Original Article

Journal of Time Series Analysis

First version received December 2009

Published online in Wiley Online Library: 27 September 2010

(wileyonlinelibrary.com) DOI: 10.1111/j.1467-9892.2010.00688.x

Robust estimation of the scale and of the autocovariance function of Gaussian short- and long-range dependent processes

Céline Lévy-Leduc[a,*,†], Hélène Boistard[b], Eric Moulines[c], Murad S. Taqqu[d] and Valderio A. Reisen[e]

A desirable property of an autocovariance estimator is to be robust to the presence of additive outliers. It is well known that the sample autocovariance, being based on moments, does not have this property. Hence, the use of an autocovariance estimator which is robust to additive outliers can be very useful for time-series modelling. In this article, the asymptotic properties of the robust scale and autocovariance estimators proposed by Rousseeuw and Croux (1993) and Ma and Genton (2000) are established for Gaussian processes, with either short- or long-range dependence. It is shown in the short-range dependence setting that this robust estimator is asymptotically normal at the rate $\sqrt{n}$, where n is the number of observations. An explicit expression of the asymptotic variance is also given and compared with the asymptotic variance of the classical autocovariance estimator. In the long-range dependence setting, the limiting distribution displays the same behaviour as that of the classical autocovariance estimator, with a Gaussian limit and rate $\sqrt{n}$ when the Hurst parameter H is less than 3/4 and with a non-Gaussian limit (belonging to the second Wiener chaos) with rate depending on the Hurst parameter when $H \in (3/4,1)$. Some Monte Carlo experiments are presented to illustrate our claims and the Nile River data are analysed as an application. The theoretical results and the empirical evidence strongly suggest using the robust estimators as an alternative to estimate the dependence structure of Gaussian processes.

Keywords: Autocovariance function; long-memory; robustness; influence function; scale estimator; Hadamard differentiability; functional Delta method.

1. INTRODUCTION

The autocovariance function of a stationary process plays a key role in time-series analysis. However, it is well known that the classical sample autocovariance function is very sensitive to the presence of additive outliers in the data. A small fraction of additive outliers, in some cases even a single outlier, can affect the classical autocovariance estimate making it virtually useless; see, for instance Deutsch *et al.* (1990) Chan (1992, 1995), Maronna *et al.* (2006, chap. 8) and the references therein. Since additive outliers are quite common in practice, the definition of an autocovariance estimator which is robust to the presence of additive outliers is an important task.

Ma and Genton (2000) proposed a robust estimator of the autocovariance function and discussed its performance on synthetic and real data sets. This estimator has later been used by Fajardo *et al.* (2009) to derive robust estimators for autoregressive moving average (ARMA) and autoregressive fractionally integrated moving average (ARFIMA) models. The autocovariance estimator proposed by Ma and Genton (2000) is based on a method due to Gnanadesikan and Kettenring (1972), which consists in estimating the covariance of the random variables X and X' by comparing the scale of two appropriately chosen linear combinations of these variables; more precisely, if a and b are any non-zero constants, then

$$\mathrm{cov}(X,X') = \frac{1}{4ab}\{\mathrm{var}(aX+bX') - \mathrm{var}(aX-bX')\}. \tag{1}$$

Assume that S is a robust scale functional; we write for short $S(X) = S(F_X)$, where F_X is the cumulative distribution function (c.d.f.) of X and assume that S is *affine equivariant* in the sense that $S(aX + b) = |a|S(X)$. Following Huber (1981), if we replace in the above expression var(·) by $S^2(\cdot)$, then (1) is turned into the definition of a robust alternative to the covariance

$$C_S(X,X') = \frac{1}{4ab}\left\{S^2(aX+bX') - S^2(aX-bX')\right\}. \tag{2}$$

[a]CNRS/LTCI/Telecom ParisTech
[b]Université Toulouse
[c]Institut Telecom/Telecom ParisTech
[d]Boston University
[e]Universidade Federal do Espírito Santo
*Correspondence to: CNRS/LTCI/Telecom ParisTech - 46, Rue Barrault, 75634 Paris Cédex 13, France.
†E-mail: celine.levy-leduc@telecom-paristech.fr

© 2010 Blackwell Publishing Ltd.

As explained in Huber (1981), if S is standardized such that $S(X) = 1$ in the case where X is standard Gaussian, then, provided that (X, X') is bivariate normal,

$$C_S(X, X') = \mathrm{cov}(X, X'). \tag{3}$$

Ma and Genton (2000) suggested to use for S the robust-scale estimator introduced in Rousseeuw and Croux (1993). This scale estimator is based on the Grassberger–Procaccia correlation integral, defined as:

$$r \mapsto U(r, F_X) = \int\int \mathbb{1}_{\{|x-x'|\leq r\}}\, \mathrm{d}F_X(x)\, \mathrm{d}F_X(x'), \tag{4}$$

which measures the probability that two independent copies X and X' distributed according to F_X fall at a distance smaller than r. The robust-scale estimator introduced in Rousseeuw and Croux (1993, p. 1277), defines the scale $Q(F_X)$ of a c.d.f. F_X as being proportional to the first quartile of $r \mapsto U(r,F_X)$, namely,

$$Q(F_X) = c(F_X)\inf\{r \geq 0, U(r, F_X) \geq 1/4\}, \tag{5}$$

where $c(F_X)$ is a constant depending only on the shape of the c.d.f. F_X. The constant $c(F_X)$ in (5) is there to ensure consistency. In the sequel, the c.d.f. F_X is assumed to belong to the Gaussian location-scale family

$$\{\Phi_{\mu,\sigma}(\cdot) = \Phi((\cdot - \mu)/\sigma),\ \mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*\}, \tag{6}$$

where Φ is the c.d.f. of a standard Gaussian random variable. The reason we focus on the Gaussian family is that if we want to use Q as the scale S in (2), we will need to compute $c(F_{aX+bX'})$ and $c(F_{aX-bX'})$. This is easily done when (X, X') is a Gaussian vector. Indeed, in view of (3), one has

$$\mathrm{cov}(X, X') = \frac{1}{4}\left[Q^2(F_{X+X'}) - Q^2(F_{X-X'})\right], \tag{7}$$

and in particular, since by (4) and (5), $Q^2(F_{2X}) = (2Q(F_X))^2$,

$$\mathrm{var}(X) = Q^2(F_X). \tag{8}$$

When $F_X = \Phi_{\mu,\sigma}$ we can then obtain the constant $c(\Phi_{\mu,\sigma})$ in (5) explicitly as noted by Rousseeuw and Croux (1993). Since $Q(\Phi_{\mu,\sigma}) = \sigma$, (5) becomes

$$\sigma = Q(\Phi_{\mu,\sigma}) = c(\Phi_{\mu,\sigma})\sigma r_0, \tag{9}$$

where r_0 is such that, in (4), $U(r_0,\Phi) = 1/4$. Hence for all $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$,

$$c(\Phi_{\mu,\sigma}) = c(\Phi) = 1/r_0 = 1/(\sqrt{2}\Phi^{-1}(5/8)) = 2.21914. \tag{10}$$

Let $(X_i)_{i\geq 1}$ be a stationary Gaussian process. Given the observations $X_{1:n} = (X_1, \ldots, X_n)$, the c.d.f. of the observations may be estimated using the empirical c.d.f. $r \mapsto F_n(r) = n^{-1}\sum_{i=1}^n \mathbb{1}_{\{X_i \leq r\}}$. Plugging F_n into (5) leads to the following robust-scale estimator

$$\mathrm{Q}_n(X_{1:n}, \Phi) = c(\Phi)\{|X_i - X_j|; 1 \leq i, j \leq n\}_{(k_n)}, \tag{11}$$

where $k_n = \lfloor n^2/4 \rfloor$. That is, up to the multiplicative constant $c(\Phi)$, $\mathrm{Q}_n(X_{1:n},\Phi)$ is the k_nth order statistics of the n^2 distances $|X_i - X_j|$ between all the pairs of observations.

Using the robust-scale estimator $\mathrm{Q}_n(\cdot,\Phi)$ in (11) and the identity (2) with $a = b = 1$, the robust autocovariance estimator of

$$\gamma(h) = \mathrm{cov}(X_1, X_{h+1}) = \frac{1}{4}\{\mathrm{var}(X_1 + X_{h+1}) - \mathrm{var}(X_1 - X_{h+1})\}$$

is

$$\hat{\gamma}_Q(h, X_{1:n}, \Phi) = \frac{1}{4}\left\{\mathrm{Q}_{n-h}^2(X_{1:n-h} + X_{h+1:n}, \Phi) - \mathrm{Q}_{n-h}^2(X_{1:n-h} - X_{h+1:n}, \Phi)\right\}. \tag{12}$$

Thus, in the sample version (12), the random variable $X_1 \pm X_{h+1}$ is replaced by the vector $X_{1:n-h} \pm X_{h+1:n}$ of length $n - h$.

In this article, we establish the asymptotic properties of $Q_n(X_{1:n},\Phi)$ and the corresponding robust autocovariance estimator $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ for Gaussian processes displaying both short- and long-range dependence. We say that the process is short-range dependent if the autocovariance function $\{\gamma(h)\}_{h\in\mathbb{Z}}$ is absolutely summable, $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$. We say that it is long-range dependent if the autocovariance function is regularly varying at infinity with exponent D, $\gamma(h) = h^{-D}L(h)$ with $0 < D < 1$ and L is a slowly varying function, that is, $\lim_{k\to\infty} L(ak)/L(k) = 1$, for any $a > 0$, and is positive for large enough k. The exponent D is related to the so-called *Hurst coefficient* by the relation $H = 1 - D/2$. See, for more details, Doukhan *et al.* (2003, pp. 5–38). The limiting distributions of these estimators are obtained by using the functional delta method; see van der Vaart (1998).

In the short memory case, the results stem directly from the weak invariance principle satisfied by the empirical process F_n under mild technical assumptions. The rate of convergence of the robust covariance estimator is $\sqrt{n}$ and the limiting distribution is Gaussian; an explicit expression of the asymptotic variance is given in Theorem 2.

In the long memory case, the situation is more involved. When $D \geq 1/2$ (or $H \leq 3/4$), the rate of convergence is still $\sqrt{n}$, the limiting distribution is Gaussian and the asymptotic variance of the covariance estimator is the same as in the short-memory case.

© 2010 Blackwell Publishing Ltd.

When $0 < D < 1/2$, the rate of convergence becomes equal to $n^D/\tilde{L}(n)$, where $\tilde{L}$ is a slowly varying function defined in (37); the limiting distribution is non-Gaussian and belongs to the second Wiener Chaos; see Theorem 4. We prove that these rates are identical to the ones of the classical autocovariance estimators. The study of the asymptotic distribution of the empirical process is not enough to derive these results. It is necessary to use results on the empirical version of the correlation integral which requires extensions of the results derived for U-processes under short-range dependence conditions by Borovkova *et al.* (2001). For this part, we use novel results on U-processes of long memory time series that are developed in a companion paper by Lévy-Leduc *et al.* (2009).

The outline of the article is as follows. In Section 2, the limiting distributions of the robust-scale estimator $Q_n(X_{1:n}, \Phi)$ in the Gaussian short- and long-range dependence settings are proved. From these results, the asymptotic distribution of $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ is derived. In Section 3, some Monte Carlo experiments are presented so as to support our theoretical claims. The Nile River data are studied as an application in Section 4. Section 5 is dedicated to the asymptotic properties of U-processes which are useful to establish the results of Section 2 in the long-range case. Sections 6 and 7 detail the proofs of the theoretical results stated in Section 2. Some concluding remarks are provided in Section 8.

Notation. For an interval I in the extended real line $[-\infty, \infty]$, we denote by $\mathrm{D}(I)$ the set of all functions $z : I \to \mathbb{R}$ that are right-continuous and whose limits from the left exist everywhere on I. We always equip $\mathrm{D}(I)$ with the uniform norm, denoted by $\|\cdot\|_\infty$. We denote by $\mathcal{M}([-\infty, \infty])$ the set of c.d.f. on $[-\infty, \infty]$ equipped with the topology of uniform convergence. For $U \in I$, let U^{-1} denote its generalized inverse, $U^{-1}(\eta) = \inf\{r \in I, U(r) \geq \eta\}$. The convergence in distribution in $(\mathrm{D}([0,\infty]), \|\cdot\|_\infty)$ is meant with respect to the σ-algebra generated by the set of open balls. We denote by $\longrightarrow_d$ the convergence in distribution and by ϕ the probability density function of the standard Gaussian random variable.

2. THEORETICAL RESULTS

Define the following mappings:

$$\begin{aligned} T_1 : \mathcal{M}([-\infty, \infty]) &\to \mathrm{D}[0, \infty] \\ F &\mapsto \left\{ r \mapsto \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{\{|x-y| \leq r\}} \, \mathrm{d}F(x) \, \mathrm{d}F(y) \right\}, \end{aligned} \tag{13}$$

$$\begin{aligned} T_2 : [0, \infty] &\to \mathbb{R} \\ U &\mapsto U^{-1}(1/4) \end{aligned} \tag{14}$$

and

$$T_0 = T_2 \circ T_1 : \mathcal{M}([-\infty, \infty])) \to \mathbb{R} \tag{15}$$

$$F \mapsto U^{-1}(1/4). \tag{16}$$

Then, the scale estimator $\mathrm{Q}_n(X_{1:n}, \Phi)$ introduced in (11) may be expressed as:

$$\mathrm{Q}_n(X_{1:n}, \Phi) = c(\Phi) T_0(F_n), \tag{17}$$

where F_n is the empirical c.d.f. based on $X_{1:n}$.

2.1. Short-range dependence setting

2.1.1. Properties of the scale estimator

The following lemma gives an asymptotic expansion for $\mathrm{Q}_n(X_{1:n}, \Phi)$, which is used for deriving a central limit theorem (Theorem 1). It supposes that the empirical c.d.f. F_n, adequately normalized, converges. Let us first define the influence function. Following Huber (1981, p. 13), the influence function $x \mapsto \mathrm{IF}(x, T, F)$ is defined for a functional T at a distribution F at point x as the limit

$$\mathrm{IF}(x, T, F) = \lim_{\varepsilon \to 0+} \varepsilon^{-1} \{T(F + \varepsilon(\delta_x - F)) - T(F)\},$$

where δ_x is the Dirac distribution at x. Influence functions are a classical tool in robust statistics used to understand the effect of a small contamination at the point x on the estimator.

