Journal of Computational Mathematics
Vol.40, No.1, 2022, 108–126.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2007-m2020-0109

ADAPTIVE AND OPTIMAL POINT-WISE ESTIMATIONS FOR DENSITIES IN GARCH-TYPE MODEL BY WAVELETS*

Cong Wu, Jinru Wang and Xiaochen Zeng[1)]
College of Mathematics, Faculty of Science, Beijing University of Technology, Beijing 100124, China
E-mails: wuc@hbut.edu.cn, wangjinru@bjut.edu.cn, zengxiaochen@bjut.edu.cn

Abstract

This paper considers adaptive point-wise estimations of density functions in GARCH-type model under the local Hölder condition by wavelet methods. A point-wise lower bound estimation of that model is first investigated; then we provide a linear wavelet estimate to obtain the optimal convergence rate, which means that the convergence rate coincides with the lower bound. The non-linear wavelet estimator is introduced for adaptivity, although it is nearly-optimal. However, the non-linear wavelet one depends on an upper bound of the smoothness index of unknown functions, we finally discuss a data driven version without any assumptions on the estimated functions.

Mathematics subject classification: 42C40, 62G07, 62G20.
Key words: Wavelets, Point-wise risk, Thresholding, Data-driven, GARCH-type model.

1. Introduction

The density estimation for the GARCH-type (Generalized Autoregressive Conditionally Heteroskedastic) model plays an important role in both statistics and econometrics [5]. In this current paper, we consider that density estimation model which can be described by the following mathematical model:

$$Y = XZ, \tag{1.1}$$

where X and Z are independent random variables. More precisely, the unknown density function f of X is to be estimated and *supp* $f \subseteq [0,1]$, while the density of Z is known. In general, we suppose that

$$Z := \prod_{i=1}^{v} U_i,$$

where v is a positive integer and $U_1, \cdots, U_v$ are independent and identically distributed (i.i.d.) random variables with uniform distribution $U_{(0,1)}$. When $v = 1$, model (1.1) reduces to the standard multiplicative concerning model, for which is sometimes called generalized multiplicative censoring model [1–3, 18, 19]. The purpose is to find an estimator $\widehat{f}_n$ based on the i.i.d. observed data $Y_1, \cdots, Y_n$ of Y approximating the unknown density f in some sense.

For the GARCH-type model, lots of literatures have been done the density estimations over L^p-risk by wavelet methods [4, 6–9, 16]. Asymptotic properties of the kernel estimators for a density derivative have been considered earlier in [15], while the performance of wavelet estimator was discussed in [16]. In 2012, Chesneau & Doosti [9] investigated the wavelet estimation of

* Received April 28, 2020 / Accepted July 20, 2020 /
Published online August 13, 2021 /
1) Corresponding author

a density in GARCH model under various dependence structures. One year later, Chesneau [8] provided the upper bounds over L^2-risk of wavelet density estimation for GARCH-type model. Rao [16] considered L^2-risk estimation for the derivative of a density in GARCH-type model over Besov balls by wavelets in 2017. After two years, Cao & Wei [4] extended Rao's result to L^p-risk ($1 \le p < \infty$).

In contrast to the above L^p-risk estimation, we consider point-wise risk estimations for model (1.1) in this paper, because it is more concerned in some applications. For a density function set Σ, the maximal point-wise risk at $x \in \mathbb{R}$ over Σ means that

$$R_{p,n}(\widehat{f}_n, \Sigma, x) := \sup_{f \in \Sigma} \left[E \left| \widehat{f}_n(x) - f(x) \right|^p \right]^{\frac{1}{p}}$$

with $1 \le p < \infty$ and EX being the expectation of X. An estimator $\widehat{f}_n^*$ is said to be the optimal over Σ, if

$$R_{p,n}(\widehat{f}_n^*, \Sigma, x) \lesssim \inf_{\widehat{f}_n} R_{p,n}(\widehat{f}_n, \Sigma, x),$$

where the infimum runs over all possible estimator of $f \in \Sigma$. Here and throughout, $A \lesssim B$ denotes $A \le cB$ for some independent constant $c > 0$; $A \gtrsim B$ means $B \lesssim A$; $A \sim B$ stands for both $A \lesssim B$ and $A \gtrsim B$. Clearly,

$$R_{p,n}(\widehat{f}_n^*, \Sigma, x) \ge \inf_{\widehat{f}_n} R_{p,n}(\widehat{f}_n, \Sigma, x)$$

holds automatically.

It is more reasonable to estimate $f(x_0)$ (for fixed $x_0 \in \mathbb{R}$) under some smoothness of f in a neighborhood Ω_{x_0} of x_0 instead of $\mathbb{R}$. For a function f on $\mathbb{R}^d$ and $x_0 \in \mathbb{R}^d$, we introduce the local Hölder condition of order s ($0 < s \le 1$) at x_0 in the sense that with $C > 0$ and Ω_{x_0},

$$|f(y) - f(x)| \le C|y - x|^s \tag{1.2}$$

holds for each $x, y \in \Omega_{x_0}$. We use $H^s(\Omega_{x_0})$ to denote all those functions satisfying (1.2) with a fixed constant $C > 0$.

For $s = N + \delta$ with $\delta \in (0, 1]$ and $N \in \mathbb{N}$, define

$$H^s(\Omega_{x_0}) := \left\{ f,\ f^{(N)} \in H^\delta(\Omega_{x_0}) \right\}.$$

More properties and advantages of the local Hölder space $H^s(\Omega_{x_0})$ can be found in Ref. [14,20]. Moreover, $H^s(\Omega_{x_0},\ M)$ stands for

$$H^s(\Omega_{x_0},\ M) := \{ f \in H^s(\Omega_{x_0}),\ f \text{ is a density and } \|f\|_\infty \le M \}.$$

Obviously, $H^s(\Omega_{x_0},\ M) \subset L^2(\mathbb{R})$, because

$$\int_{\mathbb{R}} |f(x)|^2 dx \le \|f\|_\infty \|f\|_1 \le M \|f\|_1 < \infty.$$

This paper is organized as follows. In Section 2, we shall give a lower bound estimation of a density in GARCH-type model. Let $\widehat{f}_n$ be an estimator of a function in $H^s(\Omega_{x_0},\ M)$. Then with $1 \le p < \infty$,

$$\sup_{x \in \Omega_{x_0}} \inf_{\widehat{f}_n} \sup_{f \in H^s(\Omega_{x_0},\ M)} \left[E \left| \widehat{f}_n(x) - f(x) \right|^p \right]^{\frac{1}{p}} \gtrsim n^{-\frac{s}{2s+2v+1}}.$$

Section 3 shows firstly the linear wavelet estimator $\widehat{f}_n^{lin}$ attaining the optimal convergence rate $n^{-\frac{s}{2s+2v+1}}$ on $H^s(\Omega_{x_0},\ M)$. Due to the non-adaptivity of $\widehat{f}_n^{lin}$, the non-linear wavelet estimation is then provided, which turns out to be adaptive and nearly-optimal (optimal up to a $\ln n$ factor). Because the non-linear wavelet estimator depends on an upper bound of the unknown parameter s, we finally consider a data driven version in the last section, which is motivated by the work of Rebelles [17] and Goldonshluger & Lepski [11], and has a better convergence rate than that of the non-linear wavelet one.

2. Lower Bound

We shall provide a point-wise lower bound estimation for densities in model (1.1) in this section. Firstly, the relationship between the densities of X and Y should be clarified. By the assumptions on Z, the density function of Z is

$$f_Z(z) = \frac{1}{(v-1)!}(-\ln z)^{v-1}, \qquad z \in (0,1). \tag{2.1}$$

In fact, (2.1) can be proved by induction. Clearly, $f_Z(z) = 1$ with $z \in (0,1)$ for $v = 1$ ($Z = U_1 \sim U_{(0,1)}$). Moreover, when $v \geq 2$, $Z = \prod_{i=1}^{v} U_i = \prod_{i=1}^{v-1} U_1 \cdot U_v$. Let

$$f_{U_1 \cdots U_{v-1}}(x) = \frac{1}{(v-2)!}(-\ln x)^{v-2} \qquad (0 < x < 1).$$

Then for $z \in (0,1)$,

$$\begin{aligned} P\{Z \leq z\} = P\left\{\prod_{l=1}^{v-1} U_l \cdot U_v \leq z\right\} &= \int_0^1 \int_0^z f_{U_1 \cdots U_{v-1}}(x)dxdy + \int_z^1 \int_0^{\frac{z}{x}} f_{U_1 \cdots U_{v-1}}(x)dydx \\ &= \int_0^z f_{U_1 \cdots U_{v-1}}(x)dx + \int_z^1 \frac{z}{x} f_{U_1 \cdots U_{v-1}}(x)dx. \end{aligned}$$

Hence,

$$\begin{aligned} f_Z(z) &= \int_z^1 \frac{1}{x} f_{U_1 \cdots U_{v-1}}(x)dx \\ &= -\int_z^1 \frac{1}{(v-2)!}(-\ln x)^{v-2} d(-\ln x) = \frac{1}{(v-1)!}(-\ln z)^{v-1} \end{aligned}$$

with $z \in (0,1)$.

