Asymptotic distributions of some scale estimators in nonlinear models with long memory errors having infinite variance[1]

by

Hira L. Koul and Donatas Surgailis

Michigan State University and Vilnius University

Abstract

To have scale invariant M estimators of regression parameters in regression models there is a need for having a robust scale invariant estimator of a scale parameter. The two such estimators are the median of the absolute residuals s_1 and the median of the absolute differences of pairwise residuals, s_2. The asymptotic distributions of these estimators in regression models when errors have finite variances are known in the case errors are either i.i.d. or form a long memory stationary process. Since M estimators are robust against heavy tail error distributions, it is natural to know if these scale estimators are consistent under heavy tail error distribution assumptions. This paper derives their limiting distributions when errors form a linear long memory stationary process with α-stable $(1 < \alpha < 2)$ innovations and moving average coefficients decaying as $j^{d-1}, 0 < d < 1 - 1/\alpha$. We prove that s_2 has an α_*-stable limit distribution with $\alpha_* = \alpha(1-d) < \alpha$ while the convergence rate of s_1 is generally worse than that of s_2. The proof is based on the 2nd order asymptotic expansion of the empirical process of the stated infinite variance stationary sequence derived in this paper.

1 Introduction and Summary

Let $p \wedge q \geq 1$, be fixed integers, $n \geq p$ be an integer, and Ω be an open subset of the p-dimensional Euclidean space $\mathbb{R}^p$, $\mathbb{R} = \mathbb{R}^1$. Let $\{z_{ni},\ i = 1, 2, \cdots, n\}$, be arrays of known constants and g be a known real valued function on $\Omega \times \mathbb{R}^q$. In non-linear regression model of interest here, one observes an array of random variables $\{X_{ni},\ i = 1, 2, \cdots, n\}$ such that for some $\beta_0 \in \Omega$,

$$X_{ni} = g(\beta_0, z_{ni}) + \varepsilon_i, \qquad 1 \leq i \leq n, \tag{1.1}$$

where the errors ε_i are given by the moving average

$$\varepsilon_i = \sum_{j \leq i} b_{i-j}\zeta_j, \qquad b_j \sim c_0 j^{-(1-d)}, \qquad (j \to \infty),\ d < 1/2,\ c_0 > 0. \tag{1.2}$$

---

[1] *AMS 1991 subject classification:* Primary 62G05; Secondary 62G35.

Key words and phrases. Median. Absolute residuals. Pairwise absolute differences.

When $\zeta_j, j \in \mathbb{Z} := \{0, \pm 1, \cdots\}$ are i.i.d. having zero mean and finite variance, $\varepsilon_i, i \in \mathbb{Z}$ is well-defined strictly stationary process for all $d < 1/2$ and has long memory in the sense that the sum of lagged auto-covariances diverge for all $0 < d < 1/2$.

For convenience, from now on, we shall write $g_{ni}(\beta)$ for $g(\beta, z_{ni})$ and g_{ni} for $g_{ni}(\beta_0)$. Write $\mathrm{med}\{x_i; 1 \le i \le m\}$ for the median of a given set of real numbers $\{x_i, 1 \le i \le m\}$. Let $\widehat{\beta}$ be an estimator of β_0, $r_{n,i} \equiv X_{ni} - g_{ni}(\widehat{\beta})$ and define the scale estimators

$$(1.3) \qquad s_1 \;:=\; \mathrm{med}\left\{|r_{n,i}|;\, 1 \le i \le n\right\}, \qquad s_2 := \mathrm{med}\left\{|r_{n,i} - r_{n,j}|\,;\, 1 \le i < j \le n\right\}.$$

Observe that s_1 estimates the median σ_1 of the distribution of $|\varepsilon_1|$ while s_2 the median σ_2 of the distribution of $|\varepsilon_1 - \varepsilon_1'|$, where ε_1' is an independent copy of ε_1. The fact that each of these estimators estimates a different scale parameter is not a point of concern if our goal is only to use them in arriving at scale invariant robust estimators of β_0.

A class of M estimators of β_0 is *a priori* robust against heavy tail error distributions, cf., Huber (1981). In order to make these estimators scale invariant, it is desirable to use s_1 or s_2 in their derivation. Koul (2002) derived the asymptotic distributions of these two estimators when errors are i.i.d. or form a stationary long memory moving average process with i.i.d. zero mean finite variance innovations. In the latter case it was observed that $n^{1/2-d}(s_1 - \sigma_1) = o_p(1)$, when the marginal error distribution function (d.f.) is symmetric around zero. The same fact remains true for s_2 without the symmetry assumption. The claim made in that paper that the rate of consistency of s_1 for the median σ_1 is faster than that of s_2 for σ_2 is erroneous, since the factor $\int [f(\sigma_2 + x) - f(-\sigma_2 + x)] f(x) dx$ that appears on page 9 of that paper in the analysis of $K_n(\sigma_2)$ vanishes, for any square integrable f.

In this paper we prove that under general (nonsymmetric) long memory moving average errors in (1.2) with α-stable innovations, $1 < \alpha < 2$ and $0 < d < 1 - 1/\alpha$, the convergence rate of s_2 is better than that of s_1, in the sense that $n^{1-1/\alpha-d}(s_1 - \sigma_1) = O_p(1)$ while $n^{1-1/\alpha_*}(s_2 - \sigma_2)$ has an α_*-stable limit distribution, see Theorem 2.2, where

$$(1.4) \qquad \alpha_* := \alpha(1 - d)$$

and $1 - 1/\alpha - d < 1 - 1/\alpha_*$. In the case of symmetric errors, s_1 also has an α_*-stable limit distribution and the convergence rate of s_1 and s_2 is the same. See Theorem 2.1 (ii). In view of these facts, one may prefer to use s_2 over s_1 in the computation of a scale invariant M estimator of β_0.

Proceeding a bit more precisely, we assume that the i.i.d. innovations $\zeta_j, j \in \mathbb{Z}$ of (1.2) belong to the domain of attraction of an α-stable law, $1 < \alpha < 2$, and that the d.f. G of ζ_0 has zero mean and satisfies the tail regularity condition

$$(1.5) \qquad \lim_{x \to -\infty} |x|^\alpha G(x) = c_-, \qquad \lim_{x \to \infty} x^\alpha (1 - G(x)) = c_+,$$

for some $1 < \alpha < 2$ and some constants $0 \le c_\pm < \infty$, $c_+ + c_- > 0$. From Ibragimov and Linnik (1971, Thm. 2.6.1) it follows that the above assumptions in particular imply that

$$n^{-1/\alpha} \sum_{j=1}^{n} \zeta_j \to_D Z, \tag{1.6}$$

where Z is α-stable r.v. with characteristic function

$$\begin{aligned} Ee^{\mathbf{i}\, uZ} &= e^{-|u|^\alpha \omega(\alpha,u)}, \quad u \in \mathbb{R}, \qquad \Big(\mathbf{i} := \sqrt{-1}\Big), \\ \omega(\alpha,u) &:= -\frac{\Gamma(2-\alpha)(c_+ + c_-)}{\alpha - 1} \cos(\pi\alpha/2)\Big(1 - \mathbf{i}\frac{c_+ - c_-}{c_+ + c_-}\mathrm{sgn}(u) \tan(\pi\alpha/2)\Big). \end{aligned} \tag{1.7}$$

In addition, the weights $b_j, j \ge 0$ satisfy the asymptotics (1.2) for

$$0 < d < 1 - 1/\alpha. \tag{1.8}$$

It follows, say from Hult and Samorodnitsky (2008, Remark 3.3), that under these assumptions,

$$\textstyle\sum_{j=0}^\infty |b_j| = \infty, \qquad \sum_{j=0}^\infty |b_j|^\alpha < \infty, \tag{1.9}$$

the linear process ε_i of (1.2) is well defined in the sense of the convergence in probability, and its marginal d.f. F satisfies

$$\lim_{x\to-\infty} |x|^\alpha F(x) = B_-, \qquad \lim_{x\to\infty} x^\alpha (1 - F(x)) = B_+, \tag{1.10}$$

where

$$B_- := \sum_{j=0}^{\infty} ((b_j)_+^\alpha c_- + (b_j)_-^\alpha c_+), \quad B_+ := \sum_{j=0}^{\infty} ((b_j)_+^\alpha c_+ + (b_j)_-^\alpha c_-), \quad (b_j)_\pm := \max(0, \pm b_j).$$

Note that (1.10) implies $E|\varepsilon_0|^\alpha = \infty$ and $E|\varepsilon_0|^r < \infty$, for every $r < \alpha$, in particular $E\varepsilon_0^2 = \infty$ and $E\varepsilon_0 = 0$. Because of these facts and (1.9), this process will be called *long memory* moving average process with infinite variance. In the sequel, *we refer to the assumptions in* (1.2), (1.5)*, and* (1.8)*, as the standard assumptions* about the errors in consideration. We also note that under (1.5), the upper bound $d < 1 - 1/\alpha$ in (1.8) is necessary for the convergence of the series in (1.2) almost surely and in probability, see e.g. Samorodnitsky and Taqqu (1994, Ex. 1.26). Since the interval of d in (1.8) diminishes with α decreasing, thicker tails imply that a smaller degree of memory can be considered.

The class of moving averages satisfying these assumptions includes ARFIMA (p, d, q) with α-stable innovations, where d satisfies (1.8). See Kokoszka and Taqqu (1995) for a detailed discussion of the properties of stable ARFIMA series.

Section 2 describes the asymptotic distributions of standardized s_1, s_2 along with the needed assumptions (Theorems 2.1 and 2.2) with the proofs appearing in Section 4. Section 3 discusses first and second order asymptotic expansions of the residual empirical process. Section 5 (Appendix A) contains the proofs of the asymptotic expansions of Section 3 while Section 6 (Appendix B) contains the proofs of the two auxiliary results needed in Section 5.

2 Asymptotic distributions

This section describes the asymptotic distributions of suitably standardized s_1 and s_2 under the above set up. Let

$$a := 1 - d - 1/\alpha. \tag{2.1}$$

Assume there exists a p-vector of functions $\dot{g}$ on $\mathbb{R}^p \times \mathbb{R}^q$ such that the following holds, where $\dot{g}_{ni}(s) \equiv \dot{g}(s, z_{ni})$. For every $s \in \mathbb{R}^q$ and for every $0 < k < \infty$,

$$\sup_{1\le i\le n, \|u\|\le k} n^a \Big| g_{ni}(s + n^{-a}u) - g_{ni}(s) - n^{-a}u'\dot{g}_{ni}(s) \Big| = o(1). \tag{2.2}$$

In addition assume that

$$\max_{1\le i\le n} \|\dot{g}_{ni}(\beta_0)\| = O_p(1). \tag{2.3}$$

About the d.f. G we assume that its third derivative $G^{(3)}$ exists and satisfies

$$|G^{(3)}(x)| \;\le\; C(1 + |x|)^{-\alpha}, \qquad x \in \mathbb{R} \tag{2.4}$$

$$|G^{(3)}(x) - G^{(3)}(y)| \;\le\; C\,|x - y|(1 + |x|)^{-\alpha}, \qquad x, y \in \mathbb{R}, |x - y| < 1. \tag{2.5}$$

These conditions are satisfied if G is α-stable d.f., which follows from asymptotic expansion of stable density, see e.g. Christoph and Wolf (1992, Th. 1.5) or Ibragimov and Linnik (1971, (2.4.25) and the remark at the end of ch.2 §4). In this case, (2.4)–(2.5) hold with $\alpha+2$ instead of α. Conditions (2.4)-(2.5) imply the existence and smoothness of the marginal probability density $f(x) = dF(x)/dx$, see Lemma 5.2 (5.8) below. They are not vital for our results and can be relaxed by assuming a weak decay condition at infinity of the characteristic function of G as in (GKS, (10.2.1)); however they simplify some technical derivations below.

Before stating our results we need to recall the following fact. Let $\bar{\varepsilon}_n := n^{-1}\sum_{i=1}^n \varepsilon_i$. Astrauskas (1984), Avram and Taqqu (1986, 1992), Kasahara and Maejima (1988) describe the limiting distribution of $\bar{\varepsilon}_n$ under the standard conditions as follows. Let

$$\tilde{c} = c_0 \left(\int_{-\infty}^{1} \left(\int_0^1 (t-s)_+^{-(1-d)}\, dt \right)^{\alpha} ds \right)^{1/\alpha}. \tag{2.6}$$

Then, with Z as in (1.7),

$$n^{1-d-1/\alpha}\bar{\varepsilon}_n = n^{-d-1/\alpha}\sum_{i=1}^n \varepsilon_i \to_D \tilde{c}\, Z. \tag{2.7}$$

Let $\gamma_\pm(x) := f(x) \pm f(-x), x \geq 0$. We are now ready to state the following two theorems.