Lemma 1. *Let $(X_i)_{i \geq 1}$ be a stationary Gaussian process of mean μ and variance σ^2. Assume that there exists a non-decreasing sequence (a_n) such that $a_n(F_n - \Phi_{\mu,\sigma})$ converges weakly in $(D([0,\infty]), \|\cdot\|_\infty)$. Then, $Q_n(X_{1:n}, \Phi)$ defined by* (11) *has the following asymptotic expansion:*

$$a_n(\mathrm{Q}_n(\mathrm{X}_{1:n}, \Phi) - \sigma) = \frac{\mathrm{a}_n}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{IF}(\mathrm{X}_\mathrm{i}, \mathrm{Q}, \Phi_{\mu,\sigma}) + \mathrm{o}_\mathrm{P}(1), \tag{18}$$

where, for all x in $\mathbb{R}$,

© 2010 Blackwell Publishing Ltd.

$$\mathrm{IF}(x, Q, \Phi_{\mu,\sigma}) = \sigma \mathrm{IF}((x-\mu)/\sigma, Q, \Phi), \tag{19}$$

and

$$\mathrm{IF}(x, Q, \Phi) = c(\Phi)\left(\frac{1/4 - \Phi(x + 1/c(\Phi)) + \Phi(x - 1/c(\Phi))}{\int_{\mathbb{R}} \phi(y)\phi(y + 1/c(\Phi))\,\mathrm{d}y}\right). \tag{20}$$

The proof of Lemma 1 is given in Section 6.

Remark 1. Note that IF(x, Q, Φ) has the same expression as the influence function of the functional Q evaluated at the c.d.f. Φ given by Rousseeuw and Croux (1993, p. 1277), and Ma and Genton (2000, p. 675).

We focus here on the case where the process $(X_i)_{i\geq 1}$ satisfies the following assumption: (**A1**) $(X_i)_{i\geq 1}$ is a stationary mean-zero Gaussian process with autocovariance sequence $\gamma(h) = \mathbb{E}(X_1 X_{h+1})$ satisfying:

$$\sum_{h\geq 1} |\gamma(h)| < \infty.$$

To state the results, we must first define the Hermite rank of the influence function $x \mapsto \mathrm{IF}(x,Q,\Phi)$. Let $\{H_k\}$ denote the Hermite polynomials having leading coefficient equal to 1. These are $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2-1,\ldots$. Let f be a function such that $\int f^2(z)\,\mathrm{d}\Phi(z) < \infty$. The expansion of f in Hermite polynomials is given by

$$f(z) = \sum_{q=\tau(f)}^{\infty} \frac{\alpha_q(f)}{q!} H_q(z), \tag{21}$$

where $\alpha_q(f) = \int f(z)H_q(z)\,\mathrm{d}\Phi(z)$ and where the convergence is in $L^2(\mathbb{R}, \Phi)$. The index of the first non-zero coefficient in the expansion, denoted $\tau(f)$, is called the Hermite rank of the function f. Theorem 1 in Breuer and Major (1983) shows that if

$$\sum_{h=-\infty}^{\infty} |\gamma(h)|^{\tau(f)} < \infty, \tag{22}$$

then the variance $\mathrm{var}\left(n^{-1/2}\sum_{i=1}^{n} f(X_i)\right)$ converges as n goes to infinity to a limiting value $\sigma^2(f)$, which is given by

$$\sigma^2(f) = \mathrm{var}[f(X_1)] + 2\sum_{h=1}^{\infty} \mathrm{cov}[f(X_{h+1}), f(X_1)] = \sum_{q=\tau}^{\infty} \frac{\alpha_q^2(f)}{q!}\left\{\gamma^q(0) + 2\sum_{h=1}^{\infty}\gamma^q(h)\right\}. \tag{23}$$

In addition, the renormalized partial sum is asymptotically Gaussian,

$$n^{-1/2}\sum_{i=1}^{n} f(X_i) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(f)\right). \tag{24}$$

Concerning the empirical process, Csörgő and Mielniczuk (1996) proved that if

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \tag{25}$$

then $\sqrt{n}(F_n(\cdot) - \Phi_{0,\sigma}(\cdot))$ converges in D([$-\infty$, ∞]) to a mean-zero Gaussian process $W(\cdot)$ with covariance

$$\mathbb{E}(W(r)W(r')) = \sum_{q=1}^{\infty} \frac{J_q(r)J_q(r')}{q!}\left\{\gamma^q(0) + 2\sum_{h=1}^{\infty}\gamma^q(h)\right\},$$

where $J_q(r) = \int [\mathbb{1}_{\{\sigma x \leq r\}} - \Phi_{0,\sigma}(r)]H_q(x)\,\mathrm{d}\Phi(x)$ for all r in $[-\infty, \infty]$. These results are used to prove Theorem 1 in Section 6.

Theorem 1. *Under Assumption (A1), $Q_n(X_{1:n}, \Phi)$ defined by (11), satisfies the following central limit theorem:*

$$\sqrt{n}(\mathrm{Q}_n(X_{1:n}, \Phi) - \sigma) \xrightarrow{d} \mathcal{N}(0, \tilde{\sigma}^2),$$

where $\sigma = \sqrt{\gamma(0)}$ and the limiting variance $\tilde{\sigma}^2$ is given by

$$\tilde{\sigma}^2 = \gamma(0)\mathbb{E}[\mathrm{IF}^2(X_1/\sigma, Q, \Phi)] + 2\gamma(0)\sum_{k\geq 1}\mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)\mathrm{IF}(X_{k+1}/\sigma, Q, \Phi)], \tag{26}$$

IF($\cdot$, Q, Φ) being defined in (20).

It is interesting to compare, under Assumption (A1), the asymptotic distribution of the proposed estimator $\mathrm{Q}_n(X_{1:n},\Phi)$ with that of the square root of the sample variance

© 2010 Blackwell Publishing Ltd.

$$\hat{\sigma}^2_{n,X} = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2 = \frac{1}{2n(n-1)}\sum_{1\le i\ne j\le n}(X_i - X_j)^2, \tag{27}$$

where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$.

Proposition 1. *Under Assumption (A1),*

$$\sqrt{n}\left(\hat{\sigma}_{n,X} - \sigma\right) \xrightarrow{d} \mathcal{N}(0, \tilde{\sigma}^2_{cl}),$$

where

$$\tilde{\sigma}^2_{cl} = (2\gamma(0))^{-1}\left(\gamma(0)^2 + 2\sum_{h\ge 1}\gamma(h)^2\right). \tag{28}$$

The relative asymptotic efficiency $\tilde{\sigma}^2_{cl}/\tilde{\sigma}^2$ of the estimator $Q_n(X_{1:n},\Phi)$ compared to $\hat{\sigma}_{n,X}$ is larger than 82.27%.

The index '*cl*' stands for 'classical'. The proof of Proposition 1 is given in Section 6.

2.1.2. Properties of the autocovariance estimator

We establish here the limiting behaviour of the autocovariance estimator given, for $0 \le h < n$, by (12).

Theorem 2. *Assume that (A1) holds and let h be a non-negative integer. Then, the autocovariance estimator $\hat{\gamma}_Q(\mathrm{h}, \mathrm{X}_{1:\mathrm{n}}, \Phi)$ satisfies the following central limit theorem:*

$$\sqrt{n}\left(\hat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)\right) \xrightarrow{d} \mathcal{N}(0, \breve{\sigma}^2_h),$$

where

$$\breve{\sigma}^2(h) = \mathbb{E}[\psi^2(X_1, X_{1+h})] + 2\sum_{k\ge 1}\mathbb{E}[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})], \tag{29}$$

and the function ψ is defined by

$$\psi : (x,y) \mapsto \left\{(\gamma(0)+\gamma(h))\ \mathrm{IF}\left(\frac{x+y}{\sqrt{2(\gamma(0)+\gamma(h))}}, Q, \Phi\right) - (\gamma(0)-\gamma(h))\ \mathrm{IF}\left(\frac{x-y}{\sqrt{2(\gamma(0)-\gamma(h))}}, Q, \Phi\right)\right\}, \tag{30}$$

where IF is defined in (20).

The proof of Theorem 2 is given in Section 6.

Remark 2. Note that ψ has the same expression as the influence function of $\gamma_Q(\cdot)$ given in Ma and Genton (2000, p. 675).

Remark 3. Under Assumption (A1), let us now compare the asymptotic distribution of the proposed estimator with the classical autocovariance estimator defined by

$$\hat{\gamma}(h) = n^{-1}\sum_{i=1}^{n-h}(X_i - \bar{X}_n)(X_{i+h} - \bar{X}_n), \quad 0 \le h \le n-1. \tag{31}$$

Under (A1), applying Thm 4 of Arcones (1994) to $f{:}(x, y) \mapsto xy$ and $X_j = (X_j, X_{j+h})$, where h is a non-negative integer, leads to Proposition 2.

Proposition 2. For a given non-negative integer h, as $n \to \infty$,

$$\sqrt{n}(\hat{\gamma}(h) - \gamma(h)) \xrightarrow{d} \mathcal{N}(0, \breve{\sigma}^2_{cl}(h)),$$

where

$$\breve{\sigma}^2_{cl}(h) = \gamma^2(0) + \gamma^2(h) + 2\sum_{k\ge 1}\gamma^2(k) + 2\sum_{k\ge 1}\gamma(k+h)\gamma(k-h). \tag{32}$$

Let us now compare $\breve{\sigma}^2(h)$ in (29) with $\breve{\sigma}^2_{cl}(h)$ in (32). Since the theoretical lower bound for the asymptotic relative efficiency (ARE) defined by $\mathrm{ARE}(h) = \breve{\sigma}^2_{cl}(h)/\breve{\sigma}^2(h)$ is difficult to obtain, the estimation of ARE was calculated in the case where $(X_i)_{i\ge 1}$ is an AR(1) process: $X_i = \phi_1 X_{i-1} + \varepsilon_i$, where $(\varepsilon_i)_{i\ge 1}$ is a Gaussian white noise, for $\phi_1 = 0.1$, 0.5 and 0.9. These results are given in Figure 1 which

© 2010 Blackwell Publishing Ltd.

displays ARE for $h = 1, \ldots, 60$. From this figure, we can see that ARE ranges from 0.82 to 0.90 which indicates empirically that the robust procedure has almost no loss of efficiency.

2.2. Long-range dependence setting

In this section, we study the behaviour of the robust scale and autocovariance estimators $Q_n(X_{1:n}, \Phi)$ and $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ defined in (17) and (12) respectively, in the case where the process is long-range dependent. Long-range dependent processes play a key role in many domains, and it is therefore worthwhile to understand the behaviour of such estimators in this context. (**A2**) $(X_i)_{i\geq 1}$ is a stationary mean-zero Gaussian process with autocovariance $\gamma(h) = \mathbb{E}(X_1 X_{h+1})$ satisfying:

$$\gamma(h) = h^{-D} L(h), \quad 0 < D < 1,$$

where L is slowly varying at infinity and is positive for large h.

A classical model for long memory process is the so-called ARFIMA(p,d,q), which is a natural generalization of standard autoregressive integrated moving average (ARIMA(p,d,q)) models. By allowing d to assume any value in $(-1/2,1/2)$, a fractional ARFIMA model is defined by $\Phi(B)(1-B)^d X_i = \Theta(B) Z_i$. Here $(Z_i)_{i\in\mathbb{Z}}$ is a white Gaussian noise, B denotes the backshift operator, $\Phi(B)$ defines the AR part, $\Theta(B)$ defines the MA part of the process and $(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k$ is the fractional difference operator. For $d \neq 0$, one has

$$D = 1 - 2d, \tag{33}$$

see Taqqu (1975, eqn 66) For $d = 0$, we obtain the usual ARMA model. Long memory occurs for $d > 0$. As $h \to \infty$, the autocovariance of an ARFIMA(p,d,q) decreases as $\gamma(h) = Ch^{2d-1}$. Such processes satisfy (A2) with $D = 1 - 2d$, see Doukhan *et al.* (2003, chap. 1), for more details.

Perhaps surprisingly, the proof of the asymptotic properties of $Q_n(X_{1:n}, \Phi)$ in the long-range dependence framework does not follow the same steps as in the short-range dependence case. To understand why, assume that Assumption (A2) holds with $\gamma(0) = 1$. Theorem 1.1 in Dehling and Taqqu (1989) shows that the difference between the empirical distribution function F_n and Φ, the c.d.f. of the standard Gaussian distribution $\mathcal{N}(0,1)$ renormalized by nd_n^{-1}, that is, $nd_n^{-1}(F_n - \Phi)$, converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence. The sequence d_n depends on the exponent D governing the decay of the autocorrelation function to 0 and also on the slowly varying function L appearing in (A2): more precisely,

$$d_n = \alpha(D)^{1/2} n^{1-D/2} L^{1/2}(n) \tag{34}$$

with $\alpha(D) = 2(1-D)^{-1}(2-D)^{-1}$ for D in (0,1) defined in (A2). Therefore, Lemma 1 shows that the asymptotic expansion of $a_n(Q_n(X_{1:n},\Phi) - 1)$ in (18) remains valid with $a_n = nd_n^{-1}$, and that it remains to study the convergence of $d_n^{-1} \sum_{i=1}^n \mathrm{IF}(X_i, Q, \Phi)$. This type of nonlinear functional of stationary long-memory Gaussian sequences have been studied in Taqqu (1975) and Breuer and Major (1983). The limiting behaviour of these functionals depend both on D and on the Hermite rank of the function $\mathrm{IF}(\cdot,Q,\Phi)$. According to Breuer and Major (1983) and Taqqu (1975), under Assumption (A2), two markedly different behaviours may occur, depending on the value of D. If $D \in (1/2, 1)$, then, by Breuer and Major (1983), $n^{-1/2} \sum_{i=1}^n \mathrm{IF}(X_i, Q, \Phi)$ converges to a zero-mean Gaussian random variable with finite variance. If $D \in (0, 1/2)$, then $n^{D-1} L^{-1}(n) \sum_{i=1}^n \mathrm{IF}(X_i, Q, \Phi)$ converges to a non-degenerate (non-Gaussian) random variable, see Taqqu (1975). From these two results and (34), it follows that

$$d_n^{-1} \sum_{i=1}^n \mathrm{IF}(X_i, Q, \Phi) = o_P(1),$$

for $D \neq 1/2$. Therefore, the leading term in the expansion of $Q_n(X_{1:n},\Phi) - 1$ in the short-memory setting is no longer the leading term in the long-memory case. This explains why the proof, in the long-memory case, does not follow the same line of reasoning as that in the short-range dependence case. To derive the asymptotic properties of $Q_n(X_{1:n},\Phi)$ and $\hat{\gamma}_Q(\cdot, X_{1:n}, \Phi)$ for long-memory processes, it will be necessary to carry out a careful study of the U-process

$$U_n(r) = \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} \mathbb{1}_{\{|X_i - X_j| \leq r\}} = T_1(F_n)[r] - \frac{1}{n}, \tag{35}$$

based on the class of kernels $\{\mathbb{1}_{\{|x-y|\leq r\}}, x, y \in \mathbb{R}, r \geq 0\}$. Its asymptotic properties can be derived from Propositions 5 and 6 in Section 5 which are proved in the companion paper Lévy-Leduc *et al.* (2009).

2.2.1. Properties of the scale estimator

The next theorem gives the asymptotic behaviour of the robust-scale estimator $Q_n(X_{1:n}, \Phi)$ under Assumption (A2).

