (i.e., $n_{m,n}^{-1}\Omega_{\theta_0}^{-2} = 1$ in (19)). Hence, with probability approaching 1, the minimum $\hat{\theta}_n$ of $\ell_n(\theta)$ on Θ_n cannot be an element of $\partial\Theta_n$. Hence, $\hat{\theta}_n$ must satisfy $\hat{\theta}_n \in \Theta_n \setminus \partial\Theta_n$ and $0_r = Q_{1n}(\hat{\theta}_n, t_{\hat{\theta}_n})$ in addition to

$$0_p = (2n_{m,n})^{-1}\frac{\partial\ell_n(\theta)}{\partial\theta}\Big|_{\theta=\hat{\theta}_n} = Q_{2n}(\hat{\theta}_n, t_{\hat{\theta}_n})$$

by the differentiability of $\ell_n(\theta)$.

Step 3. From the argument in Step 2, we may solve $Q_{1n}(\hat{\theta}_n, t_{\hat{\theta}_n}) = 0_r$ for $t_{\hat{\theta}_n} = b_n^d\widehat{\Sigma}_{\hat{\theta}_n}^{-1}\bar{M}_{\hat{\theta}_n} + \phi_{\hat{\theta}_n}$ or

$$b_n^{-d}t_{\hat{\theta}_n} \;=\; \widehat{\Sigma}_{\hat{\theta}_n}^{-1}\bar{M}_{\hat{\theta}_n} + b_n^{-d}\phi_{\hat{\theta}_n} \;=\; \Sigma_{\theta_0}^{-1}\left[\bar{M}_{\theta_0} + D_{\theta_0}(\theta-\theta_0)\right] + o_p(\Omega_{\hat{\theta}_n}) \tag{20}$$

by $\Omega_{\hat{\theta}_n}^{-1}b_n^{-d}\|\phi_{\hat{\theta}_n}\| = o_p(1)$, (16) and (18). Recalling also $\bar{D}_{\theta_0} = N_{\mathcal{I}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_n}\partial M_{\theta_0,\mathbf{i}}/\partial\theta \xrightarrow{p} D_{\theta_0}$ from Step 2 along with $\|\bar{D}_{\theta_0} - N_{\mathcal{I}}^{-1}\sum_{\mathbf{i}\in\mathcal{I}_n}\partial M_{\hat{\theta}_n,\mathbf{i}}/\partial\theta\| = O_p(\|\hat{\theta}_n - \theta_0\|)$, and $\max_{\mathbf{i}\in\mathcal{I}_n}|t'_{\hat{\theta}_n}M_{\hat{\theta}_n,\mathbf{i}}| \leq \|t'_{\hat{\theta}_n}\|Z_{\hat{\theta}_n} = o_p(1)$ (where again $Z_{\hat{\theta}_n} = \max_{\mathbf{i}\in\mathcal{I}_n}\|M_{\hat{\theta}_n,\mathbf{i}}\|$), we find from $Q_{2n}(\hat{\theta}_n, t_{\hat{\theta}_n}) = 0_p$ that

$$0_p \;=\; \frac{b_n^{-d}}{N_{\mathcal{I}}}\sum_{\mathbf{i}\in\mathcal{I}_n}\frac{\left(\frac{\partial M_{\hat{\theta}_n,\mathbf{i}}}{\partial\theta}\right)' t_{\hat{\theta}_n}}{1 + t'_{\hat{\theta}_n}M_{\hat{\theta}_n,\mathbf{i}}} \;=\; D'_{\theta_0}b_n^{-d}t_{\hat{\theta}_n} + o_p(\|b_n^{-d}t_{\hat{\theta}_n}\|). \tag{21}$$