Theorem 2.1 *Suppose the regression model (1.1), (1.2) holds with g satisfying (2.2) and (2.3), and the innovation d.f. G satisfying (2.4) and (2.5). In addition, suppose $\widehat{\beta}$ is an estimator of β_0 such that*

$$\left\| n^{1-d-1/\alpha}(\widehat{\beta} - \beta_0) \right\| = O_p(1). \tag{2.8}$$

(i) *If, in addition $f(\sigma_1) \neq f(-\sigma_1)$, then for every $x \in \mathbb{R}$,*

$$\begin{aligned} (2.9) \quad & P(n^{1-d-1/\alpha}(s_1 - \sigma_1) \leq x\sigma_1) \\ &= P\Big(n^{1-d-1/\alpha}\Big(\bar{\varepsilon}_n + \big(\frac{1}{n}\sum_{i=1}^n \dot{g}_{ni}\Big)'(\widehat{\beta} - \beta_0)\Big)\gamma_-(\sigma_1) \geq -x\sigma_1\gamma_+(\sigma_1)\Big) + o(1). \end{aligned}$$

(ii) *If, in addition $f(\sigma_1) = f(-\sigma_1)$, then, for every $x \in \mathbb{R}$,*

$$P\big(n^{1-1/\alpha_*}(s_1 - \sigma_1) \leq x\sigma_1\big) \quad \to \quad P\big(Z_1^* \leq x\sigma_1\gamma_+(\sigma_1)\big),$$

where $Z_1^* := \mathcal{Z}^*(\sigma_1) - \mathcal{Z}^*(-\sigma_1)$ and $\mathcal{Z}^*(x), x \in \mathbb{R}$ is α_*-stable process defined in (3.12) *below.*

Theorem 2.2 *Suppose the regression model (1.1), (1.2) and estimator $\widehat{\beta}$ satisfy the same assumptions as in Theorem 2.1. Then*

$$P\Big(n^{1-1/\alpha_*}(s_2 - \sigma_2) \leq x\sigma_2\Big) \quad \to \quad P\Big(Z_2^* \leq x\Big), \qquad \forall\, x \in \mathbb{R},$$

where Z_2^* is an α_*-stable r.v. defined in (4.28) *below.*

Remark 2.1 Koul and Surgailis (2001) verify (2.8) for a class of M-estimators when $g(\beta, z) = \beta' z$ with errors in (1.1) satisfying the above standard conditions. Using arguments similar to those appearing in Koul (1996, 1996a), (2.8) can be shown to hold for a class of M-estimators of β_0 in the general regression model (1.1) with the errors satisfying the above standard conditions under some additional smoothness conditions on g.

3 Asymptotic expansions of the residual EP

The proofs of Theorems 2.1 and 2.2 use asymptotic expansions of certain residual empirical processes, which we shall describe in this section. Accordingly, introduce the process

$$\mathcal{V}_n^{(\xi)}(x) = n^{-1}\sum_{i=1}^n \Big(I(\varepsilon_i \leq x + \xi_{ni}) - F(x + \xi_{ni}) + f(x + \xi_{ni})\varepsilon_i\Big), \qquad x \in \mathbb{R}, \tag{3.1}$$

where $(\xi) \equiv (\xi_{ni}; 1 \le i \le n)$ are non-random real-valued arrays. Write $\mathcal{V}_n^{(0)}$ for $\mathcal{V}_n^{(\xi)}$ when all $\xi_{ni} \equiv 0$. Let β_0 be as in (1.1), and define

$$
\begin{aligned}
(3.2)\ F_n(x,\beta) &:= n^{-1}\sum_{i=1}^n I(X_{ni} \le x + g_{ni}(\beta)) = n^{-1}\sum_{i=1}^n I(\varepsilon_i \le x + d_{ni}(\beta)), \quad x \in \mathbb{R}, \\
d_{ni}(\beta) &:= g_{ni}(\beta) - g_{ni}(\beta_0), \quad 1 \le i \le n, \beta \in \mathbb{R}^p.
\end{aligned}
$$

Note that under (1.1), $F_n(x, \beta_0)$ is equivalent to the first term in $\mathcal{V}_n^{(0)}(x)$.

Consider the assumptions

$$\max_{1\le i\le n} |\xi_{ni}| = O(1). \tag{3.3}$$

$$\max_{1\le i\le n} |\xi_{ni}| = o(1). \tag{3.4}$$

$$n^{-d-1/\alpha}\sum_{i=1}^n |\xi_{ni}| = O(1). \tag{3.5}$$

Let $\Rightarrow_{D(\bar{\mathbb{R}})}$ denote the weak convergence of stochastic process in the space $D(\bar{\mathbb{R}})$ with the sup-topology, where $\bar{\mathbb{R}} := [-\infty, \infty]$.

Theorem 3.1 *Suppose standard assumptions and conditions (2.4) and (2.5) hold. Then the following holds.*
(a) *If, in addition (3.3) holds, then there exists $\kappa > 0$ such that, for any $\epsilon > 0$,*

$$P\Big[\sup_{x\in\mathbb{R}} n^{1-d-1/\alpha}|\mathcal{V}_n^{(\xi)}(x)| > \epsilon\Big] \le Cn^{-\kappa}. \tag{3.6}$$

In particular, under (1.1), for any $\epsilon > 0$,

$$P\Big[\sup_{x\in\mathbb{R}} n^{1-d-1/\alpha}\Big|F_n(x,\beta_0) - F(x) + f(x)\bar{\varepsilon}_n\Big| > \epsilon\Big] \le Cn^{-\kappa}. \tag{3.7}$$

Moreover, with Z as in (1.7) *and $\tilde{c}$ is as in (2.6),*

$$n^{1-d-1/\alpha}(F_n(x,\beta_0) - F(x)) \Longrightarrow_{D(\bar{\mathbb{R}})} -\tilde{c}f(x)Z. \tag{3.8}$$

(b) *If, in addition (3.4) and (3.5) hold, then*

$$\sup_{x\in\mathbb{R}} n^{-d-1/\alpha}\Big|\sum_{i=1}^n \Big\{I(\varepsilon_i \le x + \xi_{ni}) - I(\varepsilon_i \le x) - \xi_{ni}f(x)\Big\}\Big| = o_p(1). \tag{3.9}$$

In the case of finite variance, the fact (3.6) for $\mathcal{V}_n^{(0)}$, known as the uniform reduction principle (URP), was derived in Dehling and Taqqu (1989), Ho and Hsing (1996) and Giraitis, Koul and Surgailis (1996). For more on this see Giraitis, Koul and Surgailis (2012) (GKS), and the references therein. While the first term of the expansion (3.7) is a degenerated process

$-f(x)\bar{\varepsilon}_n$ having Gaussian asymptotic distribution, higher order terms of order $[1/(1-2d)] > k \geq 2$ have a complicated limits expressed in terms of multiple Itô-Wiener integrals. Similar asymptotic expansions for $\mathcal{V}_n^{(\xi)}$ were derived in Koul and Surgailis (2002).

In the case of infinite variance, the URP (3.6) was established in Koul and Surgailis (2001) (KS) with the innovations in (1.2) having symmetric α-stable distributions. In the present paper, this result is extended to asymmetric innovations.

The above approximations extend to various statistical functionals, see GKS and KS. However, in the case when the first order approximation vanishes as in the case of s_2, a second order approximation of the residual EP by a degenerated α_*-stable process with α_* as in (1.4), can be used to obtain the limiting distribution of a suitably standardized s_2, as is seen later in this paper. In order to describe this approximation we need some more notation. Let Z_+^*, Z_-^* denote independent copies of totally skewed α_*-stable r.v. Z^* with characteristic function

$$(3.10)\quad Ee^{\mathbf{i}uZ^*} = e^{-c_*|u|^{\alpha_*}(1-\mathbf{i}\mathrm{sgn}(u)\tan(\alpha_*\pi/2))}, \quad u \in \mathbb{R}, \quad c_* := \Gamma(2-\alpha_*)\cos(\pi\alpha_*/2)/(1-\alpha_*).$$

The above choice of parameters implies $P(Z^* > x) \sim x^{-\alpha_*}$, $P(Z^* < -x) = o(x^{-\alpha_*})$, as $x \to \infty$; see Ibragimov and Linnik (1971, Ch.2). Also denote

$$(3.11)\qquad \psi_\pm^*(x) := \big(c_0^{\frac{1}{1-d}}/(1-d)\big)\int_0^\infty \big(F(x\mp s) - F(x) \pm f(x)s\big)s^{-1-\frac{1}{1-d}}ds,$$

$$(3.12)\qquad \mathcal{Z}^*(x) := c_+^{1/\alpha_*}\psi_+^*(x)Z_+^* + c_-^{1/\alpha_*}\psi_-^*(x)Z_-^*, \qquad x \in \mathbb{R}.$$

Theorem 3.2 (a) *Suppose standard assumptions, conditions (2.4), (2.5) and*

$$(3.13)\qquad \max_{1\leq i\leq n}|\xi_{ni}| = O(n^{-\epsilon}), \qquad \exists\, \epsilon > 0,$$

hold. Then

$$(3.14)\qquad n^{1-1/\alpha_*}\mathcal{V}_n^{(\xi)}(x) \Longrightarrow_{D(\bar{\mathbb{R}})} \mathcal{Z}^*(x).$$

In particular, under (1.1),

$$(3.15)\qquad n^{1-1/\alpha_*}\big(F_n(x,\beta_0) - F(x) + f(x)\bar{\varepsilon}_n\big) \Longrightarrow_{D(\bar{\mathbb{R}})} \mathcal{Z}^*(x).$$

(b) *If, in addition,*

$$(3.16)\qquad \max_{1\leq i\leq n}|\xi_{ni}| = O(n^{d+1/\alpha-1})$$

holds, then

$$(3.17)\qquad \sup_{x\in\mathbb{R}} n^{-1/\alpha_*}\Big|\sum_{i=1}^n \Big\{I(\varepsilon_i \leq x + \xi_{ni}) - I(\varepsilon_i \leq x) - \xi_{ni}f(x)\Big\}\Big| = o_p(1).$$

The proofs of Theorems 3.1 and 3.2 are given in Section 5. We remark that from the regularity properties of F (see Lemma 5.2, (5.8)) it follows that the functions $\psi^*_\pm$ in (3.11) are well-defined and continuously differentiable on $\mathbb{R}$; moreover, $|\psi^*_\pm(x)| < C(1+|x|)^{-\gamma}$ for some $\gamma > 1$ and hence $\lim_{|x|\to\infty} \psi^*_\pm(x) = 0$. We also note that $\psi^*_\pm(x)$ agree, up to a multiplicative factor, with the Marchaud (left and right) fractional derivative of $F(x)$ of order $1/(1-d) \in (1,2)$.

Remark 3.1 The α_*-stable limit in (3.15) is related to the limit results for nonlinear functions of moving averages in Surgailis (2002, 2004), and also to the limit of power variations of some Lévy driven processes with continuous time discussed in Basse-O'Connor, Lachièze-Rey and Podolskij (2016).

4 Proofs of Theorems 2.1 and 2.2

This section contains the proofs of Theorem 2.1 and 2.2. Before proceeding further we need to introduce some additional notation. With α_* as in (1.4), let

$$a_* \;=\; 1 - \frac{1}{\alpha_*}. \tag{4.1}$$

Note that

$$a_* = 1 - \frac{1}{\alpha(1-d)} = \frac{\alpha(1-d)-1}{\alpha(1-d)}, \qquad a = 1 - d - \frac{1}{\alpha} = \frac{\alpha(1-d)-1}{\alpha} = a_*(1-d).$$

Also, $0 < d < 1-(1/\alpha)$ and $\alpha < 2$ imply $0 < (2-\alpha)/\alpha < 1-2d < 1$. Therefore $a_* - 2a = a_* - 2a_*(1-d) = -a_*(1-2d) < 0$. We also have $a < a_*$. For easy reference we summarize these facts as the following inequalities.

$$a < a_* < 2a, \qquad \forall\, 0 < d < 1 - 1/\alpha. \tag{4.2}$$

Proof of Theorem 2.1. Let

$$p_1(y) \;:=\; F(y) - F(-y), \qquad y \ge 0.$$

Define σ_1 as the unique solution of $p(\sigma_1) = 1/2$. Because σ_1 is median of the distribution of $|\varepsilon_1|$, the derivation of the asymptotic distribution of s_1 is facilitated by analysing the process

$$S(y) \;:=\; \sum_{i=1}^n I\left(|r_i| \le y\right), \qquad y \ge 0, \tag{4.3}$$

where from now on, we write r_i for $r_{n,i}$ for all $1 \le i \le n$. The investigation of the asymptotic behavior of this process in turn is facilitated by that of $F_n(x,\beta)$ of (3.2). Let

$$\mu_n(x,\beta) := n^{-1} \sum_{i=1}^n F(x + d_{ni}(\beta)), \quad x \in \mathbb{R}, \beta \in \mathbb{R}^p.$$

Rewrite $r_i \equiv \varepsilon_i - d_{ni}(\widehat{\beta}\,)$, and note that because of the continuity of F,

$$n^{-1}S(y) = F_n(y, \widehat{\beta}\,) - F_n(-y, \widehat{\beta}\,), \qquad \forall\; n \geq 1,\; y \geq 0,$$

with probability (w.p.) 1.

We use the following decomposition: Let

$$\mathcal{D}_n(x, \beta) \;:=\; \sum_{i=1}^{n} \Big\{ I(\varepsilon_i \leq x + d_{ni}(\beta)) - I(\varepsilon_i \leq x) - f(x) d_{ni}(\beta) \Big\}, \quad x \in \mathbb{R},\; \beta \in \mathbb{R}^p.$$

Then, w.p.1, for all $y \geq 0$, we have

$$\begin{aligned} (4.4) \qquad S(y) - np_1(y) \;&=\; n[\mathcal{V}_n^{(0)}(y) - \mathcal{V}_n^{(0)}(-y)] - (f(y) - f(-y)) n\bar{\varepsilon}_n \\ &\quad + \mathcal{D}_n(y, \widehat{\beta}) - \mathcal{D}_n(-y, \widehat{\beta}) + (f(y) - f(-y)) \sum_{i=1}^{n} d_{ni}(\widehat{\beta}). \end{aligned}$$

In both cases (i) and (ii) of Theorem 2.1, the behavior of all terms on the r.h.s. of (4.4) except for $\mathcal{D}_n(y, \widehat{\beta})$ can be rather easily obtained from Theorem 3.1, Theorem 3.2 and the conditions in (2.2), (2.3) on the regression model.