THEOREM 3. Under Assumption (A2), $Q_n(X_{1:n}, \Phi)$ satisfies the following limit theorems as n tends to infinity:

(i) *If* $D > 1/2$,

$$\sqrt{n}(Q_n(X_{1:n}, \Phi) - \sigma) \xrightarrow{d} \mathcal{N}(0, \tilde{\sigma}^2),$$

where $\sigma = \sqrt{\gamma(0)}$,

© 2010 Blackwell Publishing Ltd.

$$\tilde{\sigma}^2 = \gamma(0)\mathbb{E}[\mathrm{IF}^2(X_1/\sigma, Q, \Phi)] + 2\gamma(0)\sum_{k\geq 1}\mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)\mathrm{IF}(X_{k+1}/\sigma, Q, \Phi)],$$

and IF(·,Q,Φ) *is defined in* (20).

(ii) *If* $D < 1/2$,

$$\beta(D)\frac{n^D}{L(n)}(\mathrm{Q}_n(X_{1:n}, \Phi) - \sigma) \xrightarrow{d} \frac{\sigma}{2}(Z_{2,D}(1) - Z_{1,D}^2(1)),$$

where $\beta(D) = B((1-D)/2,D)$, B *denoting the beta function and the processes* $Z_{1,D}(\cdot)$ *and* $Z_{2,D}(\cdot)$ *are defined in* (53) *and* (54).

Theorem 3 is proved in Section 6.

Remark 4. Note that in the case (ii) the limit distribution is not centered and is asymmetric. Moreover, it can be proved (see Lévy-Ledue *et al.*, 2009) that $\mathbb{E}[Z_{2,D}(1) - Z_{1,D}(1)^2] = -2\beta(D)/(-D+1)(-D+2)$.

Remark 5. Under Assumption (A2), it is interesting to compare the asymptotic distribution of the proposed estimator $\mathrm{Q}_n(X_{1:n},\Phi)$ with that of the square root of the sample variance $\hat{\sigma}_{n,X}^2$ defined in (27).

Proposition 3. *Suppose Assumption* (*A*2). *Then as* $n \to \infty$,

(a) *if* $D > 1/2$,

$$\sqrt{n}(\hat{\sigma}_{n,X} - \sigma) \xrightarrow{d} \mathcal{N}(0, \tilde{\sigma}_{cl}^2),$$

where $\tilde{\sigma}_{cl}^2$ *is given in (28):*

(b) *if* $D < 1/2$,

$$\beta(D)n^D L(n)^{-1}(\hat{\sigma}_{n,X} - \sigma) \xrightarrow{d} \sigma/2\Big(Z_{2,D}(1) - Z_{1,D}^2(1)\Big). \tag{36}$$

The rates of convergence of the square root of the sample variance $\hat{\sigma}_{n,X}$ *and of the robust estimator* $\mathrm{Q}_n(X_{1:n},\Phi)$ *are identical. Moreover, there is no loss of efficiency when* $D < 1/2$.

The proof of Proposition 3 is in Section 6.

2.2.2. Properties of the autocovariance estimator

In this section, we study the asymptotic properties of $\hat{\gamma}_Q(\cdot, X_{1:n}, \Phi)$ based on the asymptotic properties of $\mathrm{Q}_n(X_{1:n}, \Phi)$.

Theorem 4. *Assume that* (*A*2) *holds and that* L *has three continuous derivatives. Assume also that* $L_i(x) = x^i L^{(i)}(x)$ *satisfy:* $L_i(x)/x^\epsilon = O(1)$, *for some* ϵ *in* $(0,D)$, *as* x *tends to infinity, for all* $i = 0,1,2,3$, *where* $L^{(i)}$ *denotes the ith derivative of* L. *Let* h *be a non-negative integer. Then,* $\hat{\gamma}_Q(\mathrm{h}, \mathrm{X}_{1:\mathrm{n}}, \Phi)$ *satisfies the following limit theorems as* n *tends to infinity.*

(i) *If* $D > 1/2$,

$$\sqrt{n}\big(\hat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)\big) \xrightarrow{d} \mathcal{N}(0, \tilde{\sigma}^2(h)),$$

where

$$\tilde{\sigma}^2(h) = \mathbb{E}[\psi^2(X_1, X_{1+h})] + 2\sum_{k\geq 1}\mathbb{E}[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})],$$

ψ *being defined in* (30).

(ii) If $D < 1/2$,

$$\beta(D)\frac{n^D}{\tilde{L}(n)}\big(\hat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)\big) \xrightarrow{d} \frac{\gamma(0)+\gamma(h)}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2),$$

where $\beta(D) = B((1-D)/2,D)$, B *denotes the beta function, the processes* $Z_{1,D}(\cdot)$ *and* $Z_{2,D}(\cdot)$ *are defined in* (53) *and* (54) *and*

$$\tilde{L}(n) = 2L(n) + L(n+h)(1+h/n)^{-D} + L(n-h)(1-h/n)^{-D}. \tag{37}$$

Theorem 4 is proved in Section 6.

Remark 6. Note that the assumptions on L_i made in Theorem 4 are obviously satisfied if L is the logarithmic function or a power of it.

© 2010 Blackwell Publishing Ltd.

PROPOSITION 4. *Under Assumption (A2) with $D < 1/2$, the robust autocovariance estimator $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ has the same asymptotic behaviour as the classical autocovariance estimator (31). There is no loss of efficiency.*

Proposition 4 is proved in Section 6.

3. NUMERICAL EXPERIMENTS

In this section, we investigate the robustness properties of the estimator $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ in (12), using Monte Carlo experiments.

We shall regard the observations X_t, $t = 1, \ldots, n$, as a stationary series Y_t, $t = 1, \ldots, n$, corrupted by additive outliers of magnitude ω. Thus we set

$$X_t = Y_t + \omega W_t, \tag{38}$$

where W_t are i.i.d. random variables such that $\mathbb{P}(W = -1) = \mathbb{P}(W = 1) = p/2$ and $\mathbb{P}(W = 0) = 1 - p$, where $\mathbb{E}[W] = 0$ and $\mathbb{E}[W^2] = \mathrm{var}(W) = p$. Observe that W is the product of Bernoulli(p) and *Rademacher* independent random variables; the latter equals 1 or -1, both with probability 1/2. Y_t is a stationary time series and it is assumed that Y_t and W_t are independent random variables. The empirical study is based on 5000 independent replications with $n = 100, 500$, $p = 0$, 5%,10% and $\omega = 10$. Other cases were also simulated, for example, series with $\omega = 3, 5$ which are magnitudes that cause less impact on the estimates compared with $\omega = 10$. These additional results are available on request.

We consider first the case where Y_t follows a Gaussian AR(1) process, that is, $Y_t = \sum_{j\geq 0} \phi_1^j Z_{t-j}$ with $\phi_1 = 0.2, 0.5$ and $\{Z_t\}$ i.i.d $\mathcal{N}(0, 1)$. Then we suppose that, Y_t are Gaussian ARFIMA(0,d,0) processes, that is,

$$Y_t = (I - B)^{-d} Z_t = \sum_{j\geq 0} \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)} Z_{t-j} \tag{39}$$

with $d = 0.2$, 0.45 and $\{Z_t\}$ i.i.d. $\mathcal{N}(0, 1)$.

Classically, scale is measured by the standard deviation σ. The robust measure of scale we consider here is $Q(F_X)$, defined in (5). Recall that one has $\sigma = Q(F_X)$ in the Gaussian case [see eqn (8)]. We want to compare their respective estimators $\hat{\sigma}_{n,X}$ defined in (27) and $Q_n(X_{1:n}, \Phi)$ defined in (17).

The standard deviations of the AR(1) models are $Q(F_Y) = \sigma_Y = 1.0206$ and $Q(F_Y) = \sigma_Y = 1.1547$ for $\phi_1 = 0.2$ and $\phi_1 = 0.5$ respectively. In the case of ARFIMA processes, the standard deviations are $Q(F_Y) = \sigma_Y = 1.0481$ when $d = 0.2$ and

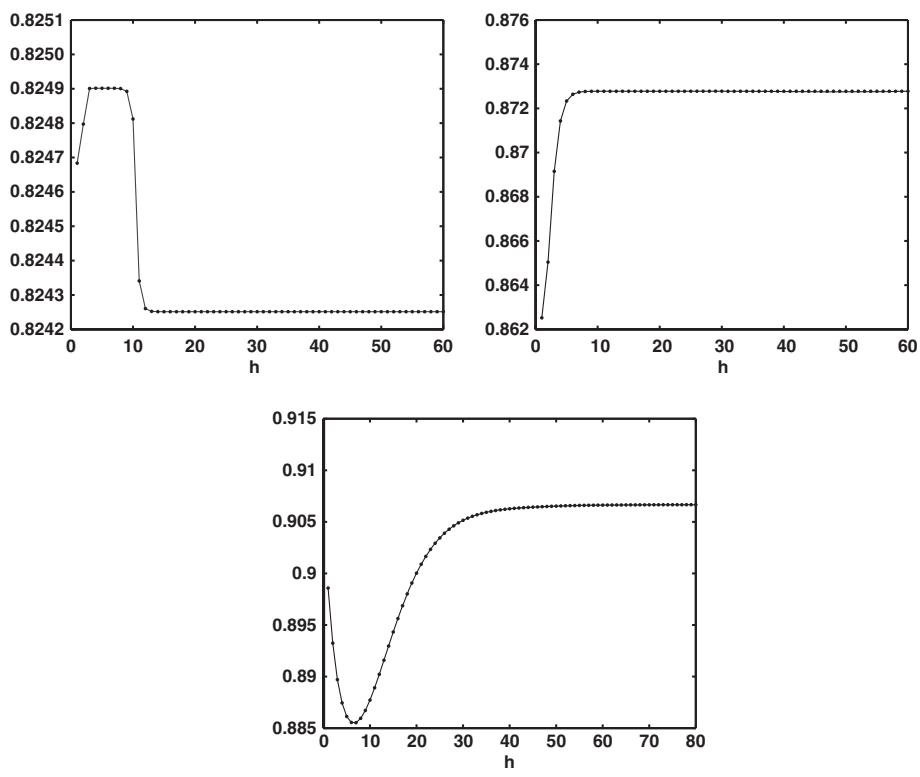

Figure 1. Asymptotic relative efficiency $\check{\sigma}^2_{cl}(h)/\check{\sigma}^2(h)$ for an AR(1) process for different values of ϕ_1: 0.1, 0.5 and 0.9 from left to right

© 2010 Blackwell Publishing Ltd.

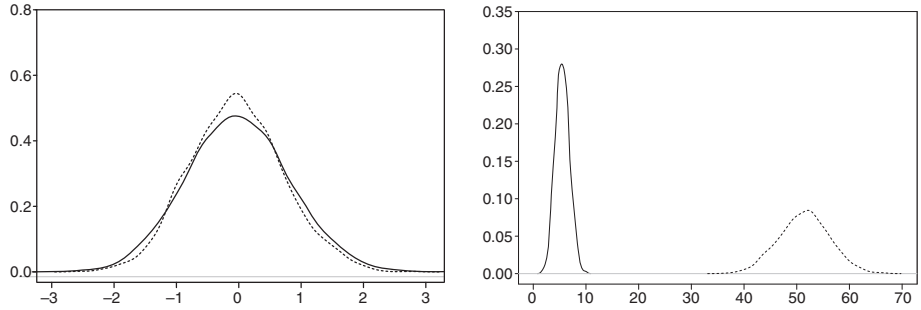

Figure 2. Empirical densities of the quantities $\sqrt{n}(\mathrm{Q}_n(X_{1:n},\Phi) - \sigma_Y)$ (plain line) and $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ (dotted line) of the AR(1) model with $\phi_1 = 0.2$, $n = 500$, without outliers (left) and with outliers with $p = 10\%$ and $\omega = 10$ (right)

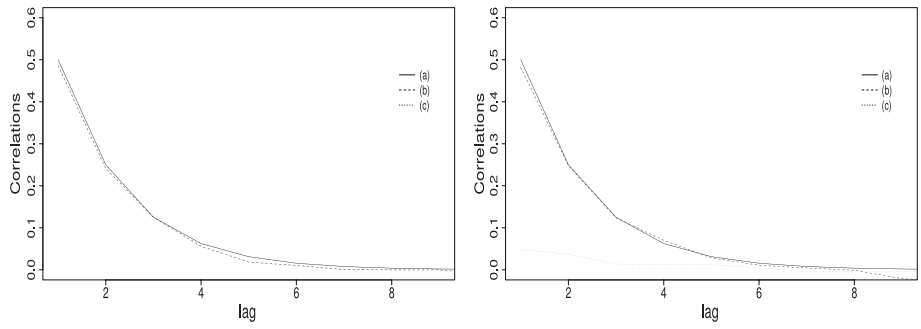

Figure 3. Sample correlations of the AR(1) model with $\phi_1 = 0.5$, $n = 500$ without outliers (left) and with outliers with $p = 10\%$ and $\omega = 10$ (right). (a) Population correlation; (b,c) robust and classical sample autocorrelations respectively

Table 1. Results for the estimation of AR(1) model with $\omega = 10$

		$p = 0$		$p = 5\%$		$p = 10\%$	
ϕ_1	n	$\hat{\phi}_{1,\gamma}$	$\hat{\phi}_{1,Q}$	$\hat{\phi}_{1,\gamma}$	$\hat{\phi}_{1,Q}$	$\hat{\phi}_{1,\gamma}$	$\hat{\phi}_{1,Q}$
0.2	100	0.1818	0.1831	0.0312	0.2212	0.01530	0.2651
		(0.0112)	(0.0128)	(0.0376)	(0.0229)	(0.0435)	(0.0388)
	500	0.1967	0.1948	0.0318	0.2381	0.0163	0.2881
		(0.0019)	(0.0025)	(0.0303)	(0.0051)	(0.0357)	(0.0150)
0.5	100	0.4767	0.4747	0.0998	0.5762	0.0495	0.6924
		(0.0084)	(0.0106)	(0.1740)	(0.0262)	(0.2142)	(0.0712)
	500	0.4967	0.4927	0.1030	0.6012	0.05647	0.7216
		(0.0015)	(0.0021)	(0.1598)	(0.0141)	(0.1988)	(0.0558)

$Q(F_Y) = \sigma_Y = 1.9085$ when $d = 0.45$. This is because the variance of AR(1) is $(1 - \phi_1^2)^{-1}$ and that of ARFIMA(0,d,0) is $\Gamma(1 - 2d)/\Gamma^2(1 - d)$ (see Brockwell and Davis, 1991). Figure 2 and Table 1 involve AR processes, and Figures 4–6 involve the ARFIMA processes.

