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 note highlights some unusual and unexpected behavior in point estimation using empirical likelihood (EL). In particular, frequency domain formulations of EL, based on the periodogram and estimating functions, have been proposed in the literature for time and spatial processes. However, in contrast to the time series case and most applications of EL, the maximum EL parameter estimator exhibits surprisingly non-standard asymptotic properties for irregularly located spatial data. In fact, a consistent normal limit cannot be guaranteed, as is typical for EL. Despite this, log-ratio EL statistics maintain standard chi-square limits with such spatial data.

Keywords: Discrete Fourier Transform, Frequency Domain Empirical Likelihood, Periodogram, Stochastic sampling.

2010 MSC: 62M30, 62G20

1. Introduction

Empirical likelihood (EL), introduced by Owen (1990), formulates a likelihood by probability profiling data in a manner that does not require a joint

*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)

distribution. The resulting EL function often shares properties with fully parametric likelihood, such as log-ratio statistics with chi-square limits (i.e., Wilks' phenomenon). Extending EL to dependent data requires caution as EL approaches for independent data (Owen, 1990; Qin and Lawless, 1994) typically fail with correlated processes. For example, Kitamura (1997) proposed a block-based EL as a device for extending EL to time series, whereby time blocks of observations are used to capture dependence. While this EL version is valid for several inference problems with time series and spatial lattice data, its performance depends on the choice of block size which is a compounding issue with spatial sampling. See Nordman and Lahiri (2014) for a review of EL for times series.

The EL approach considered here, particularly for spatial data with irregular locations, is based on a data transformation, or the discrete Fourier transform. In contrast to data-blocking, translation into the frequency domain aims to whiten or weaken dependence prior to implementing EL. Additionally, frequency domain analysis is particularly useful for examining covariance structures (cf. Bochner (1959); Gikhman and Skorokhod (1974); Yaglom (1987); Stein (1999)). For time series, Monti (1997) and Nordman and Lahiri (2006) introduced a frequency domain EL method, premised on the asymptotic independence of periodogram ordinates. Recently, Bandyopadhyay et al. (2015) (henceforth referred to as [BLN]) proposed a frequency domain EL for spatial processes having stochastic sampling designs. Similarly to the time series case, this spatial EL is based on estimating functions along with a notion of asymptotic independence of spatial periodogram values, and the method produces log-ratio statistics with chi-square limits. However, no component of EL point estimation was examined, which is our focus here.

We provide distributional results for the maximizer of the spatial EL function. Perhaps surprisingly, we find this point estimator exhibits non-standard behavior for spatial data in contrast to (equi-spaced) time series. In fact, due to complications arising in the frequency domain analysis of irregularly located spatial observations, the EL point estimator does not even necessarily have a limit distribution, where normal limits typically hold for time series and for most applications of EL (cf. Owen (2001) and references therein). Interestingly, while normal limits for the EL maximizer may fail, EL log-ratio statistics generally retain standard chi-square limits, useful for testing and calibrating confidence regions. Our goal is to outline these non-standard EL results for spatial processes.

2. Framework for Spatial Frequency Domain EL

Section 2.1 describes the spatial sampling and frequency domain inference. Section 2.2 then provides the associated spatial EL method. In our exposition, we also highlight comparisons to the more established EL counterpart for time series to help frame later comparisons.

