L^p-WAVELET REGRESSION WITH CORRELATED ERRORS AND INVERSE PROBLEMS

Rafał Kulik and Marc Raimondo

University of Ottawa and University of Sydney

Supplementary material

This note contains the proofs of Theorems 3.1 and 3.2. The proofs are based on the maxiset theorem from Kerkyacharian and Picard (2000). The steps are similar to those of Johnstone, Kerkyacharian, Picard and Raimondo (2004). The technical novelties appear in moment bounds and large deviation results for wavelet coefficients

$$\hat{\beta}_\kappa^D := \frac{1}{n}\sum_{i=1}^n \psi_\kappa(u_i)Y_i,$$

$$\hat{\beta}_\kappa^C := \int \psi_\kappa(t)dY_t.$$

which we establish under LRD assumption.

S1 Maxiset Theorem

The following theorem is borrowed from Kerkyacharian and Picard (2000). We refer to section S3 for the definition of Temlyakov property. First, we introduce some notation: μ will denote the measure such that for $j \in \mathbb{N},\ k \in \mathbb{N}$,

$$\mu\{(j,k)\} = \|\sigma_j \psi_{j,k}\|_p^p = \sigma_j^p 2^{j(\frac{p}{2}-1)}\|\psi\|_p^p, \tag{S1.2}$$

$$l_{q,\infty}(\mu) = \left\{ f,\ \sup_{\lambda>0} \lambda^q \mu\{(j,k):\ |\beta_{j,k}| > \sigma_j\lambda\} < \infty \right\}.$$

Theorem S1.1 *Let $p > 1$, $0 < q < p$, $\{\ \psi_{j,k}, j \geq -1, k = 0,1,...,2^j\}$ be a periodised wavelet basis of $L^2(\mathcal{I})$ and σ_j be a positive sequence such that the heteroscedastic basis $\sigma_j\psi_{j,k}$ satisfies Temlyakov property. Suppose that Λ_n is a set of pairs (j,k) and c_n is a deterministic sequence tending to zero with*

$$\sup_n \mu\{\Lambda_n\}\, c_n^p < \infty. \tag{S1.3}$$

If for any n and any pair $\kappa = (j,k) \in \Lambda_n$, we have

$$\mathrm{E}|\hat{\beta}_\kappa - \beta_\kappa|^{2p} \leq C\,(\sigma_j\, c_n)^{2p} \tag{S1.4}$$

$$P\Big(|\hat{\beta}_\kappa - \beta_\kappa| \geq \eta\,\sigma_j\, c_n/2\Big) \leq C\,(c_n^{2p} \wedge c_n^4) \tag{S1.5}$$

for some positive constants η and C then, the wavelet based estimator

$$\hat{f}_n = \sum_{\kappa\in\Lambda_n} \hat{\beta}_\kappa\,\psi_\kappa\,\mathbb{I}\{|\hat{\beta}_\kappa| \geq \eta\,\sigma_j\,c_n\} \tag{S1.6}$$

is such that, for all positive integers n,

$$\mathrm{E}\|\hat{f}_n - f\|_p^p \leq C\,c_n^{p-q},$$

if and only if :

$$f \in l_{q,\infty}(\mu), \quad \textit{and}, \tag{S1.7}$$

$$\sup_n c_n^{q-p} \quad \| \quad f - \sum_{\kappa\in\Lambda_n} \beta_\kappa\psi_\kappa\|_p^p < \infty. \tag{S1.8}$$

This theorem identifies the 'Maxiset' of a general wavelet estimator of the form (S1.6). This is done by using conditions (S1.7) and (S1.8) for an appropriate choice of q. In the proof of the theorems we will choose q according to the dense or sparse regime by setting:

$$q = q_d := \frac{\alpha p}{2s+\alpha}, \quad \text{when} \quad s \geq \frac{\alpha}{2}\Big(\frac{p}{\pi} - 1\Big) \tag{S1.9}$$

$$q = q_s := \frac{\frac{\alpha p}{2} - 1}{s - \frac{1}{\pi} + \frac{\alpha}{2}}, \qquad \text{when} \quad s < \frac{\alpha}{2}\Big(\frac{p}{\pi} - 1\Big). \tag{S1.10}$$

