Technical Report RM 702, Department of Statistics and Probability, Michigan State University

Goodness-of-fit tests for marginal distribution of linear random fields with long memory[1]

Hira L. Koul[2] and Donatas Surgailis[3]
Michigan State University and Vilnius University

Abstract

This paper addresses the problem of fitting a known distribution function to the marginal error distribution of a stationary long memory moving average random field observed on increasing ν-dimensional "cubic" domains when its mean is known and unknown. In the case of unknown mean, when mean is estimated by the sample mean, the first order difference between the residual empirical and null distribution functions is asymptotically degenerate at zero, and hence can not be used to fit a distribution up to an unknown mean. In this paper we show that by using a suitable class of estimators of the mean, this first order degeneracy does not occur. In a similar context, we obtain the asymptotic chi-square distribution of some test statistics based on kernel density estimators. The paper extends the recent results of Koul, Mimoto, and Surgailis (2013) from $\nu = 1$ to $\nu > 1$. As a by-product, we define consistent estimators of the long run variance and memory parameters of our spatial model which may have independent interest.

1 Introduction

Spatial statistics and data analysis has become a fast developing area of research. See the monographs of Ripley (1988), Ivanov and Leonenko (1989), Cressie (1993), Guyon (1995), and Stein (1999). While many of these works deal with rather simple autoregressive and point process models with short-range dependence, a number of empirical studies ranging from astrophysics to agriculture and atmospheric sciences indicate that spatial data may exhibit nonsummable correlations and strong dependence, see, e.g., Kashyap and Lapsa (1988), Gneiting (2000), Percival *et al.* (2008) and Carlos-Davila *et al.* (1985), among others.

[1] *AMS 1991 subject classification*: Primary 62G07; secondary 62M10.
Keyword and phrases: Kernel density estimator, chi square distribution. **June 3, 2013**
[2] Research supported in part by NSF DMS grant 1205271.
[3] This author thanks the Michigan State University for hosting his visit from April 7 to May 10, 2013, during which this project was initiated. Research of this author was also supported in part by grant MIP-063/2013 from the Research Council of Lithuania.

Many of the applied works assume Gaussian random field model, which raises the question of goodness-of-fit testing. In the case of i.i.d. observations, the goodness-of-fit testing problem has been well studied, see, e.g., Durbin (1973, 1975), and D'Agostino and Stephens (1986), among others. Koul and Surgailis (2010) and Koul, Mimoto, and Surgailis (2013) discussed the problem of fitting a known distribution function to the marginal error distribution of a stationary long memory moving average time series when its mean μ is known and unknown. In particular, Koul *et al.* (2013) provided a class of weighted least squares estimators of μ for which the weak limit of the first order difference between the residual empirical and null distribution functions is a non-degenerate Gaussian distribution, yielding a simple Kolmogorov-type test for fitting a known distribution up to an unknown mean. In the same context, the latter paper also obtained the asymptotic chi-square distribution of test statistics based on integrated square difference between kernel type estimators of the marginal density of long memory moving averages with discrete time $t \in \mathbb{Z} := \{0, \pm 1, \cdots\}$ and the expected value of the error density estimator based on errors.

The aim of the present paper is to extend the results of Koul *et al.* (2013) to spatial observations. Specifically, we consider a moving average random field

$$X_t = \sum_{s \in \mathbb{Z}^\nu} b_{t-s} \zeta_s, \quad t \in \mathbb{Z}^\nu, \tag{1.1}$$

indexed by points of ν-dimensional lattice $\mathbb{Z}^\nu := \{0, \pm 1, \pm 2, \ldots\}^\nu, \nu = 1, 2, \cdots$, where $\{\zeta_s, s \in \mathbb{Z}^\nu\}$ are i.i.d.r.v.'s with zero mean and unit variance. The moving-average coefficients $\{b_t, t \in \mathbb{Z}^\nu\}$ satisfy

$$b_t = \big(B_0(t/|t|) + o(1)\big)|t|^{-(\nu-d)}, \quad t \in \mathbb{Z}^\nu \setminus \{0\}, \qquad \text{for some} \quad 0 < d < \nu/2, \tag{1.2}$$

as $|t| \to \infty$, where $B_0(x), x \in S_{\nu-1} := \{y \in \mathbb{R}^\nu : |y| = 1\}$ is a bounded piece-wise continuous function on the unit sphere $S_{\nu-1}$. The series in (1.1) converges in mean square and defines a stationary random field $\{X_t\}$ with $EX_0 = 0$ and

$$\mathrm{Cov}(X_0, X_t) \sim R_0(t/|t|)|t|^{-(\nu-2d)}, \quad \text{as} \quad |t| \to \infty, \tag{1.3}$$

where R_0 is a strictly positive and continuous function on $S_{\nu-1}$ defined in (4.10) below, see Proposition 4.4. Since, for $0 < d < \nu/2$, $\sum_{t \in \mathbb{Z}^\nu \setminus \{0\}} |t|^{-(\nu-2d)} = \infty$, the random field $\{X_t\}$ has long memory. Let

$$A_n := [1, n]^\nu \cap \mathbb{Z}^\nu \qquad \text{and} \qquad \bar{X}_n := n^{-\nu} \sum_{t \in A_n} X_t \tag{1.4}$$

be the sample mean of $\{X_t\}$ observed on the ν-dimensional cube, A_n. Then

$$\mathrm{Var}(\bar{X}_n) \sim c(1) n^{2d-\nu} \quad \text{and} \quad \frac{n^{\nu/2-d}\bar{X}_n}{\sqrt{c(1)}} \to_D Z, \tag{1.5}$$

where

$$c(1) := \int_{[0,1]^\nu} \int_{[0,1]^\nu} R_0\Big(\frac{u-v}{|u-v|}\Big) \frac{du dv}{|u-v|^{\nu-2d}}.$$

Here, and in the sequel, Z denotes a $\mathcal{N}(0,1)$ r.v., $\to_D$ stands for the convergence in distribution, and $\to_p$ stands for the convergence in probability. For a proof of the above facts, see, e.g., Surgailis (1982).

Now, let F and f denote the marginal distribution and density functions of X_0 and F_0 be a known distribution function (d.f.) with density f_0. The problem of interest is to test the hypothesis

$$\mathcal{H}_0 : F = F_0 \qquad \text{vs.} \qquad \mathcal{H}_1 : F \neq F_0.$$

A motivation for this problem is that often in practice one uses inference procedures that are valid under the assumption of $\{X_t\}$ being a Gaussian field. The rejection of this hypothesis would cast some doubt about the validity of such inference procedures.

Now, define

$$\widehat{F}_n(x) := n^{-\nu} \sum_{t \in A_n} I(X_t \le x), \;\; x \in \mathbb{R}, \quad \theta := (c(1), d)', \quad \|f_0\|_\infty := \sup_{x \in \mathbb{R}} f_0(x).$$

A test of $\mathcal{H}_0$ is the Kolmogorov-Smirnov test based on $D_n := \sup_{x \in \mathbb{R}} |\widehat{F}_n(x) - F_0(x)|$. The limit distribution of the empirical process $\widehat{F}_n$ for long memory moving-average observations $\{X_t\}$ with one dimensional time $t \in \mathbb{Z}$ was studied in Giraitis, Koul, and Surgailis (1996), Ho and Hsing (1996), and Koul and Surgailis (2002). A similar problem for moving-average random fields in $\mathbb{Z}^\nu$, $\nu > 1$ was investigated in Doukhan, Lang, and Surgailis (2002). In particular, in this last paper it was shown, under some conditions, that

$$\mathcal{D}_n(\theta) := \frac{n^{\nu/2-d} D_n}{\sqrt{c(1)} \|f_0\|_\infty} \to_D |Z|. \tag{1.6}$$

Let $\hat{c}(1), \hat{d}$ be consistent and $\log(n)$ consistent estimators of $c(1)$ and d, under $\mathcal{H}_0$, respectively, and set $\hat{\theta} := (\hat{c}(1), \hat{d})$. Let z_α be $100(1-\alpha)$th percentile of $\mathcal{N}(0,1)$ distribution. From (1.6), we readily obtain that the test that rejects $\mathcal{H}_0$ whenever $\mathcal{D}_n(\hat{\theta}) \ge z_{\alpha/2}$, is of the asymptotic size α.

Now consider the problem of fitting F_0 to F up to an unknown location parameter. In other words now we observe Y_t's from the model $Y_t = \mu + X_t$, for some $\mu \in \mathbb{R}$, and the problem of interest is to test

$$\begin{aligned} &\mathcal{H}_{0\ell oc}: \; F(x) = F_0(x-\mu), \quad \forall\, x \in \mathbb{R}, \text{ for some } \mu \in \mathbb{R}, \text{ vs.} \\ &\mathcal{H}_{1,\ell oc}: \; \mathcal{H}_{0\ell oc} \quad \text{is not true.} \end{aligned}$$

Let $\bar{F}_n$ be the empirical d.f. based on $Y_t - \bar{Y}_n$, $t \in A_n$, and $\bar{D}_n := \sup_x |\bar{F}_n(x) - F_0(x)|$. It follows from Doukhan *et al.* (2002) that, similarly as in the case $\nu = 1$ studied by Koul and Surgailis (2010) and Koul *et al.* (2013), that $n^{\nu/2-d}\bar{D}_n \to 0$, in probability, and hence $n^{\nu/2-d}\bar{D}_n$ cannot be used asymptotically to test for $\mathcal{H}_{0\ell oc}$.