Now letting $\delta_n = \|b_n^{-d} t_{\hat{\theta}_n}\| + \Omega_{\hat{\theta}_n}$, from (20) and (21) we may from write

$$\begin{bmatrix} \Sigma_{\theta_0} & -D_{\theta_0} \\ D'_{\theta_0} & 0 \end{bmatrix} \begin{pmatrix} b_n^{-d} t_{\hat{\theta}_n} \\ \hat{\theta}_n - \theta_0 \end{pmatrix} = \begin{bmatrix} \bar{M}_{\theta_0} + o_p(\delta_n) \\ o_p(\delta_n) \end{bmatrix},$$

$$\begin{bmatrix} \Sigma_{\theta_0} & -D_{\theta_0} \\ D'_{\theta_0} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} U_{\theta_0} & \Sigma_{\theta_0}^{-1} D_{\theta_0} V_{\theta_0} \\ -V_{\theta_0} D'_{\theta_0} \Sigma_{\theta_0}^{-1} & V_{\theta_0} \end{bmatrix}.$$

By Lemma 2(ii), $n_{m,n}^{1/2} \bar{M}_{\theta_0} \xrightarrow{d} \mathcal{N}(0, \Sigma_{\theta_0})$ holds so it follows that $n_{m,n}^{1/2} \delta_n = O_p(1)$ and the limiting distribution of $\hat{\theta}_n$ is given by

$$n_{m,n}^{\frac{1}{2}} \begin{pmatrix} b_n^{-d} t_{\hat{\theta}_n} \\ \hat{\theta}_n - \theta_0 \end{pmatrix} = \begin{bmatrix} U_{\theta_0} \\ -V_{\theta_0} D'_{\theta_0} \Sigma_{\theta_0}^{-1} \end{bmatrix} n_{m,n}^{\frac{1}{2}} \bar{M}_{\theta_0} + o_p(1) \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0_r \\ 0_p \end{pmatrix}, \begin{bmatrix} U_{\theta_0} & 0 \\ 0 & V_{\theta_0} \end{bmatrix} \right) \quad (22)$$

The proof of Theorem 2 is complete.

Proof of Theorem 3. Let $P_X = X(X'X)^{-1}X'$ denote the projection matrix for a given matrix X of full column rank and let $I_{r\times r}$ denote the $r \times r$ identity matrix. Using (19) along with $\|\hat{\theta}_n - \theta_0\| = O_p(n_{m,n}^{-1/2})$ by (22) and $n_{m,n}^{-1}\Omega_{\theta_0}^{-2} = 1$ in (19), we write

$$\ell_n(\hat{\theta}_n) = n_{m,n}(\Sigma_{\theta_0}^{-\frac{1}{2}} \bar{M}_{\theta_0})'(I_{r\times r} - P_{\Sigma_{\theta_0}^{-\frac{1}{2}} D_{\theta_0}})(\Sigma_{\theta_0}^{-\frac{1}{2}} \bar{M}_{\theta_0}) + o_p(1),$$

$$\ell_n(\theta_0) = n_{m,n}(\Sigma_{\theta_0}^{-\frac{1}{2}} \bar{M}_{\theta_0})'(\Sigma_{\theta_0}^{-\frac{1}{2}} \bar{M}_{\theta_0}) + o_p(1).$$

The chi-square limit distributions in Theorem 3(i) now follow by Lemma 2(ii) as $P_{\Sigma_{\theta_0}^{-1/2} D_{\theta_0}}$, $I_{r\times r} - P_{\Sigma_{\theta_0}^{-1/2} D_{\theta_0}}$ are orthogonal idempotent matrices with ranks p, $r-p$, respectively. With Theorem 3(i) in place, Theorem 3(ii) follows from modifying arguments in Qin and Lawless (1994, Corollary 5) in the proof of Theorem 2.

A.4. Spatial empirical likelihood under parameter constraints

As a continuation of Section 3.3, here we briefly consider constrained maximum EL estimation of spatial parameters. Qin and Lawless (1995) introduced constrained EL inference for independent samples and Kitamura (1997) developed a blockwise version of constrained EL for weakly dependent time series. For spatial data, we may also consider blockwise EL estimation subject to a system of parameter constraints on a spatial parameter $\theta \in \Theta \subset \mathbb{R}^p$: $\psi(\theta) = 0_q \in \mathbb{R}^q$ where $q < p$ and $\Psi(\theta) = \partial\psi(\theta)/\partial\theta$ is of full row rank q. By maximizing the EL function in (5) under the above restrictions on θ, we find a constrained MELE $\hat{\theta}_n^{\psi}$.