We shall prove that $\mathcal{D}_n(\pm y, \widehat{\beta})$ is asymptotically negligible, uniformly in $y \geq 0$, in probability. Fix a $0 < k < \infty$ and let $t \in \mathbb{R}^p$ be such that $\|t\| \leq k$. Let

$$(4.5) \qquad \xi_{ni} := g_{ni}(\beta_0 + n^{-a}t) - g_{ni}(\beta_0).$$

Recall (2.1) and write $\xi_{ni} = (g_{ni}(\beta_0 + n^{-a}t) - g_{ni}(\beta_0) - n^{-a}t'\dot{g}_{ni}) + n^{-a}t'\dot{g}_{ni}$, where $\dot{g}_{ni} \equiv \dot{g}_{ni}(\beta_0)$. By (2.2) and (2.3), for any $\epsilon > 0$, $\exists\, N$ s.t. for all $n > N$,

$$(4.6) \qquad \sup_{1 \leq i \leq n, \|t\| \leq k} |\xi_{ni}| \;\leq\; \epsilon n^{-a} + k\, n^{-a} \max_{1 \leq i \leq n} \|\dot{g}_{ni}\| = O(n^{-a}).$$

This then verifies the condition (3.16) for the ξ_{ni} of (4.5), for each t. From (3.17) we obtain

$$n^{-1/\alpha_*} |\mathcal{D}_n(x, \beta_0 + n^{-a}t)| = o_p(1), \quad \text{for all } x \in \mathbb{R},\; \|t\| \leq k.$$

This in turn together with the monotonicity of F_n, F and the compactness of the set $\{t \in \mathbb{R}^p; \|t\| \leq k\}$ and an argument as in Koul (2002a, ch. 8) now yields that for every $0 < k < \infty$,

$$(4.7) \qquad n^{-1/\alpha_*} \sup_{\|t\| \leq k, x \in \mathbb{R}} |\mathcal{D}_n(x, \beta_0 + n^{-a}t)| = o_p(1)$$

implying by assumption (2.8) on $\widehat{\beta}$ that

$$(4.8) \qquad \sup_{x \in \mathbb{R}} |\mathcal{D}_n(x, \widehat{\beta})| = o_p(n^{1/\alpha_*}).$$

By (4.2), $a < a_*$, which is equivalent to $1/\alpha_* < d + 1/\alpha = -(a-1)$. Thus (4.8) implies

$$\sup_{x\in\mathbb{R}} n^{a-1}|\mathcal{D}_n(x,\widehat{\beta})| = n^{1/\alpha_*+a-1}\sup_{x\in\mathbb{R}} n^{-1/\alpha_*}|\mathcal{D}_n(x,\widehat{\beta})| = o_p(1). \tag{4.9}$$

Proof of (2.9) **and** (2.10). From the definition (4.3), we obtain that for any $y > 0$,

$$\begin{array}{rcll} \{s_1 \le y\} &=& \{S(y) \ge (n+1)2^{-1}\}, & n \text{ odd}, \\ \{S(y) \ge n2^{-1}\} &\subseteq& \{s_1 \le y\} \subseteq \{S(y) \ge n2^{-1} - 1\}, & n \text{ even}. \end{array} \tag{4.10}$$

Thus, to study the asymptotic distribution of s_1, it suffices to study those of $S(y),\ y \ge 0$.

Case (i): $f(\sigma_1) \ne f(-\sigma_1)$. Let $\mathcal{P}_n(x)$ denote the l.h.s. of (2.9). Let $t_n := (n^{-a}x + 1)\sigma_1$. Assume n is large enough so that $t_n > 0$. Then

$$\begin{array}{rcll} \mathcal{P}_n(x) &=& P\Big(S(t_n) \ge (n+1)/2\Big), & n \ \text{ odd} \\ P(S(t_n) \ge n/2) &\le& \mathcal{P}_n(x) \le P(S(t_n) \ge n2^{-1} - 1), & n \ \text{ even}. \end{array}$$

It thus suffices to analyze $P(S(t_n) \ge n2^{-1} + b),\ b \in \mathbb{R}$. Let

$$S_1(y) \ := \ n^a\Big[n^{-1}S(y) - p_1(y)\Big], \quad y \ge 0, \qquad s_n := n^a[2^{-1} + n^{-1}b - p_1(t_n)].$$

Then $P(S(t_n) \ge n2^{-1} + b) = P\big(S_1(t_n) \ge s_n\big)$. Recall $p_1(\sigma_1) = 1/2$ and hence

$$\begin{aligned} s_n &= -n^a[p_1(t_n) - p_1(\sigma_1)] + n^{-d-1/\alpha}b \\ &= -n^a[p_1((n^{-a}x+1)\sigma_1) - p_1(\sigma_1)] + n^{-d-1/\alpha}b \\ &= -x\sigma_1[f(\sigma_1) + f(-\sigma_1)] + o(1). \end{aligned}$$

Next, by (2.2), (3.7), (4.4), (4.9), and uniform continuity of f,

$$\begin{aligned} S_1(t_n) &= [f(t_n) - f(-t_n)]\Big(n^a\bar{\varepsilon}_n + n^{a-1}\sum_{i=1}^n d_{ni}(\widehat{\beta})\Big) \\ &\qquad + n^a[\mathcal{V}_n^{(0)}(t_n) - \mathcal{V}_n^{(0)}(-t_n)] + n^{a-1}[\mathcal{D}_n(t_n,\widehat{\beta}) - \mathcal{D}_n(-t_n,\widehat{\beta})] \\ &= [f(\sigma_1) - f(-\sigma_1)]\Big(n^a\bar{\varepsilon}_n + n^a(\hat{\beta} - \beta_0)'\big(n^{-1}\sum_{i=1}^n \dot{g}_{ni}\big)\Big) + o_p(1), \end{aligned}$$

proving (2.9).

Case (ii): $f(\sigma_1) = f(-\sigma_1)$. Let $t_n^* := (n^{-a_*}x + 1)\sigma_1$, where a_* is as in (4.1). Similarly to case (i), it suffices to analyze $P(S(t_n^*) \ge n2^{-1} + b),\ b \in \mathbb{R}$. Let

$$S_1^*(y) \ := \ n^{a_*}\Big[n^{-1}S(y) - p_1(y)\Big], \quad y \ge 0, \qquad s_n^* := n^{a_*}[2^{-1} + n^{-1}b - p_1(t_n^*)].$$

Then $P(S(t_n^*) \geq n2^{-1} + b) = P\big(S_1^*(t_n^*) \geq s_n^*\big)$, where $s_n^* = -x\sigma_1[f(\sigma_1) + f(-\sigma_1)] + o(1)$. On the other hand, using $f(t_n^*) - f(-t_n^*) = O(n^{-a_*})$ and (2.3), (3.15), (4.4), (4.8), we obtain

$$\begin{aligned} S_1(t_n^*) &= [f(t_n^*) - f(-t_n^*)]\Big(n^{a_*}\bar{\varepsilon}_n + n^{a_*-1}\sum_{i=1}^n d_{ni}(\widehat{\beta})\Big) \\ &\qquad + n^{a_*}[\mathcal{V}_n^{(0)}(t_n^*) - \mathcal{V}_n^{(0)}(-t_n^*)] + n^{a_*-1}[\mathcal{D}_n(t_n^*, \widehat{\beta}) - \mathcal{D}_n(-t_n^*, \widehat{\beta})] \\ &= \mathcal{Z}^*(\sigma_1) - \mathcal{Z}^*(-\sigma_1) + o_p(1), \end{aligned}$$

proving (2.10) and completing the proof of Theorem 2.1. $\square$

Proof of Theorem 2.2. Let

$$p_2(y) := \int [F(y+x) - F(-y+x)]dF(x), \quad y \geq 0. \tag{4.11}$$

Define σ_2 as the unique solution of

$$p_2(\sigma_2) = 1/2. \tag{4.12}$$

As noted in the introduction, σ_2 is median of the distribution of $|\varepsilon_1 - \varepsilon_1'|$, where ε_1' is an independent copy of ε_1. Recall $r_i \equiv r_{n,i} = \varepsilon_i - d_{ni}(\widehat{\beta})$. The derivation of the asymptotic distribution of s_2 is facilitated by analysing the process

$$T(y) := \sum_{1 \leq i < j \leq n} I\left(|r_i - r_j| \leq y\right), \qquad y \geq 0. \tag{4.13}$$

Because of the continuity of F,

$$2n^{-1}T(y) = n\int [F_n(y+x, \widehat{\beta}) - F_n(-y+x, \widehat{\beta})]F_n(dx, \widehat{\beta}) - 1, \qquad \forall\ n \geq 1,\ y \geq 0,$$

with probability 1. Let

$$Q_n(x) := P(n^{a_*}(s_2 - \sigma_2) \leq x\sigma_2), \qquad x_n^* := (xn^{-a_*} + 1)\sigma_2 > 0.$$

From the definition of s_2 in (1.3), we obtain

$$\begin{aligned} Q_n(x) &= P\Big(T(x_n^*) \geq (N+1)/2\Big), && N \ \text{ odd} \\ P(T(x_n^*) \geq N/2) &\leq Q_n(x) \leq P(T(x_n^*) \geq N2^{-1} - 1), && N \ \text{ even.} \end{aligned}$$

It thus suffices to analyze $P(T(x_n^*) \leq N/2 + b)$, $N := n(n-1)/2$, $b \in \mathbb{R}$. Let

$$T_1(y) \ := \ n^{a_*}[\frac{2T(y)}{n^2} + \frac{1}{n}] - n^{a_*}p_2(y), \qquad k_n^* := \frac{(N+2b)n^{a_*}}{n^2} + \frac{n^{a_*}}{n} - n^{a_*}p_2(x_n^*).$$

Then

$$P(T(x_n^*) \geq N/2 + b) = P(T_1(x_n^*) \geq k_n^*).$$

Asymptotics of k_n^*. Note $Nn^{a_*}/n^2 \sim n^{a_*}/2 = n^{a_*}p_2(\sigma_2)$. Therefore

$$k_n^* = -n^{a_*}[p_2(x_n^*) - p_2(\sigma_2)] + o(1).$$

But

$$\begin{aligned} n^{a_*}[p_2(x_n^*) - p_2(\sigma_2)] &= n^{a_*}\Big[p_2\Big((n^{-a_*}x+1)\sigma_2\Big) - p_2(\sigma_2)\Big] \\ &= n^{a_*}\int\Big[\Big\{F\big((n^{-a_*}x+1)\sigma_2 + z\big) - F\big(\sigma_2+z\big)\Big\} \\ &\qquad -\Big\{F\big(-(n^{-a_*}x+1)\sigma_2) + z\big) - F\big(-\sigma_2+z\big)\Big\}\Big]dF(z). \end{aligned}$$

Because F has uniformly continuous and bounded density f,

$$n^{a_*}\int\{F(\pm(n^{-a_*}x+1)\sigma_2)+z) - F(\pm\sigma_2+z)\}dF(z) = \pm x\sigma_2\int f(\pm\sigma_2+z)dF(z) + o(1),$$

and, hence,

$$k_n^* = -x\sigma_2\int[f(\sigma_2+z) + f(-\sigma_2+z))dF(z) + o(1). \tag{4.14}$$

Asymptotics of $T_1(x_n^*)$. Let

$$W_n^*(x,\beta) := n^{a_*}[F_n(x,\beta) - \mu_n(x,\beta)], \qquad x\in\mathbb{R},\ \beta\in\mathbb{R}^p.$$

From the definition of T_1,

$$\begin{aligned} T_1(y) &= n^{a_*}\int[F_n(y+z,\widehat{\beta}\,) - F_n(-y+z,\widehat{\beta}\,)]F_n(dz,\widehat{\beta}\,) - n^{a_*}p_2(y) \\ &= \int[W_n^*(y+z,\widehat{\beta}\,) - W_n^*(-y+z,\widehat{\beta}\,)]F_n(dz,\widehat{\beta}\,) \\ &\quad + \int[\mu_n(y+z,\widehat{\beta}\,) - \mu_n(-y+z,\widehat{\beta}\,)]W_n^*(dz,\widehat{\beta}\,) \\ &\quad + n^{a_*}\int[\mu_n(y+z,\widehat{\beta}\,) - \mu_n(-y+z,\widehat{\beta}\,)]\mu_n(dz,\widehat{\beta}\,) - n^{a_*}p_2(y) \\ &= E_1(y) + E_2(y) + E_3(y), \qquad \text{say.} \end{aligned}$$

Integration by parts and a change of variable formula shows that $E_2(y) \equiv E_1(y)$. We shall now approximate E_1 and E_3. But, first note that (2.2), (2.3) and (2.8) imply that

$$\max_{1\le i\le n}|d_{ni}(\widehat{\beta}\,)| = O_p(n^{-a}). \tag{4.15}$$

Also note

$$\int[f(y+z) - f(-y+z)]f(z)dz = 0, \qquad \forall\ y\ge 0. \tag{4.16}$$

Let $\Delta_{n,ij} \equiv d_{ni}(\widehat{\beta}\,) - d_{nj}(\widehat{\beta}\,)$. Observe that by (2.2), (2.3),

$$n^{-1-d-1/\alpha} \sum_{i=1}^{n} \sum_{j=1}^{n} |\Delta_{n,ij}| = O_p(1), \qquad \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{n,ij} = 0, \quad \text{w.p. 1.} \tag{4.17}$$