S2 Moment bounds and large deviation estimates

S2.1 FBM model

Here $\hat{\beta}_\kappa = \hat{\beta}_\kappa^C$. In what follows C denotes a generic constant which does not depends on n but may change from line to line. Recall that

$$\hat{\beta}_\kappa^C = \beta_\kappa + \varepsilon^\alpha \sigma_j z_\kappa,$$

where, as in Wang (1996), $\sigma_j^2 = \mathrm{Var}(\int \psi_\kappa(t)dB_H(t))$ and z_κ are (weakly) correlated Gaussian random variables with variance 1 and $\sigma_j = \tau 2^{-j(1-\alpha)/2}$. It follows that $\mathrm{E}\hat{\beta}_\kappa = \beta_\kappa$ and

$$\mathrm{Var}\hat{\beta}_\kappa = \mathrm{Var}\left(\varepsilon^\alpha \int \psi_\kappa(t)dB_H(t)\right) = n^{-\alpha}2^{-j(1-\alpha)}\tau^2 \leq C\sigma_j^2\, c_n^2.$$

Since the rv's $\hat\beta_\kappa - \beta_\kappa$ are Gaussian higher moments bound (S1.4) follows from the previous inequality. Similarly,

$$\Pr\left(|\hat\beta_\kappa - \beta_\kappa| > \eta\sigma_j\, c_n/2\right) \le \exp\left(-\log n \frac{\eta^2}{8}\right) \le C\,(c_n^{2p} \wedge c_n^4)$$

provided $\eta > \sqrt{8\alpha}\sqrt{p \vee 2}$. Which proves (S1.5).

S**2.2** Discrete model

Here $\hat\beta_\kappa = \hat\beta_\kappa^D$. Write

$$\begin{aligned}\hat\beta_\kappa - \beta_\kappa &= \hat\beta_\kappa - \mathrm{E}\hat\beta_\kappa + \mathrm{E}\hat\beta_\kappa - \beta_\kappa\\ &= \frac{1}{n}\sum_{i=1}^n X_i\psi_\kappa(u_i) + \left(\frac{1}{n}\sum_{i=1}^n f(u_i)\psi_\kappa(u_i) - \beta_\kappa\right).\end{aligned}$$

The main tool to derive rates of convergence is the following lemma. To establish moments bounds we do not assume that X_i's are Gaussian. These estimates may be of independent interest.

Lemma S**2.1** *For each fixed j and k, and $p > 1$,*

$$\mathrm{E}(\hat\beta_\kappa - \beta_\kappa)^2 \sim 2^{-j(1-\alpha)} n^{-\alpha}\tau_D^2, \tag{S2.11}$$

$$\mathrm{E}\left|\hat\beta_\kappa - \beta_\kappa\right|^p = O\left(n^{-\alpha p/2} 2^{-jp(1-\alpha)/2}\right). \tag{S2.12}$$

If moreover X_i's are Gaussian, then for all $\lambda > n^{-1}$,

$$\Pr\left(|\hat\beta_\kappa - \beta_\kappa| > \lambda\right) \le \frac{n^{-\alpha/2}2^{-j(1-\alpha)/2}}{\lambda}\exp\left(-\frac{\lambda^2}{2(n^{-\alpha}2^{-j(1-\alpha)}\tau_D^2)}\right). \tag{S2.13}$$

To prove this lemma we will replace β_κ with $\hat\beta_\kappa$ and use $|\mathrm{E}\hat\beta_\kappa - \beta_\kappa| = O(n^{-1})$. (Note that this just the distance between the integral $\int f(x)\psi_\kappa(x)\,\mathrm{d}x$ and the Riemann-Stjeltjes sum.