As in [4, 8, 16], we assume that there exists a positive constant C_* such that

$$\sup_{x \in [0,1]} f_Y(x) \leq C_*. \tag{2.2}$$

For any $x \in [0,1]$, $h \in C^k([0,1])$, define

$$\begin{aligned} &T(h)(x) := (xh(x))' = h(x) + xh'(x), \qquad T_k(h)(x) := T(T_{k-1}(h))(x), \\ &G(h)(x) := -xh'(x), \qquad G_k(h)(x) := G(G_{k-1}(h))(x), \end{aligned}$$

where k is a positive integer. Then the following lemma holds, which can be found in Ref. [4,8].

Lemma 2.1 ([4, 8]). *Let G and T be defined as above. Then*

(i). $f_X(x) = G_v(f_Y)(x),\ x \in [0,1]$;

(ii). *For any* $h \in C^v([0,1])$, $\displaystyle\int_0^1 f_X(x)h(x)dx = \int_0^1 f_Y(x)T_v(h)(x)dx.$

To show a lower bound estimation of density functions, we introduce and prove the following proposition.

Proposition 2.1. *Let Ψ be a density set and d be a distance function on $\Psi \times \Psi$. If there exist $f_0, f_n \in \Psi$ (n is sample size in estimator $\widehat{f}_n$) and $f_{0,Y}, f_{n,Y}$ are the corresponding densities of Y in GARCH-type model such that*

(i). $f_{0,Y}(x) > 0$ *for* $x \in (0,1)$;

(ii). $d(f_n, f_0) \geq a_n > 0$;

(iii). $\displaystyle\int_0^1 f_{0,Y}^{-1}(x) f_{n,Y}^2(x)dx \leq \sqrt[n]{\lambda}$ *with* $\lambda \in (1,5)$,

then with each estimator $\widehat{f}_n(x; Y_1, \cdots, Y_n)$ of $f \in \Psi$ and $1 \leq p < \infty$,

$$\inf_{\widehat{f}_n} \sup_{f \in \Psi} E_{f_Y} d^p(\widehat{f}_n, f) \gtrsim a_n^p.$$

Proof. It follows from Jensen's inequality and $1 \leq p < \infty$ that

$$E_{f_{0,Y}} d^p(\widehat{f}_n, f_0) + E_{f_{n,Y}} d^p(\widehat{f}_n, f_n) \geq [E_{f_{0,Y}} d(\widehat{f}_n, f_0)]^p + [E_{f_{n,Y}} d(\widehat{f}_n, f_n)]^p.$$

Moreover, by $f_0, f_n \in \Psi$,

$$2 \sup_{f \in \Psi} E_{f_Y} d^p(\widehat{f}_n, f) \geq [E_{f_{0,Y}} d(\widehat{f}_n, f_0)]^p + [E_{f_{n,Y}} d(\widehat{f}_n, f_n)]^p. \tag{2.3}$$

Since $(|a| + |b|)^p \leq 2^{p-1}(|a|^p + |b|^p)$ for $p \geq 1$, the right-hand side of (2.3) has a lower bound

$$\begin{aligned} &2^{1-p}[E_{f_{0,Y}} d(\widehat{f}_n, f_0) + E_{f_{n,Y}} d(\widehat{f}_n, f_n)]^p \\ \gtrsim &\left[E_{f_{0,Y}} |d(\widehat{f}_n, f_n) - d(f_n, f_0)| + E_{f_{n,Y}} d(\widehat{f}_n, f_n)\right]^p. \end{aligned}$$

According to Condition (ii), (2.3) reduces to

$$\begin{aligned} &\sup_{f \in \Psi} E_{f_Y} d^p(\widehat{f}_n, f) \\ \gtrsim &d^p(f_n, f_0) \left[E_{f_{0,Y}} |d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n) - 1| + E_{f_{n,Y}} [d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n)]\right]^p \\ \geq &a_n^p \left[E_{f_{0,Y}} |d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n) - 1| + E_{f_{n,Y}} [d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n)]\right]^p. \end{aligned}$$

Therefore, it suffices for the desired conclusion to show

$$I_n := E_{f_{0,Y}} |d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n) - 1| + E_{f_{n,Y}} [d^{-1}(f_n, f_0) d(\widehat{f}_n, f_n)] \gtrsim 1. \tag{2.4}$$

For $y = (y_1, \cdots, y_n) \in (0,1)^n := \overbrace{(0,1) \times \cdots \times (0,1)}^{n}$, denote

$$F_0(y) := \prod_{l=1}^n f_{0,Y}(y_l),\ F_n(y) := \prod_{l=1}^n f_{n,Y}(y_l) \ \text{ and } \ H_n(y) := \prod_{l=1}^n (f_{0,Y})^{-1}(y_l) f_{n,Y}(y_l).$$

It follows from Condition (i) that H_n is well-defined. Then

$$E_{f_{n,Y}}[d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)] = \int_{(0,1)^n} d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)F_n(y)dy$$
$$= \int_{(0,1)^n} d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)H_n(y)F_0(y)dy. \tag{2.5}$$

On the other hand,

$$E_{f_{0,Y}}|d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)-1| = \int_{(0,1)^n} |d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)-1|F_0(y)dy. \tag{2.6}$$

Combining (2.5) and (2.6), one obtains that

$$I_n \geq \int_{(0,1)^n} \Big[|d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)-1| + d^{-1}(f_n,f_0)d(\widehat{f}_n,f_n)\Big] \min\{1,H_n(y)\}F_0(y)dy$$
$$\geq \int_{(0,1)^n} \min\{1,H_n(y)\}F_0(y)dy$$

thanks to $|a-1|+|a|\geq 1$. According to the fact $\min\{x,\ y\} = \frac{1}{2}\left[(x+y)-|x-y|\right]$,

$$I_n \geq \frac{1}{2}\int_{(0,1)^n} \big[1+H_n(y)-|1-H_n(y)|\big]F_0(y)dy.$$

By $\int_0^1 f_{0,Y}(y)dy = \int_0^1 f_{n,Y}(y)dy = 1$, $\int_{(0,1)^n} F_0(y)dy = \int_{(0,1)^n} H_n(y)F_0(y)dy = 1$ and

$$I_n \geq 1-\frac{1}{2}\int_{(0,1)^n} |1-H_n(y)|F_0(y)dy =: 1-\frac{b_n}{2}. \tag{2.7}$$

It follows from Jensen's inequality that

$$b_n^2 \leq \int_{(0,1)^n} [1-H_n(y)]^2 F_0(y)dy = \int_{(0,1)^n} \big[1-2H_n(y)+H_n^2(y)\big]\, F_0(y)dy$$
$$= -1+\int_{(0,1)^n} H_n^2(y)F_0(y)dy.$$

Furthermore, by the definitions of F_0 and H_n,

$$b_n^2 \leq \left[\int_0^1 (f_{0,Y})^{-1}(y)(f_{n,Y})^2(y)dy\right]^n - 1 \leq \lambda - 1,$$

where Condition (iii) is used in the last inequality. This with (2.7) and $\lambda\in(1,5)$ shows

$$I_n \geq 1-\frac{1}{2}\sqrt{\lambda-1} > 0,$$

which reaches (2.4). The proof is complete. □

Next, we use Proposition 2.1 to give a point-wise lower bound estimation for densities in GARCH-type model on $H^s(\Omega_{x_0},\ M)$.

Theorem 2.1. *Let $\widehat{f}_n$ be an estimator of $f\in H^s(\Omega_{x_0},\ M)$ and $1\leq p<\infty$. Then*

$$\inf_{\widehat{f}_n}\ \sup_{x\in\Omega_{x_0}}\ \sup_{f\in H^s(\Omega_{x_0},\ M)} \left[E\left|\widehat{f}_n(x)-f(x)\right|^p\right]^{\frac{1}{p}} \gtrsim n^{-\frac{s}{2s+2v+1}}.$$

Proof. Clearly, it is sufficient to prove that for some $x \in \Omega_{x_0}$,

$$\sup_{f \in H^s(\Omega_{x_0},\ M)} E\left|\widehat{f}_n(x) - f(x)\right|^p \gtrsim n^{-\frac{ps}{2s+2v+1}}.$$

Without loss of generality, one takes $x_0 = 1/2$; Otherwise, one could replace the variable x with $x - x_0$ in all of the following functions f_0, f_n and g.

Define $f_0(x) := ce^{-\frac{1}{1/4-(x-1/2)^2}} I_{(0,1)}(x)$ with the constant "c" being a normalized factor such that $\|f_0\|_1 = 1$ and

$$g(x) := a\left[e^{-\frac{1}{1/16-(x-3/4)^2}} I_{(1/2,1)}(x) - e^{-\frac{1}{1/16-(x-1/4)^2}} I_{(0,1/2)}(x)\right],$$

where the constant "a" will be determined later on and $g(1/2) := 0$.

Obviously, $f_0 \in H^s(\Omega_{x_0}, M)$, *supp* $f_0 \subseteq [0,1]$ and $f_0(x) \geq M_1 > 0$ for $x \in [1/2, 3/4]$. Let C_0^∞ be the set of all infinitely many times differentiable and compactly supported functions. Then $\int_{\mathbb{R}} g(x)dx = 0$ and $g \in C_0^\infty$ with *supp* $g \subseteq [0,1]$.