We shall now describe the main results and the remaining contents of this paper. As in Koul *et al.* (2013), in Section 2 we provide a class of weighted least squares estimators $\widetilde{Y}_n$ of μ for which the normalized weak limit of the spatial empirical process $\widetilde{F}_n$ based on residuals $Y_t - \widetilde{Y}_n, t \in A_n$, has a non-degenerate Gaussian distribution under $\mathcal{H}_{0\ell oc}$ (Theorem 2.1). Testing for $\mathcal{H}_{0\ell oc}$ requires consistent estimates of the asymptotic variance of $\widetilde{Y}_n$ and the memory parameter d which are defined and studied in Sections 2 and 4, respectively; see Lemma 2.2 and Proposition 4.1. The limit distributions of test statistics based on kernel density estimators for testing for $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$ are provided in Section 3. Examples of random fields related to spatial fractional integration are discussed in Section 5. All proofs are confined to Section 6. Note also that while our study of the residual spatial empirical process relies on the results in Doukhan *et al.* (2002), we also improve some of the results of this paper by providing a better approximation rate, see Lemma 6.1.

2 Asymptotics of the spatial residual empirical process

We shall first define an estimator $\widetilde{Y}_n$ of μ. Let φ be a piece-wise continuously differentiable function on $[0,1]^\nu$ and let

$$\varphi_{nt} := n^\nu \int_{\prod_{j=1}^{\nu}\big((t_j-1)/n, t_j/n\big]} \varphi(u)du, \qquad t = (t_1, \cdots, t_\nu) \in A_n$$

be its average value on cube $\prod_{j=1}^{\nu}\big((t_j-1)/n, t_j/n\big] \subset [0,1]^\nu$. Thus, $\bar{\varphi}_n := n^{-\nu}\sum_{t\in A_n}\varphi_{nt} = \int_{[0,1]^\nu}\varphi(u)du =: \bar{\varphi}$, $\forall\, n \geq 1$. Next, define

$$\widetilde{Y}_n \;:=\; n^{-\nu}\sum_{t\in A_n} Y_t[1+\varphi_{nt}] \;=\; \mu\big(1+\bar{\varphi}\big) + \bar{X}_n + \bar{W}_n, \tag{2.1}$$

where $\{Y_t = X_t + \mu, t \in \mathbb{Z}^\nu\}$, $\{X_t\}$ is a zero-mean moving-average random field in (1.1), and

$$\bar{X}_n \;:=\; n^{-\nu}\sum_{t\in A_n} X_t, \qquad \bar{W}_n \;:=\; n^{-\nu}\sum_{t\in A_n} X_t\varphi_{nt}. \tag{2.2}$$

The following lemma establishes the asymptotic normality of $\bar{X}_n$ and $\bar{W}_n$.

Lemma 2.1 *Let $\varphi(x), x \in [0,1]^\nu$ be a piecewise continuously differentiable function and suppose $\{X_t\}$ satisfy (1.1) and (1.2). Then*

$$n^{\nu/2-d}\bar{X}_n \to_D \sqrt{c(1)}\,Z, \qquad n^{\nu/2-d}\bar{W}_n \to_D \sqrt{c(\varphi)}\,Z, \tag{2.3}$$

where

$$c(\varphi) \;:=\; \int_{[0,1]^\nu}\int_{[0,1]^\nu} \varphi(u)\varphi(v)R_0\Big(\frac{u-v}{|u-v|}\Big)\frac{du\,dv}{|u-v|^{\nu-2d}}. \tag{2.4}$$

Note $\bar\varphi = 0$ implies $\widetilde{Y}_n \to_p \mu$, in other words, $\widetilde{Y}_n$ of (2.1) is a consistent estimator of μ. Also note that when $\varphi(u) \ge -1$, $u \in [0,1]^\nu$ and $\bar\varphi = 0$, $\widetilde{Y}_n$ is a weighted least squares estimator since it minimizes the weighted sum of squares: $\widetilde{Y}_n = \mathrm{argmin}_{\mu\in\mathbb{R}} \sum_{t\in A_n}(Y_t-\mu)^2[1+\varphi_{nt}]$.

Next, we discuss the weak convergence of a suitably standardized residual empirical process. Let

$$\begin{aligned}
\widetilde{F}_n(x) \;&:=\; n^{-\nu}\sum_{t\in A_n} I(Y_t-\widetilde{Y}_n \le x) = \widehat{F}_n(x+\widetilde{\delta}_n), \quad \widetilde{\delta}_n := \widetilde{Y}_n - \mu,\\
\widetilde{D}_n(x) \;&:=\; \widetilde{F}_n(x) - F_0(x), \quad x\in\mathbb{R}, \qquad \widetilde{\mathcal{D}}_n := \sup_{x\in\mathbb{R}} |\widetilde{D}_n(x)|.
\end{aligned}$$

Let ζ be a copy of ζ_0. Assume that the innovation distribution satisfies

$$E|\zeta|^3 < C, \tag{2.5}$$

$$|Ee^{\mathbf{i}u\zeta}| \le C(1+|u|)^{-\delta}, \qquad \text{for some } 0<C<\infty,\ \delta>0, \forall\, u\in\mathbb{R}. \tag{2.6}$$

Under (2.6), it is shown in Doukhan *et al.* (2002) that the d.f. F of X_0 is infinitely differentiable and for some universal positive constant C,

$$(f(x), |f'(x)|, |f''(x)|, |f'''(x)|) \le C(1+|x|)^{-2}, \quad \forall\, x\in\mathbb{R},$$

where f'', f''' are the second and third derivatives of f, respectively. This fact in turn clearly implies f and these derivatives are square integrable.

Theorem 2.1 *Suppose (1.1), (1.2), (2.5) and (2.6) hold. Let $\varphi(x), x\in[0,1]^\nu$ be a piecewise continuously differentiable function satisfying $\bar\varphi = 0$. Then, under $\mathcal{H}_{0\ell oc}$,*

$$n^{\nu/2-d}\sup_{x\in\mathbb{R}} \big|\widetilde{F}_n(x) - F_0(x) - \bar{W}_n f_0(x)\big| \;=\; o_p(1), \tag{2.7}$$

and

$$n^{\nu/2-d}\widetilde{\mathcal{D}}_n \to_D \sqrt{c(\varphi)}\|f_0\|_\infty |Z|, \tag{2.8}$$

with $c(\varphi)$ as in (2.4).

Remark 2.1 Koul *et al.* (2013) used a slightly different version of $\widetilde{Y}_n$ with $\varphi(\frac{t}{n})$ instead of φ_{nt} in (2.1). For that version, Theorem 2.1 requires the addition condition $\bar\varphi_n = n^{-\nu}\sum_{t\in A_n}\varphi(\frac{t}{n}) = o(n^{\nu/2-\beta})$ and therefore the definition in (2.1) is preferable.

Remark 2.2 It follows from Theorem 2.1 that under $\mathcal{H}_{0\ell oc}$ and $\bar{\varphi} = 0$, the test that rejects $\mathcal{H}_{0\ell oc}$ whenever $(\sqrt{\widetilde{c}(\varphi)}\|f_0\|_\infty)^{-1} n^{\nu/2-\widetilde{d}}\ \widetilde{\mathcal{D}}_n > z_{\alpha/2}$ is asymptotically distribution free and of the asymptotic level α, where $\widetilde{c}(\varphi)$, $\widetilde{d}$ are, respectively, consistent and $\log(n)$-consistent estimators of $c(\varphi)$, d, under $\mathcal{H}_{0\ell oc}$.

Remark 2.3 Consistent estimation of long memory intensity d for some fully observable random field models was discussed in Boissy *et al.* (2005), Leonenko and Sakhno (2006), Frias *et al.* (2008), Guo *et al.* (2009). However, most of these results do not apply to the model in (1.1) - (1.2). In sec. 4 we introduce a simple variance-based estimator of d for the spatial model in (1.1) and obtain its consistency together with convergence rates under semi-parametric assumptions on the moving-average coefficients b_t in (1.2).

Next, we introduce a consistent estimator of $c(\varphi)$. Since

$$\mathrm{Var}(\bar{W}_n) = \frac{1}{n^{2\nu}} \sum_{t,s\in A_n} \varphi_{nt}\varphi_{ns} EX_t X_s \tag{2.9}$$

and $\mathrm{Var}(\bar{W}_n) \sim c(\varphi) n^{2d-\nu}$, see the proof of Lemma 2.1, a natural estimator of $c(\varphi)$ is

$$\widehat{c}(\varphi) := \frac{1}{q^{\nu+2d}} \sum_{u,v\in A_q} \varphi_{qu}\varphi_{qv}\widehat{\gamma}_n(u-v), \tag{2.10}$$

where $q \to \infty$, $q = 1, 2, \cdots$, $q = o(n)$ is a bandwidth sequence and $\widehat{\gamma}_n(u)$ is the estimator of the covariance $\gamma(u) := EX_0X_u$:

$$\widehat{\gamma}_n(u) := \frac{1}{n^\nu} \sum_{t,s\in A_n: t-s=u} (X_t - \bar{X}_n)(X_s - \bar{X}_n). \tag{2.11}$$

Note that for $\varphi(u) \equiv 1$ and $\nu = 1$, $\widehat{c}(1)$ is the well-known HAC estimator of the long-run variance $c(1)$ (see, e.g. Abadir, Distaso, and Giraitis (2009)). Note that

$$\begin{aligned} q^{\nu+2d} n^\nu \widehat{c}(\varphi) &= \sum_{u,v\in A_q} \varphi_{qu}\varphi_{qv} \sum_{t,s\in A_n: t-s=u-v} (X_t - \bar{X}_n)(X_s - \bar{X}_n) \\ &= \sum_{k\in A_n\ominus A_q} \Big(\sum_{u\in A_q: k+u\in A_n} \varphi_{qu}(X_{k+u} - \bar{X}_n) \Big)^2, \end{aligned}$$

where $A_n \ominus A_q := \{k \in \mathbb{Z}^\nu : k = t-u, t \in A_n, u \in A_q\}$. Hence, $\widehat{c}(\varphi) \geq 0$.