The smoothness of F and (4.11) together with (4.16) imply that

$$\begin{aligned}
E_3(y) &:= n^{a_*} \int [\mu_n(y+z, \widehat{\beta}\,) - \mu_n(-y+z, \widehat{\beta}\,)]\mu_n(dz, \widehat{\beta}\,) - n^{a_*} p_2(y) \\
&= n^{a_*-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \int \Big\{ F(y+z+\Delta_{n,ij}) - F(-y+z+\Delta_{n,ij}) \\
&\qquad\qquad\qquad -F(y+z) + F(-y+z) \Big\} dF(z) \\
&= n^{a_*-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \int f(z)dz \int_0^{\Delta_{n,ij}} (f(y+z+u) - f(y+z))du.
\end{aligned}$$

Using the inequalities $f(x) + |f'(x)| \le C(1+|x|)^{-\gamma}, x \in \mathbb{R}$ for any $1 < \gamma < \alpha$, see (5.8), and

$$\int_{|w|<v} (1+|w+z|)^{-\gamma} dw \le C(1+|z|)^{-\gamma}(v \vee v^{\gamma}), \qquad \forall\, z \in \mathbb{R},\ v > 0, \tag{4.18}$$

for $1 < \gamma \le 2$, from eqn. (10.2.33) in GKS which are often used in this paper, we obtain

$$\Big| \int_0^{\Delta_{n,ij}} (f(y+z+u) - f(y+z))du \Big| \le C \frac{|\Delta_{n,ij}|^{\gamma} \wedge |\Delta_{n,ij}|^2}{(1+|y+z|)^{\gamma}} \le C\Delta_{n,ij}^2 (1+|y+z|)^{-\gamma}.$$

This in turn implies

$$\begin{aligned}
\sup_{y>0} |E_3(y)| &\le Cn^{a_*-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{n,ij}^2 \sup_{y>0} \int (1+|z|)^{-\gamma}(1+|y+z|)^{-\gamma} dz \\
&\le Cn^{a_*-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{n,ij}^2 \;=\; O_p(n^{a_*-2a}) = o_p(1),
\end{aligned}$$

according to (4.15) and (4.2).

Next, let $V_n(z, \beta) := F_n(z, \beta) - F(z)$ and

$$\begin{aligned}
K_n^*(y, \beta) &:= \int [W_n^*(y+z, \beta) - W_n^*(-y+z, \beta)] dF(z), \\
U_1(y, \beta) &:= \int [W_n^*(y+z, \beta) - W_n^*(-y+z, \beta)] dV_n(z, \beta).
\end{aligned}$$

Then

$$E_1(y) = K_n^*(y, \widehat{\beta}) + U_1(y, \widehat{\beta}).$$

From (3.15) and (4.16) we have for $K_n^*(y) \equiv K_n^*(y, \beta_0)$ that

$$K_n^*(x_n^*) \ \to_D \ \int [\mathcal{Z}_*(\sigma_2 + z) - \mathcal{Z}_*(-\sigma_2 + z)] dF(z). \tag{4.19}$$

We shall prove that

$$\sup_{y>0} |K_n^*(y, \widehat{\beta}) - K_n^*(y)| \ = \ o_p(1), \tag{4.20}$$

$$\sup_{y>0} |U_1(y, \widehat{\beta})| \ = \ o_p(1). \tag{4.21}$$

To prove (4.20) note that

$$\begin{aligned} &K_n^*(y, \beta_0 + n^{-a}t) - K_n^*(y; \beta_0) \\ &= \ n^{-1/\alpha_*} \int \Big[\sum_{i=1}^n \big\{ I(\varepsilon_i \le y + z + \xi_{ni}) - I(\varepsilon_i \le y + z) - \xi_{ni} f(y+z) \big\} \\ &\qquad - \sum_{i=1}^n \big\{ I(\varepsilon_i \le -y + z + \xi_{ni}) - I(\varepsilon_i \le -y + z) - \xi_{ni} f(-y+z) \big\} \Big] dF(z), \end{aligned} \tag{4.22}$$

where ξ_{ni} are as (4.5). In view of (4.6) these ξ_{ni} satisfy (3.16). Hence (4.16) implies that for any $t \in \mathbb{R}^p$, the left hand side of (4.22) tends to zero in probability, uniformly in $y \in \mathbb{R}$. Whence, (4.20) follows using condition (2.8) and the compactness argument, similar to (4.7).

Next, consider (4.21). We shall first prove

$$\sup_{y>0} |U_1(y, \beta_0)| \ = \ o_p(1). \tag{4.23}$$

We have

$$\begin{aligned} |U_1(y, \beta_0)| \ &= \ n^{1-1/\alpha_*} \int (F_n(y+z) - F(y+z) + f(y+z)\bar{\varepsilon}_n) dV_n(z, \beta_0) \\ &\quad - n^{1-1/\alpha_*} \int (F_n(-y+z) - F(-y+z) + f(-y+z)\bar{\varepsilon}_n) dV_n(z, \beta_0) \\ &\quad + n^{1-1/\alpha_*} \bar{\varepsilon}_n \int [f(-y+z) - f(y+z)] dV_n(z, \beta_0) \\ &= \ I_1(y) - I_2(y) + I_3(y), \qquad \text{say,} \end{aligned}$$

where $\sup_{y>0} |I_i(y)| = o_p(1), i = 1, 2$ according to (3.15), and $\sup_{y>0} |I_3(y)| = O_p(n^{a_* - 2a}) = o_p(1)$ follows from (2.7), (3.6), and (4.2). Thus, $\sup_{y>0} |U_1(y, \beta_0)| = o_p(1)$.

Next, we shall prove that for every $0 < b < \infty$,

$$\sup_{y>0, \|t\| \le b} |U_1(y, \beta_0 + n^{-a}t) - U_1(y; \beta_0)| \ = \ o_p(1). \tag{4.24}$$

To this end, first fix a $\|t\| \le b$ and let $\tilde{V}_n(y,t) = V_n(y, \beta_0 + n^{-a}t)$. Similarly as in (4.22) write

$$
\begin{aligned}
&|U_1(y, \beta_0 + n^{-a}t) - U_1(y;\beta_0)| \\
\le\ & n^{-1/\alpha_*}\Big|\int \sum_{i=1}^n \big\{I(\varepsilon_i \le y + z + \xi_{ni}) - I(\varepsilon_i \le y + z) - \xi_{ni} f(y+z)\big\} d\tilde{V}_n(z,t)\Big| \\
&+ n^{-1/\alpha_*}\Big|\int \sum_{i=1}^n \big\{I(\varepsilon_i \le -y + z + \xi_{ni}) - I(\varepsilon_i \le -y + z) - \xi_{ni} f(-y+z)\big\} d\tilde{V}_n(z,t)\Big| \\
&+ n^{-1/\alpha_*}\Big|\sum_{i=1}^n \xi_{ni}\Big| \Big|\int (f(y+z) - f(-y+z)) d\tilde{V}_n(z,t)\Big| \\
=\ & J_1(y,t) + J_2(y,t) + J_3(y,t), \qquad \text{say,}
\end{aligned}
$$

where $\sup_{y>0}|J_i(y,t)| = o_p(1), i = 1,2$ follow by (3.17) and the fact that the variation of $z \mapsto \tilde{V}_n(z,t)$ on $\mathbb{R}$ does not exceed 2, $\tilde{V}_n(z,t)$ being the difference of two d.f.s. Relation $\sup_{y>0}|J_3(y,t)| = O_p(n^{a_*-2a}) = o_p(1)$ follows from (2.7), (3.6) and (4.2) as above. Uniformity with respect to $t \in \mathbb{R}^p, \|t\| \le b$ is obtained by using an argument as in Koul (2002, Ch. 8) and the monotonicity of the indicator function.

From (4.20), (4.21), (4.23), (4.24) we conclude that for $i = 1$,

$$
E_i(x_n^*) = K_n^*(x_n^*, \beta_0) + o_p(1). \tag{4.25}
$$

Then from (4.19) we finally obtain

$$
T_1(x_n^*) \to_D 2\int [\mathcal{Z}^*(\sigma_2 + z) - \mathcal{Z}^*(-\sigma_2 + z)] dF(z). \tag{4.26}
$$

Next, let $h(z) := f(\sigma_2 + z) + f(-\sigma_2 + z)$, $z \in \mathbb{R}$, and

$$
Z_2^* \ :=\ -\frac{1}{2\sigma_2 \int h(z) dF(z)} \int [\mathcal{Z}^*(\sigma_2 + z) - \mathcal{Z}^*(-\sigma_2 + z)] dF(z). \tag{4.27}
$$

The facts together with (4.14) and (4.26) in turn imply that

$$
\begin{aligned}
&P(n^{a_*}(s_2 - \sigma_2) \le x\sigma_2) = P(T_1(x_n^*) \ge k_n^*) \\
\to\ & P\Big(2\int [\mathcal{Z}^*(\sigma_2 + z) - \mathcal{Z}^*(-\sigma_2 + z)] dF(z) \ge -x\sigma_2 \int (f(\sigma_2 + z) + f(-\sigma_2 + z)) dF(z)\Big) \\
=\ & P(Z_2^* \le x).
\end{aligned}
$$

Let $Z_\pm^*$ be as in (3.10), $\psi_\pm^*(x)$ be as in (3.11), and define

$$
\varpi_\pm \ :=\ -\frac{c_\pm^{1/\alpha_*}}{2\sigma_2 \int h(z) dF(z)} \int [\psi_\pm^*(\sigma_2 + z) - \psi_\pm^*(-\sigma_2 + z)] dF(z).
$$

Then, in view of the definition of $\mathcal{Z}^*(x)$ in (3.12), the r.v. Z_2^* in (4.27) can be represented as

$$
Z_2^* = \varpi_+ Z_+^* + \varpi_- Z_-^*. \tag{4.28}
$$

Clearly, Z_2^* in (4.28) has α_*-stable distribution. This completes the proof of Theorem 2.2. □

5 Appendix A: Proofs of Theorems 3.1 and 3.2

We use several lemmas whose proofs are given in Section 6 (Appendix B).

Lemma 5.1 *Let $\eta_i, i \ge 1$ be independent r.v.'s satisfying $E\eta_i = 0$ for all i and such that for some $1 < \alpha < 2$,*

$$K := \sup_{i\ge 1} \sup_{x>0} x^\alpha P(|\eta_i| > x) < \infty. \tag{5.1}$$

Then $\sum_{i=1}^\infty a_i\eta_i$ converges in $L_p, p < \alpha$ for any real sequence $a_i, i \ge 1, \sum_{i=1}^\infty |a_i|^\alpha < \infty$ and, for $2(\alpha-1) \le p < \alpha$,

$$E|\sum_{i=1}^\infty a_i\eta_i|^p < C\Big(\sum_{i=1}^\infty |a_i|^\alpha + \Big(\sum_{i=1}^\infty |a_i|^\alpha\Big)^{p/p'}\Big), \quad p' := 2\alpha - p > \alpha, \tag{5.2}$$

where the constant $C = C(p, \alpha, K) < \infty$ depends only on p, α, K.

We need some more notation. For any integer $j \ge 0$, define the truncated moving averages

$$\varepsilon_{i,j} := \sum_{0\le k\le j} b_k\zeta_{i-k}, \quad \widetilde{\varepsilon}_{i,j} := \sum_{k>j} b_k\zeta_{i-k}, \quad \varepsilon^*_{i,j} := \sum_{k\ge 0, k\ne j} b_k\zeta_{i-k}. \tag{5.3}$$

Thus,

$$\varepsilon_{i,j} + \widetilde{\varepsilon}_{i,j} = \varepsilon_i, \qquad \varepsilon^*_{i,j} = \varepsilon_i - b_j\zeta_{i-j}.$$

Let $F_j(x) := P(\varepsilon_{i,j} \le x)$, $\widetilde{F}_j(x) := P(\widetilde{\varepsilon}_{i,j} \le x)$, $F^*_j(x) := P(\varepsilon^*_{i,j} \le x)$ be the corresponding marginal d.f.s. Also introduce

$$F_{\ne\ell,j}(x) := P\Big(\sum_{0\le k\le j: k\ne\ell} b_k\zeta_k \le x\Big) = P(\varepsilon_{i,j} - b_{i-\ell}\zeta_\ell \le x), \qquad 0 \le \ell < j. \tag{5.4}$$

W.l.g., assume that F_j and $F_{\ne\ell,j}$ are not degenerate for any $j \ge 0, 0 \le \ell < j$. Then $P(\varepsilon_i \le x + \xi_{ni}|\zeta_{i-j}) = F^*_j(x + \xi_{ni} - b_j\zeta_{i-j})$ and

$$\eta^*_{ns}(x;\zeta_s) := \textstyle\sum_{i=1\vee s}^n \Big(F^*_{i-s}(x + \xi_{ni} - b_{i-s}\zeta_s) - F(x + \xi_{ni}) + f(x + \xi_{ni})b_{i-s}\zeta_s\Big).$$

The proofs of Lemmas 5.3–5.5 use certain regularity properties of the d.f.s $F(x), F_j(x), F^*_j(x)$, $F_{\ne\ell,j}(x)$ given in the following lemma, whose proof is given in Section 6 (Appendix B) below. Related results can be found in (KS, Lemma 4.2), Surgailis (2002, Lemma 4.1), (GKS, Lemma 10.2.4).