Proof:

Note that

$$\sum_{i=1}^n \psi_\kappa^2(u_i) = 2^j \sum_{i=1}^n \psi^2\left(2^j\frac{i}{n} - k\right) = 2^j n \int_0^1 \psi^2(2^j x)dx + o(n) = n + o(n). \tag{S2.14}$$

Bearing in mind that $\mathrm{Var}(X_i) = \mathrm{E}(X_1^2) = 1$ we have:

$$\begin{aligned}\mathrm{E}(\hat\beta_\kappa - \mathrm{E}\hat\beta_\kappa)^2 &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\psi_\kappa(u_i)\right)\\ &= \frac{1}{n^2}\left(\sum_{i=1}^n \psi_\kappa^2(u_i) + \sum_{i\neq l}\psi_\kappa(u_i)\psi_\kappa(z_l)\mathrm{Cov}(X_i, X_l)\right).\end{aligned}$$

By (S2.14) above, the first part is of order $n^{-1}+o(n^{-1})$. For the second part we have

$$
\begin{aligned}
&\sum_{i\neq l}\psi_\kappa(u_i)\psi_\kappa(z_l)\mathrm{Cov}(X_i,X_l)\\
&= \sum_{i\neq l}2^j|i-l|^{-\alpha}\psi\left(2^j\frac{i}{n}-k\right)\psi\left(2^j\frac{l}{n}-k\right)\\
&= L2^jn^{-\alpha}\sum_{i\neq l}\left|\frac{i}{n}-\frac{l}{n}\right|^{-\alpha}\psi\left(2^j\frac{i}{n}-k\right)\psi\left(2^j\frac{l}{n}-k\right),
\end{aligned}
$$

which behaves asymptotically as $2^{-j(1-\alpha)}n^{2-\alpha}\tau_D^2$.

Further, the first part dominates the second one if and only if $2^j>n$, which is not possible. Thus (S2.11) follows.

To prove (S2.12), let

$$
b_r=\sum_{i=r}^n a_{i-r}\psi_\kappa(u_i),\qquad r=1,\dots,n,
$$

$$
b_r=\sum_{i=1}^n a_{i-r}\psi_\kappa(u_i),\qquad r=-\infty,\dots,0.
$$

Also, note that by (S2.11),

$$
v_n^2:=\mathrm{Var}\left(\sum_{r=-\infty}^n\epsilon_rb_r\right)=\sum_{r=-\infty}^n b_r^2=\mathrm{Var}\left(\sum_{i=1}^nX_i\psi_\kappa(u_i)\right)
$$

and thus

$$
v_n^2/(n^{2-\alpha}2^{-j(1-\alpha)}\tau_D^2)\to 1 \tag{S2.15}
$$

as $n\to\infty$.

Note now that each Gaussian sequence can be represented as

$$
X_i=\sum_{m=0}^\infty a_m\epsilon_{i-m},\quad i\geq 1,
$$

where a_m is a regularly varying sequence with index $-(\alpha+1)/2$ and $\{\epsilon_i,i\geq 1\}$ is a centered sequence of i.i.d. random variables. Via Rosenthal inequality, for $p\geq 2$

$$
\begin{aligned}
\mathrm{E}\left|\hat\beta_\kappa-\mathrm{E}\hat\beta_\kappa\right|^p &= \mathrm{E}\left|\frac{1}{n}\sum_{i=1}^nX_i\psi_\kappa(u_i)\right|^p\\
&= n^{-p}\mathrm{E}\left|\sum_{m=0}^\infty a_m\sum_{i=1}^n\epsilon_{i-m}\psi_\kappa(u_i)\right|^p=n^{-p}\mathrm{E}\left|\sum_{r=-\infty}^n\epsilon_rb_r\right|^p
\end{aligned}
$$