3.1. Short-range dependence case

Figure 2 gives some insights on Theorem 1 and Proposition 1. In the left part of Figure 2, the empirical distribution of the quantities $\sqrt{n}(\mathrm{Q}_n(X_{1:n},\Phi) - \sigma_Y)$ and $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ are displayed for the AR(1) model with $\phi_1 = 0.2$, $n = 500$, without outliers. Both present shapes close to the Gaussian density, and their standard deviations are equal to 0.8232 and 0.7377 respectively. These empirical standard deviations are close to 0.8233 and 0.7500 which are the values of the asymptotic standard deviation $\tilde{\sigma}$ in (26) and that of $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ in (28) respectively. The value 0.8233 was obtained through numerical simulations and the value 0.7500 from the fact that for an AR(1) $\gamma(h) = \phi_1^{|h|}(1 - \phi_1^2)^{-1}$ and hence $\tilde{\sigma}_{cl}^2 = (1 + 2\phi_1^2)/(2(1 - \phi_1^2))$ in (28). Hence the empirical evidence fits with the theoretical results of Theorem 1 and Proposition 1.

In the right part of Figure 2, we display the results when outliers are present. The empirical distribution of $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ is clearly located far away from zero. One can also observe the increase in the variance. The quantity $\sqrt{n}(\mathrm{Q}_n(X_{1:n},\Phi) - \sigma_Y)$ looks symmetric and is located close to zero.

We now turn to the estimation of the autocorrelations. The plots of the autocorrelations are displayed in the left and right parts of Figure 3 when $\phi_1 = 0.5$ for models without and with outliers respectively. The figures also display the population autocorrelation function $\rho(h) = \phi_1^{|h|}$ as a function of the lag h. In the absence of atypical observations (left part of Figure 3), both sample functions

© 2010 Blackwell Publishing Ltd.

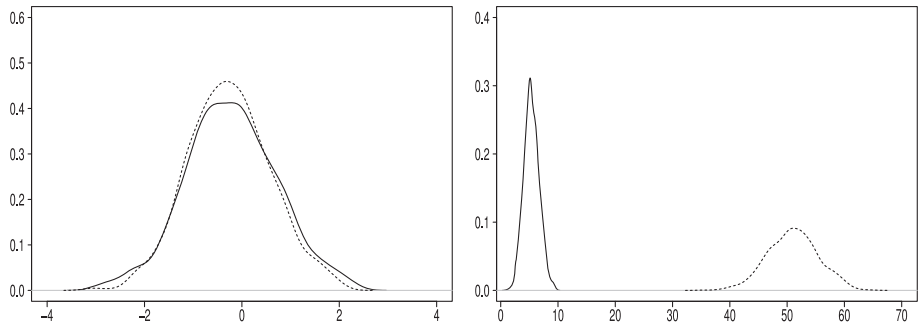

Figure 4. Empirical densities of $\sqrt{n}(\mathrm{Q}_n(X_{1:n},\Phi) - \sigma_Y)$ (plain line) and $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ (dotted line) for the ARFIMA(0, d, 0) model with $d = 0.2$, $n = 500$ without outliers (left) and with outliers with $p = 10\%$ and $\omega = 10$ (right)

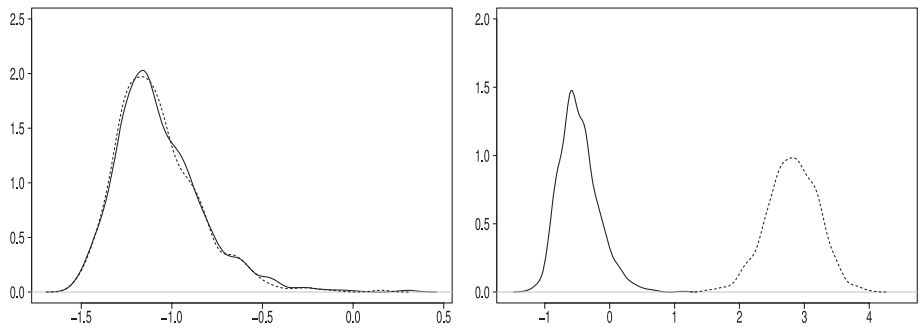

Figure 5. Empirical densities of the quantities $n^{1-2d}(\mathrm{Q}_n(X_{1:n},\Phi) - \sigma_Y)$ (plain line) and $n^{1-2d}(\hat{\sigma}_n - \sigma_Y)$ (dotted line) of the ARFIMA(0,d,0) model with $d = 0.45$, $n = 500$, without outliers (left) and with outliers $p = 10\%$ and $\omega = 10$ (right)

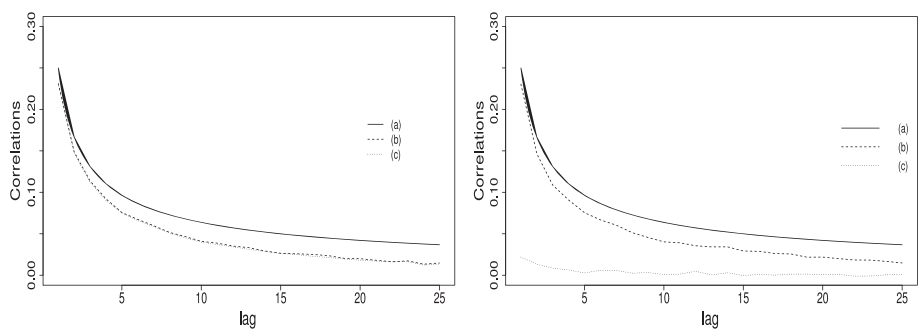

Figure 6. Sample correlations of the ARFIMA(0, d, 0) model with $d = 0.2$, $n = 500$ without outliers (left) and with outliers with $p = 10\%$ and $\omega = 10$ (right). (a) Population correlation; (b,c) robust and classical sample autocorrelations respectively

display a similar behaviour. However, when the data contain outliers (right part of Figure 3) the classical sample autocorrelation is clearly distorted.

We now want to use these autocorrelation estimators to get estimates for the AR(1) coefficient ϕ_1. The results are in Table 1. In this table, $\hat{\phi}_{1,\gamma}$ and $\hat{\phi}_{1,\mathrm{Q}}$ denote the average of the Yule-Walker estimates of the AR coefficients based on the classical estimator of the covariance γ and the robust autocovariance estimator $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ in (12) respectively. The numbers in parentheses are the corresponding square root of the sample mean-squared errors. The classical estimates were obtained using the subroutine DARMME in FORTRAN which uses a method of moments. The robust autocovariance and autocorrelation estimates were calculated using the code given in Croux and Rousseeuw (1992).

It can be seen from Table 1 that both autocovariances yield similar estimates for ϕ_1 when the process does not contain outliers. However, the picture changes significantly when the series is contaminated by atypical observations. As expected, the estimates from the classical autocovariance estimator are extremely sensitive to the presence of additive outliers. It is worth noting that the estimator based on the robust autocovariance (12) yields much more accurate estimates when the data contain outliers.

3.2. Long-range dependence case

In the case of the long-memory process ARFIMA(0,d,0) defined in (39), we choose $d = 0.2$ and $d = 0.45$, corresponding respectively, to $D = 0.6$ and $D = 0.1$ [see eqn (33)]. In the first case $D > 1/2$, in the second, $D < 1/2$, corresponding to the two cases of

© 2010 Blackwell Publishing Ltd.

Theorem 3. For $d = 0.2$, the empirical density functions of $\sqrt{n}(\mathrm{Q}_n(X_{1:n}, \Phi) - \sigma_Y)$ and $\sqrt{n}(\hat{\sigma}_n - \sigma_Y)$ are displayed in Figure 4 with and without outliers. When there is no outlier, both shapes are similar to that of the Gaussian density, and their standard deviations are equal to 0.9043 and 0.8361 respectively, corresponding to an ARE of 85.48%. As shown in the right part of Figure 4, the classical, scale estimator $\hat{\sigma}_n$ is much more sensitive to outliers than the robust one Q_n. The empirical density in the case of outliers is centered around 50.

To illustrate part (ii) of Theorem 3, we consider the empirical density functions of the quantities $n^{1-2d}(\mathrm{Q}_n(X_{1:n}, \Phi) - \sigma_Y)$ and $n^{1-2d}(\hat{\sigma}_n - \sigma_Y)$ when $d = 0.45$ ($D = 0.6$) as displayed in Figure 5. The left part of Figure 5 shows densities having means close to -1.1161 which is the value of the theoretical mean given in Remark 4. Both curves present, in fact, similar empirical standard deviation which is in accordance with Proposition 3. The impact of outliers on the estimates is clearly shown in the right side of Figure 5 where one observes a difference between the behaviour of the classical- and robust-scale estimators but not as large as in the previous cases.

Finally, the plots of the autocorrelations are displayed in the left and right parts of Figure 6 when $d = 0.2$ for models without and with outliers respectively. The figures also provide the population autocorrelation function $\rho(h) = \Gamma(1-d)\Gamma(h+d)/(\Gamma(d)\Gamma(1+h-d))$ as a function of the lag h (Hosking, 1981). Figure 7 displays the same quantities in the case where $d = 0.45$. The behaviour of the sample functions with respect to the presence of atypical observations is similar to the one observed in the short-range dependence case.

3.3. Non-Gaussian observations

We now examine the behaviour of the autocovariance estimator when it is applied to non-Gaussian observations. To do so, we generated observations $(X_t)_{1\le t\le n}$ as follows:

$$X_t = \phi_1 X_{t-1} + Z_t,$$

where $\phi_1 = 0.9$, $\varepsilon = 0.4$, $Z_t = W_t + \varepsilon Y_t^2$, where W_t and Y_t are independent random variables such that (W_t) and (Y_t) are i.i.d standard Gaussian random variables. An example of a realization of $(Z_t)_{1\le t\le n}$ is given in the histogram of Figure 8 with $n = 500$. As we can see from this figure, the presence of ε in the definition of Z_t produces an asymmetry in the data. In the right part of this figure, we

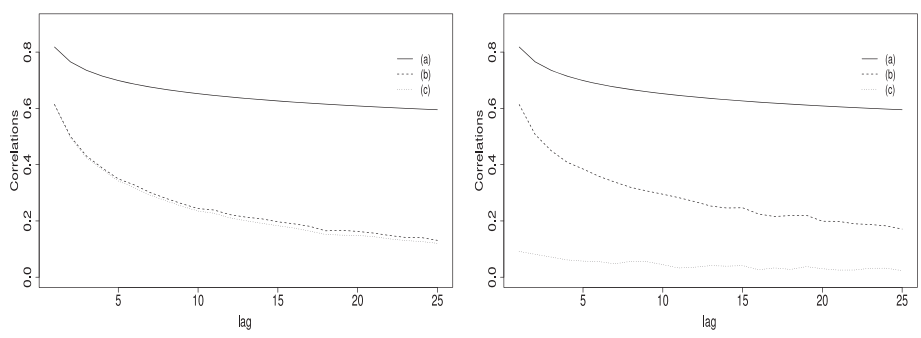

Figure 7. Sample correlations of the ARFIMA(0,d,0) model with $d = 0.45$, $n = 500$ without outliers (left) and with outliers with $p = 10\%$ and $\omega = 10$ (right). (a) Population correlation; (b,c) robust and classical sample autocorrelations respectively

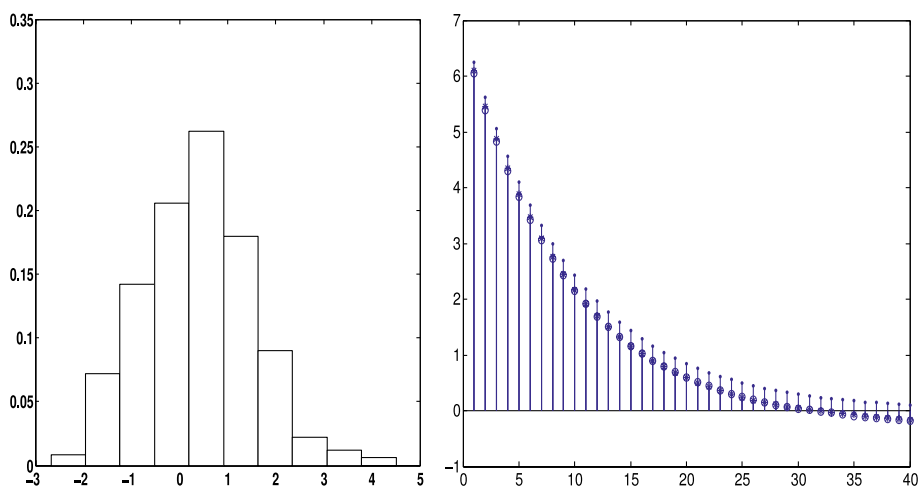

Figure 8. *Left*: Histogram of one realization of $(Z_t)_{1\le t\le 500}$. *Right*: Theoretical autocovariance(·), robust autocovariance (○) and classical autocovariance (☆) for $h = 1, \ldots, 40$

© 2010 Blackwell Publishing Ltd.

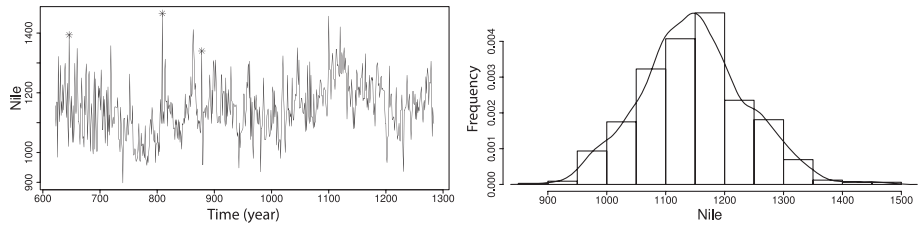

Figure 9. *Left*: The Nile River data plot. *Right*: Histogram of the Nile River data

displayed the average of the robust autocovariance $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ in (12) and the classical autocovariance $\hat{\gamma}(h)$ defined in Remark 3, for $h = 1, \ldots, 40$ and 1000 replications. From this figure, we can see that the robust autocovariance estimator does not seem to be affected by the skewness of the data.

4. AN APPLICATION

The Nile data are used here to illustrate some of the robust methodologies discussed previously. The Nile River data set is a well-known and interesting time series, which has been extensively analysed. These data are discussed in detail in the book by Beran (1994). It is first introduced in section 1.4 on p. 20, and is completely tabulated on pp. 237–239. Beran (1994) took these data from an earlier book by Toussoun (1925, pp. 366–404). The data consist of yearly minimal water levels of the Nile river measured at the Roda gauge, near Cairo, for the years 622–1284 AD (663 observations); The units for the data as presented by Beran (1994) are centimetres (presumably above some fixed reference point). The empirical mean and the standard deviation of the data are equal to 1148 and 89.05 respectively.