For $1 < \gamma \le 2$ introduce $g_\gamma(x) := (1 + |x|)^{-\gamma}, x \in \mathbb{R}$ and a finite measure μ_γ on $\mathbb{R}$ by

$$\mu_\gamma(x, y) := \textstyle\int_x^y g_\gamma(u)du, \qquad x < y. \tag{5.5}$$

Note the elementary inequalities: for any $x, y \in \mathbb{R}$

$$(5.6) \qquad g_\gamma(x+y) \le C g_\gamma(x)(1 \vee |y|)^\gamma, \qquad \left|\int_0^y g_\gamma(x+w)dw\right| \le C g_\gamma(x)(|y| \vee |y|^\gamma),$$

see (GKS, Lemma 10.2.3). We shall also use the inequality

$$(5.7) \qquad \int_x^y g_\gamma(u+v)du \le C(\mu_\gamma(x,y))^{1/\gamma}|v|, \qquad |v| \ge 1,$$

where C does not depend on x, y, v; see Surgailis (2002, p.264).

Given a function $g(x), x \in \mathbb{R}$ and points $x < y, z \in \mathbb{R}$ we use the notation $g(x,y) := g(y) - g(x)$, $g((x,y)+z) := g(y+z) - g(x+z)$.

Lemma 5.2 *Suppose standard assumptions and conditions (2.4), (2.5) hold. Then for any $p = 1, 2, 3$ the d.f. $F, F_j, F_j^*, F_{\neq \ell, j}, j \ge 0, 0 \le \ell < j$ are p times continuously differentiable. Moreover, for any $1 < \gamma < \alpha, p = 1, 2, 3$ there exists a constant $C > 0$ such that for any $x, y \in \mathbb{R}, |x-y| \le 1, j \ge 0, 0 \le \ell < j$*

$$(5.8) \qquad |F^{(p)}(x)| + |F_j^{(p)}(x)| + |(F_j^*)^{(p)}(x)| + |F_{\neq \ell, j}^{(p)}(x)| \le C g_\gamma(x),$$

$$(5.9) \qquad |F^{(p)}(x,y)| + |F_j^{(p)}(x,y)| + |(F_j^*)^{(p)}(x,y)| + |F_{\neq \ell, j}^{(p)}(x,y)| \le C|x-y| g_\gamma(x),$$

$$(5.10) \qquad \sum_{p=1}^{2} |(F_j^*)^{(p)}(x) - F^{(p)}(x)| \le C|b_j|^\alpha g_\gamma(x).$$

Moreover, for any $x < y, j \ge 0, j_2 > j_1 \ge 0, z, z_1, z_2, \xi \in \mathbb{R}, |\xi| \le 1$

$$(5.11) \qquad |F((x,y) - b_j z) - F(x,y) + f(x,y) b_j z| \le C \min\left\{\mu_\gamma(x,y)|b_j|^\gamma, (\mu_\gamma(x,y))^{1/\gamma}|b_j|\right\},$$

$$(5.12) \qquad |F_j^*((x,y) - b_j z) - F(x,y) + f(x,y) b_j z| \le C \mu_\gamma(x,y)|b_j|^\gamma (1+|z|)^\gamma,$$

$$(5.13) \qquad \left|\int_0^\xi \big(f((x,y)+v-b_j z) - Ef((x,y)+v-b_j\zeta) + f'((x,y)+v) b_j z\big) dv\right| \le C|\xi| \mu_\gamma(x,y)|b_j|^\gamma (1+|z|)^\gamma,$$

$$(5.14) \qquad \left|\int_{|w| \le |b_{j_1} z_1|} \int_{|w_2| \le |b_{j_2} z_2|} F''_{\neq j_1, j_2}((x,y) + w_1 + w_2 + z) dw_2 dw_2\right| \le C \mu_\gamma(x,y)|b_{j_1}|\,|b_{j_2}|(1+|z_1|)^\gamma (1+|z_2|)^\gamma (1+|z|)^\gamma.$$

Proof of Theorem 3.1. (a) The proof of (3.6) given in (KS, Thm. 2.1) for the case $c_+ = c_-$ of (1.5) applies *verbatim* in the general case of $c_\pm$ in (1.5). (We note that KS does not assume the distribution of ζ_0 symmetric around 0.)

(b) The sum on the l.h.s. of (3.9) can be written as $\sum_{i=1}^4 L_{ni}(x)$, where $L_{n1}(x) := n\mathcal{V}_n^{(\xi)}(x)$, $L_{n2}(x) := -n\mathcal{V}_n^{(0)}(x)$ and

$$L_{n3}(x) := \sum_{i=1}^n (F(x+\xi_{ni}) - F(x) - f(x)\xi_{ni}), \quad L_{n4}(x) := \sum_{i=1}^n (f(x) - f(x+\xi_{ni}))\varepsilon_i.$$

Then $\sup_{x\in\mathbb{R}} n^{-d-1/\alpha}|L_{ni}(x)| = o_p(1), i = 1,2$ follow from (3.6). Using $|F''(x)| \le C$, see (KS, Lemma 4.1) and (3.4), (3.5) we obtain $\sup_{x\in\mathbb{R}} n^{-d-1/\alpha}|L_{n3}(x) \le Cn^{-d-1/\alpha}\sum_{i=1}^n \xi_{ni}^2 = o(n^{-d-1/\alpha}\sum_{i=1}^n |\xi_{ni}|) = o(1)$. To evaluate $L_{n4}(x)$, let $a_{ns}(x) := n^{-d-1/\alpha}\sum_{i=1}^n (f(x) - f(x+\xi_{ni}))b_{i-s}$ and $\delta_{n,\xi} := \max_{1\le i\le n}|\xi_{ni}|$. Then $n^{-d-1/\alpha}L_{n4}(x) = \sum_{s=-\infty}^n a_{ns}(x)\zeta_s$. Moreover, using $|f(x) - f(x+\xi_{ni})| \le C\delta_{n,\xi} = o(1)$, see (3.4), and the bound in Lemma 5.1, (5.2), with $p < \alpha$ sufficiently close to α, we obtain

$$E|n^{-d-1/\alpha}L_{n4}(x)|^p \le C\big(\sum_{s=-\infty}^n |a_{ns}(x)|^\alpha + \big(\sum_{s=-\infty}^\infty |a_{ns}(x)|^\alpha\big)^{p/p'}\big). \tag{5.15}$$

But, for any fixed $x \in \mathbb{R}$,

$$\sum_{s=-\infty}^n |a_{ns}(x)|^\alpha \le C\delta_{n,\xi}^\alpha\, n^{-d\alpha-1}\sum_{s=-\infty}^n \Big(\sum_{i=1}^n (i-s)_+^{d-1}\Big)^\alpha = O(\delta_{n,\xi}^\alpha) = o(1). \tag{5.16}$$

Similarly, for any $x < y$ using $|f'(x)| \le C/(1+|x|)^r, x \in \mathbb{R}$ with $1 < r < \alpha$, see (KS, (4.4)), we obtain $E|n^{-d-1/\alpha}L_{n4}(x,y)|^p \le C\big(\sum_{s=-\infty}^n |a_{ns}(x,y)|^\alpha + \big(\sum_{s=-\infty}^\infty |a_{ns}(x,y)|^\alpha\big)^{p/p'}\big)$, and

$$\sum_{s=-\infty}^n |a_{ns}(x,y)|^\alpha \le C\big(\max_{1\le i\le n}|f(x+\xi_{ni}, y+\xi_{ni}|\big)^\alpha \le C(\mu_r(x,y))^\alpha. \tag{5.17}$$

Because $\mu_r(x,y) := \int_x^y (1+|z|)^{-r}dz$ is a finite continuous measure on $\mathbb{R}$, it follows that $E|n^{-d-1/\alpha}L_{n4}(x,y)|^p \le C(\mu_r(x,y))^{p\alpha/p'}, x < y$, where $p\alpha/p' > 1$ provided $p < \alpha$ is chosen sufficiently close to α. Therefore using the well-known tightness criterion (see e.g. GKS, Lemma 4.4.1) we conclude that the sequence of random processes $\{n^{-d-1/\alpha}L_{n4}(x), x \in \mathbb{R}\}, n \ge 1$ is tight in $D(\bar{\mathbb{R}})$, implying $\sup_{x\in\mathbb{R}} n^{-d-1/\alpha}|L_{n4}(x)| = o_p(1)$. This ends the proof of part (b) and Theorem 3.1. □

Proof of Theorem 3.2. *Proof of* (a). We follow the proof in Surgailis (2002, Thm. 2.1) and (2004, Thm. 2.1). The crucial decomposition leading to (3.14) is

$$\begin{aligned} \mathcal{R}_n(x) &:= \mathcal{V}_n^{(\xi)}(x) - \mathcal{Z}_n^*(x), \quad \text{where} \\ \mathcal{Z}_n^*(x) &:= n^{-1}\sum_{i=1}^n\sum_{s<i} E\Big[I(\varepsilon_i \le x+\xi_{ni}) - F(x+\xi_{ni}) + f(x+\xi_{ni})\varepsilon_i\Big|\zeta_s\Big]. \end{aligned} \tag{5.18}$$

We will show that $\mathcal{Z}_n^*(x)$ is the main term and $\mathcal{R}_n(x)$ is the remainder term. Note that $\mathcal{Z}_n^*(x) = n^{-1}\sum_{s<n}\eta_{ns}^*(x;\zeta_s)$, where

$$\eta_{ns}^*(x;\zeta_s) := \sum_{i=1\vee s}^n E\Big[I(\varepsilon_i \le x+\xi_{ni}) - F(x+\xi_{ni}) + f(x+\xi_{ni})\varepsilon_i\Big|\zeta_s\Big].$$

We shall approximate the above quantity by $\eta(x;\zeta_s)$, where

$$\begin{aligned} \eta(x;z) &:= \sum_{i=s}^\infty \big(F(x-b_{i-s}z) - EF(x-b_{i-s}\zeta_s) + f(x)b_{i-s}z\big) \\ &= \sum_{j=0}^\infty \big(F(x-b_jz) - EF(x-b_j\zeta_0) + f(x)b_jz\big) \end{aligned} \tag{5.19}$$

and let

$$\mathcal{Z}_n(x) := n^{-1} \sum_{s=1}^n \eta(x;\zeta_s), \qquad x \in \mathbb{R}. \tag{5.20}$$

Because $\mathcal{Z}_n(x)$ is a sum of i.i.d. r.v.'s, its α_*-stable limit will follow from the classical central limit theorem for independent r.v.'s with heavy-tailed distribution. The proof of (3.14) or of part (a) follows from the following three lemmas.

Lemma 5.3 $\quad n^{1-1/\alpha_*}\mathcal{Z}_n(x) \Longrightarrow_{D(\bar{\mathbb{R}})} \mathcal{Z}^*(x).$

Lemma 5.4 $\quad n^{1-1/\alpha_*}\sup_{x\in\mathbb{R}} |\mathcal{Z}^*_n(x) - \mathcal{Z}_n(x)| \ \to_p \ 0.$

Lemma 5.5 $\quad n^{1-1/\alpha_*}\sup_{x\in\mathbb{R}} |\mathcal{R}_n(x)| \ \to_p \ 0.$

Proof of (b). As in the proof of Theorem 3.1(b), decompose the sum on the l.h.s. of (3.17) as $\sum_{i=1}^4 L_{ni}(x)$. By (3.14), $\sup_x n^{-1/\alpha_*}|L_{n1}(x) + L_{n2}(x)| \to_p 0$. Next, by Taylor's expansion and condition (3.13),

$$\sup_x |L_{n3}(x)| = O(\textstyle\sum_{i=1}^n \xi_{ni}^2) = o(n^{1/\alpha_*}),$$

since $1/\alpha_* + 1 > 2d + 2/\alpha$, see above. The term $L_{n4}(x)$ can be estimated as in (5.15), (5.16), yielding $E|n^{-1/\alpha_*}L_{n4}(x)|^p \le C\Big(n^{d\alpha+1-\alpha/\alpha_*}\delta^\alpha_{n,\xi} + \big(n^{d\alpha+1-\alpha/\alpha_*}\delta^\alpha_{n,\xi}\big)^{p/p'}\Big)$ with $\delta_{n,\xi} \le C\max_{1\le i\le n}|\xi_{ni}| = O(n^{d+1/\alpha-1})$, see (3.13), hence $E|n^{-1/\alpha_*}L_{n4}(x)|^p = o(1)$ for any $x \in \mathbb{R}$. The proof of $\sup_x n^{-1/\alpha_*}|L_{n4}(x)| = o_p(1)$ is similar as in Theorem 3.1(b) by showing the bound $E|n^{-1/\alpha_*}L_{n4}(x,y)|^p \le C(\mu(x,y))^{p\alpha/p'}, x < y$ which follows as in (5.17) but uses a more accurate bound:

$$\begin{aligned}\textstyle\sum_{s=-\infty}^n |a_{ns}(x,y)|^\alpha &\le C\big(\max_{1\le i\le n}|f(x+\xi_{ni}, y+\xi_{ni} - f(x,y)|\big)^\alpha \\ &\le C(\mu(x,y))^\alpha \max_{1\le i\le n}|\xi_{ni}|^\alpha\end{aligned}$$

together with condition (3.13). This proves part (b) and ends the proof of Theorem 3.2. □

6 Appendix B: proofs of auxiliary lemmas

Here we present the proofs of Lemmas 5.1-5.5 used in the proofs of Theorems 3.1 and 3.2.