Corollary 1. *Suppose Theorem 2 conditions hold and, in a neighborhood of* θ_0, $\psi(\theta)$ *is continuously differentiable,* $\|\partial^2\psi(\theta)/\partial\theta\partial\theta'\|$ *is bounded, and* $\Psi(\theta_0)$ *is rank* q. *If* H_0: $\psi(\theta_0) = 0_q$ *holds, then* $r_n(\hat{\theta}_n^{\psi}) = \ell_n(\hat{\theta}_n^{\psi}) - \ell_n(\hat{\theta}_n) \xrightarrow{d} \chi^2_q$ *and* $\ell_n(\theta_0) - \ell_n(\hat{\theta}_n^{\psi}) \xrightarrow{d} \chi^2_{p-q}$ *as* $n \to \infty$.

We can then sequentially test H_0: $\psi(\theta_0) = 0_q$ with a log-likelihood ratio statistic $\ell_n(\hat{\theta}_n^{\psi}) - \ell_n(\hat{\theta}_n)$ and, if failing to reject H_0, make an approximate $100(1-\alpha)\%$ confidence region for constrained θ values $\{\theta : \psi(\theta) = 0_q,\ \ell_n(\theta) - \ell_n(\hat{\theta}_n^{\psi}) \le \chi^2_{p-q,1-\alpha}\}$.

Proof of Corollary 1. We sketch the proof which requires modifications to the proof of Theorem 2 as well as arguments from Qin and Lawless (1995) (for the i.i.d. data case); we shall employ notation used in the proof of Theorem 2. Write the functions $\psi(\theta)$, $\Psi(\theta)$ as ψ_θ, Ψ_θ in the following. To establish the existence of $\hat{\theta}_n^{\psi}$, let $Q^*_{1n}(\theta, t, \nu) = Q_{1n}(\theta, t)$, $Q^*_{2n}(\theta, t, \nu) = Q_{2n}(\theta, t) + \Psi'_\theta \nu$, and $Q^*_{3n}(\theta, t, \nu) = \psi_\theta$ and define $U_n = \{(\theta, t, \nu) \in \mathbb{R}^p \times \mathbb{R}^r \times \mathbb{R}^q : \theta \in \Theta_n,\ \|t/b_n^d\| + \|\nu\| \le n_{m,n}^{-5/12}\}$.

Step 1. It can first be shown that the system of equations:

$$Q^*_{1n}(\theta, t, \nu) = 0_r, \quad Q^*_{2n}(\theta, t, \nu) = 0_p, \quad Q^*_{3n}(\theta, t, \nu) = 0_q \tag{23}$$

has a solution $(\theta^*_n, t^*_n, \nu^*_n) \in U_n$. Uniformly in $\theta \in \Theta_n$, it holds that $b_n^{-d}\partial t_\theta/\partial\theta = \Sigma_{\theta_0}^{-1} D_{\theta_0} + o_p(1)$ (by differentiating $Q^*_{1n}(\theta, t_\theta) = 0_r$ with respect to θ) and that $(2n_{m,n})^{-1}\partial\ell_n(\theta)/\partial\theta = V_{\theta_0}^{-1}(\theta - \theta_0) + T_\theta$ where T_θ is continuous in θ and $\sup_{\theta\in\Theta_n} \|T_\theta\| = o_p(n_{m,n}^{-5/12})$ (by expanding $(2n_{m,n})^{-1}\partial\ell_n(\theta)/\partial\theta = Q_{2n}(\theta, t_\theta)$ around θ_0). For $\theta \in \Theta_n$, define $\psi_\theta - \Psi_{\theta_0}(\theta - \theta_0) = \|\theta - \theta_0\|^2 k(\theta)$, where $k(\theta)$ is continuous and bounded, and write a function $\eta(\theta)$ as

$$\eta(\theta) = \frac{1}{2n_{m,n}}\frac{\partial\ell_n(\theta)}{\partial\theta} + \Psi'_\theta(\Psi_{\theta_0} V_{\theta_0} \Psi'_\theta)^{-1}\Bigg(\|\theta - \theta_0\|^2 k(\theta) - \Psi_{\theta_0} V_{\theta_0}\left[\frac{1}{2n_{m,n}}\frac{\partial\ell_n(\theta)}{\partial\theta} - V_{\theta_0}^{-1}(\theta - \theta_0)\right]\Bigg). \tag{24}$$