2.1. Spatial Data and Spectral Inference Problem

Consider a real-valued, second-order stationary spatial process $\{Z(\boldsymbol{s} : \boldsymbol{s} \in \mathbb{R}^d\}$, where d denotes the dimension of spatial sampling. To specify the spatial sampling design, we follow Hall and Patil (1994) and [BLN]. We suppose the process $Z(\cdot)$ is observed at n irregularly located sites $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ within a sampling region $\mathcal{D}_n = \lambda_n \mathcal{D}_0 \subset \mathbb{R}^d$. Here $\mathcal{D}_0$ is an open, connected subset of $(1/2, 1/2]^d$ containing the origin, representing a "template" for the sampling region, and $\{\lambda_n\}$ is a real positive sequence such that $\lambda_n \to \infty$ as $n \to \infty$. This standard formulation permits a variety of region shapes for consideration; see [BLN] for details. To specify the sampling locations, independently of $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, let $\{\boldsymbol{X}_k\}_{k \geq 1} \subset \mathcal{D}_0$ be a sequence of independently and identically distributed $\mathbb{R}^d$-valued random vectors, with probability density function $f(\boldsymbol{x})$ on the closure of $\mathcal{D}_0$. Sites $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ are then generated as $\boldsymbol{s}_i = \lambda_n \boldsymbol{X}_i$, $i = 1, \ldots, n$ which allows for arbitrary spatial patterns in the stochastic sampling design. In contrast to the directional asymptotics of equi-spaced time series, spatial sampling is complicated by varying types of spatial asymptotic structure (Lahiri, 2003), between which the limit laws of statistics may change; see also Stein

(1989); Cressie (1993); Lahiri and Mukherjee (2004) and the references therein. Namely, the number n of sampling sites may grow at a potentially different rate than the volume λ_n^d of the sampling region $\mathcal{D}_n$, leading to different asymptotic spatial regimes (Cressie, 1993; Lahiri, 2003). Let $c = \lim_{n\to\infty} \lambda_n^d/n \in [0,\infty)$. The case $c > 0$, where the number of spatial observations is proportional to the volume of the region, corresponds to pure increasing domain (PID) asymptotics; this has parallels to standard time series sampling. On the other hand, the case $c = 0$ corresponds to spatial sampling with a heavy infill component, allowing the number of spatial observations to increase at a faster rate than the volume of the region; we refer to this as mixed increasing domain (MID) asymptotics.

To frame the inference in the spatial frequency domain, let us suppose $\sigma(\boldsymbol{h}) = \mathrm{Cov}(Z(\mathbf{0}), Z(\boldsymbol{h}))$, $\boldsymbol{h} \in \mathbb{R}^d$ and $\phi(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \mathbb{R}^d$ denote the autocovariance function and spectral density of the process $Z(\cdot)$, respectively. Consider a spatial parameter $\theta \in \Theta \subset \mathbb{R}^p$ about which information may be expressed through a system of estimating equations involving $\phi(\cdot)$. Specifically, let $G : \mathbb{R}^d \times \Theta \to \mathbb{R}^r$ be a vector of $r \geq p$ estimating functions (of both frequencies in $\mathbb{R}^d$ and parameters in $\mathbb{R}^p$) such that $G_\theta(\boldsymbol{\omega}) \equiv G(\boldsymbol{\omega}; \theta)$, $\boldsymbol{\omega} \in \mathbb{R}^d$, satisfies a spectral moment condition

$$\int_{\mathbb{R}^d} G_\theta(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \mathbf{0}_r \tag{1}$$

at a true parameter $\theta_0 \in \Theta$, where $\mathbf{0}_r \in \mathbb{R}^r$ denotes the zero vector. Parameters prescribed by the condition (2) include autocorrelations and normalized spectral distributions; for example, $r = p$ functions $G_\theta(\boldsymbol{\omega}) = \big(\cos(\boldsymbol{h}_1'\boldsymbol{\omega}), \ldots, \cos(\boldsymbol{h}_p'\boldsymbol{\omega})\big)' - \theta \in \mathbb{R}^p$ satisfies (2) for correlations $\theta = (\sigma(\boldsymbol{h}_1), \ldots, \sigma(\boldsymbol{h}_p))'/\sigma(\mathbf{0})$ at given lags $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_p \in \mathbb{R}^d$; see [BLN] for further examples.