Lemma 2.2 *Let $\varphi(x)$ and $\{X_t\}$ satisfy the conditions of Lemma 2.1. Moreover, assume that $E|\zeta_0|^3 < \infty$. Then, as $n, q, n/q \to \infty$,*

$$\widehat{c}(\varphi) \to_p c(\varphi). \tag{2.12}$$

3 Asymptotics of density based test statistics

Following Koul *et al.* (2013), we shall now study asymptotic distribution of some tests based on kernel density estimators in random field set-up. Towards this end, let K be a density kernel on $[-1,1]$, $h \equiv h_n$ be bandwidth sequence, E_0 denote the expectation under $\mathcal{H}_0$, and define

$$\begin{aligned} \widehat{f}_n(x) &:= \frac{1}{n^\nu h}\sum_{t\in A_n} K\Big(\frac{x-X_t}{h}\Big), \quad \widetilde{f}_n(x) := \frac{1}{n^\nu h}\sum_{t\in A_n} K\Big(\frac{x-(Y_t-\widetilde{Y}_n)}{h}\Big), \quad x\in\mathbb{R}, \\ T_n &:= \int \big(\widehat{f}_n(x) - E_0\widehat{f}_n(x)\big)^2 dx, \qquad \widetilde{T}_n := \int (\widetilde{f}_n(x) - E_0\widehat{f}_n(x))^2 dx, \end{aligned}$$

where $Y_t = X_t + \mu$ and $\{X_t\}, A_n, \widetilde{Y}_n$ are as in the previous section. Statistics T_n and $\widetilde{T}_n$ are the analogues of the Bickel-Rosenblatt test statistics useful in testing for $\mathcal{H}_0$ and $\mathcal{H}_{0\ell oc}$ in the current set up, see Koul *et al.* (2013). Throughout, for any square integrable function g, $\|g\|^2 := \int_{\mathbb{R}} g^2(x)dx$. The proof of the following proposition uses Lemma 6.1 below, and is similar to that of Corollary 2.1 in Koul *et al.* (2013). We omit the details.

Proposition 3.1 *Assume the conditions of Theorem 2.1 hold. Moreover, suppose that kernel K is a symmetric and continuously differentiable probability density vanishing off $(-1,1)$ and that the bandwidth satisfies $h\min(n^{2d}, n^{\nu-2d}) \to \infty$. Then, the following hold.*
(i) *Suppose $\mu = 0$, $\varphi(x) \equiv -1$, and $\mathcal{H}_0$ holds. Then*

$$n^{\nu-2d} T_n \to_D \kappa^2(1) Z^2, \qquad \kappa^2(1) := c(1)\|f_0'\|^2.$$

(ii) *Suppose $\bar{\varphi} = 0$ and $\mathcal{H}_{0\ell oc}$ holds. Then*

$$n^{\nu-2d}\widetilde{T}_n \to_D \kappa^2(\varphi) Z^2, \qquad \kappa^2(\varphi) := c(\varphi)\|f_0'\|^2.$$

Let $0 < \alpha < 1$ and k_α be the $(1-\alpha)100$th percentile of the χ^2_1 distribution. From the above proposition it follows that the test that rejects $\mathcal{H}_0$ whenever $(\hat{c}(1)\|f_0'\|^2)^{-1} n^{\nu-2\hat{d}} T_n > k_\alpha$, has the asymptotic size α. Similarly, the test that rejects $\mathcal{H}_{0\ell oc}$, whenever $n^{\nu-2\widetilde{d}}(\widetilde{T}_n/\widetilde{c}(\varphi)\|f_0'\|^2) > k_\alpha$, has the asymptotic size α. Here, $\hat{c}(\varphi)$, $\hat{d}$, $\widetilde{c}(\varphi)$, $\widetilde{d}$ are as before.

As in the case of Theorem 2.1, for $\nu = 1$ the above result, including the condition on the bandwidth, agrees with Koul *et al.* (2013). The proof of the consistency of the tests based on T_n and $\widetilde{T}_n$ is analogous as in the above mentioned paper for the case of $\nu = 1$.

4 A variance-based estimator of d

Let $\widehat{c}(1;q) \equiv \widehat{c}(1)$ denote the estimator of the long-run variance in (2.10) corresponding to $\varphi \equiv 1$,

$$\widehat{C}(q) := q^{\nu+2d}\widehat{c}(1;q) = \sum_{u,v\in A_q} \widehat{\gamma}_n(u-v), \tag{4.1}$$

and define

$$\widehat{d} \ := \ \frac{1}{2}\left\{\frac{1}{\log 2}\log\big(\frac{\widehat{C}(2q)}{\widehat{C}(q)}\big) - \nu\right\}, \quad q = 1, 2, \cdots, \ q = o(n). \tag{4.2}$$

The above estimator of d is motivated by the fact that under the assumptions of Lemma 2.2 and (2.12),

$$\frac{\widehat{C}(2q)}{\widehat{C}(q)} \ = \ \frac{(2q)^{\nu+2d}\widehat{c}(1;2q)}{q^{\nu+2d}\widehat{c}(1;q)} \ \to_p \ 2^{\nu+2d}$$

as $n, q, n/q \to \infty$. Hence it immediately follows $\widehat{d} \to_p d$, or weak consistency of the estimator (4.2) under the premisses of Lemma 2.2.

Next, we investigate the convergence rate of $\widehat{d}$. Define $V(q) := E\widehat{C}(q)$.

Proposition 4.1 *Assume the following conditions, as $n, q, n/q$ all tend to infinity: There exists sequences of real numbers $A(n,q) \to \infty$ and $B(n,q) \to \infty$, such that*

$$B(n,q)\big(\frac{V(2q)}{2^{\nu+2d}V(q)} - 1\big) \ \to 1, \tag{4.3}$$

$$A(n,q)\big(\frac{\widehat{C}(q)}{V(q)} - 1\big) \ = \ O_p(1). \tag{4.4}$$

Then

$$\widehat{d} - d \ = \ \frac{1}{2(\log 2)B(n,q)} + O_p\left(A(n,q)^{-1}\right) + o\left(B(n,q)^{-1}\right). \tag{4.5}$$

Proposition 4.1 does not require stationarity or any assumptions about $\{X_t\}$ and basically uses the identity (6.14) only. In order to apply Proposition 4.1 to concrete situations, we need to separately discuss the convergences (4.3) and (4.4).

Proposition 4.2 *Assume that (1.3) is strengthened to*

$$EX_tX_0 \ = \ R_0(t/|t|)|t|^{-(\nu-2d)}\big(1 + o(|t|^{-\lambda})\big), \qquad |t| \to \infty, \tag{4.6}$$

where $R_0(x), x \in S_{\nu-1}$ is a Lipschitz function and $0 \le \lambda < 1$. Then (4.3) holds with $B(n,q) \to \infty$ such that

$$B(n,q)\big\{o(q^{-\lambda}) + (q/n)^{\nu-2d}\big\} \ \to \ 0. \tag{4.7}$$

Proposition 4.3 *Let $\{X_t\}$ satisfy (1.1)-(1.2) and $E|\zeta_0|^{2p} < \infty$ for some $1 < p \le 2$. Then (4.4) holds with*

$$A(n,q) = q^{((1/p)-1)\nu} + (n/q)^{(\nu/2)\wedge(\nu-2d)}/\log^{1/2}(n/q) \ \to \ \infty. \tag{4.8}$$

From (4.5), (4.7), (4.8) a convergence rate $\widehat{d} - d = O_p(n^{-\epsilon})$ with some $\epsilon > 0$ can be derived. E.q., assume $E|\zeta_0|^4 < \infty$, or $p = 2$ in Proposition 4.3. Choose $q = n^a$ with $0 < a < 1$ so that the exponents of q and n/q in (4.7) and (4.8) agree. This leads to the convergence rate $\widehat{d} - d = O_p(n^{-\epsilon})$ for the estimator in (4.2) with

$$0 < \epsilon < \min\Big(\frac{\lambda(\nu - 2d)}{\lambda + \nu - 2d}, \frac{2\nu[(\nu/2) \wedge (\nu - 2d)]}{2\nu + [(\nu/2) \wedge (\nu - 2d)]}\Big)$$

where ϵ is arbitrarily close to the upper bound.

Condition (4.6) of Proposition 4.2 can be specified in terms of the moving-average coefficients, see below.

Proposition 4.4 *Let $\{X_t, t \in \mathbb{Z}^\nu\}$ be a moving-average field in (1.1) with coefficients*

$$b_t = \big(B_0(t/|t|) + o(|t|^{-\kappa})\big)|t|^{-(\nu-d)}, \quad t \in \mathbb{Z}^\nu, \ |t| \to \infty, \tag{4.9}$$

where $0 \le \kappa < d, 0 < d < \nu/2$ and where $B_0(x), x \in S_{\nu-1}$ is a nonnegative Lipschitz function on $S_{\nu-1}$ not identically zero $B_0(x) \not\equiv 0$. Then EX_tX_0 satisfies (4.6) with

$$R_0(x) := \int_{\mathbb{R}^\nu} B_0(y/|y|)B_0((x+y)/|x+y|)|y|^{-(\nu-d)}|x+y|^{-(\nu-d)}dy \tag{4.10}$$

a Lipschitz function on $S_{\nu-1}$, and any $\lambda \ge 0$ satisfying

$$\lambda < d \wedge 1 \wedge \kappa. \tag{4.11}$$

5 Fractionally integrated random fields

Let $\nu = 2$ and $L_1X_{t,s} = X_{t-1,s}, L_2X_{t,s} = X_{t,s-1}$, $(t,s) \in \mathbb{Z}^2$ be backward shift operators on $\mathbb{Z}^2$. Consider a stationary fractionally integrated random field

$$(1 - pL_1 - qL_2)^d X_{t,s} \ = \ \zeta_{t,s}, \tag{5.1}$$

where $\{\zeta_{t,s}, (t,s) \in \mathbb{Z}^2\}$ are standard i.i.d. r.v.'s, $p, q \ge 0, p+q = 1$ are parameters, $0 < d < 1$ is the order of fractional integration, and $(1-z)^d = \sum_{j=0}^\infty \psi_j(d)z^j$, $\psi_j(d) := \Gamma(j-d)/\Gamma(j+1)\Gamma(-d)$. More explicitly,