Proof of Lemma 5.1. W.l.g. let $a_i \ne 0, i \ge 1$. Split $\eta_i = \eta_i^+ + \eta_i^-, \eta_i^+ := \eta_i I(|\eta_i| \le 1/|a_i|) - E\eta_i I(|\eta_i| \le 1/|a_i|), \eta_i^- := \eta_i I(|\eta_i| > 1/|a_i|) - E\eta_i I(|\eta_i| > 1/|a_i|)$. Then $(\eta_i^\pm, i \ge 1)$ are sequences of independent zero mean r.v.'s and $E|\sum_{i=1}^\infty a_i\eta_i|^p \le CE|\sum_{i=1}^\infty a_i\eta_i^+|^p + CE|\sum_{i=1}^\infty a_i\eta_i^-|^p$. By the well-known moment inequality, see e.g. (GKS, Lemma 2.5.2)

$$\begin{aligned}E\Big|\sum_{i=1}^\infty a_i\eta_i\Big|^p &\le CE\Big|\sum_{i=1}^\infty a_i\eta_i^-\Big|^p + C\Big(E\Big|\sum_{i=1}^\infty a_i\eta_i^+\Big|^{p'}\Big)^{p/p'} \\ &\le C\sum_{i=1}^\infty |a_i|^pE|\eta_i^-|^p + C\Big(E\Big|\sum_{i=1}^\infty |a_i|^{p'}E|\eta_i^+|^{p'}\Big)^{p/p'}.\end{aligned}$$

Then (5.2) follows from

$$E|\eta_i^-|^p \le C|a_i|^{\alpha-p}, \qquad E|\eta_i^+|^{p'} \le C|a_i|^{\alpha-p'}, \quad \forall\, i\ge 1. \tag{6.1}$$

Note $E|\eta_i^-|^p \le 2E|\eta_i|^p I(|\eta_i|>1/|a_i|), E|\eta_i^+|^{p'} \le 2E|\eta_i|^{p'} I(|\eta_i|\le 1/|a_i|)$. Hence (5.1) implies

$$\begin{aligned} E|\eta_i^-|^p &\le -2\int_{1/|a_i|}^\infty x^p dP(|\eta_i|>x) = 2|a_i|^{-p}P(|\eta_i|>1/|a_i|) + 2p\int_{1/|a_i|}^\infty x^{p-1}P(|\eta_i|>x)dx \\ &\le 2K|a_i|^{\alpha-p} + 2pK\int_{1/|a_i|}^\infty x^{p-1-\alpha}dx \le C|a_i|^{\alpha-p}, \quad \forall\, i\ge 1. \end{aligned}$$

Similarly,

$$\begin{aligned} E|\eta_i^+|^{p'} &\le -2\int_0^{1/|a_i|} x^{p'} dP(|\eta_i|>x) \\ &= -2|a_i|^{-p'}P(|\eta_i|>1/|a_i|) + 2p'\int_0^{1/|a_i|} x^{p'-1}P(|\eta_i|>x)dx \\ &\le 2K|a_i|^{\alpha-p'} + 2p'K\int_0^{1/|a_i|} x^{p'-1-\alpha}dx \le C|a_i|^{\alpha-p'}, \quad \forall\, i\ge 1. \end{aligned}$$

proving (6.1) and the lemma, too. □

Proof of Lemma 5.2. The first part of the proposition including (5.8), (5.9) follows by the argument in (KS, Lemma 4.2) with minor changes. The proof of (5.10) also proceeds as in the proof (KS, (4.6)), as follows. Since $F(x) = \int F_j^*(x-b_jy)dG(y)$ and $E\zeta = \int ydG(y) = 0$ so $F^{(p)}(x) - (F_j^*)^{(p)}(x) = \int((F_j^*)^{(p)}(x-yb_j) - (F_j^*)^{(p)}(x) + yb_j(F_j^*)^{(p+1)}(x))dG(y)$ and $|F^{(p)}(x) - (F_j^*)^{(p)}(x)| \le \sum_{i=1}^4 |J_i(x)|$ where

$$\begin{aligned} J_1(x) &:= \textstyle\int_{|yb_j|\le 1} dG(y)\int_0^{-yb_j}((F_j^*)^{(p+1)}(x+u) - (F_j^*)^{(p+1)}(x))du, \\ J_2(x) &:= \textstyle\int_{|yb_j|>1}(F_j^*)^{(p)}(x-yb_j)dG(y), \qquad J_3(x) := \int_{|yb_j|>1}(F_j^*)^{(p)}(x)dG(y), \\ J_4(x) &:= \textstyle(f_j^*)'(x)b_j\int_{|yb_j|>1} ydG(y). \end{aligned}$$

To evaluate $J_1(x)$ for $p=1,2$ use (5.9) yielding $|\int_0^{-yb_j}(F_j^*)^{(p+1)}(x,x+u)du| \le Cg_\gamma(x)\int_0^{|yb_j|} udu \le Cg_\gamma(x)(yb_j)^2$ and then

$$|J_1(x)| \le Cb_j^2 g_\gamma(x)\textstyle\int_{|yb_j|\le 1} y^2 dG(y) \le C|b_j|^\alpha g_\gamma(x)$$

follows from $P(|\zeta|>x)\le Cx^{-\alpha}$. Next, using (5.8) and (5.6) we obtain

$$\begin{aligned} |J_i(x)| &\le \textstyle C\int_{|yb_j|>1}(g_\gamma(x) + g_\gamma(x-b_jy))dy \\ &\le \textstyle Cg_\gamma(x)\int_{|yb_j|>1}(1+|b_jy|^\gamma)dG(y) \le C|b_j|^\alpha g_\gamma(x), \quad i=2,3, \\ |J_4(x)| &\le \textstyle Cg_\gamma(x)|b_j|\int_{|yb_j|>1}|y|dG(y) \le C|b_j|^\alpha g_\gamma(x), \end{aligned}$$

thereby proving (5.10).

Consider (5.11). Write the l.h.s. of (5.11) as

$$\big|\int_0^{-b_j z} dw \int_x^y (f'(v+w) - f'(v))dv\big| \le \int_{|w|\le|b_j z|} dw \int_x^y |f'(v+w) - f'(v)|dv =: I_j(x,y;z).$$

Let $|b_j z| \le 1$ then by (5.8) $\int_x^y |f'(v+w) - f'(v)|dv \le C\mu_\gamma(x,y)|w|$ and hence $I_j(x,y;z) \le C\mu_\gamma(x,y)\int_{|w|\le|b_j z|} |w|dw \le Cb_j^2 z^2 \mu_\gamma(x,y) \le C|b_j z|^\gamma \mu_\gamma(x,y)$. Next, let $|b_j z| > 1$ then the l.h.s. of (5.11) does not exceed $|F((x,y) - b_j z) - F(x,y)| + E|F((x,y) - b_j\zeta) - F(x,y)| + f(x,y)|b_j z|$, where $f(x,y)|b_j z| \le C\mu_\gamma(x,y)|b_j z| \le C(\mu_\gamma(x,y))^{1/\gamma}|b_j z|$ follows from (5.9), $\gamma > 1$ and the boundedness of μ_γ. Next, using (5.8) and (5.7) for $|b_j z| > 1$ we obtain $|F((x,y) - b_j z) - F(x,y)| \le \int_x^y (f(u - b_j z) + f(u))du \le C((\mu_\gamma(x,y))^{1/\gamma}|b_j z| + \mu_\gamma(x,y)) \le C(\mu_\gamma(x,y))^{1/\gamma}|b_j z|$. This proves (5.11).

Consider (5.12). Since $F(x) = EF_j^*(x - b_j\zeta)$ the l.h.s. of (5.12) can be written as $|L_1(x,y;z) + L_2(x,y;z)|$, where $|L_2(x,y;z)| := |(f_j^*(x,y) - f(x,y))b_j z| \le C\mu_\gamma(x,y)|b_j|^{\alpha+1}|z| \le C\mu_\gamma(x,y)|b_j|^\gamma(1+|z|)^\gamma$ according to (5.10), and

$$|L_1(x,y;z)| := \big|E\int_{-b_j\zeta}^{-b_j z} dw \int_x^y ((f_j^*)'(v+w) - (f_j^*)'(v))dv\big| \le \bar L_1(x,y;z) + E\bar L_1(x,y;\zeta),$$

where $\bar L_1(x,y;z) := \int_0^{|b_j z|} dw \int_x^y |(f_j^*)'(v+w) - (f_j^*)'(v)|dv$. Let $|b_j z| \le 1$ then $\bar L_1(x,y;z) \le C\mu_\gamma(x,y)\int_{|w|\le|b_j z|} |w|dw \le Cb_j^2 z^2\mu_\gamma(x,y) \le C|b_j z|^\gamma\mu_\gamma(x,y)$ as in the proof of (5.11) above. Next, let $|b_j z| > 1$ then $\bar L_1(x,y;z) \le C\int_x^y dv \int_0^{|b_j z|}(g_\gamma(v+w) + g_\gamma(w))dw \le C|b_j z|^\gamma \int_x^y g_\gamma(v)dv = C\mu_\gamma(x,y)|b_j z|^\gamma$ follows from (5.8) and (5.6). Hence $\bar L_1(x,y;z) \le C\mu_\gamma(x,y)|b_j z|^\gamma$, $E\bar L_1(x,y;\zeta) \le C\mu_\gamma(x,y)|b_j|^\gamma E|\zeta|^\gamma \le C\mu_\gamma(x,y)|b_j|^\gamma$ since $\gamma < \alpha$, proving (5.12).

Next, the l.h.s. of (5.13) can be written as $|T_j(x,y;z,\xi) - ET_j(x,y;\zeta,\xi)|$, where $T_j(x,y;z,\xi) := \int dG(u)\int_0^\xi dv \int_{-b_j u}^{-b_j z} dw \int_x^y (f''(t+v+w) - f''(v+w))dt$, and hence

$$|T_j(x,y;z,\xi)| \le \bar T_j(x,y;z,\xi) + E\bar T_j(x,y;\zeta,\xi),$$

where

$$\bar T_j(x,y;z,\xi) := \int_{|v|\le|\xi|} dv \int_{|w|\le|b_j z|} dw \int_x^y |f''(t+v+w) - f''(t+v)|dt.$$

Let $|b_j z| \le 1$ then by (5.9) and (5.6), $\int_x^y |f''(t+v, t+v+w)|dt \le C|w|\mu_\gamma(x-v, y-v) \le C|w|\mu_\gamma(x,y)$ for $|v| \le |\xi| \le 1$ implying $\bar T_j(x,y;z,\xi) \le C|\xi|(b_j z)^2\mu_\gamma(x,y) \le C|\xi||b_j z|^\gamma\mu_\gamma(x,y)$. Next, let $|b_j z| > 1$ then by (5.8) and (5.6) we obtain that $\bar T_j(x,y;z,\xi) \le C\int_0^\xi dv \int_{|w|\le|b_j z|} dw \big\{\int_x^y g_\gamma(t+v+w)dt + \int_x^y g_\gamma(t+v)dt\big\} \le C|\xi||b_j z|^\gamma\mu_\gamma(x,y)$, proving (5.13).

Finally, consider (5.14). In view of (5.8), the l.h.s. of (5.14) can be bounded by $C\int_x^y d\xi \int_{|w|\le|b_{j_1}z_1|} dw_1 \int_{|w_2|\le|b_{j_2}z_2|} g_\gamma(\xi + w_1 + w_2 + z)dw_2 =: U_{j_1,j_2}(x,y;z_1,z_2,z)$. Then using repeatedly inequality (5.6) we get

$$U_{j_1,j_2}(x,y;z_1,z_2,z) \;\le\; C(|b_{j_2}z_2| \vee |b_{j_2}z_2|^\gamma)\int_x^y d\xi \int_{|w|\le|b_{j_1}z_1|} g_\gamma(\xi + w_1 + z)dw_1$$

$$\leq C(|b_{j_1}z_1| \vee |b_{j_1}z_1|^\gamma)(|b_{j_2}z_2| \vee |b_{j_2}z_2|^\gamma) \int_x^y g_\gamma(\xi + z)d\xi$$
$$\leq C(|b_{j_1}z_1| \vee |b_{j_1}z_1|^\gamma)(|b_{j_2}z_2| \vee |b_{j_2}z_2|^\gamma)(1 \vee |z|^\gamma)\mu_\gamma(x,y),$$

proving (5.14) and the lemma, too. □

Proof of Lemma 5.3. From (5.11) we get $|F(x - b_j z) - EF(x - b_j\zeta_0) + f(x)b_j z| \leq C|b_j|^r(1 + |z|)^r$ with any $1 < r < \alpha$ sufficiently close to α so that the series in (5.19) absolutely converges for any $z \in \mathbb{R}$:

$$|\eta(x;z)| \leq C(1+|z|)^r \textstyle\sum_{j=0}^\infty |b_j|^r \leq C(1+|z|)^r.$$

We shall next prove the existence of the limits

$$\lim_{z\to\pm\infty} |z|^{-\frac{1}{1-d}}\eta(x;z) \;=\; \psi^*_\pm(x), \tag{6.2}$$

with $\psi^*_\pm(x)$ given in (3.11). Note the integrals in (3.11) converge for $0 < d < 1/2$ since F is twice differentiable. To prove (6.2), let

$$\widetilde{\eta}(x;z) := \textstyle\sum_{j=0}^\infty \big(F(x - b_j z) - F(x) + f(x)b_j z\big), \tag{6.3}$$

where $|F(x - b_j z) - F(x) + f(x)b_j z| \leq \int_{|u|<|b_j z|} |f(x+u) - f(x)|du \leq Cg_r(x)|b_j z|^r$, for any $1 < r < \alpha$ follows from (5.9), (5.6), and hence the series in (6.3) converges for any $x, z \in \mathbb{R}$ and $E|\widetilde{\eta}(x;\zeta)| < C$. Since $\eta(x;z) = \widetilde{\eta}(x;z) - E\widetilde{\eta}(x;\zeta)$ it suffices to prove (6.2) for $\eta(x;z)$ replaced by $\widetilde{\eta}(x;z)$. We have for $z > 0$ that $z^{-\frac{1}{1-d}}\widetilde{\eta}(x;z) = \psi^*_+(x) + \phi(x;z)$, where

$$\phi(x;z) \;=\; \textstyle\int_0^\infty \big\{F(x - zb_{[uz^{1/(1-d)}]}) - F(x - \frac{c_0}{u^{1-d}}) + f(x)\big(zb_{[uz^{1/(1-d)}]} - \frac{c_0}{u^{1-d}}\big)\big\}du \to 0$$

as $z \to \infty$, by the dominated convergence theorem. The limit as $z \to -\infty$ in (6.2) follows analogously. Relations (6.2) and (1.5) imply tail relations $P(\eta(x,\zeta_0) > y) \sim \gamma_+(x)y^{-\alpha_*}$, $P(\eta(x,\zeta_0) < -y) \sim \gamma_-(x)|y|^{-\alpha_*}$, as $y \to \infty$ with $\gamma_\pm(x) \geq 0$ written in terms of $c_\pm, \psi^*_\pm(x)$. See Surgailis (2004, Lemma 3.2), (2002, Lemma 3.1) for details. Hence, by the classical CLT for i.i.d. r.v.'s and the Cramér-Wold device, finite-dimensional distributions of $n^{1-1/\alpha_*}\mathcal{Z}_n(x)$ converge weakly to those of $\mathcal{Z}^*$.