The question has been raised as to whether the Nile time series contains outliers; see, for example Beran (1992), Robinson (1995), Chareka *et al.* (2006) and Fajardo *et al.* (2009). The test procedure developed by Chareka *et al.* (2006), suggests the presence of outliers at 646 AD (*p*-value: 0.0308) and at 809 (*p*-value: 0.0007). Another possible outlier is at 878 AD. A plot of the time series where the observations which have been judged to be outliers are marked, is shown in the left part of Figure 9, and the right part of this figure displays the histogram of the data. Although the theory developed in this article is related to Gaussian processes, we believe that the small asymmetry of the data does not compromise the use of this series as an illustration of our robust methodology. A way to avoid this asymmetry is to consider the logarithm of the data. However, this does not make a significant difference in the estimates.

The left part of Figure 10 displays plots of the classical and robust sample autocorrelation functions of the original data. The autocorrelation values from the former are smaller than those of the latter one. However, the difference between the autocorrelations may not be large enough to suggest the presence of outliers. Thus, to better understand the influence of outliers on the sample autocorrelation functions in practical situations a new data set with artificial outliers was generated. We replaced the presumed outliers detected by Chareka *et al.* (2006) by the mean plus 5 or 10 standard deviations. The sample autocorrelations (robust and classical ones) were again calculated, see the right part of Figure 10. As expected, the values of the robust autocorrelations remained stable. However, the classical autocorrelations were significantly affected by the increase of the size of the observation. This is in accordance with the results presented in the simulation section.

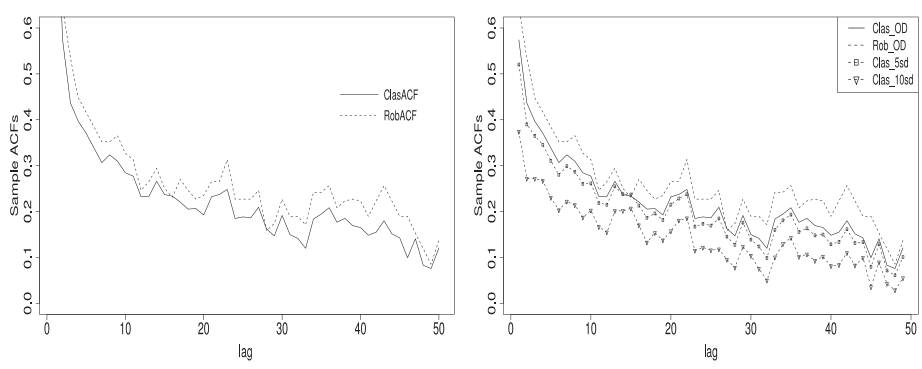

Figure 10. *Left*: Classical (plain line) and robust (dotted line) sample autocorrelation functions of the Nile River data. *Right*: Classical (plain line) and robust (dotted line) sample autocorrelation functions of the Nile River data and classical sample autocorrelation functions of the Nile River data with artificial outliers at the times detected by Chareka *et al.* (2006; original data plus 5 standard deviations with '□' and original data plus 10 standard deviations with '▽')

© 2010 Blackwell Publishing Ltd.

5. ASYMPTOTIC BEHAVIOUR OF *U*-PROCESSES

Consider the *U*-process $\{U_n^G(r), r \in I\}$ satisfying

$$U_n^G(r) = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \mathbb{1}_{\{G(X_i,X_j) \le r\}}, \quad r \in I \tag{40}$$

based on the class of kernels

$$k_G(x,y,r) = \mathbb{1}_{\{G(x,y) \le r\}}, \tag{41}$$

where I is an interval included in $\mathbb{R}$, G is a symmetric function that is $G(x,y) = G(y,x)$ for all x,y in $\mathbb{R}$, and the process $(X_i)_{i \ge 1}$ satisfies Assumption (A2) with $\gamma(0) = 1$.

The asymptotic properties of these *U*-processes have been studied in Lévy-Ledue *et al.* (2009). They are based on the computation of the Hermite rank of the class of functions $\{\mathbb{1}_{\{G(\cdot,\cdot) \le r\}} - U^G(r), r \in I\}$, where

$$U^G(r) = \int_{\mathbb{R}^2} \mathbb{1}_{\{G(x,y) \le r\}} \phi(x)\phi(y)\, \mathrm{d}x\, \mathrm{d}y, \quad \text{for all } r \in I. \tag{42}$$

The Hermite rank of the class of functions $\{\mathbb{1}_{\{G(\cdot,\cdot) \le r\}} - U(r), r \in I\}$ is obtained by expanding the function $\mathbb{1}_{\{G(\cdot,\cdot) \le r\}}$ on the basis of Hermite polynomials with leading coefficient equal to 1:

$$\mathbb{1}_{\{G(x,y) \le r\}} = \sum_{p,q \ge 0} \frac{\alpha_{p,q}(r)}{p!q!} H_p(x) H_q(y), \quad \text{for all } x, y \text{ in } \mathbb{R}, \tag{43}$$

where $\alpha_{p,q}(r) = \mathbb{E}\left[\mathbb{1}_{\{G(X,Y) \le r\}} H_p(X) H_q(Y)\right]$, X and Y being independent standard Gaussian random variables. The first few Hermite polynomials are $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$. Note that $\alpha_{0,0}(r)$ is equal to $U^G(r)$ for all r, where $U^G(r)$ is defined in (42). The previous expansion can also be rewritten as:

$$\mathbb{1}_{\{G(x,y) \le r\}} - U^G(r) = \sum_{\substack{p,q \ge 0 \\ p+q \ge m}} \frac{\alpha_{p,q}(r)}{p!q!} H_p(x) H_q(y), \tag{44}$$

where $m = m(r)$ is called the Hermite rank of the function $\mathbb{1}_{\{G(\cdot,\cdot) \le r\}} - U^G(r)$ when r is fixed.

We state the results for family of kernels having Hermite rank equal to $m = 2$ (this is all we need here) and refer to Lévy-Ledue *et al.* (2009) for other cases.

PROPOSITION 5. *Let I be a compact interval of $\mathbb{R}$, let $k_G(\cdot,\cdot,r)$ be defined in (41), and let*

$$k_{G,1}(x,r) = \mathbb{E}[k_G(x,Y,r)], \quad x \in \mathbb{R}, \quad r \in I, \tag{45}$$

where Y is a standard Gaussian variable. Suppose that the Hermite rank of the class of functions $\{k_G(\cdot,\cdot,r) - U^G(r), r \in I\}$ is $m = 2$ and that Assumption (A2) is satisfied with $\gamma(0) = 1$ and $1/2 < D < 1$. Assume that k_G satisfies the following three conditions:

(i) *There exists a positive constant C such that for all s, t in I, u, v in $\mathbb{R}$,*

$$\mathbb{E}[|k_G(X+u, Y+v, s) - k_G(X+u, Y+v, t)|] \le C|t-s|, \tag{46}$$

where(X,Y) is a standard Gaussian random vector.

(ii) *There exists a positive constant C such that for all $\ell \ge 1$ and s,t in I, u, v in $\mathbb{R}$,*

$$\mathbb{E}[|k_G(X_1+u, X_{1+\ell}+v, t) - k_G(X_1, X_{1+\ell}, t)|] \le C|u-v|, \tag{47}$$

$$\mathbb{E}[|k_G(X_1, X_{1+\ell}, s) - k_G(X_1, X_{1+\ell}, t)|] \le C|t-s|. \tag{48}$$

(iii) *There exists a positive constant C such that for all t in I, and x, u, v in $\mathbb{R}$,*

$$|k_{G,1}(x+u,t) - k_{G,1}(x+v,t)| \le C|u-v|, \tag{49}$$

$$|k_{G,1}(x,s) - k_{G,1}(x,t)| \le C|t-s|. \tag{50}$$

Then the U-process $\{\sqrt{n}(U_n^G(r) - U^G(r)), r \in I\}$ defined in (40) and (42) converges weakly in the space of cadlag functions on I, $\mathcal{D}(I)$, equipped with the topology of uniform convergence to the zero mean Gaussian process $\{W_G(r), r \in I\}$ with covariance structure given by

$$\begin{aligned}\mathbb{E}[W_G(s)W_G(t)] &= 4\,\mathrm{cov}(k_{G,1}(X_1,s), k_{G,1}(X_1,t)) \\ &\quad + 4\sum_{\ell \ge 1} \mathrm{cov}(k_{G,1}(X_1,s), k_{G,1}(X_{\ell+1},t)) + \mathrm{cov}(k_{G,1}(X_1,t), k_{G,1}(X_{\ell+1},s)).\end{aligned} \tag{51}$$

Moreover, for a fixed r in I, as n tends to infinity,

© 2010 Blackwell Publishing Ltd.

$$\sqrt{n}(U_n^G(r) - U^G(r)) = \frac{2}{\sqrt{n}}\sum_{i=1}^{n}\left[k_{G,1}(X_i, r) - U^G(r)\right] + o_P(1). \tag{52}$$

We now consider the case where $D < 1/2$. In this case, the normalization depends, as expected, on D and the slowly varying function L and the limiting distribution is no longer a Gaussian process. Let $(Z_{1,D}(t))_{t\in\mathbb{R}_+}$ denote the standard fractional Brownian motion (fBm) and $(Z_{2,D}(t))_{t\in\mathbb{R}_+}$ the Rosenblatt process. They are defined through multiple Wiener–Itô integrals and given by

$$Z_{1,D}(t) = \int_{\mathbb{R}}\left[\int_0^t (u-x)_+^{-(D+1)/2}\,\mathrm{d}u\right]\mathrm{d}B(x), \quad 0 < D < 1 \tag{53}$$

and

$$Z_{2,D}(t) = \int_{\mathbb{R}^2}'\left[\int_0^t (u-x)_+^{-(D+1)/2}(u-y)_+^{-(D+1)/2}\,\mathrm{d}u\right]\mathrm{d}B(x)\,\mathrm{d}B(y), \quad 0 < D < 1/2, \tag{54}$$

where B is the standard Brownian motion, see Fox and Taqqu (1987). The symbol $\int'$ means that the domain of integration excludes the diagonal. Introduce also the beta function

$$B(\alpha,\beta) = \int_0^\infty y^{\alpha-1}(1+y)^{-\alpha-\beta}\mathrm{d}y = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha > 0, \beta > 0. \tag{55}$$

Proposition 6. *Let I be a compact interval of $\mathbb{R}$. Suppose that the Hermite rank of the class of functions $\{k_G(\cdot,\cdot,r) - U^G(r), r \in I\}$ is $m = 2$ and that Assumption (A2) is satisfied with $\gamma(0) = 1$ and $D < 1/2$. Assume the following:*

(i) *There exists a positive constant C such that, for all $\ell \geq 1$ and for all s,t in I,*

$$\mathbb{E}[|k_G(X_1, X_{1+\ell}, s) - k_G(X_1, X_{1+\ell}, t)|] \leq C|t-s|. \tag{56}$$

(ii) *U^G is a Lipschitz function.*
(iii) *The function $\tilde{\Lambda}$ defined, for all s in I, by*

$$\tilde{\Lambda}(s) = \mathbb{E}[k_G(X, Y, s)(|X| + |XY| + |X^2 - 1|)], \tag{57}$$

where X and Y are independent standard Gaussian random variables, is also a Lipschitz function.

Then, the U-process $\{U_n^G(r) - U^G(r), r \in I\}$ defined in (40) and (42) has the following asymptotic properties:

$$\left\{n^D L^{-1}(n)\left(U_n^G(r) - U^G(r)\right); r \in I\right\}$$

converges weakly in the space of cadlag functions $\mathcal{D}(I)$, equipped with the topology of uniform convergence, to

$$\left\{\beta(D)^{-1}\left[\alpha_{1,1}(r)Z_{1,D}(1)^2 + \alpha_{2,0}(r)Z_{2,D}(1)\right]; r \in I\right\},$$

where the fractional Brownian motion $Z_{1,D}(\cdot)$ and the Rosenblatt process $Z_{2,D}(\cdot)$ are defined in (53) and (54) respectively, and where $\beta(D) = B((1-D)/2, D)$, B denoting the beta function, defined in (55).

Propositions 5 and 6 will be applied to the U-process $U_n(r)$ in (35) with

$$U(r) = \int_{\mathbb{R}^2} \mathbb{1}_{\{\|x-y|\leq r\}}\,\mathrm{d}F(x)\,\mathrm{d}F(y) = T_1(F)[r], \tag{58}$$

with T_1 given in (13). By Lemma 4, the Hermite rank of the class of functions $\{\mathbb{1}_{|x-y|\leq r} - U(r), x, y \in \mathbb{R}, r \in I\}$ is equal to 2 where $I = [r_0 - \eta, r_0 + \eta]$ for some positive η defined in Lemma 4 and where $r_0 = 1/c(\Phi)$ [see eqn (10)] is such that

$$T_1(\Phi)[r_0] = T_1(\Phi)[1/c(\Phi)] = 1/4. \tag{59}$$

6. PROOFS

Proof of Lemma 1. Denote by F the c.d.f. $\Phi_{\mu,\sigma}$ of X_1. Since $a_n(F_n - F)$ converges in distribution in the space of cadlag functions equipped with the topology of uniform convergence, the asymptotic expansion (18) can be deduced from the functional delta method stated, for example in Thm 20.8 of van der Vaart (1998). To show this, we have to prove that $T_0 = T_1 \circ T_2$ is Hadamard differentiable, where T_1 and T_2 are defined in (13) and (14) respectively, and that the corresponding Hadamard differential is defined

© 2010 Blackwell Publishing Ltd.

and continuous on the whole space of cadlag functions. For a definition of Hadamard differentiability, we refer to chap. 20 in van der Vaart (1998).