$$
\begin{aligned}
&\leq n^{-p}\left(\sum_{r=-\infty}^{n} b_r^2\right)^{p/2} + n^{-p}\sum_{r=-\infty}^{n}|b_r|^p \\
&\leq n^{-p}\left(\sum_{r=-\infty}^{n} b_r^2\right)^{p/2} + n^{-p}\sup_r |b_r|^{p-2}\sum_{r=-\infty}^{n} b_r^2 \\
&= n^{-p}O\left(\left(n^{2-\alpha}2^{-j(1-\alpha)}\right)^{p/2}\right) + n^{-p}n^{p/2-1}O\left(n^{2-\alpha}2^{-j(1-\alpha)}\right) \\
&= O\left(n^{-\alpha p/2}2^{-jp/2(1-\alpha)} + n^{1-\alpha-p/2}2^{-j(1-\alpha)}\right).
\end{aligned}
$$

The second term is negligible for all j such that $2^j \leq n$.

To prove (S2.13) note that $\sum_{i=1}^{n} X_i\psi_\kappa(u_i) \sim \mathcal{N}(0, v_n^2)$. Thus,

$$
\Pr\left(|\hat{\beta}_\kappa - \mathrm{E}\hat{\beta}_\kappa| > \lambda\right) \leq C\frac{v_n}{n\lambda}\exp\left(-\frac{n^2\lambda^2}{2v_n^2}\right).
$$

and the result follows by (S2.15). $\odot$

Consequently,

$$
\mathrm{E}\left|\hat{\beta}_\kappa - \beta_\kappa\right|^p = O\left(n^{-\alpha p/2}2^{-jp/2(1-\alpha)}\right) = O(\sigma_j^p c^p(n))
$$

and taking $\lambda = \lambda_j = \eta\sigma_j c_n$,

$$
\Pr\left(|\hat{\beta}_\kappa - \beta_\kappa| > \eta\sigma_j c_n/2\right) \leq \exp\left(-\log n\frac{\eta^2}{8}\right) = O(c_n^{2p})
$$

provided $\eta > \sqrt{8p\alpha}$. The similar argument applies to $1 < p < 2$. In this case we require $\eta > \sqrt{16\alpha}$.

S3 Temlyakov property

As seen in Johnstone, Kerkyacharian, Picard and Raimondo (2004, appendix A), the basis $(\sigma_j\psi_{j,k}(.))$ satisfies Temlyakov property as soon as

$$
\sum_{\Lambda_n} 2^j\,\sigma_j^2 \leq C\sup_{\Lambda_n}\left(2^j\sigma_j^2\right),
$$

and

$$
\sum_{\Lambda_n} 2^{jp/2}\,\sigma_j^p \leq C\sup_{\Lambda_n}\left(2^{jp/2}\sigma_j^p\right), \quad 1 \leq p < 2,
$$

which is clearly satisfied when $\sigma_j^2 = \tau^2\,2^{-j(1-\alpha)}$.

S4 Fine resolution tuning

Here we check that condition (S1.3) is satisfied. Using (S1.2),

$$\mu(\Lambda_n) = \sum_{j \le j_1} \sum_{k=0}^{2^j-1} \mu(j,k) = \sum_{j \le j_1} 2^j \mu(j,k) = \tau^p \sum_{j \le j_1} 2^j \, 2^{j(\frac{p}{2}-1-\frac{p(1-\alpha)}{2})} = O(2^{\frac{j_1}{2}}),$$

which with the choice of j_1 and $p > 1$ yields

$$\mu(\Lambda_n) c_n^p = \Big(\frac{n}{\log n}\Big)^{\frac{\alpha}{2}} \Big(\frac{(\log n)^{\frac{1}{2}}}{n^{\frac{\alpha p}{2}}}\Big) = O\Big(c_n^{p-1} (\frac{\log n}{(\log n)^{\alpha}})^{\frac{1}{2}}\Big) = o(1),$$

which shows that condition (S1.3) is satisfied.