We shall now prove the tightness of the process $n^{1-1/\alpha_*}\mathcal{Z}_n$. By the well-known tightness criterion in (Billingsley, 1968, Thm. 15.6), it suffices to show that there exist $r > 1$ and a finite continuous measure μ such that for all $\epsilon > 0, x < y$

$$P(|\mathcal{Z}_n(x,y)| > \epsilon n^{1/\alpha_* - 1}) \leq \epsilon^{-\alpha_*}(\mu(x,y))^r, \tag{6.4}$$

where $\mathcal{Z}_n(x,y) = \mathcal{Z}_n(y) - \mathcal{Z}_n(x)$. Let $\eta(x,y;z) = \eta(y;z) - \eta(x;z), x < y$. By the definition of $\mathcal{Z}_n$ in (5.20), $P(|\mathcal{Z}_n(x,y)| > \epsilon n^{1/\alpha_*-1}) \leq nP(|\eta(x,y;\zeta)| > \epsilon n^{1/\alpha_*})$. Hence (6.4) follows from $P(|\eta(x,y;\zeta| > \epsilon n^{1/\alpha_*}) \leq (n\epsilon^{\alpha_*})^{-1}(\mu(x,y))^r$, or

$$P(|\eta(x,y;\zeta)| > w) \leq w^{-\alpha_*}(\mu(x,y))^r, \qquad \forall\, w > 0. \tag{6.5}$$

In turn, (6.5) follows from $P(|\zeta| > x) \le C/x^{\alpha}$, $\forall\, x > 0$ and the fact that there exist $1 < \gamma < \alpha, C < \infty$ such that for all $x < y, |z| > 1$, with μ_γ as in (5.5),

$$|\eta(x, y; z)| \le C|z|^{1/(1-d)}(\mu_\gamma(x, y))^{1/\gamma(1-d)}. \tag{6.6}$$

Indeed, (6.6) entails $P(|\eta(x, y; \zeta)| > w) \le P(C|\zeta|^{1/(1-d)}(\mu_\gamma(x, y))^{1/\gamma(1-d)} > w) = P(|\zeta| > Cw^{1-d}/(\mu_\gamma(x, y))^{1/\gamma}) \le Cw^{-\alpha_*}(\mu_\gamma(x, y))^{\alpha/\gamma}$ implying (6.5) with $r = \alpha/\gamma > 1$. See Surgailis (2002, proof of Lemma 3.2).

Let us show (6.6). Since $|\eta(x, y; z)| = |\widetilde{\eta}(x, y; z) - E\widetilde{\eta}(x, y; \zeta)| \le |\widetilde{\eta}(x, y; z)| + E|\widetilde{\eta}(x, y; \zeta)|$ and $E|\zeta|^{1/(1-d)} < \infty$ for $1/(1-d) < \alpha$, it suffices to prove (6.6) for $\widetilde{\eta}(x, y; z)$ instead of $\eta(x, y; z)$ as this implies $E|\widetilde{\eta}(x, y; z)| \le C(\mu_\gamma(x, y))^{1/\gamma(1-d)}E|\zeta|^{1/(1-d)} \le C(\mu_\gamma(x, y))^{1/\gamma(1-d)}$ $|z|^{1/(1-d)}$, $|z| \ge 1$. Using (5.11) with $\chi := |z|(\mu_\gamma(x, y))^{1/\gamma}$ we obtain

$$\begin{aligned}
|\widetilde{\eta}(x, y; z)| &\le C\sum_{j=0}^{\infty} \min\left\{|b_j z|(\mu_\gamma(x, y))^{1/\gamma}, |b_j z|^{\gamma}\mu_\gamma(x, y)\right\} \\
&\le C\chi^{\gamma} \sum_{j > \chi^{1/(1-d)}} j^{-(1-d)\gamma} + C\chi \sum_{0 \le j \le \chi^{1/(1-d)}} j_+^{-(1-d)} \\
&\le C(\mu_\gamma(x, y))^{1/\gamma(1-d)}|z|^{1/(1-d)},
\end{aligned}$$

proving (6.6), (6.4) and the required tightness of the process $n^{1-1/\alpha_*}\mathcal{Z}_n(x)$. Lemma 5.3 is proved. □

Proof of Lemma 5.4. We shall prove that there exist $1 < r < \alpha, \kappa > 0$ and a finite measure μ on $\mathbb{R}$ such that

$$E|\mathcal{Z}_n^*(x, y) - \mathcal{Z}_n(x, y)|^r \le \mu(x, y)n^{r(1/\alpha_* - 1) - \kappa}, \qquad \forall\, x < y. \tag{6.7}$$

Lemma 5.4 follows from (6.7) using the chaining argument as in KS and Surgailis (2002). To prove (6.7), similarly to Surgailis (2004, proof of (3.6)) decompose

$$n(\mathcal{Z}_n^*(x) - \mathcal{Z}_n(x)) = \textstyle\sum_{i=1}^4 V_{ni}(x), \tag{6.8}$$

where $V_{n1}(x) := \sum_{s \le 0} \phi_{n1,s}(x; \zeta_s)$, $V_{ni}(x) := \sum_{s=1}^n \phi_{ni,s}(x; \zeta_s), i = 2, 3, 4$ with

$$\begin{aligned}
\phi_{n1,s}(x; z) &:= \sum_{i=1}^{n} \left\{F_{i-s}^*(x + \xi_{ni} - b_{i-s}z) - F(x + \xi_{ni}) + f(x + \xi_{ni})b_{i-s}z\right\}, \\
\phi_{n2,s}(x; z) &:= -\sum_{i=n+1}^{\infty} \left\{F(x - b_{i-s}z) - EF(x - b_{i-s}\zeta_0) + f(x)b_{i-s}z\right\}, \\
\phi_{n3,s}(x; z) &:= \textstyle\sum_{i=s}^{n} \left\{F_{i-s}^*(x + \xi_{ni} - b_{i-s}z) - F(x + \xi_{ni} - b_{i-s}z)\right. \\
&\qquad\qquad \left. - F(x + \xi_{ni}) + EF(x + \xi_{ni} - b_{i-s}\zeta_0)\right\}, \\
\phi_{n4,s}(x; z) &:= \textstyle\sum_{i=s}^{n} \left\{F(x + \xi_{ni} - b_{i-s}z) - F(x - b_{i-s}z) - EF(x + \xi_{ni} - b_{i-s}\zeta_0)\right. \\
&\qquad\qquad \left. + EF(x - b_{i-s}\zeta_0) + (f(x + \xi_{ni}) - f(x))b_{i-s}(z - E\zeta_0)\right\}.
\end{aligned}$$

Note the for each $1 \le i \le 4$, $\phi_{ni,s}(x;\zeta_s)$, $s \in \mathbb{Z}$, are zero mean independent r.v.'s. It suffices to prove that for some $r_i > 1, \gamma_i > 1, \kappa_i > 0$,

$$(6.9) \qquad E|V_{ni}(x,y)|^{r_i} \ \le\ C(\mu_{\gamma_i}(x,y))^{r_i} n^{r_i/\alpha_* - \kappa_i}, \quad x<y, \quad 1 \le i \le 4.$$

Indeed, (6.9) imply (6.7) with $r = \min\{r_i, 1 \le i \le 4\} > 1, \kappa = \min\{\kappa_i r/r_i, 1 \le i \le 4\}$ and $\mu = C\mu_\gamma, \gamma = \min\{\gamma_i, 1 \le i \le 4\}$ which follows from $E|V_{ni}(x,y)|^r \le (E|V_{ni}(x,y)|^{r_i})^{r/r_i} \le C(\mu_{\gamma_i}(x,y))^r (n^{r_i/\alpha_* - \kappa_i})^{r/r_i} \le C\mu_\gamma(x,y) n^{r/\alpha_* - \kappa}$. Because of the last fact we will not indicate the subscript in r_i, γ_i, κ_i in the subsequent proof of (6.9).

Consider (6.9) for $i = 1$. Then $E|V_{n1}(x,y)|^r \le C\sum_{s\le 0} E|\phi_{n1}(x,y;\zeta_s)|^r$, $1 \le r \le 2$. Moreover, w.l.g. we can restrict the proof to $\xi_{ni} \equiv 0$. Hence, using Minkowski's inequality,

$$\begin{aligned} E|V_{n1}(x,y)|^r &\le C\sum_{s=1}^{\infty} E\Big|\sum_{j=s}^{n+s}\big\{F_j^*((x,y)-b_j\zeta_0) - F(x,y) + f(x,y)b_j\zeta_0\big\}\Big|^r \\ &\le C\sum_{s=1}^{\infty}\Big(\sum_{j=s}^{n+s} E^{1/r}\big|F_j^*((x,y)-b_j\zeta_0) - F(x,y) + f(x,y)b_j\zeta_0|^r\Big)^r. \end{aligned}$$

Whence and from (5.12) with $\gamma = \alpha'$ and $1 < r < \alpha_* < \alpha' < \alpha < \alpha r$ and α' sufficiently close to α so that $(1-d)\alpha'/r > 1$ follows from $r < \alpha_*$, we obtain

$$\begin{aligned} E|V_{n1}(x,y)|^r &\le C(\mu_\gamma(x,y))^r \sum_{s=1}^{\infty}\big(\sum_{j=s}^{n+s}|b_j|^{\alpha'/r}\big)^r \\ &\le C(\mu_\gamma(x,y))^r\Big\{\sum_{s=1}^{n}\big(\sum_{j=s}^{\infty} j^{-(1-d)\alpha'/r}\big)^r + \sum_{s>n}\big(ns^{-(1-d)\alpha'/r}\big)^r\Big\} \\ &\le C(\mu_\gamma(x,y))^r n^{1+r-\alpha'(1-d)}, \end{aligned}$$

implying (6.9) for $i = 1$ since $1 + r - \alpha'(1-d) < r/\alpha_*$ follows by noting that for $\alpha' = \alpha$ the last inequality becomes $r(1 - 1/\alpha_*) < \alpha_* - 1$, or $r < \alpha_*$, and hence it holds also for all $\alpha' < \alpha$ sufficiently close to α.

Next, consider (6.9) for $i = 2$. Use (5.11) with $\gamma = \alpha'$ and r, α' as above to obtain

$$\begin{aligned} &E|V_{n2}(x,y)|^r \\ &\le C(\mu_\gamma(x,y))^r\sum_{s=1}^{n}\Big(\sum_{j=s}^{\infty} E^{1/r}\big|F((x,y)-b_j\zeta_0) - EF((x,y)-b_j\zeta_0) + f(x,y)b_j\zeta_0|^r\Big)^r \\ &\le C(\mu_\gamma(x,y))^r\sum_{s=1}^{n}\big(\sum_{j=s}^{\infty} j^{-(1-d)\alpha'/r}\big)^r \ \le\ C(\mu_\gamma(x,y))^r n^{1+r-\alpha'(1-d)}, \end{aligned}$$

proving (6.9) for $i = 2$.

To prove (6.9) for $i = 3$, note that (5.13) implies

$$
\begin{aligned}
|F_j^*((x,y)-b_jz+\xi_{ni})-F((x,y)-b_jz+\xi_{ni})| &\le C|b_j|^\alpha \int_x^y (1+|u-b_jz+\xi_{ni}|)^{-\gamma}du \\
&\le C|b_j|^\alpha \mu_\gamma(x,y)(1+|b_jz|)^\gamma
\end{aligned}
$$

and hence $E|F_j^*((x,y)-b_jz+\xi_{ni})-F((x,y)-b_jz+\xi_{ni})|^r \le C|b_j|^{\alpha r}(\mu_\gamma(x,y))^r(1+E|\zeta|^{\gamma r}) \le C(\mu_\gamma(x,y))^r|b_j|^{\alpha r}$ for $r,\gamma > 1$ satisfying $r\gamma < \alpha$. Moreover, we may choose $\alpha_* < r < \alpha$. Then, since $\sum_{j=0}^\infty |b_j|^\alpha < \infty$, we obtain

$$
E|V_{n3}(x,y)|^r \le C(\mu_\gamma(x,y))^r \textstyle\sum_{s=1}^n \left(\sum_{j=1}^\infty |b_j|^\alpha\right)^r \le C(\mu_\gamma(x,y))^r n,
$$

proving (6.9) for $i=3$, because $r/\alpha_* > 1$.