Take $a_j = 2^{-j(s+v+\frac{1}{2})}$ with $2^j \sim n^{\frac{1}{2s+2v+1}}$ and

$$f_n(x) := f_0(x) + a_j G_v(g_j)(x),$$

where $g_j(x) = 2^{\frac{j}{2}} g(2^j x - 2^{j-1})$.

Note that *supp* $g_j \subseteq [1/2, 3/4]$ with large j because of *supp* $g \subseteq [0,1]$. Then $f_n(x) \geq 0$ for $x \notin [1/2, 3/4]$. It is easy to calculate that

$$G_v(g_j)(x) = (-1)^v \sum_{u=1}^{v} C_u x^u (g_j)^{(u)}(x),$$

where $C_u > 0$ $(u = 1, 2, \cdots, v)$ are some constants. Then for $x \in [1/2, 3/4]$ and large j,

$$\begin{aligned} f_n(x) &\geq M_1 - \left|a_j \sum_{u=1}^{v} C_u x^u (g_j^{(u)})(x)\right| \geq M_1 - a_j 2^{j/2} \sum_{u=1}^{v} C_u 2^{uj} \|g^{(u)}(2^j \cdot - 2^{j-1})\|_\infty \\ &\geq M_1 - 2^{-js} \sum_{u=1}^{v} C_u \|g^{(u)}\|_\infty \geq 0 \end{aligned} \tag{2.8}$$

thanks to $f_0(x)|_{[1/2,3/4]} \geq M_1$ and $a_j = 2^{-j(s+1/2+v)}$. On the other hand, $\int_0^1 g(x)dx = 0$ and *supp* $g_j \subseteq [1/2, 1/2 + 2^{-j}]$ by *supp* $g \subseteq [0,1]$. These with integration by parts shows that

$$\begin{aligned} \int_0^1 x^u (g_j)^{(u)}(x)dx &= x^u (g_j)^{(u-1)}(x)\Big|_{1/2}^{1/2+2^{-j}} - u\int_0^1 x^{u-1}(g_j)^{(u-1)}(x)dx \\ &= \cdots = (-1)^m \frac{u!}{(u-m)!} \int_0^1 x^{u-m}(g_j)^{(u-m)}(x)dx \\ &= \cdots = (-1)^u u! \int_0^1 g_j(x)dx = 0 \end{aligned}$$

for any $u \in \{1, \cdots, v\}$. Therefore,

$$\int_{\mathbb{R}} f_n(x)dx = \int_0^1 f_0(x)dx + (-1)^v a_j \sum_{u=1}^{v} C_u \int_0^1 x^u (g_j)^{(u)}(x)dx = 1.$$

Combining this with (2.8), one concludes that f_n is a density function.

To prove $f_n \in H^s(\Omega_{x_0})$, one needs to check $a_j G_v(g_j) \in H^s(\Omega_{x_0})$. Denote $s = N + \delta$ with $\delta \in (0, 1]$. Then for $x,\ y \in \Omega_{x_0}$,

$$\begin{aligned}
&\Big|[a_j G_v(g_j)]^{(N)}(x) - [a_j G_v(g_j)]^{(N)}(y)\Big| \\
=&a_j \left|(-1)^v \sum_{u=0}^{v} C_u x^u (g_{j,l})^{(u)}(x) - (-1)^v \sum_{u=0}^{v} c_u x^u (g_j)^{(u+N)}(y)\right| \\
\le&a_j \sum_{u=0}^{v} 2^{j/2} 2^{j(u+N)} C_u |x^u g^{(u+N)}[2^j(x-l) - y^u g^{(u+N)}[2^j(y-l)]| \\
\lesssim&a_j 2^{j/2} 2^{j(v+N)} \sum_{u=0}^{v} |x^u g^{(u+N)}[2^j(x-l)] - y^u g^{(u+N)}[2^j(y-l)]|.
\end{aligned}$$

Then with $a_j = 2^{-j(s+v+\frac{1}{2})}$ and $x^k g^{(k')} \in H^\delta(\mathbb{R})$ holds for all positive integers k, k', the above inequality reduces to

$$\Big|[a_j G_v(g_j)]^{(N)}(x) - [a_j G_v(g_j)]^{(N)}(y)\Big| \lesssim |x-y|^\delta,$$

which shows $a_j G_v(g_j) \in H^s(\Omega_{x_0})$. Then $f_n \in H^s(\Omega_{x_0})$.

Obviously, $f_0(x) > 0$ for $x \in (0, 1)$ and

$$|f_n(x) - f_0(x)| = a_j 2^{j/2} \left|\sum_{u=1}^{v} C_u \left[2^j(x - \frac{1}{2} + \frac{1}{2})\right]^u g^{(u)}\left(2^j(x - \frac{1}{2})\right)\right|.$$

Note that $[2^j(x - \frac{1}{2} + \frac{1}{2})]^u = \sum_{l=0}^{u} \binom{u}{l} [2^j(x - \frac{1}{2})]^{u-l} (\frac{1}{2})^l 2^{lj}$. Then with $|a + b| \ge |a| - |b|$, we have

$$\begin{aligned}
&|f_n(x) - f_0(x)| \\
\ge&a_j 2^{j/2} \Bigg\{ C_v \left(\frac{1}{2}\right)^v 2^{vj} |g^{(v)}\left(2^j(x - \frac{1}{2})\right)\Big| - 2^{(v-1)j} \sum_{u=1}^{v} C_u \left(\frac{1}{2}\right)^u |g^{(u)}\left(2^j(x - \frac{1}{2})\right)\Big| - \\
&- 2^{(v-1)j} \sum_{u=1}^{v-1} C_u \sum_{l=0}^{u-1} \tbinom{u}{l} \left(\frac{1}{2}\right)^l \left|\left(x - \frac{1}{2}\right)^{u-l} g^{(u)}\left(2^j(x - \frac{1}{2})\right)\right| \Bigg\}. \qquad (2.9)
\end{aligned}$$

Choosing $x = \frac{1}{2} + \frac{2^{-j}}{16}$ with j large enough, which implies $x \in \Omega_{x_0}$ ($x_0 = \frac{1}{2}$) and $g^{(v)}(2^j(x - \frac{1}{2})) \neq 0$. Combining these with (2.9), one obtains that for large j,

$$|f_n(x) - f_0(x)| \gtrsim a_j 2^{j/2} 2^{jv} \gtrsim 2^{-js} \qquad (2.10)$$

because of $a_j = 2^{-j(s+v+\frac{1}{2})}$. Then with $2^j \sim n^{\frac{1}{2s+2v+1}}$, (2.10) reduces to

$$|f_n(x) - f_0(x)| \gtrsim n^{-\frac{s}{2s+2v+1}},$$

which concludes Condition (ii) of Proposition 2.1.

Finally, according to the formula (46) in Cao & Wei's work ([4]), we have

$$f_{n,Y} - f_{0,Y} = a_j g_j(x) \quad \text{and} \quad f_{0,Y} \ge \frac{M_1}{v!}\Big(\ln 3/4 - \ln(1/2 + 2^{-j})\Big)^v \ge M_2 > 0,$$

where M_1 and M_2 are two positive constants. Thus,

$$\begin{aligned}\int_0^1 f_{0,Y}^{-1}(x)f_{n,Y}^2(x)dx &= \int_0^1 f_{0,Y}^{-1}(x)[f_{0,Y}(x)+a_jg_j(x)]^2dx\\ =&1+\int_0^1 f_{0,Y}^{-1}(x)a_j^2g_j^2(x)dx \le 1+M_2^{-1}a_j^2\|g_j\|_2^2.\end{aligned}$$

It follows from the definition of a_j and $2^j \sim n^{\frac{1}{2s+2v+1}}$ that $a_j^2 \lesssim n^{-1}$ and

$$\left[\int_0^1 f_{0,Y}^{-1}(x)f_{n,y}^2(x)dx\right]^n \le \left[1+\widetilde{M_2}^{-1}\|g\|_2^2n^{-1}\right]^n \le \exp\left\{\widetilde{M_2}^{-1}\|g\|_2^2\right\}$$

thanks to $(1+x)^n \le e^{nx}$ for $x \ge 0$. By choosing the constant "a" in the definition of g such that $\|g\|_2$ is sufficient small,

$$\left(\int_0^1 f_{0,Y}^{-1}(x)f_{n,Y}^2(x)dx\right)^n \le \lambda$$

with $\lambda \in (1,5)$, which reaches Condition (iii) of Proposition 2.1. This completes the proof. □

3. Wavelet Estimations

This section is devoted to show the upper bounds of linear and non-linear wavelet estimators over point-wise risk respectively.

Let ϕ and ψ be orthonormal scaling and wavelet function of $L^2(\mathbb{R})$. Then for $f \in L^2(\mathbb{R})$,

$$f = \sum_k \alpha_{j_0k}\phi_{j_0k} + \sum_{j=j_0}^{\infty}\sum_k \beta_{jk}\psi_{jk}$$

holds in L^2-sense, where $\alpha_{j_0k} := \langle f, \phi_{j_0k}\rangle$ and $\beta_{jk} := \langle f, \psi_{jk}\rangle$. As in wavelet analysis, P_j stands for the orthogonal projection operator from $L^2(\mathbb{R})$ onto the scaling space V_j. Then $P_jf = \sum_k \alpha_{jk}\phi_{jk}$. All of these claims can be found in Ref. [12].