$$\begin{aligned}(1 - pL_1 - qL_2)^d X_{t,s} \ &= \ \sum_{j=0}^\infty \psi_j(d) \sum_{k=0}^j \binom{j}{k} p^k q^{j-k} L_1^k L_2^{j-k} X_{t,s} \\ &= \ \sum_{u,v \ge 0} a_{u,v} X_{t-u,s-v}, \qquad a_{u,v} := \psi_{u+v}(d)\mathrm{bin}(u, u+v; p)\end{aligned}$$

and where $\mathrm{bin}(k,j;p) := \binom{j}{k}p^kq^{j-k}$, $0 \le k \le j$ are binomial probabilities. Note $\sum_{u,v\ge 0}|c_{u,v}| = \sum_{j=0}^\infty |\psi_j(d)| < \infty$, $d > 0$ and therefore the l.h.s. of (5.1) is well-defined for any stationary random field $\{X_{t,s}\}$ with $E|X_{0,0}| < \infty$. A stationary solution of (5.1) with zero-mean and finite variance can be defined as a moving-average random field:

$$X_{t,s} = (1 - pL_1 - qL_2)^{-d}\zeta_{t,s} \ = \ \sum_{u,v\ge 0} b_{u,v}\zeta_{t-u,s-v}, \tag{5.2}$$

where $b_{u,v} := \psi_{u+v}(-d)\,\mathrm{bin}(u, u+v; p)$. The random field in (5.2) is well-defined for any $0 < d < 3/4$ since

$$\begin{aligned}\sum_{u,v\ge 0} b^2_{u,v} \ &:= \ \sum_{j=0}^{\infty} \psi^2_j(-d)\sum_{k=0}^{j}(\mathrm{bin}(k,j;p))^2 \le \sum_{j=0}^{\infty}\psi^2_j(-d)\max_{0\le k\le j}\mathrm{bin}(k,j;p)\\ &\le \ C\sum_{j=0}^{\infty} j_+^{2(d-1)} j_+^{-1/2} \ < \ \infty, \qquad 0 < d < 3/4.\end{aligned} \tag{5.3}$$

It easily follows from the Moivre-Laplace theorem (Feller, 1966, ch.7, §2, Thm.1) that the result in (5.3) cannot be improved, in the sense that for any $d \ge 3/4$

$$\sum_{u,v\ge 0} b^2_{u,v} \ \ge \ \sum_{j=0}^{\infty}\psi^2_j(-d) \sum_{0\le k\le j: |k-pj|\le c/\sqrt{j}} (\mathrm{bin}(k,j;p))^2 \ \ge \ c\sum_{j\ge j_0} j^{2(d-1)} j^{-1/2} \ = \ \infty,$$

where $c > 0, j_0 > 0$ are some constants. The moving average coefficients $b_{u,v}$ in (5.2) do not satisfy the assumption (1.2) since they are very much "concentrated" along the line $uq - vp = 0$ and exponentially decay if $u, v \to \infty$ so that $|uq - vp| > c > 0$. The random field in (5.2) exhibits strongly anisotropic long memory behavior different from the random fields in (1.1)-(1.2). See Puplinskaitė and Surgailis (2012). Obviously, the results in the previous sections do not apply to (5.2).

Assume now that $p \in [0,1]$ is random and has a bounded probability density $\ell(p)$ on $[0,1]$. Consider a moving-average random field

$$\widetilde{X}_{t,s} \ = \ \sum_{u,v\ge 0} \widetilde{b}_{u,v}\zeta_{t-u,s-v}, \qquad (t,s) \in \mathbb{Z}^2, \tag{5.4}$$

where

$$\widetilde{b}_{u,v} := Eb_{u,v} = \psi_{u+v}(-d)\binom{u+v}{u}\int_0^1 p^u(1-p)^v\ell(p)dp. \tag{5.5}$$

It easily follows that

$$\begin{aligned}\widetilde{b}_{u,v} \ &\le \ C\psi_{u+v}(-d)\binom{u+v}{u}\int_0^1 p^u(1-p)^v dp\\ &= \ C\psi_{u+v}(-d)\binom{u+v}{u}B(u+1, v+1) = C\psi_{u+v}(-d)(u+v)^{-1}\end{aligned}$$

and therefore

$$\sum_{u,v\geq 0} \widetilde{b}^2_{u,v} \ \leq \ C\sum_{j=0}^{\infty} \psi_j^2(-d)(j+1)^{-1} \ < \ \infty, \quad \text{for any} \ \ 0<d<1. \tag{5.6}$$

The random field in (5.4) is of interest since it arises by aggregating random-coefficient autoregressive random fields with common innovations following Granger's (1980) contemporaneous aggregation scheme, as explained below. Consider a nearest-neighbor autoregressive random field

$$(1-a(pL_1+qL_2))Y_{t,s} \ = \ \zeta_{t,s}, \qquad (t,s)\in\mathbb{Z}^2, \tag{5.7}$$

where $a\in[0,1), p\in[0,1], q=1-p$ are random coefficients, a is independent of p and having a beta distribution with density

$$\phi(x) := B(d,1-d)^{-1}x^{d-1}(1-x)^{-d}, \qquad 0<x<1, \quad 0<d<1. \tag{5.8}$$

Let $(a_i,p_i,q_i=1-p_i), i=1,\cdots,N$ be independent copies of $(a,p,q=1-p)$ and

$$Y^{(i)}_{t,s} \ = \ (1-a_i(p_iL_1+q_iL_2))^{-1}\zeta_{t,s} = \sum_{j=0}^{\infty} a_i^j(p_iL_1+q_iL_2)^j\zeta_{t,s} \tag{5.9}$$

be solution of (5.7) with (a,p,q) replaced by (a_i,b_i,q_i). The limit aggregated random field is defined as the limit in probability: $\mathcal{Y}_{t,s} := \lim_{N\to\infty} N^{-1}\sum_{i=1}^N Y^{(i)}_{t,s}$. It easily follows by the law of large numbers that

$$\mathcal{Y}_{t,s} \ = \ \sum_{j=0}^{\infty} Ea^jE(pL_1+qL_2)^j\zeta_{t,s} \ = \ \sum_{u,v\geq 0} Ea^{u+v}E\mathrm{bin}(u,u+v;p)\zeta_{t-u,s-v} \tag{5.10}$$

In the case of beta density in (5.8), $Ea^j=\Gamma(j+d)/\Gamma(j+1)\Gamma(d)=\psi_j(-d)$ and therefore the moving average coefficients $Ea^{u+v}E\mathrm{bin}(u,u+v;p)$ in (5.10) coincide with $\widetilde{b}_{u,v}$ of (5.5), implying $\{\mathcal{Y}_{t,s}\}=\{\widetilde{X}_{t,s}\}$.

The following proposition shows that under some regularity conditions on the density ℓ of $p\in[0,1]$, the random field in (5.4) belongs to the class of random fields (1.1) discussed in this paper. Let $S_2 := \{(u,v)\in\mathbb{R}^2 : |(u,v)| = \sqrt{u^2+v^2}=1\}$, $S_2^+ := \{(u,v)\in S_2 : u\geq 0, v\geq 0\}$.

Proposition 5.1 *Assume that $\ell(x), x\in[0,1]$ is a continuous function with support in $(0,1)$. Then*

$$\widetilde{b}_{u,v} \ \sim \ \frac{1}{\Gamma(d)}\ell(\frac{u}{u+v})\frac{1}{(u+v)^{2-d}}, \qquad u+v\to\infty. \tag{5.11}$$

In particular, $\widetilde{b}_{u,v}$ in (5.5) satisfy (1.2) with

$$B_0(u,v) \ := \ \begin{cases} \frac{1}{\Gamma(d)}\ell(\frac{u}{u+v})\frac{1}{(u+v)^{2-d}}, & (u,v)\in S_2^+, \\ 0, & (u,v)\in S_2\setminus S_2^+. \end{cases}$$

Remark 5.1 Boissy *et al.* (2005), Guo *et al.* (2009) discuss fractionally integrated random fields satisfying

(5.12) $$(1-L_1)^{d_1}(1-L_2)^{d_2}Y_{t,s} = \zeta_{t,s}, \qquad (t,s)\in\mathbb{Z}^2$$

with possibly different parameters $|d_i| < 1/2, i = 1,2$. We note that (5.12) form a distinct class from (1.1) - (1.2) and also from (5.1). Extension of Theorem 2.1 to fractionally integrated spatial models in (5.12) and (5.1) remains open.

6 Proofs

Proof of Lemma 2.1. Since $\bar{X}_n$ is a particular case of $\bar{W}_n$, it suffices to consider $\bar{W}_n$ only. From (1.3), (2.9) and the dominated convergence theorem it easily follows $\mathrm{Var}(n^{\nu/2-d}\bar{W}_n) \to c(\varphi) = E(\sqrt{c(\varphi)}Z)^2$. The asymptotic normality of $\bar{W}_n$ can be established following the scheme of discrete stochastic integrals, see e.g. Surgailis (1982), Koul and Surgailis (2002, Lemma 2.4 (iii)), Giraitis *et al.* (2012, Prop.14.3.1). Details are omitted for the sake of brevity.