S5 Besov embedding and Maxiset condition

S5.1 Part I

For both the dense (S1.9) and sparse (S1.10) regime, we look for a Besov scale δ such that

$$\mathcal{B}^{\delta}_{\pi,r} \subseteq l_{q,\infty}.$$

As usual we note that it is easier to work with

$$l_q(\mu) = \left\{ f \in L_p : f = \sum_{j,k \in A_j} \frac{|\beta_{jk}|^q}{\sigma_j^q} \, \|\sigma_j \psi_{j,k}\|_p^p < \infty \right\},$$

where A_j is a set of cardinality proportional to 2^j. Since $\|\sigma_j \psi_{j,k}\|_p^p = \sigma_j^p \, 2^{j(\frac{p}{2}-1)} = 2^{j(\frac{\alpha p}{2}-1)}$, we see that $f \in l_q(\mu)$ if

$$\sum_{j \ge 0} 2^{j \frac{(\alpha p - 2 + (1-\alpha) q)}{2}} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^q = \sum_{j \ge 0} 2^{jq \left[\frac{\alpha(p-q)}{2q} + \frac{1}{2} - \frac{1}{q}\right]} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^q < +\infty.$$

The latter condition holds when $f \in \mathcal{B}^{\delta}_{q,q}$ for

$$\delta = \frac{\alpha}{2}\Big(\frac{p}{q} - 1\Big). \tag{S5.16}$$

Now depending on whether we are in the dense (S1.9) or sparse phase (S1.10) we look for s and π such that

$$\mathcal{B}^{s}_{\pi,r} \subseteq \mathcal{B}^{\delta}_{q,q}. \tag{S5.17}$$

The dense phase. By definition (S1.9) of $q = q_d$ we have $\pi \ge q_d$. Hence (S5.17) follows from (S5.21) as long as $s \ge \delta = \frac{\alpha}{2}(\frac{p}{q} - 1)$ which is always true under the dense regime where $q = q_d$. Note that $\delta = \frac{\alpha}{2}(\frac{p}{q_d} - 1) = s$, thus automatically $\delta > 0$.

The sparse phase. Take $q = q_s$ and $\delta = \frac{\alpha}{2}\left(\frac{p}{q_s} - 1\right) = \alpha\frac{sp - \frac{p}{\pi} + 1}{\alpha p - 2}$. We consider two cases. If $\pi > q_s$ we use the embedding (S5.21). We have to check that $s > \alpha\frac{sp-\frac{p}{\pi}+1}{\alpha p-2}$ which is equivalent to $s < \frac{\alpha}{2}\left(\frac{p}{\pi} - 1\right)$, which is true in the sparse case. Further, we must guarantee that $\delta > 0$ which leads to the two conditions i) $p > 2/\alpha$ and $s > \frac{1}{\pi} - \frac{1}{p}$ or ii) $p < 2/\alpha$ and $s < \frac{1}{\pi} - \frac{1}{p}$. However, the last one is not relevant since we have $s > \frac{1}{\pi}$. Thus we established (S5.17) for $q_s < \pi < q_d$.

If $\pi < q_s$ we introduce a new Besov scale s' and index $q = q_s$ such that

$$s - \frac{1}{\pi} = s' - \frac{1}{q}, \quad s' = \frac{\alpha}{2}\Big(\frac{p}{q} - 1\Big). \tag{S5.18}$$

In this case, (S5.22) and (S5.16) ensures that

$$\mathcal{B}^s_{\pi,r} \subseteq \mathcal{B}^{s'}_{q,q} \equiv l_q(\mu),$$

as had to be proved. Solving (S5.18) yields definition (S1.10) of q under the sparse regime.