The following proposition is important for wavelet estimations over point-wise risk.

Proposition 3.1 ([21]). *Let ϕ be a compactly supported and continuous scaling function of $L^2(\mathbb{R})$ and ψ be the corresponding wavelets with vanishing moments of order N. If $f \in H^s(\Omega_{x_0},\ M)$ with $s = N+\delta$ $(0<\delta\le 1)$ and $\beta_{jk} = \int_{\mathbb{R}} f(y)\psi_{jk}(y)dy$, then for sufficiently large j and j_0,*

(i). $\sup\limits_{x\in\Omega_{x_0}} \sup\limits_{f\in H^s(\Omega_{x_0},\ M)} \sum\limits_k |\beta_{jk}\psi_{jk}(x)| \lesssim 2^{-js}$;

(ii). $f(x) = \sum\limits_k \alpha_{j_0k}\phi_{j_0k}(x) + \sum\limits_{j=j_0}^{\infty}\sum\limits_k \beta_{jk}\psi_{jk}(x), \quad x \in \Omega_{x_0}$;

(iii). $\sup\limits_{x\in\Omega_{x_0}} \sup\limits_{f\in H^s(\Omega_{x_0},\ M)} |f(x) - P_{j_0}f(x)| \lesssim 2^{-j_0s}$.

Before giving main results of this section, we recall the definitions of linear and non-linear (hard thresholding) wavelet estimators firstly, which can be found in Ref. [4,16]. Define

$$\widehat{f}_n^{lin}(x) := \sum_k \widehat{\alpha}_{j_0k}\phi_{j_0k}(x), \quad \widehat{f}_n^{non}(x) := \widehat{f}_n^{lin}(x) + \sum_{j=j_0}^{j_1}\sum_k \widetilde{\beta}_{jk}\psi_{jk}(x), \tag{3.1}$$

where

$$\widehat{\alpha}_{jk} := \frac{1}{n}\sum_{i=1}^{n} T_v(\phi_{jk})(Y_i), \quad \widehat{\beta}_{jk} := \frac{1}{n}\sum_{i=1}^{n} T_v(\psi_{jk})(Y_i), \quad \widetilde{\beta}_{jk} := \widehat{\beta}_{jk} I_{\{|\widehat{\beta}_{jk}|>\tau_{j,n}\}} \tag{3.2}$$

with $\tau_{j,n} := \gamma 2^{vj}\sqrt{j/n}$ and γ being a positive constant. By the work of [8], $E\widehat{\alpha}_{jk} = \alpha_{jk}$ and $E\widehat{\beta}_{jk} = \beta_{jk}$. To use Proposition 3.1, we choose ϕ and ψ satisfying all conditions, such as Daubechies' functions D_{2N} and ψ_{2N} with large N.

The next lemma is useful in later proofs.

Lemma 3.1. *Let $\widehat{\alpha}_{jk}$ and $\widehat{\beta}_{jk}$ be defined in* (3.2). *Then for $2^j \le n$ and $1 \le p < \infty$,*

$$E|\widehat{\alpha}_{jk} - \alpha_{jk}|^p \lesssim n^{-\frac{p}{2}} 2^{vjp} \quad \text{and} \quad E|\widehat{\beta}_{jk} - \beta_{jk}|^p \lesssim n^{-\frac{p}{2}} 2^{vjp},$$

where the constants in "$\lesssim$" only depend on ϕ, ψ, C_* and p.

Proof. Clearly, one only needs to prove the first inequality and the second one is similar. According to (3.2) and $E\widehat{\alpha}_{jk} = \alpha_{jk}$, one obtains that

$$\widehat{\alpha}_{jk} - \alpha_{jk} = \frac{1}{n}\sum_{i=1}^{n} T_v(\phi_{jk})(Y_i) - \frac{1}{n}\sum_{i=1}^{n} E[T_v(\phi_{jk})(Y_i)] = \frac{1}{n}\sum_{i=1}^{n} \xi_i \tag{3.3}$$

with $\xi_i := T_v(\phi_{jk})(Y_i) - E[T_v(\phi_{jk})(Y_i)]$. Then $E\xi_i = 0$ and $\xi_1, \cdots, \xi_n$ are i.i.d. samples, because $Y_1, \cdots, Y_n$ have the same property. These with (3.3) shows

$$E|\widehat{\alpha}_{jk} - \alpha_{jk}|^p = n^{-p} E\left|\sum_{i=1}^{n} \xi_i\right|^p. \tag{3.4}$$

By the definition of the operator T,

$$T_v(\phi_{jk})(x) = \sum_{u=0}^{v} C_u x^u (\phi_{jk})^{(u)}(x), \tag{3.5}$$

where C_u $(u = 0, 1, \cdots, v)$ are positive constants. Then with $M_0 = \max\{C_u,\ u = 0, 1, \cdots, v\}$,

$$\sup_{x\in[0,1]} |T_v(\phi_{jk})(x)| \le M_0 \sum_{u=0}^{v} \sup_{x\in[0,1]} |(\phi_{jk})^{(u)}(x)|.$$

Hence, it follows from the smoothness and compact support of ϕ that

$$\sup_{x\in[0,1]} |T_v(\phi_{jk})(x)| \lesssim 2^{(v+1/2)j}.$$

Moreover, combining it with $|\xi_i| \le |T_v(\phi_{jk})(Y_i)| + |E[T_v(\phi_{jk})(Y_i)]| \le 2\sup_{x\in[0,1]} |T_v(\phi_{jk})(x)|$, one concludes that

$$|\xi_i| \lesssim 2^{(v+1/2)j}. \tag{3.6}$$

By $E|\xi_i|^2 \le E[T_v(\phi_{jk})(Y_i)]^2$, (3.5) tells that

$$E|\xi_i|^2 \le E\left[\sum_{u=0}^{v} C_u Y_i^u (\phi_{jk})^{(u)}(Y_i)\right]^2. \tag{3.7}$$

On the other hand, due to (2.2),

$$E\left[(\phi_{jk})^{(u)}(Y_i)\right]^2 \lesssim 2^{(2u+1)j}\int_0^1 [\phi^{(u)}(2^jx-k)]^2dx \lesssim 2^{2uj}.$$

This with (3.7) and $Y_i \in [0,1]$ implies that

$$E|\xi_i|^2 \lesssim 2^{2vj}. \tag{3.8}$$

According to Rosenthal's inequality [12],

$$E\left|\sum_{i=1}^n \xi_i\right|^p \lesssim \sum_{i=1}^n E|\xi_i|^p I_{\{p>2\}} + \left(\sum_{i=1}^n E|\xi_i|^2\right)^{\frac{p}{2}}.$$

Then with (3.6) and (3.8), the above inequality reduces to

$$E\left|\sum_{i=1}^n \xi_i\right|^p \lesssim n2^{(v+1/2)j(p-2)}2^{2vj}I_{\{p>2\}} + n^{\frac{p}{2}}2^{vjp} = n^{\frac{p}{2}}2^{vjp}\left[n^{1-\frac{p}{2}}2^{j(\frac{p}{2}-1)}I_{\{p>2\}}+1\right].$$

Combining this with $2^j \le n$ and (3.4), one knows that

$$E|\widehat{\alpha}_{jk}-\alpha_{jk}|^p \lesssim n^{-p}n^{\frac{p}{2}}2^{vjp} = n^{-\frac{p}{2}}2^{vjp},$$

which completes the proof. □

Now, we are in a position to prove the linear wavelet estimation on $H^s(\Omega_{x_0},\ M)$.

Theorem 3.1. *Let $2^{j_0} \sim n^{\frac{1}{2s+2v+1}}$ and $1 \le p < \infty$. Then*

$$\sup_{x\in\Omega_{x_0}}\ \sup_{f\in H^s(\Omega_{x_0},\ M)}\left[E\left|\widehat{f}_n^{lin}(x)-f(x)\right|^p\right]^{\frac{1}{p}} \lesssim n^{-\frac{s}{2s+2v+1}}.$$

Proof. Clearly, it follows from Jensen's inequality that

$$E\left|\widehat{f}_n^{lin}(x)-f(x)\right|^p \lesssim E\left|\widehat{f}_n^{lin}(x)-P_{j_0}f(x)\right|^p + |P_{j_0}f(x)-f(x)|^p. \tag{3.9}$$

By Proposition 3.1 (iii) and $2^{j_0} \sim n^{\frac{1}{2s+2v+1}}$,

$$|P_{j_0}f(x)-f(x)|^p \lesssim 2^{-j_0sp} \lesssim n^{-\frac{sp}{2s+2v+1}}. \tag{3.10}$$