We prove first that the Hadamard differentiability of the functional T_1 defined in (13). Let (g_t) be a sequence of cadlag functions with bounded variations such that $\|g_t - g\|_\infty \to 0$, as $t \to 0$, where g is a cadlag function. For any non-negative r, we consider

$$t^{-1}\{T_1(F + tg_t)[r] - T_1(F)[r]\} = 2\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|x-y|\le r\}}\,\mathrm{d}F(x)\,\mathrm{d}g_t(y) + t\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|x-y|\le r\}}\,\mathrm{d}g_t(x)\,\mathrm{d}g_t(y).$$

Since

$$\begin{aligned}&\left|\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|x-y|\le r\}}\,\mathrm{d}F(x)\,\mathrm{d}g_t(y) - \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|x-y|\le r\}}\,\mathrm{d}F(x)\,\mathrm{d}g(y)\right| \\ &\quad = \left|\int_{\mathbb{R}}(g_t(x+r) - g(x+r))\,\mathrm{d}F(x) - \int_{\mathbb{R}}(g_t(x-r) - g(x-r))\,\mathrm{d}F(x)\right| \le 2\|g_t - g\|_\infty \to 0,\end{aligned}$$

as t tends to zero, the Hadamard differential of T_1 at g is given by:

$$(DT_1(F).g)(r) = 2\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|x-y|\le r\}}\,\mathrm{d}F(x)\,\mathrm{d}g(y) = 2\int_{\mathbb{R}}\{g(x+r) - g(x-r)\}\,\mathrm{d}F(x).$$

By Lem 21.3 in van der Vaart (1998), T_2 is Hadamard differentiable. Finally, using the chain rule (Thm 20.9 in van der Vaart, 1998), we obtain the Hadamard differentiability of T_0 with the following Hadamard differential:

$$DT_0(F).g = -\frac{(DT_1(F).g)(T_0(F))}{(T_1(F))'[T_0(F)]} = -\frac{2\int_{\mathbb{R}}\{g(x+T_0(F)) - g(x-T_0(F))\}\,\mathrm{d}F(x)}{(T_1(F))'[T_0(F)]}. \tag{60}$$

In view of the last expression, $DT_0(F)$ is a continuous function of g and is defined on the whole space of cadlag functions. Thus, by Thm 20.8 of van der Vaart (1998), we obtain:

$$a_n(\mathrm{Q}_n(X_{1:n}, \Phi) - Q(F)) = c(\Phi)DT_0(F).\{a_n(F_n - F)\} + o_P(1), \tag{61}$$

where $c(\Phi)$ is the constant defined in (10). By (13), $T_1(F)[r] = \int_{\mathbb{R}}[F(x+r) - F(x-r)]\,\mathrm{d}F(x)$ and since $F(\cdot) = \Phi_{\mu,\sigma}(\cdot) = \Phi((\cdot - \mu)/\sigma)$, we get

$$(T_1(F))'[r] = \frac{2}{\sigma}\int_{\mathbb{R}} \Phi\left(y + \frac{r}{\sigma}\right)\Phi(y)\,\mathrm{d}y.$$

Since $\sigma = Q(\Phi_{\mu,\sigma}) = c(\Phi)T_0(F)$ by (5), we get

$$(T_1(F))'[T_0(F)] = 2\sigma^{-1}\int \phi(y)\phi(y + 1/c(\Phi))\,\mathrm{d}y. \tag{62}$$

Applying (60) with $T_0(F) = \sigma/c(\Phi)$, using (62), and setting $g = a_n(F_n - F)$, we get

$$DT_0(F).\{a_n(F_n - F)\} = A_n - B_n, \tag{63}$$

where

$$A_n = a_n\big((T_1(F))'[T_0(F)]\big)^{-1}\int_{\mathbb{R}}\left\{F\left(x + \frac{\sigma}{c(\Phi)}\right) - F\left(x - \frac{\sigma}{c(\Phi)}\right)\right\}\mathrm{d}F(x) \tag{64}$$

and B_n has the same expression with F replaced by F_n. The integral in A_n equals

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{|y-x|\le\sigma/c(\Phi)\}}\,\mathrm{d}F(x)\,\mathrm{d}F(y) = T_1(F)[T_0(F)] = \frac{1}{4}, \tag{65}$$

by definition [see eqn(5)]. The corresponding integral in B_n equals

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\left\{F\left(X_i + \frac{\sigma}{c(\Phi)}\right) - F\left(X_i - \frac{\sigma}{c(\Phi)}\right)\right\} \\ &\quad = \frac{1}{n}\sum_{i=1}^{n}\left\{\Phi\left(\frac{X_i - \mu}{\sigma} + \frac{1}{c(\Phi)}\right) - \Phi\left(\frac{X_i - \mu}{\sigma} - \frac{1}{c(\Phi)}\right)\right\}.\end{aligned}$$

The result follows from (58), (57), (60), (59) and the above expressions for A_n and B_n. □

Proof of Theorem 1. Assumption (A1) and the theorem of Csörgő and Mielniczuk (1996) implies that $\sqrt{n}(F_n - \Phi_{0,\sigma})$ converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence. Thus, the asymptotic expansion of $a_n(\mathrm{Q}_n(X_{1:n}, \Phi) - \sigma)$ obtained in (18) is valid with $a_n = \sqrt{n}$. We thus have to prove a central limit theorem for

© 2010 Blackwell Publishing Ltd.

$n^{-1/2}\sum_{i=1}^{n}\mathrm{IF}(X_i/\sigma, Q, \Phi)$. Using Lemma 2, we note that the Hermite rank of IF(·,Q,Φ) is equal to 2 and the conclusion follows by applying Thm 1 of Breuer and Major (1983). □

Proof of Proposition 1. Note that $\hat{\sigma}^2_{n,X} = \gamma(0)\hat{\sigma}^2_{n,Y}$, where $(Y_i)_{i\geq 1}$ satisfies (A20) with $\gamma(0) = 1$. Observe that $\hat{\sigma}^2_{n,Y} - 1$ is a U-statistic with kernel $k(x,y) = (x-y)^2/2 - 1$. The Hoeffding decomposition of this kernel is given by $k(x,y) = (x^2-1)/2 + (y^2-1)/2 - xy$. From this, we obtain the corresponding Hoeffding decomposition of $\hat{\sigma}^2_{n,Y} - 1$ as:

$$\hat{\sigma}^2_{n,Y} - 1 = \frac{1}{n}\sum_{i=1}^{n} H_2(Y_i) - \frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n} Y_iY_j. \tag{66}$$

Under Assumption (A1), the first term of this decomposition is the leading one. Then, using Thm 1 of Breuer and Major (1983), we get that $\sqrt{n}(\hat{\sigma}^2_{n,X} - \sigma^2)$ converges to a zero-mean Gaussian random variable having a variance equal to $2(\gamma(0)^2 + 2\sum_{k\geq 1}\gamma(k)^2)$. Using the delta method to go from $\hat{\sigma}^2_{n,X}$ to $\hat{\sigma}_{n,X}$, setting $f(x) = \sqrt{x}$, so that $f'(\sigma^2) = 1/(2\sqrt{\sigma^2}) = 1/(2\sigma)$, we get that the asymptotic variance of $\sqrt{n}(\hat{\sigma}_{n,X} - \sigma)$ is thus equal to (28).

By Lemma 2, the Hermite rank of IF(.,Q,Φ) is equal to 2, hence using Lem 1 of Arcones (1994), $\tilde{\sigma}^2$ defined in (26) satisfies

$$\tilde{\sigma}^2 \leq \gamma(0)^{-1}\mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)^2]\left\{\gamma(0)^2 + 2\sum_{k\geq 1}\gamma(k)^2\right\}.$$

Finally, in this case, using that $\mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)^2] \approx 0.6077$ (Rousseeuw and Croux, 1993, p. 1278), the relative asymptotic efficiency $\tilde{\sigma}^2_{cl}/\tilde{\sigma}^2$ of $Q_n(X_{1:n},\Phi)$ compared to $\hat{\sigma}_{n,X}$ is larger than 82.27% since

$$\frac{(2\gamma(0))^{-1}(\gamma(0)^2 + 2\sum_{k\geq 1}\gamma(k)^2)}{\gamma(0)^{-1}\mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)^2]\{\gamma(0)^2 + 2\sum_{k\geq 1}\gamma(k)^2\}} \approx 0.5/0.6077 \approx 82.27\%.$$

□

Proof of Theorem 2. Let $\Phi_{\sigma,+}$ and $\Phi_{\sigma,-}$ denote the c.d.f. of $(X_i + X_{i+h})_{i\geq 1}$ and $(X_i - X_{i+h})_{i\geq 1}$, respectively. Let also denote by $F_{+,n-h}$ and $F_{-,n-h}$ the empirical c.d.f. of $(X_i + X_{i+h})_{1\leq i\leq n-h}$ and $(X_i - X_{i+h})_{1\leq i\leq n-h}$ respectively. Since $(X_i)_{i\geq 1}$ satisfies Assumption (A1), it is the same for $(X_i + X_{i+h})_{i\geq 1}$ and $(X_i - X_{i+h})_{i\geq 1}$ with scales equal to $Q(\Phi_{\sigma,+})$ and $Q(\Phi_{\sigma,-})$ respectively. Thus, using the theorem of Csörgő and Mielniczuk (1996), we obtain that $\sqrt{n-h}(F_{+,n-h} - \Phi_{\sigma,+})$ converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence and that the same holds for $\sqrt{n-h}(F_{-,n-h} - \Phi_{\sigma,-})$. As a consequence, expansion (18) is valid for $Q_{n-h}(X_{1:n-h} + X_{h+1:n},\Phi)$ and $Q_{n-h}(X_{1:n-h} - X_{h+1:n},\Phi)$ with $a_{n-h} = \sqrt{n-h}$, that is,

$$\sqrt{n-h}[Qn - h(X_{1:n-h} \pm X_{h+1:n}, \Phi) - Q(\Phi_{\sigma,\pm})] = \frac{1}{\sqrt{n-h}}\sum_{i=1}^{n-h}\mathrm{IF}(X_i \pm X_{i+h}, Q, \Phi_{\sigma,\pm}) + o_P(1).$$

Then, applying the delta method (Thm 3.1 of van der Vart, 1998) with the transformation $b(x) = x^2$, $b'(x) = 2x$, we get

$$\sqrt{n-h}[Qn - h(X_{1:n-h} \pm X_{h+1:n}, \Phi)^2 - Q^2(\Phi_{\sigma,\pm})] = \frac{2Q(\Phi_{\sigma,\pm})}{\sqrt{n-h}}\sum_{i=1}^{n-h}\mathrm{IF}(X_i \pm X_{i+h}, Q, \Phi_{\sigma,\pm}) + o_P(1).$$

Hence, $\hat{\gamma}_Q(h, X_{1:n}, \Phi)$ in (12) satisfies the following asymptotic expansion:

$$\sqrt{n-h}(\hat{\gamma}_Q(h, X_{1:n}, \Phi) - \{Q^2(\Phi_{\sigma,+}) - Q^2(\Phi_{\sigma,-})\}/4) = \frac{1}{\sqrt{n-h}}\sum_{i=1}^{n-h}\psi(X_i, X_{i+h}) + o_P(1), \tag{67}$$

where for all x and y,

$$\psi(x,y) = \frac{1}{2}\{Q(\Phi_{\sigma,+})\ \mathrm{IF}(x+y, Q, \Phi_{\sigma,+}) - Q(\Phi_{\sigma,-})\ \mathrm{IF}(x-y, Q, \Phi_{\sigma,-})\}.$$

Using identity (19), ψ has the expression given in (30). We now have to prove a central limit theorem for $(n-h)^{-1/2}\sum_{i=1}^{n-h}\psi(X_i, X_{i+h})$. Using Lemma 3, the definition of the Hermite rank given on p. 2245 of Arcones (1994) and Assumption (A1), we obtain that condition (2.40) of Thm 4 in Arcones (1994, p. 2256), is satisfied with $\tau = 2$. This concludes the proof of the theorem by observing that $\{Q^2(\Phi_{\sigma,+}) - Q^2(\Phi_{\sigma,-})\}/4 = \mathbb{E}[X_1X_{1+h}] = \gamma(h)$ [see eqn (7)]. □

Proof of Theorem 3. Since, by scale invariance, $Q_n(X_{1:n}, \Phi) - \sigma = \sigma(Qn(X_{1:n}/\sigma, \Phi) - 1)$, we shall focus in the sequel on the case $\gamma(0) = 1$. First, note that using Lemma 4 the Hermite rank of the class of functions $\{\mathbb{1}_{\{|\cdot-\cdot|\leq r\}} - U(r), r \in [r_0 - \eta, r_0 + \eta]\}$ is $m = 2$, where U is defined in (58) and r_0 in (59).

(i) Suppose first $D > 1/2$. Let us verify that the assumptions of Proposition 5 hold. Conditions (46) and (47) are easily verified. Let us check condition (48). Note that for all $\ell \geq 1$, $X_1 - X_{1+\ell} \sim \mathcal{N}(0, 2(1-\gamma(\ell)))$, thus if $t \leq s$, there exists a positive constant C such that

© 2010 Blackwell Publishing Ltd.

$$\mathbb{E}[k(X_1, X_{1+\ell}, s) - k(X_1, X_{1+\ell}, t)] = \mathbb{P}(t \le |X_1 - X_{1+\ell}| \le s) \le \frac{2}{\sqrt{4\pi(1-\gamma(\ell))}}|t-s| \le C|t-s|,$$

where $k(x, y, r) = \mathbb{1}_{\{|x-y| \le r\}}$. Since $\gamma(\ell) \to 0$ as $\ell \to \infty$, we obtain (48).

Conditions (49) and (50) are satisfied since

$$k_1(x, r) = \mathbb{E}[\mathbb{1}_{\{|x-Y| \le r\}}] = \Phi(x+r) - \Phi(x-r). \tag{68}$$

Now, consider the process

$$\{\sqrt{n}(T_1(F_n)[r] - T_1(F)[r]), \quad r \in [r_0 - \eta, r_0 + \eta]\}, \tag{69}$$

where $F = \Phi$ and

$$T_1(F)[r] = \int_{\mathbb{R}^2} \mathbb{1}_{\{|y-x| \le r\}}\, \mathrm{d}\Phi(x)\, \mathrm{d}\Phi(y) = \int_{\mathbb{R}} [\Phi(x+r) - \Phi(x-r)]\, \mathrm{d}\Phi(x).$$

By Proposition 5, process (69) converges weakly to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence for some $\eta > 0$ when $D > 1/2$.

(ii) Suppose now $D < 1/2$. Let us check that the assumptions of Proposition 6 hold. Condition (56) holds since it is the same as condition (48). Since k_1 is a Lipschitz function, so is U defined in (58). Let us now check condition (57). If $s \le t$,

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_{\{s<x-y\le t\}}(|x| + |xy| + |x^2-1|)\phi(x)\phi(y)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{R}}\left(\int_{x-t}^{x-s} \phi(y)\, \mathrm{d}y\right)|x|\phi(x)\, \mathrm{d}x$$
$$+ \int_{\mathbb{R}}\left(\int_{x-t}^{x-s} |y|\phi(y)\, \mathrm{d}y\right)|x|\phi(x)\, \mathrm{d}x + \int_{\mathbb{R}}\left(\int_{x-t}^{x-s} \phi(y)\, \mathrm{d}y\right)|x^2-1|\phi(x)\, \mathrm{d}x.$$

Since $\phi(\cdot)$ and $|\cdot|\phi(\cdot)$ are bounded and that the moments of Gaussian random variables are all finite, we get (57). Then, applying Proposition 6 and Lemma 4 leads to the weak convergence of the process $\{\beta(D)n^D/L(n)(T_1(F_n)[r] - T_1(F)[r]), r \in [r_0 - \eta, r_0 + \eta]\}$ to $\{\dot{\phi}(r/\sqrt{2})(Z_{2,D}(1) - Z_{1,D}(1)^2); r \in [r_0 - \eta, r_0 + \eta]\}$.

We now want to use the functional delta method as in the proof of Lemma 1 in both cases (i) and (ii).