Considering a regular time process $\{Z_t : t \in \mathbb{Z}\}$ (i.e., sampled on the integer grid), the frequency domain formulations of EL use similar estimating equations (cf. Monti, 1997; Nordman and Lahiri, 2006). In this case, the analogous moment condition to (1) becomes

$$\int_{-\pi}^{\pi} G_{\theta_0}(\omega)\phi(\omega) d\omega = \mathbf{0}_r \in \mathbb{R}^r \tag{2}$$

where $\phi(\omega)$, $\omega \in [-\pi, \pi]$, denotes the spectral density of the (second-order) stationary process $\{Z_t\}$ and $G_\theta(\omega) \equiv G(\omega, \theta)$ represents $r \geq p$ estimating functions. For time processes, Monti (1997) and Yau (2012) have considered types of $G_\theta(\cdot)$ for EL estimation of spectral density models, and functions may also be prescribed for inference about ratios of spectral means, such as autocorrelations (Nordman and Lahiri, 2006). However, for irregularly located spatial data, the difference, and indeed challenge, is that the frequency domain corresponds to $\mathbb{R}^d$ (e.g., in the condition (1)) rather than a compact set for regular time series (e.g., $[-\pi, \pi]$ in (2)). We later show that this feature, combined with the differing spatial asymptotic designs, contributes to non-standard distributional results with point estimation in spatial EL (Section 3).

2.2. Frequency Domain Version of EL

The spatial EL approach uses the periodogram for frequency domain inference. Define a raw spatial periodogram of the data $\{Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)\}$ at a frequency $\boldsymbol{\omega} \in \mathbb{R}^d$ as

$$I_n(\boldsymbol{\omega}) = \left| \lambda_n^{d/2} n^{-1} \sum_{j=1}^{n} Z(\boldsymbol{s}_j) \exp\left(\imath \boldsymbol{\omega}' \boldsymbol{s}_j\right) \right|^2, \quad \imath \equiv \sqrt{-1}.$$

This data transformation serves to pre-whiten spatial dependence and, similarly to time series, the resulting spatial periodogram has independence properties in large samples (Bandyopadhyay and Lahiri, 2009). Namely, along two sequences of frequencies $\{\boldsymbol{\omega}_{1n}\}_{n\geq 1}, \{\boldsymbol{\omega}_{2n}\}_{n\geq 1} \subset \mathbb{R}^d$, ordinates $I_n(\boldsymbol{\omega}_{1n}), I_n(\boldsymbol{\omega}_{2n})$ are asymptotically independent if and only if $\{\boldsymbol{\omega}_{1n}\}$ and $\{\boldsymbol{\omega}_{2n}\}$ are asymptotically distant (i.e., $\|\lambda_n(\boldsymbol{\omega}_{1n} - \boldsymbol{\omega}_{2n})\| \to \infty$ as $n \to \infty$). However, unlike regular time series (Brillinger, 1981), The spatial periodogram can have a nontrivial bias, depending on spatial asymptotics (Bandyopadhyay and Lahiri, 2009; Matsuda and Yajima, 2009). Namely,

$$\lim_{n\to\infty} \mathrm{E} I_n(\boldsymbol{\omega}) = c\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega}) \quad \text{for } \boldsymbol{\omega} \in \mathbb{R}^d, \tag{3}$$

holds, where $\phi(\cdot)$ and $\sigma(\cdot)$ denote the spectral density and covariance function of the process $Z(\cdot)$, $K = (2\pi)^d \int_{\mathbb{R}^d} f^2$ involves the spatial sampling density f, and

$\lim_{n\to\infty} \lambda_n^d/n = c$. Under PID sampling ($c \in (0,\infty)$), there exists a non-trivial bias component, which disappears in the MID case ($c = 0$). To address this, we use a bias-corrected periodogram as

$$\widetilde{I}_n(\boldsymbol{\omega}) = I_n(\boldsymbol{\omega}) - n^{-1}\lambda_n^d \widehat{\sigma}_n(\mathbf{0}), \ \boldsymbol{\omega} \in \mathbb{R}^d$$

where $\widehat{\sigma}_n(\mathbf{0}) = n^{-1}\sum_{j=1}^n (Z(\boldsymbol{s}_j) - \bar{Z}_n)^2$ is the sample variance with $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z(\boldsymbol{s}_i)$.