On the other hand,

$$\begin{aligned}&\left|\widehat{f}_n^{lin}(x)-P_{j_0}f(x)\right|\\ =&\left|\sum_k(\widehat{\alpha}_{j_0k}-\alpha_{j_0k})\phi_{j_0k}(x)\right| \le \sum_k|\widehat{\alpha}_{j_0k}-\alpha_{j_0k}||\phi_{j_0k}(x)|^{\frac{1}{p}}|\phi_{j_0k}(x)|^{\frac{1}{p'}}\end{aligned}$$

with $\frac{1}{p}+\frac{1}{p'}=1$. Then due to the Hölder inequality,

$$E\left|\widehat{f}_n^{lin}(x)-P_{j_0}f(x)\right|^p \le \sum_k E|\widehat{\alpha}_{j_0k}-\alpha_{j_0k}|^p|\phi_{j_0k}(x)|\left(\sum_k|\phi_{j_0k}(x)|\right)^{\frac{p}{p'}}.$$

Combining this with Lemma 3.1 and $\sum_k |\phi(x-k)| \lesssim 1$, one obtains that

$$\begin{aligned} & E\left|\widehat{f}_n^{lin}(x) - P_{j_0} f(x)\right|^p \\ \lesssim & n^{-\frac{p}{2}} 2^{vj_0 p} \left(\sum_k |\phi_{j_0 k}(x)|\right)^p \lesssim \left(\frac{2^{(2v+1)j_0}}{n}\right)^{\frac{p}{2}} \lesssim n^{-\frac{sp}{2s+2v+1}} \end{aligned} \tag{3.11}$$

thanks to $1 + \frac{p}{p'} = p$ and $2^{j_0} \sim n^{\frac{1}{2s+2v+1}}$. This with (3.9)–(3.10) leads to the desired conclusion and the proof is done. □

Remark 3.1. Compared with Theorem 2.1, the linear wavelet estimator $\widehat{f}_n^{lin}$ attains the optimal convergence rate $n^{-\frac{s}{2s+2v+1}}$.

The linear wavelet estimator is not adaptive, because $\widehat{f}_n^{lin}$ depends on the unknown parameter s. The non-linear wavelet estimator avoids this disadvantage, although it is nearly-optimal (optimal up to a $\ln n$ factor).

Before giving the proof of non-linear one, we introduce a classical inequality.

Lemma 3.2 (Bernstein inequality, [12]). *Let $X_1, \cdots, X_n$ be i.i.d. with $EX_i = 0$ and $|X_i| \le \|X\|_\infty$ $(i = 1, \cdots, n)$. Then for each $\epsilon > 0$,*

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i\right| \ge \epsilon\right\} \le 2\exp\left(-\frac{n\epsilon^2}{2(EX_i^2 + \|X\|_\infty \epsilon/3)}\right).$$

Theorem 3.2. *Let $2^{j_0} \sim n^{\frac{1}{2m+2v+1}}$ $(s < m)$ and $2^{j_1} \sim \left(\frac{n}{\ln n}\right)^{\frac{1}{2v+1}}$. Then with $1 \le p < \infty$,*

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},\ M)} \left[E\left|\widehat{f}_n^{non}(x) - f(x)\right|^p\right]^{\frac{1}{p}} \lesssim (\ln n)^{1-\frac{1}{p}} \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}.$$

Proof. Denote $J_n := \left[E\left|\widehat{f}_n^{non}(x) - f(x)\right|^p\right]^{\frac{1}{p}}$. Then

$$J_n \le \left[E\left|\widehat{f}_n^{lin}(x) - P_{j_0} f(x)\right|^p\right]^{\frac{1}{p}} + \left[E\left|\sum_{j=j_0}^{j_1}\sum_k (\widetilde{\beta}_{jk} - \beta_{jk})\psi_{jk}(x)\right|^p\right]^{\frac{1}{p}} + |P_{j_1+1} f(x) - f(x)|$$

thanks to the definition of $\widehat{f}_n^{non}$ and Proposition 3.1 (ii). By (3.11) and $2^{j_0} \sim n^{\frac{1}{2m+2v+1}}$ with $s < m$,

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},M)} \left[E\left|\widehat{f}_n^{lin}(x) - P_{j_0} f(x)\right|^p\right]^{\frac{1}{p}} \lesssim \left(\frac{2^{(2v+1)j_0}}{n}\right)^{\frac{1}{2}} \lesssim n^{-\frac{m}{2m+2v+1}} \lesssim n^{-\frac{s}{2s+2v+1}}.$$

On the other hand, Proposition 3.1 (iii) with the choice $2^{j_1} \sim \left(\frac{n}{\ln n}\right)^{\frac{1}{2v+1}}$ tells that

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},\ M)} |P_{j_1+1} f(x) - f(x)| \lesssim 2^{-j_1 s} \lesssim \left(\frac{\ln n}{n}\right)^{\frac{s}{2v+1}} \le \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}.$$

Hence, it suffices to show

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},M)} E\left|\sum_{j=j_0}^{j_1}\sum_k (\widetilde{\beta}_{jk} - \beta_{jk})\psi_{jk}(x)\right|^p \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\frac{sp}{2s+2v+1}}. \tag{3.12}$$

in order to conclude Theorem 3.2. Similar to [10, 13], define

$$A_j := \{k,\ |\widehat{\beta}_{jk}| > \tau_{j,n}\}, \qquad B_j := \left\{k,\ |\beta_{jk}| > \frac{\tau_{j,n}}{2}\right\} \text{ and } C_j := \{k,\ |\beta_{jk}| > 2\tau_{j,n}\}.$$

Then

$$E\left|\sum_{j=j_0}^{j_1}\sum_k(\widetilde{\beta}_{jk}-\beta_{jk})\psi_{jk}(x)\right|^p \lesssim E|e_1|^p + E|e_2|^p + E|e_3|^p + E|e_4|^p,$$

where

$$e_1 := \sum_{j=j_0}^{j_1}\sum_k(\widehat{\beta}_{jk}-\beta_{jk})\psi_{jk}(x)I_{\{k\in A_j\cap B_j^c\}},$$

$$e_2 := \sum_{j=j_0}^{j_1}\sum_k(\widehat{\beta}_{jk}-\beta_{jk})\psi_{jk}(x)I_{\{k\in A_j\cap B_j\}},$$

$$e_3 := \sum_{j=j_0}^{j_1}\sum_k\beta_{jk}\psi_{jk}(x)I_{\{k\in A_j^c\cap C_j\}}, \qquad e_4 := \sum_{j=j_0}^{j_1}\sum_k\beta_{jk}\psi_{jk}(x)I_{\{k\in A_j^c\cap C_j^c\}}.$$

In order to estimate $E|e_1|^p$, one observes that

$$\widehat{\beta}_{jk}-\beta_{jk} = \frac{1}{n}\sum_{i=1}^n\eta_i$$

with $\eta_i := T_v(\psi_{jk})(Y_i) - ET_v(\psi_{jk})(Y_i)$. Obviously, $E\eta_i = 0$ and $\eta_1, \cdots, \eta_n$ are i.i.d. samples. Similar to (3.6) and (3.8), one knows

$$\|\eta_i\|_\infty \le M_\infty 2^{(v+1/2)j}, \qquad E\eta_i^2 \le M_2 2^{2vj}.$$

Here, M_∞ and M_2 are two positive constants. Combining these with Lemma 3.2 and $\tau_{j,n} = \gamma 2^{vj}\sqrt{\frac{j}{n}}$, one obtains that

$$\begin{aligned} P\left\{\left|\widehat{\beta}_{jk}-\beta_{jk}\right| > \frac{\tau_{j,n}}{2}\right\} &\le 2\exp\left(-\frac{n\tau_{j,n}^2}{8[E\eta_i^2 + \|\eta_i\|_\infty\tau_{j,n}/6]}\right) \\ &\le 2\exp\left(-\frac{\gamma^2 j 2^{2vj}}{8[M_2 2^{2vj} + M_\infty 2^{(v+1/2)j}\gamma 2^{vj} j^{\frac{1}{2}} n^{-\frac{1}{2}}/6]}\right) \\ &\le 2\exp(-c\gamma j) \end{aligned}$$

with $\gamma > 0$ large enough and some $c > 0$, where one uses $2^j \le 2^{j_1} \sim (\frac{n}{\ln n})^{\frac{1}{2v+1}}$ in last inequality. Furthermore,

$$EI_{\{k\in A_j\cap B_j^c\}} \le EI_{\left\{|\widehat{\beta}_{jk}-\beta_{jk}|>\frac{\tau_{j,n}}{2}\right\}} = P\left\{|\widehat{\beta}_{jk}-\beta_{jk}| > \frac{\tau_{j,n}}{2}\right\} \lesssim 2^{-c\gamma j}. \tag{3.13}$$

Since $|\widehat{\beta}_{jk}| = |\frac{1}{n}\sum_{i=1}^n T_v(\psi_{jk})(Y_i)| \lesssim 2^{(v+1/2)j}$, $|\beta_{jk}| \le E|\widehat{\beta}_{jk}| \lesssim 2^{(v+1/2)j}$ and

$$E|e_1|^p \lesssim j_1^{p-1}E\sum_{j=j_0}^{j_1}2^{(v+\frac{1}{2})jp}\left[\sum_k|\psi_{jk}(x)|I_{\{k\in A_j\cap B_j^c\}}\right]^p.$$

By the Hölder inequality with $\frac{1}{p}+\frac{1}{p'}=1$, the above inequality reduces to

$$E|e_1|^p \lesssim j_1^{p-1} E \sum_{j=j_0}^{j_1} 2^{(v+\frac{1}{2})jp} \sum_k |\psi_{jk}(x)| I_{\{k\in A_j\cap B_j^c\}} \left(\sum_k |\psi_{jk}(x)|\right)^{\frac{p}{p'}}.$$