By Lem 21.3 in van der Vaart (1998), T_2 defined in (14) is Hadamard differentiable with the following Hadamard differential: $DT_2(T_1(\Phi)) \cdot g = -g(r_0)/(T_1(\Phi))'[r_0]$. Thus, $DT_2(T_1(\Phi))$ is a continuous function with respect to g. By the functional delta method, with $T_0 = T_2 \circ T_1$, we obtain the following expansion:

$$a_n(\mathrm{Q}_n(X_{1:n}, \Phi) - Q(\Phi)) = c(\Phi)a_n(T_0(F_n) - T_0(\Phi)) = -c(\Phi)a_n\frac{(T_1(F_n) - T_1(\Phi))[r_0]}{(T_1(\Phi))'[r_0]} + o_P(1), \tag{70}$$

where $a_n = \sqrt{n}$ in the case (i) and $a_n = \beta(D)n^D/L(n)$ in the case (ii). In case (i),

$$-c(\Phi)\sqrt{n}\frac{(T_1(F_n) - T_1(\Phi))[r_0]}{(T_1(\Phi))'[r_0]} \xrightarrow{d} \mathcal{N}(0, \sigma_1^2),$$

where σ_1^2 is given by (51) in Proposition 5:

$$\sigma_1^2 = 4\mathrm{var}\left[-\frac{c(\Phi)}{(T_1(\Phi))'[r_0]}k_1(X_1, r_0)\right] + 8\sum_{k\ge1}\mathrm{cov}\left[-\frac{c(\Phi)}{(T_1(\Phi))'[r_0]}k_1(X_1, r_0), -\frac{c(\Phi)}{(T_1(\Phi))'[r_0]}k_1(X_{k+1}, r_0)\right], \tag{71}$$

where k_1 is defined in (68). Since

$$\mathbb{E}[k_1(X_1, r_0)] = \mathbb{E}[\Phi(X_1 + r_0) - \Phi(X_1 - r_0)] = \int_{\mathbb{R}^2} \mathbb{1}_{\{|y-x|\le r\}}\, \mathrm{d}\Phi(x)\, \mathrm{d}\Phi(y) = 1/4$$

by (59) and $r_0 = 1/c(\Phi)$ by (9), we get using (62) that

$$\frac{-2c(\Phi)[k_1(X_1, r_0) - \mathbb{E}(k_1(X_1, r_0))]}{(T_1(\Phi))'[r_0]} = c(\Phi)\frac{1/4 + \Phi(X_1 - 1/c(\Phi)) - \Phi(X_1 + 1/c(\Phi))}{2\int_{\mathbb{R}}\phi(y)\phi(y + 1/c(\Phi))\mathrm{d}y} = \mathrm{IF}(X_1, Q, \Phi), \tag{72}$$

where IF(·,Q,Φ) is defined in (20). Using (71), (72) and (82) in Lemma 2, we get that

$$\sigma_1^2 = \mathbb{E}[\mathrm{IF}(X_1, Q, \Phi)] + 2\sum_{k\ge1}\mathbb{E}[\mathrm{IF}(X_1, Q, \Phi)\mathrm{IF}(X_{k+1}, Q, \Phi)],$$

which concludes the proof of (i). In the case (ii), in view of (70), it is sufficient to show that

$$-c(\Phi)\beta(D)\frac{n^D}{L(n)}\frac{(T_1(F_n) - T_1(\Phi))[r_0]}{(T_1(\Phi))'[r_0]} \xrightarrow{d} \frac{1}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2). \tag{73}$$

© 2010 Blackwell Publishing Ltd.

This result follows from the convergence in distribution of $\beta(D)n^D/L(n)(T_1(F_n)[r_0]-T_1(\Phi)[r_0])$ to $\dot{\phi}(r_0/\sqrt{2})(Z_{2,D}(1)-Z_{1,D}(1)^2)$, (62) and the identity $-c(\Phi)\dot{\phi}(r_0/\sqrt{2})(2\int_{\mathbb{R}}\phi(y)\phi(y+r_0)\,dy)^{-1} = 1/2$. This identity follows from $\dot{\phi}(r_0/\sqrt{2}) = -(2\sqrt{\pi})^{-1}\exp(-r_0^2/4)r_0$ and $r_0 = 1/c(\Phi)$. □

Proof of Proposition 3. Using the same arguments as those used in the proof of Proposition 1, we get that $\hat{\sigma}^2_{n,Y}-1$ satisfies the Hoeffding decomposition (66), where $(Y_i)_{i\geq 1}$ satisfies (A2) with $\gamma(0) = 1$.

(a) If $D > 1/2$, using Dehling and Taqqu (1991), the first term in decomposition (66) is the leading one, then using the same arguments as those used in the proof of Proposition 1, we get that the asymptotic variance of $\sqrt{n}(\hat{\sigma}_{n,X}-\sigma)$ is equal to

$$(2\gamma(0))^{-1}\left(\gamma(0)^2+2\sum_{k\geq 1}\gamma(k)^2\right).$$

Usingthe same upper bound as the one used in the proof of Proposition 1, we get that the relative efficiency of the robust scale estimator is, in this case, larger than 82.27%.

(b) If $D < 1/2$, we can apply the results of Dehling and Taqqu (1991) and the classical delta method to show that

$$\beta(D)n^D L(n)^{-1}(\hat{\sigma}_{n,X}-\sigma)\xrightarrow{d}\sigma/2(Z_{2,D}(1)-Z^2_{1,D}(1)).$$

□

Proof of Theorem 4. Let $\Phi_{\sigma,+}$ and $\Phi_{\sigma,-}$ denote the c.d.f. of $(X_i+X_{i+h})_{i\geq 1}$ and $(X_i-X_{i+h})_{i\geq 1}$ respectively. Since $(X_i)_{i\geq 1}$ satisfies Assumption (A2), a straightforward application of a Taylor formula shows that the same holds for $(X_i+X_{i+h})_{i\geq 1}$ with a scale equal to $Q(\Phi_{\sigma,+})$ and L replaced by some slowly varying function $\tilde{L}$. Thus, in the case (i), where $D > 1/2$, we obtain that $\mathrm{Q}_{n-h}(\{X_{1:n-h}+X_{h+1:n}\}/Q(\Phi_{\sigma,+}),\Phi)$ satisfies (70) with $a_n = \sqrt{n}$ as proved in the proof of Theorem 3. Using (52), we get that

$$\sqrt{n}\{\mathrm{Q}_{n-h}(X_{1:n-h}+X_{h+1:n},\Phi)-Q(\Phi_{\sigma,+})\} = -\frac{c(\Phi)Q(\Phi_{\sigma,+})}{(T_1(\Phi))'[r_0]}\frac{2}{\sqrt{n}}\sum_{i=1}^{n}[k_1(\{X_i+X_{i+h}\}/Q(\Phi_{\sigma,+}),r_0)-U(r_0)]+o_P(1),$$

where k_1 and U are defined in (68) and (58) respectively. Thus, using (72) and (19), we obtain

$$\sqrt{n}\{\mathrm{Q}_{n-h}(X_{1:n-h}+X_{h+1:n},\Phi)-Q(\Phi_{\sigma,+})\} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{IF}(X_i+X_{i+h},Q,\Phi_{\sigma,+})+o_P(1). \tag{74}$$

In the case (ii), where $D < 1/2$, we get from expansion (70) that

$$\beta(D)\frac{(n-h)^D}{\tilde{L}(n-h)}(\mathrm{Q}_{n-h}(\{X_{1:n-h}+X_{h+1:n}\}/Q(\Phi_{\sigma,+}),\Phi)-1)) = -c(\Phi)\beta(D)\frac{(n-h)^D}{\tilde{L}(n-h)}\frac{(T_1(F_{+,n-h})-T_1(\Phi))(r_0)}{(T_1(\Phi))'[r_0]}+o_P(1), \tag{75}$$

where $F_{+,n-h}$ denotes the empirical c.d.f. of $(\{X_i+X_{i+h}\}/Q(\Phi_{\sigma,+}))_{1\leq i\leq n-h}$.

Let us now focus on the autocovariances and consider first the case (i) where $D > 1/2$. Let us denote by $\gamma_-(k)$ the autocovariance of the process $(X_i-X_{i+h})_{i\geq 1}$ computed at lag k. Using a Taylor formula, $\gamma_-(k) = O(k^{-2-D+\epsilon})$, for ϵ in $(0,D)$ such that $L_i(x)/x^{\epsilon} = O(1)$, as x tends to infinity, for all $i = 0,1,2,3$. Let $F_{-,n-h}$ denote the empirical c.d.f. of $(X_i-X_{i+h})_{1\leq i\leq n-h}$. Since $\sum_k|\gamma_-(k)| < \infty$, the process $(X_i-X_{i+h})_{i\geq 1}$ satisfies Assumption (A1) implying that $\sqrt{n}(F_{-,n-h}-\Phi_{\sigma,-})$ converges in distribution to a Gaussian process in the space of cadlag functions equipped with the topology of uniform convergence (Csörgő and Mielniczuk, 1996). As a consequence, by Lemma 1, expansion (18) is valid for $\mathrm{Q}_{n-h}(X_{1:n-h}-X_{h+1:n},\Phi)$ with $a_n = \sqrt{n}$ where IF is defined in (20).

Then, in the case (i), using the delta method (Thm 3.1, p. 26, in van der Vaart, 1998) $\hat{\gamma}_Q(h,X_{1:n},\Phi)$ satisfies the following asymptotic expansion as in (67):

$$\sqrt{n-h}(\hat{\gamma}_Q(h,X_{1:n},\Phi)-\{Q^2(\Phi_{\sigma,+})-Q^2(\Phi_{\sigma,-})\}/4) = \frac{1}{\sqrt{n-h}}\sum_{i=1}^{n-h}\psi(X_i,X_{i+h})+o_P(1), \tag{76}$$

where ψ is defined in (30). Hence, we have to establish a central limit theorem for $(n-h)^{-1/2}\sum_{i=1}^{n-h}\psi(X_i,X_{i+h})$. Using Lemma 3, the definition of the Hermite rank given on p. 2245 in Arcones (1994) and Assumption (A2) with $D > 1/2$, we obtain that condition (2.40) of Thm 4 (p. 2256) in Arcones (1994) is satisfied with $\tau = 2$. This concludes the proof of (i) by observing from (7) that $\{Q^2(\Phi_{\sigma,+})-Q^2(\Phi_{\sigma,-})\}/4 = \mathbb{E}[X_1X_{1+h}] = \gamma(h)$.

Consider the case (ii) where $D < 1/2$. Using (12) and $\gamma(h) = [Q^2(\Phi_{\sigma,+})-Q^2(\Phi_{\sigma,-})]/4$, one has

$$\hat{\gamma}_Q(h,X_{1:n},\Phi)-\gamma(h) = A_n^+-A_n^-, \tag{77}$$

where

© 2010 Blackwell Publishing Ltd.

$$A_n^{\pm} = \frac{1}{4}[\mathrm{Q}_{n-h}(X_{1:n-h} \pm X_{h+1:n}, \Phi)^2 - Q^2(\Phi_{\sigma,\pm})].$$

We first show that the contribution of A_n^- is negligeable. Since the expansion (18) holds for $\sqrt{n-h}(\mathrm{Q}_{n-h}(X_{1:n-h} - X_{h+1:n}, \Phi) - Q(\Phi_{\sigma,-}))$, we conclude by arguing as in the proof of Theorem 1, that this expression is $O_P(1)$. Applying the delta method, we get the same type of result for Q_{n-h}^2, namely $\sqrt{n-h}(\mathrm{Q}_{n-h}(X_{1:n-h} - X_{h+1:n}, \Phi)^2 - Q^2(\Phi_{\sigma,-})) = O_P(1)$ and therefore, since $D < 1/2$,

$$\beta(D)\frac{(n-h)^{D-1/2}}{\tilde{L}(n-h)}\sqrt{n-h}\, A_n^- = o_P(1). \tag{78}$$

We now turn to A_n^+. Applying the delta method with the transformation $b(x) = x^2$ to (75) and using (73) yields

$$\beta(D)\frac{(n-h)^D}{\tilde{L}(n-h)}A_n^+ = -\frac{c(\Phi)\beta(D)}{2}\frac{(n-h)^D}{\tilde{L}(n-h)}Q^2(\Phi_{\sigma,+})\frac{(T_1(F_{+,n-h}) - T_1(\Phi))[r_0]}{(T_1(\Phi))'[r_0]} + o_P(1) \xrightarrow{d} \frac{Q^2(\Phi_{\sigma,+})}{4}(Z_{2,D}(1) - Z_{1,D}(1)^2).$$

The result follows from (77), (78) and $Q^2(\Phi_{\sigma,+}) = \mathrm{var}(X_1+X_h) = 2(\gamma(0) + \gamma(h))$. □

Proof of Proposition 4. The classical autocovariance estimator can be obtained from the classical scale estimator $\hat{\sigma}_{n,X}$ as in (12). More precisely, a straightforward calculation leads to

$$\hat{\gamma}(h) = \frac{1}{4}\left(\hat{\sigma}^2_{n-h,X_{1:n-h}+X_{h+1:n}} - \hat{\sigma}^2_{n-h,X_{1:n-h}-X_{h+1:n}}\right)(1 + o(1)) + O_P\left(\frac{1}{n^2}\right). \tag{79}$$

To alleviate the notations, $\hat{\sigma}_{n-h,X_{1:n-h}+X_{h+1:n}}$ will now be denoted by $\hat{\sigma}_+$ and $\hat{\sigma}_{n-h,X_{1:n-h}-X_{h+1:n}}$ by $\hat{\sigma}_-$.
On one hand, using Proposition 3 and the same arguments as in the beginning of the proof of Theorem 4, we have

$$\beta(D)\frac{n^D}{\tilde{L}(n)}(\hat{\sigma}_+ - \sigma_+) \xrightarrow{d} \frac{\sigma_+}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2),$$

where σ_+ denotes the standard deviation of $X_1 + X_{1+h}$ and $\tilde{L}(n)$ is defined in Theorem 4. Note that $\sigma_+^2 = 2(\gamma(0) + \gamma(h))$. By the classical delta method, we thus obtain

$$\beta(D)\frac{n^D}{\tilde{L}(n)}(\hat{\sigma}_+^2 - \sigma_+^2) \xrightarrow{d} 2(\gamma(0) + \gamma(h))(Z_{2,D}(1) - Z_{1,D}(1)^2). \tag{80}$$

On the other hand, by the same arguments as in Theorem 4, the process $(X_i - X_{i+h})_{i\geq 1}$ satisfies Assumption (A21). Let $\sigma_-^2 = 2(\gamma(0) - \gamma(h))$ denote the variance of $X_1 - X_{1+h}$. Then as in the proof of Proposition 1, $\sqrt{n}(\hat{\sigma}_-^2 - \sigma_-^2)$ converges in distribution. This implies that

$$\beta(D)\frac{n^D}{\tilde{L}(n)}\left(\hat{\sigma}_-^2 - \sigma_-^2\right) = o_P(1). \tag{81}$$

Using (79), (80) and (81), we get:

$$\beta(D)\frac{n^D}{\tilde{L}(n)}(\hat{\gamma}(h) - \gamma(h)) \xrightarrow{d} \frac{\gamma(0) + \gamma(h)}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2).$$

□

7. TECHNICAL LEMMAS

Lemma 2. *Let X be a standard Gaussian random variable. The influence function defined in* (20) *has the following properties:*

$$\mathbb{E}[\mathrm{IF}(X, Q, \Phi)] = 0, \tag{82}$$

$$\mathbb{E}[X\ \mathrm{IF}(X, Q, \Phi)] = 0, \tag{83}$$

$$\mathbb{E}[X^2\ \mathrm{IF}(X, Q, \Phi)] = (2\sqrt{\pi}\beta)^{-1}\exp(-1/(4c^2)) \neq 0, \tag{84}$$

where Φ is the c.d.f. of a standard Gaussian random variable, $c = c(\Phi)$ is defined in (17) *and $\beta = \int \phi(y)\phi(y + 1/c)\, dy$.*

© 2010 Blackwell Publishing Ltd.