S5.2 Part II

First we look for a Besov scale δ such that for any $f \in \mathcal{B}^\delta_{p,r}$ the maxiset condition (S1.8) is satisfied. We have

$$c_n^{q-p}\|f - \sum_{\kappa\in\Lambda_n}\beta_\kappa\Psi_\kappa\|_p^p = c_n^{q-p}\,2^{-j_1\delta p}\,\|f\|_{\mathcal{B}^\delta_{p,r}} = O\Big(c_n^{q-p+2\delta p}\,(\frac{(\log n)^\alpha}{\log n})^{\delta p}\Big).$$

Thus condition (S1.7) holds for any $f \in \mathcal{B}^\delta_{p,r}$ if

$$\delta = \frac{1}{2}(1 - \frac{q}{p}). \tag{S5.19}$$

Now we look for s and π such that

$$\mathcal{B}^s_{\pi,r} \subseteq \mathcal{B}^\delta_{p,r}. \tag{S5.20}$$

To answer this question, we will use two different types of Besov embedding, depending on whether $\pi \geq p$ or $\pi < p$. We recall that

$$\mathcal{B}^s_{\pi,r} \subseteq \mathcal{B}^{s''}_{p,r}, \quad \text{provided that } \pi \geq p, \text{ and } s \geq s''. \tag{S5.21}$$

$$\mathcal{B}^s_{\pi,r} \subseteq \mathcal{B}^{s''}_{p,r}, \quad \text{provided that } \pi < p, \text{ and } s - \frac{1}{\pi} = s'' - \frac{1}{p}. \tag{S5.22}$$

The case $\pi \geq p$. We note that in this case we are always in the dense phase since s must be non-negative. Here we use (S5.21) with $s'' = \delta$ at (S5.19). Hence we see that (S5.20) holds as long as $s \geq \frac{1}{2}(1 - \frac{q}{p})$. Using definition (S1.9) of $q = q_d$ this will happen when $s \geq \frac{1-\alpha}{2}$.

The dense case when $\pi < p$. Here we introduce a new Besov scale s'' such that $s - \frac{1}{\pi} = s'' - \frac{1}{p}$ and use embedding (S5.22). For (S5.20) to hold in the dense case we need $s'' \geq \delta$ for $q = q_d$ at (S1.9), we obtain the following condition:

$$s \geq \frac{2}{2s+\alpha} + \frac{1}{\pi} - \frac{1}{p}.$$

The sparse case when $\pi < p$. Here we introduce a new Besov scale s'' such that $s - \frac{1}{\pi} = s'' - \frac{1}{p}$ and use embedding (S5.22). For (S5.20) to hold in the sparse case we need $s'' \geq \delta$ for $q = q_s$ at (S1.10), we obtain the following condition:

$$s > \frac{1}{\pi} - \frac{\alpha}{2}$$

which is always true since $s > \frac{1}{\pi}$.

S6 Final step

Recall

$$\gamma = \frac{\alpha s p}{2(s + \frac{\alpha}{2})}, \quad \text{if } s \geq \frac{\alpha}{2}(\frac{p}{\pi} - 1), \tag{S6.23}$$

$$\gamma = \frac{\alpha p(s - \frac{1}{\pi} + \frac{1}{p})}{2(s - \frac{1}{\pi} + \frac{\alpha}{2})}, \quad \text{if } \frac{1}{\pi} < s < \frac{\alpha}{2}(\frac{p}{\pi} - 1). \tag{S6.24}$$

The proof(s) are a direct application of Theorem S1.1 with our choice of σ_j, c_n and η. Combining results of sections S3,..., S5.2 we see that all the assumptions Theorem S1.1 are satisfied. Using the embedding results of Section S5.1 we derive rate exponent (S6.23) for any $f \in \mathcal{B}^s_{\pi,r}$ from definition (S1.9) of q when $s \geq \frac{\alpha}{2}(\frac{p}{\pi} - 1)$. Finally we derive rate exponent (S6.24) for any $f \in \mathcal{B}^s_{\pi,r}$ using definition (S1.10) of q when $\frac{1}{\pi} - \frac{1}{p} < s < \frac{\alpha}{2}(\frac{p}{\pi} - 1)$.