This with (3.13) and $\sum_k |\psi(x-k)| \lesssim 1$ leads to

$$E|e_1|^p \lesssim j_1^{p-1} \sum_{j=j_0}^{j_1} 2^{(v+\frac{1}{2})jp} 2^{\frac{j}{2}(1+\frac{p}{p'})} 2^{-c\gamma j} \lesssim j_1^{p-1} \sum_{j=j_0}^{j_1} 2^{(vp+p-c\gamma)j}.$$

Similarly, $E|e_3|^p \lesssim j_1^{p-1} \sum_{j=j_0}^{j_1} 2^{(2v+1)jp-c\gamma j}$. Then for sufficiently large $\gamma > \frac{vp+p+mp}{c}$,

$$E|e_1|^p + E|e_3|^p \lesssim j_1^{p-1} 2^{(vp+p-c\gamma)j_0}.$$

Moreover, it follows from $2^{j_0} \sim n^{\frac{1}{2m+2v+1}}$, $2^{j_1} \sim \left(\frac{n}{\ln n}\right)^{\frac{1}{2v+1}}$ and $s<m$ that

$$E|e_1|^p + E|e_3|^p \lesssim (\ln n)^{p-1} n^{-\frac{mp}{2m+2v+1}} \le (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{-\frac{sp}{2s+2v+1}}. \tag{3.14}$$

Next, $E|e_4|^p$ is investigated as follows. Put

$$2^{j^*} \sim \left(\frac{n}{\ln n}\right)^{\frac{1}{2s+2v+1}}.$$

Then $j_0 \le j^* \le j_1$ holds because of $\frac{1}{2m+2v+1} \le \frac{1}{2s+2v+1} \le \frac{1}{2v+1}$ and $0<s<m$. Furthermore,

$$e_4 = \left(\sum_{j=j_0}^{j^*} + \sum_{j=j^*+1}^{j_1}\right) \sum_k \beta_{jk}\psi_{jk}(x) I_{\{k\in A_j^c\cap C_j^c\}} := e_{41} + e_{42}.$$

According to Proposition 3.1 (i), $2^{j^*} \sim (\frac{n}{\ln n})^{\frac{1}{2s+2v+1}}$ and $2^{j_1} \sim (\frac{n}{\ln n})^{\frac{1}{2v+1}}$,

$$\begin{aligned} E|e_{42}|^p &\lesssim j_1^{p-1} \sum_{j=j^*}^{j_1} \left[\sum_k |\beta_{jk}\psi_{jk}(x)|\right]^p \lesssim j_1^{p-1} \sum_{j=j^*}^{j_1} 2^{-jsp} \\ &\le j_1^{p-1} 2^{-j^* sp} \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\frac{sp}{2s+2v+1}}. \end{aligned} \tag{3.15}$$

On the other hand, $|\beta_{jk}| \le 2\tau_{j,n} \lesssim 2^{vj}\sqrt{\frac{j}{n}}$ for $k\in C_j^c$,. This with $\sum_k |\psi(x-k)| \lesssim 1$ shows

$$\begin{aligned} E|e_{41}|^p &\lesssim E j_1^{p-1} \sum_{j=j_0}^{j^*} \left|\sum_k \beta_{jk}\psi_{jk}(x) I_{\{k\in A_j^c\cap C_j^c\}}\right|^p \\ &\lesssim j_1^{p-1} \sum_{j=j_0}^{j^*} 2^{(v+\frac{1}{2})jp} \left(\frac{j}{n}\right)^{\frac{p}{2}} \le j_1^{p-1} n^{-\frac{p}{2}} \left(j^* 2^{(2v+1)j^*}\right)^{\frac{p}{2}}. \end{aligned}$$

By the choices 2^{j^*} and 2^{j_1}, the above inequality reduces to

$$E|e_{41}|^p \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\frac{sp}{2s+2v+1}}.$$

Combining this with $e_4 = e_{41} + e_{42}$ and (3.15), one concludes that

$$E|e_4|^p \lesssim (\ln n)^{p-1}\Big(\frac{\ln n}{n}\Big)^{\frac{sp}{2s+2v+1}}. \tag{3.16}$$

Finally, it remains to estimate $E|e_2|^p$, in which one writes down

$$e_2 = \left(\sum_{j=j_0}^{j^*} + \sum_{j=j^*+1}^{j_1}\right)\sum_k (\widehat{\beta}_{jk} - \beta_{jk})\psi_{jk}(x) I_{\{k\in A_j\cap B_j\}} := e_{21} + e_{22}.$$

It follows from the Hölder inequality with $\frac{1}{p} + \frac{1}{p'} = 1$ and Lemma 3.1 that

$$\begin{aligned} E|e_{21}|^p &\lesssim j_1^{p-1} E\sum_{j=j_0}^{j^*}\sum_k |\widehat{\beta}_{jk} - \beta_{jk}|^p |\psi_{jk}(x)| \left(\sum_k |\psi_{jk}(x)|\right)^{\frac{p}{p'}} \\ &\lesssim n^{-\frac{p}{2}} j_1^{p-1} \sum_{j=j_0}^{j^*} 2^{vjp}\left(\sum_k |\psi_{jk}(x)|\right)^p. \end{aligned}$$

Then with $\sum_k |\psi(x-k)| \lesssim 1$, $j_1 \lesssim \ln n$ and $2^{j^*} \sim (\frac{n}{\ln n})^{\frac{1}{2s+2v+1}}$ that

$$E|e_{21}|^p \lesssim j_1^{p-1}\sum_{j=j_0}^{j^*}\left(\frac{2^{(2v+1)j}}{n}\right)^{\frac{p}{2}} \le j_1^{p-1}\left(\frac{2^{(2v+1)j^*}}{n}\right)^{\frac{p}{2}} \lesssim (\ln n)^{p-1}\Big(\frac{\ln n}{n}\Big)^{\frac{sp}{2s+2v+1}}.$$

For $E|e_{22}|^p$, one observes that $|\beta_{jk}| > \frac{\tau_{j,n}}{2}$ for $k \in B_j$ and

$$E|e_{22}|^p \lesssim j_1^{p-1}\sum_{j=j^*+1}^{j_1}\sum_k E|\widehat{\beta}_{jk} - \beta_{jk}|^p |\psi_{jk}(x)||\beta_{jk}\tau_{j,n}^{-1}| \left(\sum_k |\psi_{jk}(x)\beta_{jk}|\tau_{j,n}^{-1}\right)^{\frac{p}{p'}},$$

where the Hölder inequality with $\frac{1}{p} + \frac{1}{p'} = 1$ is used in last step. According to Lemma 3.1, Proposition 3.1 (i) and $\tau_{j,n} \sim 2^{vj}\sqrt{j/n}$,

$$\begin{aligned} E|e_{22}|^p \lesssim& j_1^{p-1}\sum_{j=j^*+1}^{j_1} n^{-\frac{p}{2}} 2^{vjp}\tau_{j,n}^{-p}\left(\sum_k |\beta_{jk}\psi_{jk}(x)|\right)^p \\ \lesssim& j_1^{p-1}\sum_{j=j^*+1}^{j_1} j^{-\frac{p}{2}} 2^{-jsp} \lesssim j_1^{p-1} 2^{-j^* sp}. \end{aligned}$$

Moreover, $E|e_{22}|^p \lesssim (\ln n)^{p-1}(\frac{\ln n}{n})^{\frac{sp}{2s+2v+1}}$ thanks to the choices 2^{j_1} and 2^{j^*}. Therefore,

$$E|e_2|^p \lesssim (\ln n)^{p-1}\Big(\frac{\ln n}{n}\Big)^{\frac{sp}{2s+2v+1}}. \tag{3.17}$$

Furthermore, $\sum_{\ell=1}^4 E|e_\ell|^p \lesssim (\ln n)^{p-1}(\frac{\ln n}{n})^{\frac{sp}{2s+2v+1}}$ holds due to (3.14), (3.16) and (3.17). Since all constants in "$\lesssim$" of above arguments do not depend on $x \in \Omega_{x_0}$ and $f \in H^s(\Omega_{x_0}, M)$, the desired conclusion (3.12) is concluded. This finishes the proof of Theorem 3.2. □

Remark 3.2. In contract to the L^p-risk estimation [4, 6–9, 16], we do not use the wavelet characterization theorem in the proofs of Theorem 3.1 and Theorem 3.2, because the local Hölder condition is only assumed in this paper. In fact, we do not know whether that theorem holds on $H^s(\Omega_{x_0})$ or not. Instead, Proposition 3.1 helps and the Hölder inequality is frequently used.

On the other hand, the convergence rates of Theorem 3.1 and Theorem 3.2 are exact the same up to a $\ln n$ factor. Compared with Theorem 2.1, the non-linear wavelet estimator not only attains a nearly-optimal convergence rate but also gets the adaptivity.

We need know an upper bound m of the unknown parameter s for the non-linear wavelet estimation. The next section considers a data driven version for densities in GARCH-type model, which dose not need any information on s.