Proof of Lemma 2.2. We follow the argument in Lavancier, Philippe, and Surgailis (2010, proof of Prop. 4.1) in the case $\nu = 1$. Write $\widehat{c}(\varphi) = \widehat{c}_1(\varphi) + \widehat{c}_2(\varphi)$, $\widehat{c}_i(\varphi) := q^{-\nu-2d}\sum_{t,s\in A_q}\varphi_{qt}\varphi_{qs}\widehat{\gamma}_{ni}(t-s), i = 1,2$, where

(6.1) $$\widehat{\gamma}_{n1}(t-s) := \frac{1}{n^\nu}\sum_{u,v\in A_n: u-v=t-s} X_u X_v$$

is the empirical covariance from noncentered observations, and $\widehat{\gamma}_{n2}(t-s) := \widehat{\gamma}_n(t-s) - \widehat{\gamma}_{n1}(t-s)$. Then (2.12) follows from

(6.2) $$\widehat{c}_1(\varphi) \to_p c(\varphi), \qquad \widehat{c}_2(\varphi) = o_p(1).$$

To prove the first relation of (6.2), write $\widehat{c}_1(\varphi) = \sum_{i=1}^3 k_i(\varphi)$, where $k_1(\varphi)$ is obtained by replacing X_uX_v in (6.1) by $EX_uX_v = EX_tX_s = \gamma(t-s)$, viz.,

$$\begin{aligned} k_1(\varphi) &:= q^{-\nu-2d}\sum_{t,s\in A_q}\varphi_{qt}\varphi_{qs}\gamma(t-s) \\ &= \frac{1}{q^{2\nu}}\sum_{t,s\in A_q}\varphi_{qt}\varphi_{qs}R_0\left(\frac{\frac{t}{q}-\frac{s}{q}}{|\frac{t}{q}-\frac{s}{q}|}\right)\frac{1}{|\frac{t}{q}-\frac{s}{q}|^{\nu-2d}} \to c(\varphi) \end{aligned} \tag{6.3}$$

as $q\to\infty$. Terms $k_i(\varphi), i = 2,3$ correspond to the decomposition $X_uX_v - EX_uX_v = \sum_{w\in\mathbb{Z}^\nu} b_{u+w}b_{v+w}\eta_w + \sum_{w_1,w_2\in\mathbb{Z}^\nu, w_1\neq w_2} b_{u+w_1}b_{v+w_2}\zeta_{w_1}\zeta_{w_2}$ of X_uX_v in (6.1) with $\eta_w := \zeta_w^2 - E\zeta_w^2$, yielding

$$
\begin{aligned}
(6.4)\qquad k_2(\varphi) &:= q^{-\nu-2d}\sum_{w\in\mathbb{Z}^\nu}\eta_w\sum_{t,s\in A_q}\varphi_{qt}\varphi_{qs}\,\frac{1}{n^\nu}\sum_{u,v\in A_n:u-v=t-s}b_{u+w}b_{v+w},\\
k_3(\varphi) &:= q^{-\nu-2d}\sum_{w_1\neq w_2}\zeta_{w_1}\zeta_{w_2}\sum_{t,s\in A_q}\varphi_{qt}\varphi_{qs}\,\frac{1}{n^\nu}\sum_{u,v\in A_n:u-v=t-s}b_{u+w_1}b_{v+w_2}.
\end{aligned}
$$

To estimate $k_2(\varphi)$, we use the fact that the η_u's are i.i.d.r.v.'s, the well-known inequality $E|\sum_i\xi_i|^p \le 2\sum_i E|\xi_i|^p$ for independent zero mean random variables ξ_i with $E|\xi_i|^p<\infty$ and $1\le p\le 2$ (see, e.g., Giraitis *et al.*, 2012, Lemma 2.5.2), the fact $E|\eta_s|^p = C<\infty$ for $1<p\le 3/2$ due to $E|\zeta_s|^3<\infty$, and the Minkowski inequality. Using these facts, we obtain

$$
\begin{aligned}
E|k_2(\varphi)|^p &\le Cq^{-p\nu-2pd}n^{-\nu p}\sum_{w\in\mathbb{Z}^\nu}\Big(\sum_{t,s\in A_q}\sum_{u,v\in A_n:u-v=t-s}|b_{u+w}b_{v+w}|\Big)^p\\
&\le Cq^{-p\nu-2pd}n^{-\nu p}\Big(\sum_{t,s\in A_q}\sum_{u,v\in A_n:u-v=t-s}\big(\sum_{w\in\mathbb{Z}^\nu}|b_{u+w}b_{v+w}|^p\big)^{1/p}\Big)^p\\
&\le Cq^{-p\nu-2pd}n^{-\nu p}\Big(\sum_{t,s\in A_q}\sum_{u,v\in A_n:u-v=t-s}\big(\sum_{w\in\mathbb{Z}^\nu}|u+w|_+^{-p(\nu-d)}|v+w|_+^{-p(\nu-d)}\big)^{1/p}\Big)^p\\
&\le Cq^{-p\nu-2pd}n^{-\nu p}\big(\sum_{t,s\in A_q}\sum_{u,v\in A_n:u-v=t-s}\big(|u-v|_+^{\nu-2p(\nu-d)}\big)^{1/p}\big)^p\\
&\le Cq^{-p\nu-2pd}\big(\sum_{t,s\in A_q}|t-s|_+^{(\nu/p)-2(\nu-d)}\big)^p\\
(6.5)\qquad &\le Cq^{-p\nu-2pd}\big(q^{2\nu+(\nu/p)-2(\nu-d)}\big)^p \;=\; Cq^{-(p-1)\nu}\;\to\;0,
\end{aligned}
$$

since $p>1$. Finally, using the fact that for $u_1\neq u_2$, the r.v.s $\zeta_{u_1}\zeta_{u_2}$ have zero mean, finite variance and are mutually uncorrelated, we obtain

$$
\begin{aligned}
E|k_3(\varphi)|^2 &\le Cq^{-2\nu-4d}n^{-2\nu}\sum_{w_1,w_2\in\mathbb{Z}^\nu}\Big(\sum_{t,s\in A_q}\sum_{u\in A_n}b_{u+w_1}b_{u-t+s+w_2}\Big)^2\\
&\le Cq^{-2\nu-4d}n^{-2\nu}\sum_{t,s,t',s'\in A_q}\sum_{u,u'\in A_n}\sum_{w_1,w_2}|b_{u+w_1}b_{u-t+s+w_2}b_{u'+w_1}b_{u'-t'+s'+w_2}|\\
&\le Cq^{-2\nu-4d}n^{-2\nu}\sum_{t,s,t',s'\in A_q}\sum_{u,u'\in A_n}|u-u'|_+^{2d-\nu}|u-u'+s-s'-t+t'|_+^{2d-\nu}\\
(6.6)\qquad &\le Cq^{\nu-4d}n^{-\nu}\sum_{|u|<n}|u|_+^{2d-\nu}\sum_{|t|<2q}|u+t|_+^{2d-\nu}\;\le\;C(J_1+J_2),
\end{aligned}
$$

where

$$\begin{aligned} J_1 &:= q^{\nu-4d} n^{-\nu} \sum_{|u|<4q} |u|_+^{2d-\nu} \sum_{|t|<6q} |t|_+^{2d-\nu} \\ &\le Cq^{\nu-4d} n^{-\nu} q^{4d} \ = \ O((q/n)^\nu) \ = \ o(1), \\ J_2 &:= Cq^{\nu-4d}(q/n)^\nu \sum_{4q\le|u|<n} |u|^{4d-2\nu} \ \le \ Cq^{\nu-4d}(q/n)^\nu \begin{cases} n^{4d-\nu}, & 2\nu-4d<\nu \\ q^{4d-\nu}, & 2\nu-4d>\nu \\ \log(n/q), & 2\nu-4d=\nu \end{cases} \end{aligned}$$

and so $J_2 = o(1)$ as $q, n, n/q \to \infty$ in all three cases (in the last case where $2\nu - 4d = \nu$, this follows from the fact that $x \to 0$ entails $x^\nu \log(1/x) \to 0$). Clearly, (6.3)-(6.6) prove the first relation in (6.2).

It remains to show the second relation in (6.2). It follows from

$$q^{-2d} \sum_{|t|\le q} E|\widehat{\gamma}_{n2}(t)| = o(1). \tag{6.7}$$

Using the definition of $\widehat{\gamma}_{n2}(t) = \widehat{\gamma}_n(t) - \widehat{\gamma}_{n1}(t)$, the Cauchy-Schwarz inequality, and (1.5), we obtain

$$\begin{aligned} &(E|\widehat{\gamma}_n(t) - \widehat{\gamma}_{n1}(t)|)^2 \\ &\le \ E\bar{X}_n^2\, E\Big(n^{-\nu} \sum_{u,v\in A_n: u-v=t} X_v\Big)^2 + E\bar{X}_n^2\, E\Big(n^{-\nu} \sum_{u,v\in A_n: u-v=t} X_u\Big)^2 + \big(E\bar{X}_n^2\big)^2 \\ &\le \ Cn^{4d-2\nu}, \end{aligned}$$

with C independent of $|t| < n/2$. Hence, (6.7) reduces to $(q/n)^{\nu-2d} = o(1)$ which is a consequence of $d < \nu/2$ and $q/n \to 0$. This proves (6.2) and completes the proof of Lemma 2.2.

Proofs of Theorem 2.1 and Proposition 3.1 rely on Lemma 6.1, below. It improves the bound in Doukhan *et al.* (2002, Lemma 1.4) (DLS for brevity). For reader's convenience, in the rest of the paper we use the parametrization $\beta := \nu - d$ from DLS. Let us note although (DLS, (1.5)) assume a more restrictive form of moving average coefficients b_t (with term $o(1)$) missing in (1.2), the proofs in DLS easily extend to b_t in (1.2) since they rely on the upper bound $|b_t| \le C|t|^{-(\nu-d)}$ only. Let

$$\begin{aligned} H_t(x) &:= I(X_t \le x) - F(x) + f(x)X_t, \quad H_t(x,y) := H_t(y) - H_t(x), \ \ t \in \mathbb{Z}^\nu, \\ g(x) &:= (1+|x|)^{-3/2}, \qquad x, y \in \mathbb{R}. \end{aligned}$$

We also need to define

$$a \ := \ \begin{cases} 4\beta - 2\nu, & \nu/2 < \beta < 3\nu/4, \\ 2\beta - \nu/2, & 3\nu/4 < \beta < \nu, \end{cases} \tag{6.8}$$

Lemma 6.1 *For all $t, t' \in \mathbb{Z}^\nu$, $x, y, x', y', \in \mathbb{R}$,*

$$\big|\mathrm{Cov}\big(H_t(x,y), H_{t'}(x',y')\big)\big| \ \le \ C\Big(\int_x^y g(u)du \int_{x'}^{y'} g(v)dv\Big)^{1/2} |t-t'|_+^{-a}. \tag{6.9}$$

Remark 6.1 Lemma 6.1 and its extensions play an important role in long memory inference, see Giraitis *et al.* (2012, Chs. 10, 11). Note that in the case $\nu = 1$ the exponent a in (6.9) agrees with that given in Koul *et al.* (2013, (4.1)) or Giraitis *et al.* (2012, Lemma 10.2.5). Lemma 6.1 extends the above mentioned works also in the case $\nu = 1$ since it applies to noncausal moving averages, in contrast to causal processes discussed in most of the long memory literature.