Frequency domain versions of EL for time series use the periodogram and estimating functions to approximate a spectral moment condition (e.g., (2)) along a discrete set of frequencies (i.e., the usual Fourier frequencies $2\pi j/n$, integer $|j| < n/2$) and exploit features of asymptotic independence in periodogram ordinates. The spatial EL is similarly formulated using the (bias-corrected) spatial periodogram to mimic the spectral mean (1) along a discretized set of frequencies, chosen to ensure the periodogram evaluations are approximately uncorrelated; see [BLN]. For $\kappa \in (0,1)$, $\eta \in (\kappa,\infty)$, and $\mathcal{C} \in (0,\infty)$, define a set of Fourier frequencies as

$$\mathcal{N}_n = \left\{ \boldsymbol{j}\lambda_n^{-\kappa} : \boldsymbol{j} \in \mathbb{Z}^d, \boldsymbol{j} \in [-\mathcal{C}\lambda_n^{\eta}, \mathcal{C}\lambda_n^{\eta}]^d \right\}.$$

Let $N = |\mathcal{N}_n|$ be the cardinality of $\mathcal{N}_n$, and let $\boldsymbol{\omega}_{kn}, k = 1,\ldots,N$ (with arbitrary ordering) denote the elements of $\mathcal{N}_n$. The frequencies $\{\boldsymbol{\omega}_{kn}\}_{k=1}^N$ form a regular lattice over the set $[-\mathcal{C}\lambda_n^{\eta-\kappa}, \mathcal{C}\lambda_n^{\eta-\kappa}] \uparrow \mathbb{R}^d$ as $n \to \infty$, which is important for covering the entire frequency domain $\mathbb{R}^d$ asymptotically. Also, any pair of frequencies $\boldsymbol{\omega}_{kn}, \boldsymbol{\omega}_{jn} \in \mathcal{N}$ in the set is asymptotically distant (i.e., $\lambda_n\|\boldsymbol{\omega}_{kn} - \boldsymbol{\omega}_{jn}\| \geq \lambda_n^{1-\kappa} \to \infty$), which ensures that associated periodogram values are approximately independent; see also [BLN].

To numerically assess the plausibility of a parameter θ, the (normalized) EL function for θ is

$$\mathcal{R}_n(\theta) = \sup\left\{ \prod_{k=1}^N Np_k : \sum_{k=1}^N p_k = 1, p_k \geq 0, \sum_{k=1}^N p_k G_\theta(\boldsymbol{\omega}_{kn})\widetilde{I}_n(\boldsymbol{\omega}_{kn}) = \mathbf{0}_r \right\} \in [0,1] \tag{4}$$

based on estimating functions $G_\theta(\cdot)$. The EL function is a likelihood found by probability profiling the (approximately independent) periodogram variants under a linear constraint that imitates the moment condition (1). Similarly to parametric likelihood, (4) quantifies the strength of evidence in support of θ. Maximizing $\mathcal{R}_n(\theta)$ over the parameter space $\Theta \subset \mathbb{R}^p$ produces an EL point estimator $\widehat{\theta}_n \in \Theta$ of interest here. note that EL functions and point estimators for time series have the same formulation, but based on the standard Fourier frequencies in (4) (cf. Monti, 1997; Nordman and Lahiri, 2006). However, the large-sample properties of the maximizer $\widehat{\theta}_n$ in the spatial case differ dramatically from the time series version of EL, as considered next.

3. Main Results

To study the spatial EL point estimator $\widehat{\theta}_n$ in frequency domain inference, Section 3.1 first outlines a critical spatial regularity condition which impacts distributional results for the estimator, described in Section 3.2. The latter section concludes with summary remarks.