It can be shown that $\eta(\theta) = V_{\theta_0}^{-1}(\theta - \theta_0) + \tilde{T}_\theta$, where $\tilde{T}_\theta$ is continuous in θ and $\sup_{\theta\in\Theta_n} \|\tilde{T}_\theta\| = o_p(n_{m,n}^{-5/12})$, which implies that there exists $\hat{\theta}^*_n \in \Theta_n \setminus \partial\Theta_n$ such that $-\eta(\hat{\theta}^*_n) = 0_p$. This root $\hat{\theta}^*_n$ of $\eta(\theta)$ inside $\Theta_n \setminus \partial\Theta_n$ is deduced from Lemma 2 of Aitchison and Silvey (1958); this result entails that because, for large n, $-\sigma_1^{-1}\eta(\theta)$ maps Θ_n into $\{(\theta - \theta_0) : \theta \in \Theta_n\}$ and $(\theta - \theta_0)'\{-\sigma_1^{-1}\eta(\theta)\} < -\sigma_0/(2\sigma_1)$ holds for $\theta \in \partial\Theta_n$ (i.e., $(\theta - \theta_0)'\{-\sigma_1^{-1}\eta(\theta)\}$ is negative for $\|\theta - \theta_0\| = n_{m,n}^{-5/12}$), where σ_1 and $\sigma_0 > 0$ respectively denote the largest and smallest eigenvalues of $V_{\theta_0}^{-1}$, it must follow that $-\sigma_1^{-1}\eta(\hat{\theta}^*_n) = 0$ for some $\|\hat{\theta}^*_n - \theta_0\| <$

$n_{m,n}^{-5/12}$ by Brouwer's fixed point theorem. From this root, we have that $0_q = \Psi_{\theta_0} V_{\theta_0} \eta(\hat\theta_n^*) = \|\hat\theta_n^* - \theta_0\|^2 k(\hat\theta_n^*) + \Psi_{\hat\theta_n^*}(\hat\theta_n^* - \theta_0) = \psi_{\hat\theta_n^*}$ from (24) as well as

$$\frac{1}{2n_{m,n}} \frac{\partial \ell_n(\hat\theta_n^*)}{\partial \theta} = \Psi'_{\hat\theta_n^*}(\Psi_{\theta_0} V_{\theta_0} \Psi'_{\hat\theta_n^*})^{-1} \Psi_{\theta_0} V_{\theta_0} \frac{1}{2n_{m,n}} \frac{\partial \ell_n(\hat\theta_n^*)}{\partial \theta}. \tag{25}$$

This yields that $\hat\theta_n^*$, the EL Lagrange multiplier $t_{\hat\theta_n^*}$ for $\hat\theta_n^*$ defined by $Q_{1n}(\hat\theta_n^*, t_{\hat\theta_n^*}) = 0_r$, and $\nu_n^* = -(\Psi_{\theta_0} V_{\theta_0} \Psi'_{\hat\theta_n^*})^{-1} \Psi_{\theta_0} V_{\theta_0} (2n_{m,n})^{-1} \partial \ell_n(\hat\theta_n^*)/\partial\theta$ satisfy (23) jointly.

Step 2. We now show that any solution of (23) in U_n, say $(\tilde\theta, \tilde t, \tilde\nu)$, must minimize $\ell_n(\theta)$ on Θ_n subject to the condition $\psi_\theta = 0_q$. To see this, note if $\theta \in \Theta_n$ with $\psi_\theta = 0_q$, then we make a Taylor expansion around $\tilde\theta$:

$$\frac{1}{2n_{m,n}}\Big[\ell_n(\theta) - \ell_n(\tilde\theta)\Big] = \frac{1}{2n_{m,n}} \frac{\partial \ell_n(\tilde\theta)}{\partial \theta'}(\theta - \tilde\theta) + \frac{1}{4n_{m,n}}(\theta - \tilde\theta)' \frac{\partial^2 \ell_n(\theta^*)}{\partial\theta\partial\theta'}(\theta - \tilde\theta),$$
$$\theta^* \text{ between } \theta,\, \tilde\theta.$$