Proof of Lemma 6.1. Using (DLS, (1.10)) and the notation from DLS, we have $H_t(x,y) = \sum_{s\succeq 0} U_{t,s}(x,y)$ and hence

$$\big|\mathrm{Cov}\big(H_t(x,y), H_{t'}(x',y')\big)\big| \ \le \ \sum_{s,s'\succeq 0} \big|\mathrm{Cov}\big(U_{t,s}(x,y), U_{t',s'}(x',y')\big)\big| \ \le \ \sum_{i=1}^{3} \Gamma_{t,t',i},$$

where

$$\begin{aligned}
\Gamma_{t,t',1} &:= \sum_{s,s'\succeq 0: t+s=t'+s'} \big|\mathrm{Cov}\big(U_{t,s}(x,y), U_{t',s'}(x',y')\big)\big|,\\
\Gamma_{t,t',2} &:= \sum_{s,s'\succeq s_1: t+s\ne t'+s'} \big|\mathrm{Cov}\big(U_{t,s}(x,y) - U_{t,s;t',s'}(x,y), U_{t',s'}(x',y') - U_{t',s';t,s}(x',y')\big)\big|,\\
\Gamma_{t,t',3} &:= \sum_{0\preceq s,s'\preceq s_1: t+s\ne t'+s'} \big|\mathrm{Cov}\big(U_{t,s}(x,y) - U_{t,s;t',s'}(x,y), U_{t',s'}(x',y') - U_{t',s';t,s}(x',y')\big)\big|.
\end{aligned}$$

Terms $\Gamma_{t,t',2}$ and $\Gamma_{t,t',2}$ are evaluated exactly as in DLS. In particularly, from (DLS, (3.13), (3.10)) we obtain

$$\begin{aligned}
|\Gamma_{t,t',2}| &\le \sum_{s,s'\succeq s_1: t+s\ne t'+s'} E^{1/2}\big(U_{t,s}(x,y) - U_{t,s;t',s'}(x,y)\big)^2 E^{1/2}\big(U_{t',s'}(x',y') - U_{t',s';t,s}(x',y')\big)^2\\
&\le C(\mu(x,y)\mu(x',y'))^{1/2} \sum_{s,s'\in\mathbb{Z}^\nu} |s|_+^{-\beta}|s'|_+^{-\beta}|t'+s'-t|_+^{-\beta}|t+s-t'|_+^{-\beta}\\
&\le C(\mu(x,y)\mu(x',y'))^{1/2}|t-t'|_+^{2\nu-4\beta}, \qquad (6.10)
\end{aligned}$$

where $\mu(x,y) := \int_x^y g(u)du$ is a finite measure on $\mathbb{R}$. Also, (DLS, (3.14)) with $\delta = 1$ extends to possibly different intervals $(x,y), (x',y')$, see also (DLS, (5.1)), yielding

$$|\Gamma_{t,t',3}| \ \le \ C(\mu(x,y)\mu(x',y'))^{1/2}|t-t'|_+^{\nu-3\beta}. \tag{6.11}$$

Let us check that for any $t, s \in \mathbb{Z}^\nu$, $x < y$,

$$EU_{t,s}^2(x,y) \ \le \ C\mu(x,y)|s|_+^{\nu-4\beta}, \tag{6.12}$$

which improves the exponent $\alpha := \min(3\beta, 3\beta - \nu/2) = 3\beta - \nu/2$ in the corresponding bound (DLS, Lemma 3.3) due to $\beta > \nu/2$. The proof of (6.12) follows the decomposition $U_{t,s}(x,y) = \sum_{i=1}^3 V_i$ as in (DLS, p.888), where $EV_1^2 \le C\mu(x,y)|b_s|^3 \le C\mu(x,y)|s|_+^{-3\beta}$ and $EV_3^2 \le C\mu(x,y)b_s^2|s|_+^{\nu-2\beta} \le C\mu(x,y)|s|_+^{\nu-4\beta}$ as in (DLS, (4.12), (4.15), (4.16)). However, $V_3 = b_s\zeta_{t+s}\int_x^y[f'(z) - f'(z - Y_{t,s})]dz$ is now estimated using (DLS, (4.4)) with $\gamma = 3/2$ which leads to $|V_2| \le C|b_s\zeta_{t+s}|\mu(x,y)(|Y_{t+s}| + |Y_{t,s}|^{3/2})$ and

$$EV_2^2 \le C\mu(x,y)b_s^2(E|Y_{t,s}|^2 + E|Y_{t,s}|^3) \le C\mu(x,y)|s|_+^{\nu-4\beta},$$

using the fact that $Y_{t,s} = \sum_{u\succ s} b_u\zeta_{t+s}$ is a sum of independent r.v.'s, hence $E|Y_{t,s}|^p \le C\big(\sum_{u\succ s} b_s^2(E|\zeta_{t+s}|^p)^{2/p}\big)^{p/2} \le C\big(\sum_{u\succ s} b_s^2\big)^{p/2} \le CB_s$ for $2 \le p \le 3$ by Rosenthal's inequality (see, e.g., Giraitis et al., 2012, Lemma 2.5.2), where $B_s \le C|s|_+^{\nu-2\beta}$ is estimated in (DLS, (4.16)). This proves (6.12). Using (6.12) and the Cauchy-Schwarz inequality, term $\Gamma_{t,t',1}$ can be easily estimated as

$$\begin{aligned} |\Gamma_{t,t',1}| &\le C(\mu(x,y)\mu(x',y'))^{1/2}\sum_{s\in\mathbb{Z}^d}|s|_+^{\nu/2-2\beta}|t+s-t'|_+^{\nu/2-2\beta} \\ &\le C(\mu(x,y)\mu(x',y'))^{1/2}|t-t'|_+^{-a} \end{aligned} \tag{6.13}$$

with a given in (6.8). The statement of the lemma now follows from (6.10), (6.11), and (6.13).

Proof of Theorem 2.1. To prove (2.7), write $\widetilde{F}_n(x) - F_0(x) - \bar{W}_n f_0(x) = U_{n1}(x) + U_{n2}(x)$, where

$$\begin{aligned} U_{n1}(x) &:= \widehat{F}_n(x+\widetilde{\delta}_n) - F_0(x+\widetilde{\delta}_n) + f_0(x+\widetilde{\delta}_n)\bar{X}_n, \\ U_{n2}(x) &:= F_0(x+\widetilde{\delta}_n) - F_0(x) - f_0(x+\widetilde{\delta}_n)\bar{X}_n - f_0(x)\bar{W}_n, \qquad \widetilde{\delta}_n := \widetilde{Y}_n - \mu. \end{aligned}$$

According to (DLS, Cor.1.2), $\|U_{n1}\|_\infty = o_p(n^{\nu/2-\beta})$. Next, by Taylor's expansion, $U_{n2}(x) = f_0(x)\widetilde{\delta}_n + o_p(\widetilde{\delta}_n) - f_0(x)\bar{X}_n - f_0(x)\bar{W}_n = f_0(x)\mu\bar{\varphi} + o_p(\widetilde{\delta}_n) = o(n^{\nu/2-\beta})$ uniformly in $x \in \mathbb{R}$. Recall $\beta = \nu - d$. These facts and $\widetilde{\delta}_n = \bar{X}_n + \bar{W}_n = O_p(n^{\nu/2-\beta})$ entail (2.7), from which (2.8) immediately follows.

Proof of Proposition 4.1. From definition (4.2) follows the immediate identity

$$\begin{aligned} \widehat{d} - d &= \frac{1}{2\log 2}\log\Big(\frac{\widehat{C}(2q)}{V(2q)}\,\frac{V(q)}{\widehat{C}(q)}\,\frac{V(2q)}{2^{\nu+2d}V(q)}\Big) \\ &= \frac{1}{2\log 2}\Big\{\log\big(1 + \frac{\widehat{C}(2q) - V(2q)}{V(2q)}\big) - \log\big(1 + \frac{\widehat{C}(q) - V(q)}{V(q)}\big) \\ &\qquad + \log\big(1 + \big(\frac{V(2q)}{2^{\nu+2d}V(q)} - 1\big)\big)\Big\}. \end{aligned} \tag{6.14}$$

It is clear from (4.3) that $\log\big(1+\big(\frac{V(2q)}{2^{\nu+2d}V(q)}-1\big)\big) = B(n,q)^{-1}+o\left(B(n,q)^{-1}\right)$. Therefore, (4.5) follows from (6.14) and the fact $\log\big(1+\frac{\widehat{C}(2q)-V(2q)}{V(2q)}\big)-\log\big(1+\frac{\widehat{C}(q)-V(q)}{V(q)}\big) = O_p\left(A(n,q)^{-1}\right)$, which is an easy consequence of (4.4). Proposition 4.1 is proved.