Consider (6.9) for $i=4$. Note $\phi_{n4,s}(x,y;z)$ can be rewritten as

$$
\sum_{i=s}^n \Big\{ \int dG(u) \int_0^{\xi_{ni}} dv \big(f((x,y)+v-b_{i-s}z)-f((x,y)+v-b_{i-s}u)+f'((x,y)+v)b_{i-s}(z-u)\big)\Big\}.
$$

Hence by (5.13) with $\gamma = 1/(1-d) \in (1,\alpha)$ and $1 < r < \alpha_*$ we obtain

$$
\begin{aligned}
E|\phi_{n4,s}(x,y;\zeta)|^r &\le C\Big(\textstyle\sum_{i=s}^n |\xi_{ni}|\mu_\gamma(x,y)|b_{i-s}|^\gamma E^{1/r}(1+|\zeta|)^{\gamma r}\Big)^r \\
&\le C(\mu_\gamma(x,y))^r \max_{1\le i\le n}|\xi_{ni}|^r \Big(\textstyle\sum_{i=s}^n |b_{i-s}|^{1/(1-d)}\Big)^r,
\end{aligned}
$$

by noting that $E|\zeta|^{\gamma r} < \infty$ since $\gamma r = r/(1-d) < \alpha$ is equivalent to $r < \alpha_*$. Note also that for the above choice $|b_j|^\gamma = O(j^{-1})$. Then

$$
\begin{aligned}
E|V_{n4}(x,y)|^r &\le C(\mu_\gamma(x,y))^r \max_{1\le i\le n}|\xi_{ni}|^r \textstyle\sum_{s=1}^n \left(\sum_{j=1}^n j^{-1}\right)^r \\
&\le C(\mu_\gamma(x,y))^r \max_{1\le i\le n}|\xi_{ni}|^r n(\log n)^r.
\end{aligned}
$$

Thus, (6.9) for $i=4$ holds due to condition (3.13) since one can choose $r < \alpha_*$ and $\kappa > 0$ so that $r/\alpha_* - \kappa$ is arbitrary close to 1. This ends the proof of (6.9) and that of Lemma 5.4. □

Proof of Lemma 5.5. Rewrite

$$
\begin{aligned}
(6.10)\qquad \mathcal{R}_n(x) &= \mathcal{V}_n^{(\xi)}(x) - \mathcal{Z}_n^*(x) = n^{-1}\sum_{i=1}^n h_{ni}(x), \qquad \text{where} \\
h_{ni}(x) &:= I(\varepsilon_i \le x+\xi_{ni}) - F(x+\xi_{ni}) + f(x+\xi_{ni})\varepsilon_i \\
&\qquad - \sum_{s\le i} E\big[I(\varepsilon_i \le x+\xi_{ni}) - F(x+\xi_{ni}) + f(x+\xi_{ni})\varepsilon_i \big| \zeta_s\big]
\end{aligned}
$$

Similarly as in the proof of the previous lemma, it suffices to prove that there exist $1 < r < \alpha, \kappa > 0$ and a finite measure μ on $\mathbb{R}$ such that for all $n \ge 1$ and $x < y$

$$
(6.11)\qquad E\Big|\sum_{i=1}^n h_{ni}(x,y)\Big|^r \le \mu(x,y)n^{(r/\alpha_*)-\kappa}.
$$

Recall the notation $\varepsilon_{i,j}, \widetilde{\varepsilon}_{i,j}, \varepsilon^*_{i,j}$, in (5.3) and $F_j(x) := P(\varepsilon_{i,j} \le x), \widetilde{F}_j(x) := P(\widetilde{\varepsilon}_{i,j} \le x), F^*_j(x) := P(\varepsilon^*_{i,j} \le x)$. Then

$$
\begin{aligned}
h_{ni}(x) &= \sum_{s\le i} U_{ni,s}(x), \qquad \text{where} \\
U_{ni,s}(x) &:= F_{i-s-1}(x+\xi_{ni} - b_{i-s}\zeta_s - \widetilde{\varepsilon}_{i,i-s}) - F_{i-s}(x+\xi_{ni} - \widetilde{\varepsilon}_{i,t-s}) \\
&\quad - \widehat{F}_{i-s}(x+\xi_{ni} - b_{i-s}\zeta_s) + F(x+\xi_{ni})
\end{aligned} \tag{6.12}
$$

with the convention $F_{-1}(x) := I(x \ge 0)$. Then $\sum_{i=1}^n h_{ni}(x,y) = \sum_{s\le n} M_{ns}(x,y)$, where

$$
M_{ns}(x,y) := \sum_{i=1\vee s}^{n} U_{ni,s}(x,y), \quad s \le n \tag{6.13}
$$

is a martingale difference sequence with $E[M_{ns}(x,y)|\zeta_u, u\le s] = 0$. Hence by the well-known martingale moment inequality, see e.g. (GKS, Lemma 2.5.2),

$$
E\big|\sum_{i=1}^n h_{ni}(x,y)\big|^r \le C\sum_{s\le n} E|M_{ns}(x,y)|^r \ \le\ C\sum_{s\le n}\Big(\sum_{i=1\vee s}^{n} E^{1/r}|U_{ni,s}(x,y)|^r\Big)^r. \tag{6.14}
$$

To proceed further, we need a 'good' bound of $E|U_{ni,s}(x,y)|^r$. Towards this end, noting that $E[U_{ni,s}(x,y)|\zeta_s] = 0$, we expand $U_{ni,s}(x,y)$ similarly as in (6.12) but according to the *conditional* probability $P[\cdot|\zeta_s]$, see Surgailis (2004, (4.8)):

$$
\begin{aligned}
U_{ni,s}(x,y) &= \sum_{u<s}\big(E[U_{ns}(x,y)|\zeta_s,\zeta_v, v\le u] - E[U_{ns}(x,y)|\zeta_s,\zeta_v, v\le u-1]\big) \\
&= \sum_{u<s} W_{n,i,s,u}(x,y),
\end{aligned}
$$

where

$$
\begin{aligned}
&W_{n,i,s,u}(x) \\
&:= F_{\ne i-s, i-u-1}(x+\xi_{ni} - b_{i-s}\zeta_s - b_{i-u}\zeta_u - \widetilde{\varepsilon}_{i,i-u}) - F_{\ne i-s,i-u}(x+\xi_{ni} - b_{i-s}\zeta_s - \widetilde{\varepsilon}_{i,t-u}) \\
&\quad - F_{i-s-1}(x+\xi_{ni} - b_{i-u}\zeta_u - \widetilde{\varepsilon}_{i,i-u}) + F_{i-s}(x+\xi_{ni} - \widetilde{\varepsilon}_{i,i-u}).
\end{aligned}
$$

Similarly to Surgailis (2004, (4.12)) $W_{n,i,s,u}(x,y)$ can be rewritten as

$$
\begin{aligned}
W_{n,i,s,u}(x,y) &= E^0\big[H(b_{i-s}\zeta + b_{i-u}\eta + \widetilde{\varepsilon}) - H(b_{i-s}\zeta^0 + b_{i-s}\eta + \widetilde{\varepsilon}) \\
&\qquad - H(b_{i-s}\zeta + b_{i-s}\eta^0 + \widetilde{\varepsilon}) + H(b_{i-s}\zeta^0 + b_{i-s}\eta^0 + \widetilde{\varepsilon})\big] \\
&= E^0\Big[\int_{b_{i-s}\zeta^0}^{b_{i-s}\zeta}\int_{b_{i-u}\eta^0}^{b_{i-u}\eta} H^{(2)}(z_1+z_2+\widetilde{\varepsilon})dz_1dz_2\Big],
\end{aligned}
$$

where, for fixed $n, u<s<i, x<y$,

$$
H(z) := F_{\ne i-s,i-u-1}((x,y)+\xi_{ni} - z), \quad \zeta := \zeta_s, \quad \eta := \zeta_u, \quad \widetilde{\varepsilon} := \widetilde{\varepsilon}_{i,i-u},
$$

(ζ^0, η^0) is an independent copy of (ζ, η) and E^0 denotes the expectation w.r.t. (ζ^0, η^0) only; $H^{(2)}(z) = d^2 H(z)/dz^2$. From (5.14) for any $1 < \gamma < \alpha$ we get

$$|W_{n,i,s,u}(x,y)| \;\le\; C\mu_\gamma(x,y)|b_{i-s}|\,|b_{i-u}|(|\zeta|^\gamma + 1)(|\eta|^\gamma + 1), \tag{6.15}$$

with probability 1 and nonrandom $C < \infty$, and hence for any $1 < r < \alpha, 1 < \gamma < \alpha, r\gamma < \alpha$

$$E[|W_{n,i,s,u}(x,y)|^r|\zeta_s] \;\le\; C(\mu_\gamma(x,y))^r|b_{i-s}|^r|b_{i-u}|^r(|\zeta_s| + 1)^{r\gamma}.$$

Then using the above mentioned martingale moment inequality w.r.t. $P[\cdot|\zeta_s]$ we obtain

$$\begin{aligned}
E[|U_{ni,s}(x,y)|^r|\zeta_s] &\le C\sum_{u<s} E[|W_{n,i,s,u}(x,y)|^r|\zeta_s] \\
&\le C(\mu_\gamma(x,y))^r|b_{i-s}|^r(|\zeta_s|+1)^{r\gamma}\sum_{u<s}(i-u)^{-r(1-d)} \\
&\le C(\mu_\gamma(x,y))^r(i-s)^{1-2r(1-d)}(|\zeta_s|+1)^{r\gamma}.
\end{aligned}$$

Hence

$$E^{1/r}|U_{ni,s}(x,y)|^r \le C\mu_\gamma(x,y)(i-s)^{1/r-2(1-d)}. \tag{6.16}$$

Substituting (6.16) into (6.14) we obtain

$$\begin{aligned}
E\Big|\sum_{i=1}^n h_{ni}(x,y)\Big|^r &\le C\mu_\gamma(x,y)\Big\{\sum_{s\le -n}\Big(\sum_{i=1}^n (i-s)^{1/r-2(1-d)}\Big)^r + n\Big(\sum_{i=1}^n i^{1/r-2(1-d)}\Big)^r\Big\} \\
&\le C\mu_\gamma(x,y)n^{2+r-2r(1-d)},
\end{aligned} \tag{6.17}$$

for $r < \alpha$ sufficiently close to α. (Note that for any $r < \alpha$ we can find $\gamma > 1$ such that $r\gamma < 1$.) Also note for such r

$$2 + r - 2r(1-d) < r/\alpha_* = r/\alpha(1-d) \tag{6.18}$$

Indeed, (6.18) holds for $r = \alpha$ since $2 + \alpha - 2\alpha(1-d) < 1/(1-d)$ is equivalent to $(1-2d)(\alpha(1-d)-1) > 0$, and therefore (6.18) holds for $r < \alpha$ sufficiently close to α by continuity. Relations (6.17) and (6.18) prove (6.11) and thereby complete the proof of Lemma 5.5. $\square$

7 References

Astrauskas, A. (1983). Limit theorems for sums of linearly generated random variables. *Lithuanian Math. J.*, **23**, 127–134.

Avram, F. and Taqqu, M.S. (1986). Weak convergence of moving averages with infinite variance. In: Eberlein, E. and M.S. Taqqu (eds), *Dependence in Probability and Statistics*, pp. 399-415. Birkhäuser, Boston.

Avram, F. and Taqqu, M.S. (1992). Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.*, **20**, 483–503.

Basse-O'Connor, A., Lachièze-Rey, R. and Podolskij, M. (2016). Power variation for a class of stationary increments Lévy driven moving averages. Preprint.

Christoph, G. and Wolf, W. (1992). *Convergence Theorems with Stable Limit Law.* Akademie-Verlag, Berlin.

Dehling, H. and Taqqu, M.S. (1989). The empirical process of some long range dependent sequences with an application to U-statistics. *Ann. Statist.*, **17**, 1767–1783.

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

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

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.

Huber, P.J. (1981). *Robust Statistics.* Wiley, New York.

Hult, H. and Samorodnitsky, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli*, **14**, 838–864.

Ibragimov, I.A. and Linnik, Yu.V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.

Kasahara, Y. and Maejima, M. (1988). Weighted sums of i.i.d. random variables attracted to integrals of stable processes. *Probab. Th. Rel. Fields*, **78**, 75–96.

Kokoszka, P.S. and Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stoch. Proc. Appl.* **60**, 19–47.

Koul, H.L. (2002). Asymptotic distributions of some scale estimators in nonlinear models. *Metrika*, **55**, 75–90.

Koul, H.L. (2002a). *Weighted Empirical Processes in Dynamic Nonlinear Models. 2nd Edition.* **Lecture Notes Series in Statistics**, **166**, Springer, New York, N.Y., USA.

Koul, H.L. and Surgailis, D. (2001). Asymptotics of empirical processes of long memory moving averages with infinite variance. *Stochastic Process. Appl.*, **91**, 309–336.

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

Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

Surgailis, D. (2002). Stable limits of empirical processes of long memory moving averages with infinite variance. *Stochastic Process. Appl.*, **100**, 255–274.

Surgailis, D. (2004). Stable limits of sums of bounded functions of long memory moving averages with finite variance. *Bernoulli*, **10**, 327–355.