Since $\tilde\theta$ satisfies (23), it follows from some algebra that $\tilde\theta$ also satisfies (25) after substituting $\tilde\theta$ for $\hat\theta_n^*$. Using $0_q = \psi_\theta - \psi_{\tilde\theta} = \Psi_{\tilde\theta}(\theta - \tilde\theta) + o(\|\theta - \tilde\theta\|^2)$, we find $(2n_{m,n})^{-1}\partial\ell_n(\tilde\theta)/\partial\theta'(\theta - \tilde\theta) = o_p(\|\theta - \tilde\theta\|^2)$ for $\tilde\theta$ fulfilling (25); it may also be shown that $(2n_{m,n})^{-1}\partial^2\ell_n(\theta^*)/\partial\theta\partial\theta' = V_{\theta_0}^{-1} + o_p(1)$ (by expanding $(2n_{m,n})^{-1}$ $\partial\ell_n(\theta)/\partial\theta = Q_{2n}(\theta, t_\theta)$ around θ_0). Hence, $\ell_n(\theta) - \ell_n(\tilde\theta) \geq \{\sigma_0/2 + o_p(1)\} n_{m,n} \|\theta - \tilde\theta\|^2$, where the $o_p(1)$ term is uniform for $\theta \in \Theta_n$, $\psi_\theta = 0$.

Step 3. By the first two steps, we have therefore established that there exists a consistent MELE $\hat\theta_n^\psi$ of θ_0, given by $\hat\theta_n^\psi = \hat\theta_n^* \in \Theta_n \setminus \partial\Theta_n$, that satisfies the condition $\psi(\hat\theta_n^\psi) = 0$; we may denote $t_{\hat\theta_n^\psi} = t_{\hat\theta_n^*}$ and $\nu_n^\psi = \nu_n^*$. We now show

$$n_{m,n}^{\frac{1}{2}} \begin{pmatrix} \hat\theta_n^\psi - \theta_0 \\ \nu_n^\psi \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0_{r+p+q}, \begin{bmatrix} P_{\theta_0} & 0 \\ 0 & R_{\theta_0} \end{bmatrix}\right), \quad \begin{aligned} P_{\theta_0} &= V_{\theta_0}\Big(I_{p\times p} - \Psi'_{\theta_0} R_{\theta_0} \Psi_{\theta_0} V_{\theta_0}\Big), \\ R_{\theta_0} &= \Big(\Psi_{\theta_0} V_{\theta_0} \Psi'_{\theta_0}\Big)^{-1}. \end{aligned} \tag{26}$$

Expanding $Q_{in}^*(\theta, t, \nu)$ at $(\theta_0, 0, 0)$ and using that $(\hat\theta_n^\psi, t_{\hat\theta_n^\psi}, \nu_n^\psi)$ satisfies (23), we have:

$$\begin{pmatrix} -Q_{1n}(\theta_0, 0_r) + o_p(\delta_n^*) \\ o_p(\delta_n^*) \\ o_p(\delta_n^*) \end{pmatrix} = \Sigma_n^* \begin{pmatrix} \frac{t_{\hat\theta_n^\psi}}{b_n^d} \\ \hat\theta_n^\psi - \theta_0 \\ \nu_n^\psi \end{pmatrix}, \quad \Sigma_n^* = \begin{bmatrix} \frac{\partial Q_{1n}(\theta_0, 0_r)}{\partial t} & \frac{\partial Q_{1n}(\theta_0, 0_r)}{\partial \theta} & 0 \\ \frac{\partial Q_{2n}(\theta_0, 0_r)}{\partial t} & 0 & \Psi'_{\theta_0} \\ 0 & \Psi_{\theta_0} & 0 \end{bmatrix},$$

where $Q_{1n}(\theta_0, 0_r) = \bar M_{\theta_0}$, $b_n^d \partial Q_{1n}(\theta_0, 0_r)/\partial t = -\widehat\Sigma_{\theta_0}$, $\partial Q_{1n}(\theta_0, 0_r)/\partial\theta = \bar D_{\theta_0} = [b_n^d \partial Q_{2n}(\theta_0, 0_r)/\partial t]'$ and $\delta_n^* = \|\hat\theta_n^\psi - \theta_0\| + \|t_{\hat\theta_n^\psi}/b_n^d\| + \|\nu_n^\psi\|$. Using Lemma 2(iii) and $\bar{D}_{\theta_0} \xrightarrow{p} D_{\theta_0}$ from the proof of Theorem 2, we have