3.1. A Spatial Frequency-Average Condition

We require some mild assumptions on the dependence of the second-order stationary process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, expressed in terms of mixing/moment conditions; see also [BLN]. These regularity conditions (denoted as Conditions (C.1)-(C.8)) are described in the Supplementary Materials. For our purposes, we need only emphasize one condition (Condition (C.3)) regarding the r spectral estimating functions $G_\theta(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \mathbb{R}^d$ in the spatial EL method. Recall that the EL function (4) uses a frequency grid $\{\boldsymbol{\omega}_{jn}\}_{j=1}^N$ for evaluating the spatial periodogram as well as estimating functions $\{G_\theta(\boldsymbol{\omega}_{jn})\}_{j=1}^N$. Let $\theta_0 \in \mathbb{R}^p$ denote the true parameter solving (1).

We suppose that, given any subsequence $\{n_j\} \subset \{n\}$ or path of spatial sample sizes, one may extract a further subsequence $\{k \equiv n_k\} \subset \{n_j\}$ such that the frequency-based average

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{j=1}^{N_k} G_{\theta_0}(\boldsymbol{\omega}_{jk}) G'_{\theta_0}(\boldsymbol{\omega}_{jk}) = \Gamma^* \equiv \Gamma^*(\{n_k\}) \tag{5}$$

has a limit for some non-singular $r \times r$ matrix Γ^* that *may change* with the subsequence $k = n_k$. Here $\{G_\theta(\boldsymbol{\omega}_{jk})\}_{j=1}^{N_k}$ represents an evaluation of the frequency grid along the subsequence $\{k \equiv n_k\}$. Note that, in the regular time series setting, the analog of (5) would be

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{|j|<n_k/2} G_{\theta_0}(2\pi j/n_k) G'_{\theta_0}(2\pi j/n_k) = \Gamma^*$$

using the usual discrete Fourier frequencies (with $N_k \approx n_k$). In the time case and under mild conditions on $G_{\theta_0}(\cdot)$ (e.g., Riemann integrable on $[-\pi, \pi]$), such limits hold trivially with a unified limit point $\Gamma^* \equiv (2\pi)^{-1} \int_{-\pi}^{\pi} G_{\theta_0}(\omega) G'_{\theta_0}(\omega) d\omega$ that *does not* depend on the subsequence $\{n_k\}$. However, as a complication with irregularly located spatial data, the spatial integral counterparts

$$\int_{\mathbb{R}^d} G_{\theta_0}(\boldsymbol{\omega}) G'_{\theta_0}(\boldsymbol{\omega}) d\boldsymbol{\omega}$$

are often divergent due to the unbounded frequency domain $\mathbb{R}^d$ and, while the condition (5) is not generally stringent, there is no guarantee that the limit $\Gamma^*(\{n_k\})$ is consistent throughout all subsequences. For example, a real-valued estimating function $G_{\theta_0}(\boldsymbol{\omega}) = \mathbb{I}(\boldsymbol{\omega} \leq \boldsymbol{t}) - \theta_0$, for the normalized spectral distribution $\theta_0 = \int_{\boldsymbol{\omega}\leq \boldsymbol{t}} \phi(\boldsymbol{\omega}) d\boldsymbol{\omega} / \int \phi(\boldsymbol{\omega}) d\boldsymbol{\omega}$ at a given $\boldsymbol{t} \in \mathbb{R}^d$, has a limit $\Gamma^* = -2^{-d} + \theta_0^2$ regardless of the subsequence $\{n_k\}$. on the other hand, for illustration, a real estimating function $G_{\theta_0}(\cdot) : \mathbb{R}^d \to \mathbb{R}$ may be chosen so that

$$\frac{1}{N} \sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jn})^2 \propto \exp[\sin(\pi N/4)] + 2 \quad \text{as } n \to \infty,$$

in which case the limit $\Gamma^* \equiv \Gamma^*(\{n_k\})$ in (5) has five possible values: $\exp(x) + 2$ for $x \in \{0, \pm 1, \pm 1/\sqrt{2}\}$. This potential variation in the spatial limit (5) , along with the different asymptotic regimes possible with irregularly located spatial data, translates into unusual asymptotic behavior for the spatial EL estimator $\widehat{\theta}_n$ that does not occur in the EL analog for time series.