Proof of Lemma 2. Let us first prove that $\mathbb{E}[\mathrm{IF}(X,Q,\Phi)] = 0$. It is enough to prove that $\mathbb{E}[\Phi(X+1/c) - \Phi(X-1/c)] = 1/4$. Using the definition of c, namely (65) or (59), we get:

$$\begin{aligned}\mathbb{E}[\Phi(X+1/c) - \Phi(X-1/c)] &= \int_{\mathbb{R}} (\Phi(x+1/c) - \Phi(x-1/c))\phi(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{\{|y-x|\le 1/c\}}\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y = T_1(\Phi)[1/c] = 1/4.\end{aligned} \tag{85}$$

Then, let us prove that $\mathbb{E}[X\mathrm{IF}(X,Q,\Phi)] = 0$. Since X has a standard Gaussian distribution, it suffices to prove that $\mathbb{E}[X\{\Phi(X+1/c) - \Phi(X-1/c)\}] = 0$. By symmetry of ϕ, we obtain:

$$\begin{aligned}\mathbb{E}[X\Phi(X+1/c)] = \int_{\mathbb{R}} x\Phi(x+1/c)\phi(x)\,\mathrm{d}x &= \int_{\mathbb{R}} x(1-\Phi(-x-1/c))\phi(x)\,\mathrm{d}x \\ &= -\int_{\mathbb{R}} x\Phi(-x-1/c)\phi(x)\,\mathrm{d}x = \mathbb{E}[X\Phi(X-1/c)].\end{aligned}$$

Finally, let us compute: $\mathbb{E}[X^2\mathrm{IF}(X,Q,\Phi)]$. Set $\beta = \int \phi(y)\phi(y+1/c)\,\mathrm{d}y$. By integrating by parts, using (85) and finally the symmetry of ϕ, we get

$$\begin{aligned}(\beta/c)\mathbb{E}[X^2\mathrm{IF}(X,Q,\Phi)] &= -\int_{\mathbb{R}}\left(\int_{y-1/c}^{y+1/c} x^2\phi(x)\,\mathrm{d}x\right)\phi(y)\,\mathrm{d}y + 1/4 \\ &= -\int_{\mathbb{R}}\{(y-1/c)\phi(y-1/c) - (y+1/c)\phi(y+1/c)\}\phi(y)\,\mathrm{d}y - \int_{\mathbb{R}}\left(\int_{y-1/c}^{y+1/c}\phi(x)\,\mathrm{d}x\right)\phi(y)\,\mathrm{d}y + 1/4 \\ &= \int_{\mathbb{R}}\{-(y-1/c)\phi(y-1/c) + (y+1/c)\phi(y+1/c)\}\phi(y)\,\mathrm{d}y,\end{aligned}$$

where the last equality comes from $\int_{\mathbb{R}}\left(\int_{y-1/c}^{y+1/c}\phi(x)\,\mathrm{d}x\right)\phi(y)\,\mathrm{d}y = T_1(\Phi)(1/c) = 1/4$. By symmetry of ϕ,

$$\begin{aligned}\int_{\mathbb{R}}\{-(y-1/c)\phi(y-1/c) + (y+1/c)\phi(y+1/c)\}\phi(y)\,\mathrm{d}y &= -2\int_{\mathbb{R}}(y-1/c)\phi(y-1/c)\phi(y)\,\mathrm{d}y \\ &= (2c\sqrt{\pi})^{-1}\exp(-1/(4c^2)),\end{aligned}$$

which concludes the proof. □

Lemma 3. *Let (X,Y) be a standard Gaussian random vector such that cov(X,Y) = 0 and let Φ_+ and Φ_- denote the c.d.f. of X + Y and X − Y respectively. The influence function ψ defined, for all x and y in $\mathbb{R}$, by*

$$\psi(x,y) = \frac{1}{2}\{Q(\Phi_+)\mathrm{IF}(x+y,Q,\Phi_+) - Q(\Phi_-)\mathrm{IF}(x-y,Q,\Phi_-)\},$$

satisfies the following properties:

$$\mathbb{E}[\psi(X,Y)] = 0, \tag{86}$$

$$\mathbb{E}[X\psi(X,Y)] = \mathbb{E}[Y\psi(X,Y)] = 0, \tag{87}$$

$$\mathbb{E}[XY\psi(X,Y)] \neq 0. \tag{88}$$

Proof of Lemma 3. Using (19), (82) and $Q(\Phi_\pm)^2 = \mathrm{var}(X \pm Y)$ [see eqn (8)], we get that

$$\mathbb{E}[\psi(X,Y)] = \frac{1}{2}\{Q(\Phi_+)^2 - Q(\Phi_-)^2\}\mathbb{E}[\mathrm{IF}(U,Q,\Phi)] = 0,$$

where U is a standard Gaussian random variable, which gives (86). Let us now prove (87).
First note that,

$$\mathbb{E}[X\psi(X,Y)] = \frac{1}{2}\{\mathbb{E}[(X+Y)\psi(X,Y)] + \mathbb{E}[(X-Y)\psi(X,Y)]\}.$$

But,

© 2010 Blackwell Publishing Ltd.

$$\begin{aligned}\mathbb{E}[(X+Y)\psi(X,Y)] &= \frac{1}{2}\mathbb{E}[Q(\Phi_+)^2(X+Y)\mathrm{IF}((X+Y)/Q(\Phi_+),Q,\Phi) - Q(\Phi_-)^2(X+Y)\mathrm{IF}((X-Y)/Q(\Phi_-),Q,\Phi)]\\ &= \frac{1}{2}[Q(\Phi_+)^3\mathbb{E}[U\mathrm{IF}(U,Q,\Phi) - Q(\Phi_-)^2Q(\Phi_+)\mathbb{E}[U\mathrm{IF}(V,Q,\Phi)],\end{aligned}$$

where $U = (X+Y)/Q(\Phi_+)$ and $V = (X-Y)/Q(\Phi_-)$ are independent standard Gaussian random variables. By (83), $\mathbb{E}[(X+Y)\psi(X,Y)] = 0$. In the same way, $\mathbb{E}[(X-Y)\psi(X,Y)] = 0$, which gives (87). Let us now prove (88). Using that $4XY = (X+Y)^2 - (X-Y)^2$, we get

$$\begin{aligned}8\mathbb{E}[XY\psi(X,Y)] &= \mathbb{E}[(X+Y)^2Q(\Phi_+)\mathrm{IF}(X+Y,Q,\Phi_+) + (X-Y)^2Q(\Phi_-)\mathrm{IF}(X-Y,Q,\Phi_-)]\\ &\quad - \mathbb{E}[(X-Y)^2Q(\Phi_+)\mathrm{IF}(X+Y,Q,\Phi_+) + (X+Y)^2Q(\Phi_-)\mathrm{IF}(X-Y,Q,\Phi_-)]\\ &= (Q(\Phi_+)^4 + Q(\Phi_-)^4)\mathbb{E}[U^2\mathrm{IF}(U,Q,\Phi)] - Q(\Phi_+)^2Q(\Phi_-)^2\left(\mathbb{E}[V^2\mathrm{IF}(U,Q,\Phi)] + \mathbb{E}[U^2\mathrm{IF}(V,Q,\Phi)]\right),\end{aligned} \tag{89}$$

where U and V are as above. The first term is non-zero by (84) whereas the second term is zero by independence of U and V and (82). This yields (88).

Lemma 4. *Let $\alpha_{p,q}(r) = \mathbb{E}[\mathbb{1}_{\{|X-Y|\le r\}}H_p(X)H_p(Y)]$ where X and Y are independent standard Gaussian random variables and H_p is the pth Hermite polynomial with leading coefficient equal to 1. Then,*

(i) $\alpha_{1,0}(r) = 0, \forall r \in \mathbb{R}$.
(ii) $\alpha_{2,0}(r) = -\alpha_{1,1}(r) = \dot{\phi}(r/\sqrt{2}), \forall r \in \mathbb{R}$.
(iii) *Moreover, there exists some positive η such as $\alpha_{2,0}(r) = -\alpha_{1,1}(r)$ is different from 0 when r is in $[r_0 - \eta; r_0 + \eta]$, where r_0 is defined in (10).*

Proof of Lemma 4. The proof of (i) follows from the symmetry of the Gaussian distribution and the proof of (ii) relies on the following identity: for all $r \in \mathbb{R}$,

$$\int_{\mathbb{R}} (\phi(x+r) - \phi(x-r))x\phi(x)\,\mathrm{d}x = \dot{\phi}(r/\sqrt{2}).$$

Let us now turn to the proof of (iii). $\dot{\phi}(r/\sqrt{2})$ is equal to zero only if $r = 0$. By (10), r_0 is such that $\Phi(r_0/\sqrt{2}) = 5/8$, and hence is different from 0. The existence of η follows from the continuity of $\dot{\phi}$. □

8. CONCLUSION

In this article, we studied the asymptotic properties of the robust-scale estimator Q_n (Rousseeuw and Croux, 1993) and of the robust autocovariance estimator $\hat{\gamma}_Q$ (Ma and Genton, 2000), for short-and long-range dependent processes. We showed that the asymptotic variance of these estimators is optimal, or close to it, and we verified, by using simulations, that these estimators are indeed robust in the presence of outliers. Complete proofs of the asymptotic properties of the robust-scale Q_n and the covariance estimator $\hat{\gamma}_Q$ are provided for Gaussian stationary processes. The central limit theorems for Q_n and $\hat{\gamma}_Q$ were obtained. In all cases, the rate of convergence of the estimators is $\sqrt{n}$, except for long-range dependent processes with $D \in (0,1/2)$, for which the rate is $n^D L(n)^{-1}$. Empirical Monte Carlo experiments were conducted so as to illustrate the finite sample size properties of the estimators. The robustness of Q_n and $\hat{\gamma}_Q$ were also investigated when the process contained outliers. The theoretical results and the empirical evidence strongly suggest the use of these estimators as an alternative to estimate the scale and the autocovariance structure of the process. The classical scale and autocovariance estimators were also considered as means of comparison. All estimators showed similar empirical accuracy when the data did not contain outliers. However, the classical estimators were significantly affected when additive outliers are present. The robust ones, however, were much less affected.

REFERENCES

Arcones, M. (1994) Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Annals of Probability* **22**, 2242–74.
Beran, J. (1992) Statistical methods for data with long-range dependence. *Statistical Science* **7**, 404–16.
Beran, J. (1994) *Statistics for Long-Memory Processes*, volume 61 of *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.
Borovkova, S., Burton, R. and Dehling, H. (2001) Limit theorems for functionals of mixing processes with applications to *U*-statistics and dimension estimation. *Transactions of the American Mathematical Society* **353**, 4261–318.
Breuer, P. and Major, P. (1983) Central limit theorems for nonlinear functionals of Gaussian fields. *Journal of Multivariate Analysis* **13**, 425–41.
Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods*, 2nd edn., Springer Series in Statistics. New York: Springer-Verlag.
Chan, W. (1992) A note on time series model specification in the presence of outliers. *Journal of Applied Statistics* **19**, 117–24.
Chan, W. (1995) Outliers and financial time series modelling: a cautionary note. *Mathematics and Computers in Simulation* **39**, 425–30.
Chareka, P., Matarise, F. and Turner, R. (2006) A test for additive outliers applicable to long-memory time series. *J. Econ. Dynam. Control* **30**, 595–21.
Croux, C. and Rousseeuw, P. (1992) Time-efficient algorithms for two highly robust estimators of scale. *Computational Statistics* **1**, 411–28.

© 2010 Blackwell Publishing Ltd.

Csörgő, S. and Mielniczuk, J. (1996) The empirical process of a short-range dependent stationary sequence under Gaussian subordination. *Probability Theory Related Fields* **104**, 15–25.

Dehling, H. and Taqqu, M. S. (1989) The empirical process of some long-range dependent sequences with an application to *U*-statistics. *Annals of Statistics* **17**, 1767–83.

Dehling, H. and Taqqu, M. S. (1991) Bivariate symmetric statistics of long-range dependent observations. *Journal of Statistical Planning and Inference* **28**, 153–65.

Deutsch, S., Richards, J. and Swain, J. (1990) Effects of a single outlier on ARMA identification. *Communications in Statistics: Theory and Methods* **19**, 2207–27.

Doukhan, P. and Oppenheim, G. and Taqqu, M. S. (eds) (2003) *Theory and Applications of Long-Range Dependence*. Boston: Birkhäuser Boston Inc.

Fajardo, M. F., Reisen, V. A. and Cribari-Neto, F. (2009) Robust estimation in long-memory processes under additive outliers. *Journal of Statistical Planning and Inference* **139**, 2511–2525.

Fox, R. and Taqqu, M. S. (1987) Multiple stochastic integrals with dependent integrators. *Journal of multivariate analysis* **21**, 105–27.

Gnanadesikan, R. and Kettenring, J. R. (1972) Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* **28**, 81–124.

Hosking, J. R. (1981) Fractional differencing. *Biometrika* **68**, 165–76.

Huber, P. J. (1981) *Robust Statistics*, Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons Inc.

Lévy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S. and Reisen, V. A. (2009) Asymptotic properties of *U*-processes under long-range dependence. Technical Report, arXiv:0912.4688v1.

Ma, Y. and Genton, M. (2000) Highly robust estimation of the autocovariance function. *Journal of Time Series Analysis* **21**, 663–84.

Maronna, R. A., Martin, R. D. and Yohai, V. J. (2006) *Robust Statistics*, Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd.

Robinson, P. M. (1995) Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* **23**, 1630–61.

Rousseeuw, P. and Croux, C. (1993) Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88**, 1273–83.

Taqqu, M. S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **31**, 287–302.

Toussoun, O. (1925) Mémoire sur l'Histoire du Nil, Mémoires présentées á l'Institut d'Egypte.

van der Vaart, A. W. (1998) *Asymptotic Statistics, vol. 3 of Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press.

© 2010 Blackwell Publishing Ltd.