4. Data-driven Estimation

Motivated by the work of Rebelles [17] and Goldonshluger & Lepski [11], we first introduce selection rules, then provide a totally adaptive point-wise estimation on $H^s(\Omega_{x_0},\ M)$ in this section.

Let $\widehat{f}_{j_0}$ be the linear wavelet estimator $\widehat{f}_n^{lin}$ given in (3.1). Define

$$\widehat{f}_{j,j^*}(x) := \sum_k \widehat{\alpha}_{j\wedge j^*,k}\phi_{j\wedge j^*,k}(x)$$

with $j \wedge j^* = \min\{j,\ j^*\}$ and

$$\widehat{\nu}_j := \nu\sqrt{\frac{(1+\frac{p}{2})2^{(2v+1)j}\max\{1,\ (\ln 2)(2v+1)j\}}{n}}, \tag{4.1}$$

where the constant ν will be specified later on.

Let $\mathcal{H} := \{0, 1, 2, \cdots, \lfloor \frac{1}{2v+1}\log_2(\frac{n}{\ln n})\rfloor\}$ with $\lfloor x\rfloor$ denoting the largest integer smaller or equal to x. Then define the integer $j_0 \in \mathbb{N}$ by the following selection rules. For $x \in \Omega_{x_0}$ and $x_+ := \max\{x, 0\}$,

(i). $\widehat{\xi}_j(x) := \max\limits_{j^*\in\mathcal{H}}\left[|\widehat{f}_{j,j^*}(x) - \widehat{f}_{j^*}(x)| - \widehat{\nu}_{j^*} - \widehat{\nu}_j\right]_+$;

(ii). $\widehat{\xi}_{j_0}(x) + 2\widehat{\nu}_{j_0} := \min\limits_{j\in\mathcal{H}}\left[\widehat{\xi}_j(x) + 2\widehat{\nu}_j\right]$.

We need a classical lemma in real analysis, in order to show the data driven estimation.

Lemma 4.1. *Let $(X, \mathcal{F}, \mu)$ be a measurable space and $f \in L^p(X, \mathcal{F}, \mu)$ with $0 < p < \infty$. Then with $\lambda(t) := \mu\{x \in X,\ |f(x)| > t\}$,*

$$\int_X |f|^p d\mu = p\int_0^{+\infty} t^{p-1}\lambda(t)dt.$$

Theorem 4.1. *Let j_0 be given by the above selection rules. Then the linear wavelet estimator $\widehat{f}_{j_0}$ satisfies that for $1 \le p < \infty$,*

$$\sup_{x\in\Omega_{x_0}}\ \sup_{f\in H^s(\Omega_{x_0},\ M)}\left[E\left|\widehat{f}_{j_0}(x) - f(x)\right|^p\right]^{\frac{1}{p}} \lesssim \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}.$$

Proof. With the choice $j_1 := \lfloor \frac{1}{2s+2v+1}\log_2(\frac{n}{\ln n})\rfloor$, $2^{j_1} \sim (\frac{n}{\ln n})^{\frac{1}{2s+2v+1}}$ and $j_1 \in \mathcal{H}$ thanks to $s > 0$. According to the proof of Theorem 3.1,

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},\ M)} \left[E\left|\widehat{f}_{j_1}(x) - f(x)\right|^p\right]^{\frac{1}{p}} \lesssim \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}. \tag{4.2}$$

By selection rules (i) and (ii),

$$\begin{aligned}|\widehat{f}_{j_1,j_0} - \widehat{f}_{j_0}| + |\widehat{f}_{j_1,j_0} - \widehat{f}_{j_1}| &\le (\widehat{\xi}_{j_1} + \widehat{\nu}_{j_0} + \widehat{\nu}_{j_1}) + (\widehat{\xi}_{j_0} + \widehat{\nu}_{j_0} + \widehat{\nu}_{j_1})\\ &= (\widehat{\xi}_{j_0} + 2\widehat{\nu}_{j_0}) + (\widehat{\xi}_{j_1} + 2\widehat{\nu}_{j_1}) \le 2(\widehat{\xi}_{j_1} + 2\widehat{\nu}_{j_1}).\end{aligned}$$

Furthermore, the above inequality leads to

$$\begin{aligned}&|\widehat{f}_{j_0}(x) - f(x)|^p\\ &\lesssim \left[|\widehat{f}_{j_0}(x) - \widehat{f}_{j_1,j_0}(x)| + |\widehat{f}_{j_1,j_0}(x) - \widehat{f}_{j_1}(x)|\right]^p + |\widehat{f}_{j_1}(x) - f(x)|^p\\ &\lesssim [\widehat{\xi}_{j_1}(x)]^p + \widehat{\nu}_{j_1}^p + |\widehat{f}_{j_1}(x) - f(x)|^p.\end{aligned}$$

Clearly, $\widehat{\nu}_{j_1}^p \lesssim (\frac{\ln n}{n})^{\frac{sp}{2s+2v+1}}$ due to (4.1) and $2^{j_1} \sim (\frac{n}{\ln n})^{\frac{1}{2s+2v+1}}$. These with (4.2) show that

$$\begin{aligned}&\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},\ M)} \left[E\left|\widehat{f}_{j_0}(x) - f(x)\right|^p\right]^{\frac{1}{p}}\\ &\lesssim \sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},\ M)} \left(E[\widehat{\xi}_{j_1}(x)]^p\right)^{\frac{1}{p}} + \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}},\end{aligned}$$

where $(|a|+|b|+|c|)^\theta \lesssim |a|^\theta + |b|^\theta + |c|^\theta$ for $\theta > 0$ and $\sup(|x|+|y|+|z|) \le \sup|x| + \sup|y| + \sup|z|$ are used. Hence, it suffices for concluding the desired conclusion to show

$$\sup_{x\in\Omega_{x_0}} \sup_{f\in H^s(\Omega_{x_0},M)} \left[E\widehat{\xi}_{j_1}(x)^p\right]^{\frac{1}{p}} \lesssim \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}. \tag{4.3}$$

It follows from $\widehat{f}_{j_1,j^*} = \widehat{f}_{j^*}$ for $j^* \le j_1$ that

$$\begin{aligned}\widehat{\xi}_{j_1} :&= \max_{j^*\in\mathcal{H}} \left[|\widehat{f}_{j_1,j^*} - \widehat{f}_{j^*}| - \widehat{\nu}_{j^*} - \widehat{\nu}_{j_1}\right]_+ = \max_{j^*>j_1} \left[|\widehat{f}_{j_1} - \widehat{f}_{j^*}| - \widehat{\nu}_{j^*} - \widehat{\nu}_{j_1}\right]_+\\ &\le \max_{j^*>j_1} \left[|\widehat{f}_{j_1} - E\widehat{f}_{j_1}| + |E\widehat{f}_{j_1} - f| + |f - E\widehat{f}_{j^*}| + |E\widehat{f}_{j^*} - \widehat{f}_{j^*}| - \widehat{\nu}_{j^*} - \widehat{\nu}_{j_1}\right]_+.\end{aligned} \tag{4.4}$$

Combining $E\widehat{f}_j(x) = P_j f(x)$ with Proposition 3.1 (iii) and $j^* > j_1$, one obtains

$$|E\widehat{f}_{j_1}(x) - f(x)| + |f(x) - E\widehat{f}_{j^*}(x)| \lesssim 2^{-j_1 s} + 2^{-j^* s} \lesssim 2^{-j_1 s}.$$

This with (4.4) leads to

$$[\widehat{\xi}_{j_1}(x)]^p \lesssim 2^{-j_1 sp} + \left[|\widehat{f}_{j_1}(x) - E\widehat{f}_{j_1}(x)| - \widehat{\nu}_{j_1}\right]_+^p + \max_{j^*>j_1} \left[|E\widehat{f}_{j^*}(x) - \widehat{f}_{j^*}(x)| - \widehat{\nu}_{j^*}\right]_+^p.$$

Moreover, by $j^* \in \mathcal{H}$ and $j^* \le \lfloor \frac{1}{2v+1}\log_2(\frac{n}{\ln n})\rfloor$,

$$E[\widehat{\xi}_{j_1}(x)]^p \lesssim 2^{-j_1 sp} + \sum_{j=j_1}^{\lfloor \frac{1}{2v+1}\log_2(\frac{n}{\ln n})\rfloor} E\left[|\widehat{f}_j(x) - E\widehat{f}_j(x)| - \widehat{\nu}_j\right]_+^p. \tag{4.5}$$

For each $t \geq 0$,

$$P\left\{\left[|\widehat{f}_j(x) - E\widehat{f}_j(x)| - \widehat{\nu}_j)\right]_+ > t\right\} = P\left\{|\widehat{f}_j(x) - E\widehat{f}_j(x)| - \widehat{\nu}_j > t\right\}.$$

This with Lemma 4.1 implies that

$$\begin{aligned} E\left[|\widehat{f}_j(x) - E\widehat{f}_j(x)| - \widehat{\nu}_j\right]_+^p &= p\int_0^\infty t^{p-1}P\left\{|\widehat{f}_j(x) - E\widehat{f}_j(x)| > \widehat{\nu}_j + t\right\}dt \\ &= p\widehat{\nu}_j^p\int_0^\infty t^{p-1}P\left\{|\widehat{f}_j(x) - E\widehat{f}_j(x)| > (1+t)\widehat{\nu}_j\right\}dt. \end{aligned} \tag{4.6}$$

According to the definition of $\widehat{f}_j$, $\widehat{f}_j(x) - E\widehat{f}_j(x) = \frac{1}{n}\sum_{i=1}^n W_i$ with

$$W_i := \sum_k \phi_{jk}(x)[T_v(\phi_{jk})(Y_i) - ET_v(\phi_{jk})(Y_i)].$$

Clearly, $W_1, W_2, \cdots, W_n$ are i.i.d. and $EW_i = 0$. Similar to the proofs of (3.6) and (3.8), $\|W_i\|_\infty \leq L_\infty 2^{(v+1)j}$ and $EW_i^2 \leq L_2 2^{(2v+1)j}$ due to $\sum_k |\phi(x-k)| \lesssim 1$, where L_∞ and L_2 are two positive constants.