3.2. Non-standard Distribution of Spatial EL Point Estimator

To present results for $\widehat{\theta}_n$, we note first that [BLN] have shown that limit distributions for certain spatial EL statistics (i.e., $-2a \log \mathcal{R}_n(\theta_0)$ based on (4) without point estimation) change slightly between PID/MID spatial sampling regimes, corresponding to $\lim_{n\to\infty} c_n \in (0,\infty)$ or $\lim_{n\to\infty} c_n = 0$ for $c_n = \lambda_n^d/n$ (namely, the proper scaling a may be 1 or 1/2). A further dichotomy exists in the MID case according how much faster n grows relative to the volume λ_n^d of the spatial sampling region, which is quantified by the volume $N\lambda_n^{-\kappa d} \to \infty$ of the frequency grid in the EL method (cf. Section 2.2). Recall N and $\lambda_n^{-\kappa d}$ denote the number of, and volume between, such frequency spacings. By theory in [BLN], "slow infill rate" MID involves a sample size n bounded by a slow increase to the size of the sampling region (i.e., $c_n^{-2} = n^2/\lambda_n^{2d} \ll N\lambda_n^{-\kappa d}$) while "fast infill rate" MID involves n remaining larger than such an increase (i.e., $N\lambda_n^{-\kappa d} \ll c_n^{-2} = n^2/\lambda_n^{2d}$). Essentially, these definitions or types of spatial asymptotics relate to how fast a term $c_n = \lambda_n^d/n$, controlling bias in the spatial periodogram (3), decays to zero, with the fastest decay rate occurring under the "fast infill" MID case. We show that the spatial sampling design also impacts the EL point estimator here, with further, more serious, complications due to the frequency average (5). Let $P(\cdot) \equiv P(\cdot|\boldsymbol{X})$ denote probability conditional on the collection of random vectors $\boldsymbol{X} \equiv \{\boldsymbol{X}_i\}_{i\geq 1} \subset \mathbb{R}^d$ (defining spatial locations in Section 2.1), and let $P_{\boldsymbol{X}}$ denote the joint distribution of $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$. Write $s_n \sim t_n$ for two positive sequences where $s_n/t_n \to 1$ as $n \to \infty$.

Theorem 1. *Suppose Conditions (C.1)-(C.8) hold,* $D_{\theta_0} \equiv \int_{\mathbb{R}} [\partial G_{\theta_0}(\boldsymbol{\omega})/\partial\theta]\phi(\boldsymbol{\omega})d\boldsymbol{\omega}$ *has full column rank* p*, and let* $c_n = \lambda_n^d/n$, $b_n^2 = Nc_n^2 + \lambda_n^{\kappa d}$*. Given any subsequence* $\{n_j\} \subset \{n\}$*, extract a further subsequence* $\{k \equiv n_k\} \subset \{n_j\}$ *such that (5) holds with a limit* Γ^**. Then, for the spatial EL estimator* $\widehat{\theta}_k$ *along the subsequence, it follows that* $\lambda_k^{\kappa d}/b_k \to \infty$ *and*

$$\frac{\lambda_k^{\kappa d}}{b_k}(\widehat{\theta}_k - \theta_0) \xrightarrow{d} N\left(\mathbf{0}_p, (D_{\theta_0}' V^{-1} D_{\theta_0})^{-1}\right) \quad a.s.\ (P_{\mathbf{X}}) \tag{6}$$