$$\Sigma_n^* \xrightarrow{p} \begin{bmatrix} -\Sigma_{\theta_0} & D_{\theta_0} & 0 \\ D'_{\theta_0} & 0 & \Psi'_{\theta_0} \\ 0 & \Psi_{\theta_0} & 0 \end{bmatrix} \equiv \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \equiv \tilde{C}, \qquad \begin{matrix} C_{12} = \begin{bmatrix} D_{\theta_0} & 0 \end{bmatrix}, \ C_{21} = C'_{12} \\ C_{11} = -\Sigma_{\theta_0}, \ C_{22} = \begin{bmatrix} 0 & \Psi'_{\theta_0} \\ \Psi_{\theta_0} & 0 \end{bmatrix}. \end{matrix}$$

Note that $\det(\tilde{C}) = \det(C_{11})\det(Q_c) = \det(-\Sigma_{\theta_0})\det(V_{\theta_0}^{-1})\det(-R_{\theta_0}^{-1}) \neq 0$, for $Q_c = C_{22} - C_{21}C_{11}^{-1}C_{12}$, and

$$\tilde{C}^{-1} = \begin{bmatrix} -\Sigma_{\theta_0}^{-1} + \Sigma_{\theta_0}^{-1}C_{12}Q_c^{-1}C_{21}\Sigma_{\theta_0}^{-1} & \Sigma_{\theta_0}^{-1}C_{12}Q_c^{-1} \\ Q_c^{-1}C_{21}\Sigma_{\theta_0}^{-1} & Q_c^{-1} \end{bmatrix}, \ Q_c^{-1} = \begin{bmatrix} P_{\theta_0} & V_{\theta_0}\Psi'_{\theta_0}R_{\theta_0} \\ R_{\theta_0}\Psi_{\theta_0}V_{\theta_0} & -R_{\theta_0} \end{bmatrix}.$$

Since, by Lemma 2(ii), $n_{m,n}^{1/2}Q_{1n}(\theta_0, 0_r) = n_{m,n}^{1/2}\bar{M}_{\theta_0} \xrightarrow{d} \mathcal{N}(0_r, \Sigma_{\theta_0})$, it follows that $\delta_n^* = O_p(n_{m,n}^{-1/2})$. Then,

$$\begin{aligned} n_{m,n}^{\frac{1}{2}} \begin{pmatrix} \hat{\theta}_n^{\psi} - \theta_0 \\ \nu_n^{\psi} \end{pmatrix} &= -n_{m,n}^{\frac{1}{2}} Q_c^{-1}C_{21}\Sigma_{\theta_0}^{-1}Q_{1n}(\theta_0, 0_r) + o_p(1) \\ &\xrightarrow{d} \mathcal{N}\left(0_{p+q}, \begin{bmatrix} P_{\theta_0} & 0 \\ 0 & R_{\theta_0} \end{bmatrix}\right). \end{aligned}$$

Step 4. As in the proof of Theorem 2, we can then expand by (19)

$$\begin{aligned} \ell_n(\hat{\theta}_n^{\psi}) &= n_{m,n}\Big(\bar{M}_{\theta_0} + D_{\theta_0}(\hat{\theta}_n^{\psi} - \theta_0)\Big)'\Sigma_{\theta_0}^{-1}\Big(\bar{M}_{\theta_0} + D_{\theta_0}(\hat{\theta}_n^{\psi} - \theta_0)\Big) + o_p(1) \\ &= n_{m,n}Q'_{1n}(\theta_0, 0_r)\Big(I_{r\times r} - D_{\theta_0}P_{\theta_0}D'_{\theta_0}\Sigma_{\theta_0}^{-1}\Big)'\Sigma_{\theta_0}^{-1}\Big(I_{r\times r} - D_{\theta_0}P_{\theta_0}D'_{\theta_0}\Sigma_{\theta_0}^{-1}\Big)\bar{M}_{\theta_0} \\ &\qquad + o_p(1) \\ &= \Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big]'\Big[I_{r\times r} - (P_{\Sigma_{\theta_0}^{-\frac{1}{2}}D_{\theta_0}} - P_{H_{\theta_0}})\Big]\Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big] + o_p(1), \end{aligned}$$