Using Lemma 3.2, one knows that

$$\begin{aligned} &P\left\{|\widehat{f}_j(x) - E\widehat{f}_j(x)| > (1+t)\widehat{\nu}_j\right\} \\ \leq &2\exp\left\{-\frac{n(1+t)^2\widehat{\nu}_j^2}{2[L_2 2^{(2v+1)j} + L_\infty 2^{(v+1)j}\cdot(1+t)\widehat{\nu}_j/3]}\right\}. \end{aligned} \tag{4.7}$$

For $j \leq \lfloor \frac{1}{2v+1}\log_2(\frac{n}{\ln n})\rfloor$, $j2^{(2v+1)j} \leq n$ and $\widehat{\nu}_j \lesssim 1$ thanks to (4.1). Furthermore,

$$2\left[L_2 2^{(2v+1)j} + \frac{L_\infty}{3}2^{(v+1)j}\cdot(1+t)\widehat{\nu}_j\right] \leq L_0 2^{(2v+1)j}(1+t) \quad (L_0 > 0).$$

This with (4.1) shows that the right hand side of (4.7) is bounded by

$$2\exp\left\{-\frac{1+t}{L_0}\nu^2\Big(1+\frac{p}{2}\Big)\max\{1,\ (\ln 2)(2v+1)j\}\right\}.$$

With the choice $\nu \geq \sqrt{L_0}$, (4.7) reduces to

$$\begin{aligned} &P\left\{|\widehat{f}_j(x) - E\widehat{f}_j(x)| > (1+t)\widehat{\nu}_j\right\} \\ \leq &2\exp\left\{-(1+t)\Big(1+\frac{p}{2}\Big)\max\{1,\ (\ln 2)(2v+1)j\}\right\} \\ \leq &2\exp\left\{-\Big(1+\frac{p}{2}\Big)t\right\}\exp\left\{-(\ln 2)\Big(1+\frac{p}{2}\Big)j(2v+1)\right\} \\ \leq &2e^{-t}\cdot 2^{-(1+\frac{p}{2})(2v+1)j}. \end{aligned} \tag{4.8}$$

Substituting (4.8) into (4.6), one obtains that

$$\begin{aligned} E\left[|\widehat{f}_j(x) - E\widehat{f}_j(x)| - \widehat{\nu}_j\right]_+^p &\lesssim \left(\int_0^\infty t^{p-1}e^{-t}dt\right)\widehat{\nu}_j^p 2^{-(1+\frac{p}{2})(2v+1)j} \\ &\lesssim 2^{-(2v+1)j}\left(\frac{j}{n}\right)^{\frac{p}{2}} \lesssim 2^{-(2v+1)j}\left(\frac{\ln n}{n}\right)^{\frac{p}{2}} \end{aligned} \tag{4.9}$$

thanks to (4.1) and $j \leq \lfloor \frac{1}{2v+1} \log_2(\frac{n}{\ln n}) \rfloor$. Finally, it follows from (4.5) and (4.9) that

$$E[\widehat{\xi}_{j_1}(x)]^p \lesssim 2^{-j_1 sp} + 2^{-(2v+1)j} \left(\frac{\ln n}{n}\right)^{\frac{p}{2}} \lesssim 2^{-j_1 sp} + \left(\frac{\ln n}{n}\right)^{\frac{p}{2}} \lesssim \left(\frac{\ln n}{n}\right)^{\frac{sp}{2s+2v+d}}$$

due to the choice of j_1 and $\frac{s}{2s+2v+1} \leq \frac{1}{2}$.

Because all constants in "$\lesssim$" of the above agruments are independent of $x \in \Omega_{x_0}$ and $f \in H^s(\Omega_{x_0},\ M)$, the desired (4.3) is concluded. The proof is done. □

Remark 4.1. For $p > 1$, the upper bound $(\frac{\ln n}{n})^{\frac{s}{2s+2v+1}}$ in Theorem 4.1 is little better than

$$(\ln n)^{1-\frac{1}{p}} \left(\frac{\ln n}{n}\right)^{\frac{s}{2s+2v+1}}$$

in Theorem 3.2. In addition, we do not have to know an upper bound m of s in this case. On the other hand, we need to pay more computational prices for finding j_0 in selection rules (i) and (ii) than the traditional non-linear wavelet method.

Acknowledgments. The authors would like to thank Prof. Youming Liu for his important comments and suggestions. This work is supported by the National Natural Science Foundation of China (No. 11901019), and the Science and Technology Program of Beijing Municipal Commission of Education (No. KM202010005025).

References

[1] M. Abbaszadeh, C. Chesneau, H. Doosti, Multiplicative censoring: estimation of a density and its derivatives under the L^p-risk, *REVSTAT*, **11**:3 (2013), 255–276.

[2] K.E. Andersen, M. B. Hansen, Multiplicative censoring: density estimation by a series expansion approach, *J. Statist. Plann. Inference*, **98**:1-2 (2001), 137–155.

[3] M. Asgharian, M. Carone, V. Fakoor, Large-sample study of the kernel density estimators under multiplicative censoring, *Ann. Statist.*, **40**:1 (2012), 159–187.

[4] K.K. Cao, J.Y. Wei, Adaptive wavelet estimations for the derivative of a density in GARCH-type model, *J. Inequal. Appl.*, **2019**:106 (2019), 21pp.

[5] M. Carrasco, X. H. Chen, Mixing and moment properties of various GARCH and stochastic volatility models, *Economet. Theory*, **18**:1 (2002), 17–39.

[6] Y. P. Chaubey, C. Chesneau, H. Doosti, On linear wavelet density estimation: some recent developments, *J. Indian Soc. Agricultural Statist.*, **65**:2 (2011), 169-179.

[7] Y. P. Chaubey, C. Chesneau, H. Doosti, Adaptive wavelet estimation of a density from mixtures under multiplicative censoring, *Statistics*, **49**:3 (2015), 638–659.

[8] C. Chesneau, Wavelet estimation of a density in a GARCH-type model, *Comm. Statist. Theory Methods*, **42**:1 (2013), 98–117.

[9] C. Chesneau, H. Doosti, Wavelet linear density estimation for a GARCH model under various dependence structures, *J. Iran. Stat. Soc.*, **11**:1 (2012), 1–21.

[10] D. L. Donoho, I.M. Johnstone, G. Kerkyacharian, D. Picard, Density estimation by wavelet thresholding, *Ann. Statist.*, **24**:2 (1996), 508–539.

[11] A. Goldenshluger, O. Lepski, On adaptive minimax density estimation on $\mathbb{R}^d$, *Probab. Theory Related Fields*, **159**:3–4 (2014), 479–543.

[12] W. Härdle, G. Kerkyacharian, D. Picard, A. Tsybakov. Wavelets, Approximation and Statistical Applications. New York: Springer-Verlag, 1998.

[13] R. Li, Y.M. Liu, Wavelet optimal estimations for a density with some additive noises, *Appl. Comput. Harmon. Anal.*, **36**:3 (2014), 416–433.

[14] Y. M. Liu, C. Wu, Point-wise estimation for anisotropic densities, *J. Multivariate Anal.*, **171** (2019), 112–125.

[15] B.L.S. Prakasa Rao, Nonparametric Functional Estimation, Academic Press, New York, 1983.

[16] B.L.S. Prakasa Rao, Wavelet estimation for derivative of a density in a GARCH-type model, *Comm. Statist. Theory Methods*, **46**:5 (2017), 2396–2410.

[17] G. Rebelles, Pointwise adaptive estimation of a multivariate density under independence hypothesis, *Bernoulli*, **21**:4 (2015), 1984–2023.

[18] Y. Vardi, Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation, *Biometrika*, **76**:4 (1989), 751–761.

[19] Y. Vardi, C.H. Zhang, Large sample study of empirical distributions in a random-multiplicative censoring model, *Ann. Statist.*, **20**:2 (1992), 1022–1039.

[20] C. Wu, J.R. Wang, X.C. Zeng, A lower bound for generalized deconvolution density estimation, *Appl. Anal.*, (2021), 100: 8, 1806-1815.

[21] C. Wu, X.C. Zeng, J.R. Wang, Wavelet adaptive pointwise density estimations with super-smooth noises, *Acta Math. Sinica* (*Chin. Ser.*), **62**:5 (2019), 687–702.