Proof of Proposition 4.2. We have $V(q) = \sum_{t,s\in A_q} E\widehat{\gamma}_n(u-v) = \sum_{i=0}^{3} V_i(q)$, where $V_0(q) := c(1)q^{\nu+2d}$ is the main term (which satisfies $\frac{V_0(2q)}{2^{\nu+2d}V_0(q)}-1=0$ by definition), and

$$
\begin{aligned}
V_1(q) &:= q^{\nu+2d}\Big\{\frac{1}{q^{2\nu}}\sum_{t,s\in A_q}\frac{R_0\big(\frac{\frac{t}{q}-\frac{s}{q}}{|\frac{t}{q}-\frac{s}{q}|}\big)}{|\frac{t}{q}-\frac{s}{q}|^{\nu-2d}}-c(1)\Big\},\\
V_2(q) &:= \sum_{t,s\in A_q}\Big(EX_tX_s-\frac{R_0(t-s)}{|t-s|_+^{\nu-2d}}\Big),\\
V_3(q) &= O_p(q^{2\nu}E\bar{X}_n^2) = O_p(q^{2\nu}n^{2d-\nu}). \qquad (6.15)
\end{aligned}
$$

are remainder terms. By Lipschitz continuity of R_0 [**(6.16) needs to be double checked - the exponent may be incorrect!**],

$$V_1(q) = O(q^{\nu+2d-1}). \qquad (6.16)$$

Assumption (4.6) implies

$$V_2(q) = o(q^{\nu+2d-\lambda}). \qquad (6.17)$$

From (6.15)-(6.17), we obtain

$$V(q) = c(1)q^{\nu+2d}\big(1+O(q^{-\lambda})+O_p((q/n)^{\nu-2d})\big)$$

and hence

$$\frac{V(2q)}{2^{\nu+2d}V(q)}-1 = \frac{1+O(q^{-\lambda})+O_p((q/n)^{\nu-2d})}{1+O(q^{-\lambda})+O_p((q/n)^{\nu-2d})}-1 = O(q^{-\lambda})+O_p((q/n)^{\nu-2d}),$$

or the statement of the proposition.

Proof of Proposition 4.3. We follow the proof of Lemma 2.2. With (4.1) in mind, write

$$\widehat{c}(1;q) = q^{-\nu-2d}\sum_{u,v\in A_q}\widehat{\gamma}_n(u-v) = \widehat{c}_1(1;q)+\widehat{c}_2(1;q),$$

where $\widehat{c}_1(1;q) := q^{-\nu-2d}\sum_{u,v\in A_q}\widehat{\gamma}_{n1}(u-v)$ and $\widehat{\gamma}_{n1}$ are defined in (6.1). Then, similarly as in the proof of Lemma 2.2, split $\widehat{c}_1(1;q) = \sum_{i=1}^{3} k_i(1;q)$, where $k_1(1;q) = q^{-\nu-2d}V(q)$ and $k_i(1;q), i=2,3$ are defined as in (6.4) with $\varphi_{qt}=\varphi_{qs}\equiv 1$. Obviously, the bounds (6.5) and (6.6) apply, yielding

$$E|k_2(1;q)|^p = O(q^{(p-1)\nu}), \qquad E|k_3(1;q)|^2 = O((q/n)^{\nu\wedge(2\nu-4d)}\log(n/q)).$$

From the proof of (6.7) we also have $E|\widehat{c}_2(1;q)| \le C(q/n)^{\nu-2d}$. Combining the above facts yields

$$\frac{\widehat{C}(q)}{V(q)} \;=\; 1 + O_p(q^{(1-(1/p))\nu}) + O_p((q/n)^{(\nu/2)\wedge(\nu-2d)} \log^{1/2}(\nu/q)) + O_p((q/n)^{\nu-2d}).$$

Proposition 4.3 is proved.

Proof of Proposition 4.4. Note first that $R_0(x)$ in (4.10) is strictly positive and continuous on $S_{\nu-1}$. Indeed, since $B_0(y) \ge 0$ and $B_0 \not\equiv 0$, there exist $w_0 \in S_{\nu-1}, \epsilon > 0$ such that $\inf_{w\in U_\epsilon(w_0)} B_0(w) > \epsilon$, where $U_\epsilon(w_0) := \{w \in S_{\nu-1} : |w-w_0| < \epsilon\}$. Let $y = rw \in \mathbb{R}^\nu, r > 0, w \in U_\epsilon(w_0)$. Then for any $x \in S_{\nu-1}$, $\lim_{r\to\infty}(x+rw)/|x+rw| = \lim_{r\to\infty}(x/r+w)/|x/r+w| = w \in U_\epsilon(w_0)$ and the last convergence is uniform in $w \in U_\epsilon(w_0)$. In particular, $B_0(y/|y|) > 0$ and $B_0((x+y)/|x+y|) > 0$ for any $y = rw \in \mathbb{R}^\nu, r > r_0, w \in U_\epsilon(w_0)$, where $r_0 > 0$ is large enough. Therefore the integral $R_0(x) > 0$.

Denote $b^0_t := B_0(t/|t|)|t|^{-(\nu-d)}, b^1_t := b_t - b^0_t$. For $t = x|t| \in \mathbb{Z}^\nu, x \in S_{\nu-1}$, $|t| > 0$ consider

$$R_{|t|}(x) \;:=\; |t|^{\nu-2d} E X_t X_0 \;=\; |t|^{\nu-2d} \sum_{s\in\mathbb{Z}^\nu} b_s b_{t+s}.$$

Since $\inf_{x\in S_{\nu-1}} R_0(x) > 0$ (see above), (4.6) follows from

$$|R_{|t|}(x) - R_0(x)| \;=\; o(|t|^{-\lambda}), \qquad |t| \to \infty. \tag{6.18}$$

We have $R_{|t|}(x) = R^0_{|t|}(x) + R^1_{|t|}(x)$, where

$$\begin{aligned} R^0_{|t|}(x) \;&:=\; |t|^{\nu-2d} \sum_{s\in\mathbb{Z}^\nu} b^0_s b^0_{t-s} \\ &=\; |t|^{\nu-2d} \sum_{s\in\mathbb{Z}^\nu} B_0(s/|s|) B_0((t-s)/|t-s|) |s|^{-(\nu-d)} |t-s|^{-(\nu-d)} \\ &=\; |t|^{-\nu} \sum_{y=s/|t|\in\mathbb{R}^\nu, s\in\mathbb{Z}^\nu} B_0\big(\frac{y}{|y|}\big) B_0\big(\frac{x-y}{|x-y|}\big) |y|^{-(\nu-d)} |x-y|^{-(\nu-d)}, \end{aligned}$$

with the convention $B_0(0/|0|)|0|^{-(\nu-d)} := b_0$. Let $F_x(y) := B_0(y/|y|)B_0((x-y)/|x-y|)|y|^{-(\nu-d)}|x-y|^{-(\nu-d)}$. Then

$$|R^0_{|t|}(x) - R_0(x)| \;\le\; \frac{C}{|t|^\nu} + \int_{|y|>\frac{2\nu}{|t|}, |x-y|>\frac{2\nu}{|t|}} \big\{ \sup_{z:|y-z|\le\frac{\nu}{|t|}} |F_x(y) - F_x(z)| \big\} dy$$

Using $|\frac{y}{|y|} - \frac{z}{|z|}| = |\frac{y-z}{|y|} + \frac{z(|z|-|y|)}{|y||z|}| \le \frac{|y-z|}{|y|} + \frac{||z|-|y||}{|y|} \le 2\frac{|y-z|}{|y|}$ and $|\frac{1}{|y|^{\nu-d}} - \frac{1}{|z|^{\nu-d}}| \le C\frac{|y-z|}{|y|^{1+\nu-d}}$ ($|y-z| < |y|/2$) together with the Lipschitz condition for B_0 we obtain

$$\begin{aligned} \Big|\frac{B_0(y/|y|)}{|y|^{\nu-d}} - \frac{B_0(z/|z|)}{|z|^{\nu-d}}\Big| \;&\le\; \Big|\frac{B_0(y/|y|) - B_0(z/|z|)}{|y|^{\nu-d}}\Big| + B_0(z/|z|)\Big|\frac{1}{|y|^{\nu-d}} - \frac{1}{|z|^{\nu-d}}\Big| \\ &\le\; C\frac{|y-z|}{|y|^{1+\nu-d}}. \end{aligned} \tag{6.19}$$

This implies

$$
\begin{aligned}
|F_x(y)-F_x(z)| &\le \Big|\frac{B_0(y/|y|)}{|y|^{\nu-d}}-\frac{B_0(z/|z|)}{|z|^{\nu-d}}\Big|\times\frac{B_0((x-y)/|x-y|)}{|x-y|^{\nu-d}}\\
&+ \frac{B_0(z/|z|)}{|z|^{\nu-d}}\times\Big|\frac{B_0((x-y)/|x-y|)}{|x-y|^{\nu-d}}-\frac{B_0((x-z)/|x-z|)}{|x-z|^{\nu-d}}\Big|\\
&\le C\frac{|y-z|}{|y|^{1+\nu-d}|x-y|^{\nu-d}}+C\frac{|y-z|}{|y|^{\nu-d}|x-y|^{1+\nu-d}}
\end{aligned}
$$

and consequently

$$
\begin{aligned}
|R^0_{|t|}(x)-R_0(x)| &\le \frac{C}{|t|^\nu}+\frac{C}{|t|}\Big(\int_{|y|>\frac{2\nu}{|t|}}\frac{dy}{|y|^{1+\nu-d}|x-y|^{\nu-d}}+\int_{|x-y|>\frac{2\nu}{|t|}}\frac{dy}{|y|^{\nu-d}|x-y|^{1+\nu-d}}\Big)\\
&\le \frac{C}{|t|^\nu}+\frac{C}{|t|}\Big(\int_{\frac{2\nu}{|t|}}^1\frac{r^{\nu-1}dr}{r^{1+\nu-d}}+1\Big)\\
&\le C\begin{cases}|t|^{-d}, & 0<d<1,\\ |t|^{-1}, & d>1,\\ |t|^{-1}\log|t|, & d=1.\end{cases} \qquad (6.20)
\end{aligned}
$$

Next, with $\delta_{|t|}=o(1)$

$$
\begin{aligned}
|R^1_{|t|}(x)| &\le C|t|^{\nu-2d}\sum_{s\in\mathbb{Z}^\nu}\big(\delta_{|s|}|s|_+^{-(\nu-d+\kappa)}|t-s|_+^{-(\nu-d)}+|s|_+^{-(\nu-d)}\delta_{|t-s|}|t-s|_+^{-(\nu-d+\kappa)}\big)\\
&\le o(|t|^{-\kappa}). \qquad (6.21)
\end{aligned}
$$

Combining (6.20) and (6.21) we obtain (6.18), or relation (4.6) with λ as in (4.11).