where $H_{\theta_0} = \Sigma_{\theta_0}^{-\frac{1}{2}}D_{\theta_0}(D'_{\theta_0}\Sigma_{\theta_0}^{-1}D_{\theta_0})^{-1}\Psi'_{\theta_0}$. Then,

$$\begin{aligned} \ell_n(\hat{\theta}_n^{\psi}) - \ell_n(\hat{\theta}_n) &= \Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big]'P_{H_{\theta_0}}\Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big] + o_p(1), \\ \ell_n(\theta_0) - \ell_n(\hat{\theta}_n^{\psi}) &= \Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big]'(P_{\Sigma_{\theta_0}^{-\frac{1}{2}}D_{\theta_0}} - P_{H_{\theta_0}})\Big[n_{m,n}^{\frac{1}{2}}\Sigma_{\theta_0}^{-\frac{1}{2}}\bar{M}_{\theta_0}\Big] + o_p(1). \end{aligned}$$

Note now that $n_{m,n}^{1/2}\Sigma_{\theta_0}^{-1/2}Q_{1n}(\theta_0, 0_r)\bar{M}_{\theta_0} \xrightarrow{d} \mathcal{N}(0, I_{r\times r})$ by Lemma 2(ii), $P_{H_{\theta_0}}$ and $P_{\Sigma_{\theta_0}^{-1/2}D_{\theta_0}} - P_{H_{\theta_0}}$ are idempotent matrices with

$$\mathrm{rank}\Big(P_{H_{\theta_0}}\Big) = \mathrm{rank}\big(H_{\theta_0}\big) = \mathrm{rank}(\Psi_{\theta_0}) = q;$$

$$\text{rank}\Big(P_{\Sigma_{\theta_0}^{-1/2}D_{\theta_0}} - P_{H_{\theta_0}}\Big) = p - \text{trace}\big[P_{H_{\theta_0}}\big] \;=\; p - \text{rank}\big[P_{H_{\theta_0}}\big] \;=\; p - q.$$

For $\text{rank}(P_{H_{\theta_0}}) = q$ above, we used $\text{rank}(H_{\theta_0}) \le \text{rank}(\Psi_{\theta_0})$, $\text{rank}(\Psi_{\theta_0}) = \text{rank}(D'_{\theta_0} \Sigma_{\theta_0}^{-1/2} H_{\theta_0}) \le \text{rank}(H_{\theta_0})$. Corollary 1 now follows.

A.5. Spatial block bootstrap algorithm

Here we outline a spatial block bootstrap method for generating bootstrap version $\mathcal{Y}_n^*$ of the original vectorized spatial data $\mathcal{Y}_n = \{Y_{\mathbf{s}} : \mathbf{s} \in \mathcal{R}_{m,n} \cap \mathbb{Z}^d\}$ on $\mathcal{R}_{m,n} \subset \mathbb{R}^d$. Bootstrap replicates $\mathcal{Y}_n^*$ of spatial data, on a bootstrap sampling region $\mathcal{R}_{m,n}^*$, are used to formulate the empirical Bartlett correction for the spatial EL method as described in Section 4.