for a positive definite matrix V *defined as*

$V = 2\sigma^2(\mathbf{0})\Gamma^*$ *for PID:* $c_n \to c \in (0,\infty)$, $b_n^2 \sim Nc_n^2$;
$V = 2\sigma^2(\mathbf{0})\Gamma^*$ *for "slow infill rate" MID:* $c_n \to 0$, $b_n^2 \sim Nc_n^2$;
$V = 2\Gamma$ *for "fast infill rate" MID:* $c_n \to 0$, $b_n^2 \sim \lambda_n^{\kappa d}$,

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

Distributional convergence behavior in the spatial EL estimator $\widehat{\theta}_n$ holds regardless of the outcome of spatial sampling locations $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ (i.e., a.s. $(P_{\mathbf{X}})$). While $\widehat{\theta}_n$ is indeed consistent for the true spatial parameter θ_0, an unusual issue is that the variance V of an associated normal limit depends on the asymptotic spatial sampling structure, particularly the dichotomy of "fast infill" MID or not (i.e., alternatively "slow infill" MID or PID regimes). Furthermore, for "slow infill" MID or PID spatial cases, this limiting variance V of $\widehat{\theta}_n$ can also depend on the sample subsequence $\{k \equiv n_k\}$ determining Γ^* from (5). Note that this is not an issue with a "fast infill" MID structure, as the limiting variance does not involve Γ^*. However, because Γ^* may in fact change across subsequences (Section 3.1), there is no guarantee that the spatial EL estimator $\widehat{\theta}_n$ actually converges in distribution in the "slow infill" MID and PID spatial cases. This is atypical for EL methods generally and for the frequency domain version with time series in particular. For (equi-spaced) time series, the same point estimator has a standard normal limit (cf. Nordman and Lahiri, 2006):

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}_p, V_{\theta_0}),$$

where V_{θ_0} has the same form as V in Theorem 1 under the "fast infill rate" MID case (with the convention that the domain of integrals is switched to $[-\pi, \pi]$ and $K = 2\pi$ with $d = 1$ there). This variance correspondence is also, at first glance, surprising as asymptotic sampling scheme for time series (i.e., $Z_1, \ldots, Z_n$ with $n \to \infty$) most closely resembles PID spatial sampling (cf. Cressie, 1993). The explanation for this correspondence, as well as for the different cases in Theorem 1 by spatial asymptotic structures, lies in the fact the spatial periodogram has bias (3) that change depending on the spatial asymptotic regime. Under "fast infill" MID, the spatial periodogram is essentially unbiased which then closely matches periodogram properties for standard time series (Brillinger, 1981). In comparison, the bias of the spatial periodogram does not decay as quickly under "slow infill" MID or PID regimes as a function of $c_n = \lambda_n^d / n$ (cf. (3)), which creates the limiting cases in Theorem 1.

The Theorem 1 behavior in point estimation may be traced to a few key and complicating factors in frequency domain EL inference for irregularly located spatial data. The spatial periodogram provides a data transformation which is helpful for weakening spatial dependence prior to implementing EL. But, as a consequence, one must then treat spectral estimating functions and their moment condition (1) in the spatial EL. This step loses estimating functions framed in terms of process probability distributions, which are more commonly and broadly applied in EL formulations for independent and time series data; see Owen (1990, 2001), Qin and Lawless (1994), Kitamura (1997). Then upon translating EL into the frequency domain, a second complicating feature is the unbounded frequency domain $\mathbb{R}^d$ inherent to such spatial data. In the spatial EL, this aspect creates additional complexities for spectral estimating functions, based on spatial frequencies, because of the averaging condition (5). Non-standard asymptotics then result for EL point estimators in combination with the spatial asymptotic structure (and bias issues (3) in the spatial periodogram).