Finally, let us prove that R_0 in (4.10) is Lipschitz. Using the notation and the argument in (6.19) we have that

$$
\begin{aligned}
|R_0(x)-R_0(w)| &\le \int_{\mathbb{R}^\nu}\frac{B_0(y/|y|)}{|y|^{\nu-d}}\Big|\frac{B_0((x+y)/|x+y|)}{|x+y|^{\nu-d}}-\frac{B_0((w+y)/|w+y|)}{|w+y|^{\nu-d}}\Big|dy\\
&\le C\int_{\mathbb{R}^\nu}\frac{|x-w|}{|y|^{\nu-d}|x+y|^{\nu-d}}dy \ \le\ C|x-w|.
\end{aligned}
$$

Proposition 4.4 is proved.

Proof of Proposition 5.1. Let $j=u+v$ and $u/(u+v)=k/j$. Then

$$\widetilde{b}_{u,v} \;=\; \psi_j(-d)\int_0^1 \mathrm{bin}(k,j;p)\ell(p)dp.$$

We shall use the following version of the Moivre-Laplace theorem (Feller, 1966, ch.7, §2, Thm.1): *There exists a constant C such that when $j\to\infty$ and $k\to\infty$ vary in such a way that*

$$\frac{(k-pj)^3}{j^2} \;\to\; 0, \qquad (6.22)$$

then

$$(6.23)\qquad \left|\frac{\mathrm{bin}(k,j;p)}{\frac{1}{\sqrt{2\pi jp(1-p)}}\exp\{-\frac{(k-jp)^2}{2jp(1-p)}\}}-1\right| \;<\; \frac{C}{j}+\frac{C|k-pj|^3}{j^2}.$$

Note that the constant C in (6.23) does not depend on $p \in (\epsilon, 1-\epsilon), \epsilon > 0$ separated from 0 and 1. Next, for a small $\delta > 0$, split $E\mathrm{bin}(k,j;p) = E\mathrm{bin}(k,j;p)I(|k-pj|^3/j^2 \le \delta) + E\mathrm{bin}(k,j;p)I(|k-pj|^3/j^2 > \delta) =: \beta_1(k,j) + \beta_2(k,j)$. Using (6.23), we can write $\beta_1(k,j) = \gamma_1(k,j) + \gamma_2(k,j)$, where

$$\gamma_1(k,j) \;:=\; \int_{\{|p-k/j|^3\le\delta/j\}} \frac{1}{\sqrt{2\pi jp(1-p)}}\exp\{-\frac{(k-jp)^2}{2jp(1-p)}\}\ell(p)dp$$

and

$$\begin{aligned}|\gamma_2(k,j)| &\le C(\delta+j^{-1})\int_{\{|p-k/j|^3\le\delta/j\}} \frac{1}{\sqrt{2\pi jp(1-p)}}\exp\{-\frac{(k-jp)^2}{2jp(1-p)}\}\ell(p)dp\\ &\le C(\delta+j^{-1})\frac{1}{\sqrt{2\pi j\epsilon}}\int_{\mathbb{R}}\exp\{-(j/2)(p-k/j)^2\}dp\\ &\le C(\delta+j^{-1})j^{-1} \;=\; o(1/j), \qquad (6.24)\end{aligned}$$

where we used the facts that $1 \ge p(1-p) > \epsilon$ on $\{p \in [0,1] : \ell(p) > 0\}$ and that $\delta > 0$ can be chosen arbitrarily small. Next, using the continuity of $\ell(p)$ and $1/p(1-p)$ we see that $\gamma_1(k,j) = \tilde\gamma_1(k,j)(1+o(1))$, $j \to \infty$, where

$$\begin{aligned}\tilde\gamma_1(k,j) &:= \frac{\ell(k/j)}{\sqrt{2\pi j(k/j)(1-(k/j))}}\int_{\{|p-k/j|^3\le\delta/j\}}\exp\{-\frac{(k-jp)^2}{2(k/j)(1-(k/j))}\}dp\\ &= \frac{\ell(k/j)}{j}\big(1+o(1)\big). \qquad (6.25)\end{aligned}$$

To estimate $\beta_2(k,j)$, we use Hoeffding's inequality (Hoeffding, 1963), according to which

$$(6.26)\qquad \beta_2(k,j) \le \sup_{\epsilon<p<1-\epsilon} b(k,j;p)I(|k-pj|^3/j^2 > \delta) \le 2e^{-2\delta^{1/3}j^{1/6}} = o(1/j).$$

Relations (6.24), (6.25), and (6.26) entail (5.11), hence the proposition.

References

Abadir, K., Distaso, W. and Giraitis, L. (2009). Two estimators of the long-run variance: beyond short memory. *J. Econometrics*, **150**, 56–70.

Boissy, Y., Bhattacharyya, B.B., Li, X. and Richardson, G.D. (2005). Parameter estimates for fractional autoregressive spatial processes. *Ann. Statist.*, **33**, 2533–2567.

Carlos-Davila, E., Mejia-Lira, F. and Moran-Lopez, J.L. (1985). Ferromagnetism and spatial long range order in binary alloys: Systems with face-centered cubic structure. *J. Phys. C: Solid State Phys.*, **18**, 1217–1224.

Cressie, N.A.C. (1993). *Statistics for spatial data.* Wiley, New York.

D'Agostino, R. B. and Stephens, M. A. (Eds.) (1986). *Goodness-of-fit techniques.* **Statistics: Textbooks and Monographs 68**. Dekker, New York.

Doukhan, P., Lang, G. and Surgailis, D. (2002). Asymptotics of weighted empirical processes of linear random fields with long range dependence. *Ann. d'Inst. de H. Poincaré*, **38**, 879–896.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, **1**, 279–290.

Durbin, J. (1975). *Distribution theory for test based on the sample d.f.* SIAM, Philadelphia.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, vol. 1. Wiley, New York.

Frias, M.P., Alonso, F.J., Ruiz-Medina, M.D. and Angulo, J.M. (2008). Semiparametric estimation of spatial long-range dependence. *J. Statist. Plann. Inf.*, **138**, 1479–1495.

Giraitis, L., Koul, H.L. and Surgailis, D. (1996). Asymptotic normality of regression estimators with long memory errors. *Statist. Probab. Lett.*, **29**, 317–335.

Giraitis, L., Robinson, P.M. and Surgailis, D. (1999). Variance-type estimators of long memory. *Stoch. Proces. Appl.*, **80**, 1–24.

Giraitis, L., Koul, H.L., Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes*. Imperial College Press, London.

Gneiting, T. (2000). Power-law correlations, related models for long range dependence, and their simulation. *J. Appl. Prob.*, **37**, 1104–1109.

Granger, C.W.J. (1980). Long memory relationship and the aggregation of dynamic models. *J. Econometrics*, **14**, 227–238.

Guo, H., Lim, C. and Meerschaert, M. (2009). Local Whittle estimator for anisotropic random fields. *J. Multiv. Anal.*, **100**, 993–1028.

Guyon, X. (1995). *Random fields on a network. Modeling, statistics and applications.* Springer, New York.

Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long memory moving averages. *Ann. Statist.*, **24**, 992–1024.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, **58**, 13–30.

Ivanov, A. V. and Leonenko, N. N. (1989). *Statistical analysis of random fields*. Kluwer Academic Publishers Group, Dordrecht.

Kashyap, R.L. and Lapsa, P.M. (1984). Synthesis and estimation of random fields using long correlations. *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 800–809.

Koul, H.L. and Surgailis, D. (2002). Asymptotic expansion of the empirical process of long memory moving averages. In: Dehling, H., Mikosch, T. and Sørensen, M., eds., *Empirical process techniques for dependent data*, pp. 213–239, Birkhäuser, Boston.

Koul, H. and Surgailis, D. (2010). Goodness-of-fit testing under long memory. *J. Statist. Plann. Inf.*, **140**, 3742-3753.

Koul, H., Mimoto, N. and Surgailis, D. (2013). Goodness-of-fit tests for long memory moving average marginal density. *Metrika*, **76**, 205–224.

Lavancier, F., Philippe, A. and Surgailis, D (2010). A two-sample test for comparison of long memory parameters. *J. Multiv. Anal.*, **101**, 2118–2136.

Leonenko, N. and Sakhno, L. (2006). On the Whittle estimators for some classes of continuous parameter random processes and fields. *Statist. Probab. Lett.*, **76**, 781-795.

Percival, D.B., Rothrock, D.A., Thorndike, A.S., and Gneiting, T. (2008). The variance of mean sea-ice thickness: Effect of long range dependence. *J. Geophys. Research:* **Oceans 113**.

Puplinskaitė, D. and Surgailis, D. (2012). Aggregation of autoregressive random fields and anisotropic long memory. Preprint.

Ripley, B.D. (1988). *Statistical Inference for Spatial Processes*. Cambridge Univ. Press, New York.

Stein, M.L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Series in Statistics, Springer, New York.

Surgailis, D. (1982). Domains of attraction of self-similar multiple integrals. *Lithuanian Math. J.*, **22**, 185–201.