Let $\mathcal{Y}_n A = \{Y_{\mathbf{s}} : \mathbf{s} \in A \cap \mathbb{Z}^d\}$ denote the observed spatial data at $\mathbb{Z}^d$ points lying inside a set $A \subset \mathcal{R}_{m,n}$. The block bootstrap requires a block scaling factor, denoted by $b_{n,bt}$, satisfying $b_{n,bt}^{-1} + b_{n,bt}^d / n_{m,n} = o(1)$. Suppose this bootstrap block scaling is used to make the blocks of size $b_{n,bt}(-1/2, 1/2]^d$ in $\mathcal{R}_{m,n}$ appearing in Figure 2(b)$-$(c). As a first step, we divide the sampling region $\mathcal{R}_{m,n}$ into NOL blocks of size $b_{n,bt}(-1/2, 1/2]^d$ that fall entirely inside $\mathcal{R}_{m,n}$, as depicted in Figure 2(b). In the notation of Section 2.2, $\{\mathcal{B}_{b_{n,bt}}(\mathbf{i}) : \mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}\}$ represents a collection of $b_{n,bt}$-scaled NOL "complete blocks" partitioning $\mathcal{R}_{m,n}$. These complete NOL blocks inside $\mathcal{R}_{m,n}$, when taken together, form a bootstrap sampling region $\mathcal{R}_{m,n}^*$ as $\mathcal{R}_{m,n}^* \equiv \{\mathcal{B}_{b_{n,bt}}(\mathbf{i}) : \mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}\}$, as shown in Figure 2(d) based on complete NOL blocks in Figure 2(b). In place of the original data $\mathcal{Y}_n$ observed on $\mathcal{R}_{m,n}$, we aim to create a bootstrap sample $\mathcal{Y}_n^*$ on $\mathcal{R}_{m,n}^*$. Each block $\mathcal{B}_{b_{n,bt}}(\mathbf{i}) = \mathbf{i} + b_{n,bt}(-1/2, 1/2]^d$, $\mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}$, that constitutes a part of $\mathcal{R}_{m,n}^*$ also corresponds to a piece of $\mathcal{R}_{m,n}$, where we originally observed the data $\mathcal{Y}_n \mathcal{B}_{b_{n,bt}}(\mathbf{i})$, $\mathcal{B}_{b_{n,bt}}(\mathbf{i}) \subset \mathcal{R}_{m,n}$. For a fixed $\mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}$, we then create a bootstrap rendition $\mathcal{Y}_n^* \mathcal{B}_{b_{n,bt}}(\mathbf{i})$ of $\mathcal{Y}_n \mathcal{B}_{b_{n,bt}}(\mathbf{i})$ by independently resampling some size $b_{n,bt}(-1/2, 1/2]^d$ block of $Y_{\mathbf{s}}$-observations from the region $\mathcal{R}_{m,n}$ (as in Figure 2(c)) and pasting this observational block into the position of $\mathcal{B}_{b_{n,bt}}(\mathbf{i})$ within $\mathcal{R}_{m,n}^*$. To make the resampling scheme precise, for each $\mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}$, we define the bootstrap version as $\mathcal{Y}_n^* \mathcal{B}_{b_{n,bt}}(\mathbf{i}) \equiv \mathcal{Y}_n \mathcal{B}_{b_{n,bt}}(\mathbf{i}^*)$ where $\mathbf{i}^* \in \mathbb{Z}^d$ is random vector selected uniformly from the collection of OL block indices given by $\mathcal{I}_{b_{n,bt}}^{OL}$ in the notation of Section 2.2; that is, we resample from all OL $b_{n,bt}$-scaled blocks within $\mathcal{R}_{m,n}$ (as depicted in Figure 2(c)) to produce a spatial block of observations $\mathcal{Y}_n^* \mathcal{B}_{b_{n,bt}}(\mathbf{i})$. We then concatenate the resampled block observations for each $\mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}$ into a single spatial bootstrap sample $\mathcal{Y}_n^* = \{\mathcal{Y}_n^* \mathcal{B}_{b_{n,bt}}(\mathbf{i}) : \mathbf{i} \in \mathcal{I}_{b_{n,bt}}^{NOL}\}$ on $\mathcal{R}_{m,n}^*$ with $n_{m,n}^* = |\mathcal{I}_{b_{n,bt}}^{NOL}| \cdot b_{n,bt}^d$ sampling sites at $\mathcal{R}_{m,n}^* \cap \mathbb{Z}^d$. In Section 4, the bootstrap EL version ℓ_n^* may be computed as in (6) after replacing $\mathcal{Y}_n$, $\mathcal{R}_{m,n}$, $n_{m,n}$ with $\mathcal{Y}_n^*$, $\mathcal{R}_{m,n}^*$, $n_{m,n}^*$. See Chapter 12.3 of Lahiri (2003a) for more details on the spatial block bootstrap.