Remark: While the EL point estimator $\widehat{\theta}_n$ may not converge in distribution under the regularity assumption (5), log-ratio statistics $-2\log[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)]$, under the same condition, generally can have chi-square limits based on the spatial EL function (4) (at the true parameter θ_0). In other words, despite the non-standard behavior in spatial point estimators $\widehat{\theta}_n$, EL log-ratio statistics based on $\widehat{\theta}_n$ have more standard behavior for potential application in parameter tests and confidence regions. Hence, the spatial EL may provide frequency domain inference for spatial processes without stringent assumptions about the underlying process distribution, the stochastic pattern of locations, or even the type of spatial asymptotic structure (PID/MID). This provides an improvement upon the EL theory and methodology in [BLN] who considered simplified test statistics $-2\log[\mathcal{R}_n(\theta_0)]$ without maximum EL estimation. The potential uses of test statistics involving the EL maximizer $\widehat{\theta}_n$ will be considered elsewhere; see (Van Hala et al., 2015).

Acknowledgements. SB's research is partially supported by NSF DMS-1406622. SNL's research is partially supported by NSF DMS-1007703 and DMS-1310068, and DJN's research is partially supported by NSF DMS-1406747.

References

Bandyopadhyay, S., and Lahiri, S. N. (2009). Asymptotic properties of discrete Fourier transforms for spatial data. *Sankhya, Ser. A* **71**, 221-259.

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

Bochner, S. (1959). *Lectures on Fourier integrals.* Princeton University Press.

Brillinger, D. R. (1981). *Time series: data analysis and theory.* Holden-Day, San Francisco.

Cressie, N. (1993). *Statistics for spatial data.* John Wiley & Sons, New York.

Gikhman, I. I., and Skorokhod, A. V. (1974). *The theory of stochastic processes I.* Springer-Verlag, Berlin.

Hall, P., and Patil, P. (1994). Properties of nonparametric estimators of autocovariance functions for stationary random fields. *Probab. Theory Related Fields* **99**, 399-424.

Kitamura, Y. (1997). Empirical likelihood methods for weakly dependent processes. *Ann. Statist.* **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, Ser. A* **65**, 356-388.

Lahiri, S. N., and Mukherjee, K. (2004). Asymptotic distributions of M-estimators in a spatial regression model under some fixed and stochastic spatial sampling designs. *Ann. Inst. Statist. Math.* **56**, 225-250.

MATSUDA, Y. AND YAJIMA, Y. (2009). Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *J. Roy. Statist. Soc. Ser. B* **71**, 191-217

MONTI, A. C. (1997). Empirical likelihood confidence regions in time series models. *Biometrika* **84**, 395-405.

NORDMAN, D. J., AND LAHIRI, S. N. (2006). A frequency domain empirical likelihood for short- and long-range dependence. *Ann. Statist.***34**, 3019-3050.

NORDMAN, D. J., AND LAHIRI, S. N. (2014). A review of empirical likelihood methods for time series. *J. Statist. Plann. Inference* **155**, 1-18.

OWEN, A. B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90-120.

OWEN, A. B. (2001). *Empirical likelihood.* Chapman & Hall/CRC, Boca Raton, Florida.

QIN, J., AND LAWLESS, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300-325.

STEIN, M. L. (1989). Asymptotic distributions of minimum norm quadratic estimators of the covariance function of a Gaussian random field. *Ann. Statist.* **17**, 980-1000.

STEIN, M. L. (1999). *Interpolation of spatial data: some theory for kriging.* Springer-Verlag, New York.

VAN HALA, M, BANDYOPADHYAY, S., LAHIRI, S. N., AND NORDMAN, D. J.(2015) A general frequency domain method for assessing spatial covariance structure. *Preprint.*

YAGLOM, A. M. (1987). *Correlation theory of stationary and related random functions.* Springer-Verlag, New York.

YAU, C. Y.(2012) Empirical likelihood in long-memory time series models. *J. Time Ser. Anal.* **33**, 269-275.