MINIMAX RATES OF CONVERGENCE AND OPTIMALITY OF BAYES FACTOR WAVELET REGRESSION ESTIMATORS UNDER POINTWISE RISKS

Natalia Bochkina and Theofanis Sapatinas

University of Edinburgh and University of Cyprus

Supplementary Material

This note contains detailed proofs for Theorems 1-6 and Propositions 1-2. Proofs of most auxiliary statements are also included for completeness. Throughout this note, we use a generic positive constant C which is not necessarily the same, even within a single equation.

S1. Non-adaptive and adaptive minimax rates of convergence under pointwise l^u-risks ($1 \le u < \infty$) in the standard nonparametric regression model.

First we prove Theorem 1 (non-adaptive case).

Proof of Theorem 1.

[*Lower bound*] $R_n^{*,u}(\tilde{f}, B_{p,q}^r(A), t_0) \ge Cn^{-\frac{u(r-1/p)}{2(r-1/p)+1}}$ as $n \to \infty$ is equivalent to the following: for any estimator $\tilde{f}$ and any sequence $B_n \to \infty$ as $n \to \infty$,

$$\limsup \left[n^{\frac{u(r-1/p)}{2(r-1/p)+1}} B_n \sup_{f \in B_{p,q}^r(A)} \mathbb{E}|\tilde{f}(t_0) - f(t_0)|^u \right] = \infty. \tag{7.1}$$

Suppose now that (7.1) does not hold for some estimator $\tilde{f}$ and some sequence B_n such that $B_n \to \infty$ and $n/\log(B_n) \to \infty$ as $n \to \infty$, i.e.,

$$\limsup \left[n^{\frac{u(r-1/p)}{2(r-1/p)+1}} B_n \sup_{f \in B_{p,q}^r(A)} \mathbb{E}|\tilde{f}(t_0) - f(t_0)|^u \right] < \infty. \tag{7.2}$$

Hence, for any $f_0 \in B_{p,q}^r(A')$, $A' < A$,

$$\limsup \left[n^{\frac{u(r-1/p)}{2(r-1/p)+1}} \; B_n \, \mathbb{E}|\tilde{f}(t_0) - f_0(t_0)|^u \right] < \infty.$$

Thus, Theorem S1.2 yields that, for $B_n \to \infty$ and $n/\log(B_n) \to \infty$ as $n \to \infty$,

$$\liminf \left[\left(n/\log(B_n)\right)^{\frac{u(r-1/p)}{2(r-1/p)+1}} \sup_{f \in B_{p,q}^r(A)} \mathbb{E}|\tilde{f}(t_0) - f(t_0)|^u \right] > 0$$

and, hence,

$$\limsup \left[n^{\frac{u(r-1/p)}{2(r-1/p)+1}} \sup_{f \in B^r_{p,q}(A)} \mathbb{E}|\tilde{f}(t_0) - f(t_0)|^u \right] = \infty,$$

which contradicts the assumption made in (7.2). This completes the proof for the lower bound of Theorem 1.

[*Upper bound*] The validity of the upper bounds follows from the following theorem. (Note that Theorems S1.1 and S1.3 are proved under a slightly more general condition on the error distribution.)

Theorem S1.1. *Assume (S1), (S2) and (S3),* $1 \le u < \infty$, *and* $f \in B^r_{pq}(A)$ *with* $1 \le p, q \le \infty$, $A > 0$ *and* $1/p < r < s$. *Let* φ_j *be the pdf of* $N(0, \sigma_j^2/2)$, $0 < \underline{\sigma} \leqslant \sigma_j \leqslant \bar{\sigma} < \infty$, *and let* $\hat{f}$ *be a hard thresholding wavelet estimator with threshold* t_{jn}*:*

$$t_{jn} = \begin{cases} \sigma_j n^{-1/2} & \textit{for} \quad j = L, L+1, \ldots, j_1, \\ \sigma_j n^{-1/2} \left[\log_2(2uj)\right]^{1/2} & \textit{for} \quad j = j_1+1, \ldots, J-1, \end{cases}$$

where $j_1 = \frac{1}{2(r-1/p)+1} \log_2 n$. *Then, for any* $t_0 \in (0,1)$,

$$R^u_n(\hat{f}, B^r_{pq}(A), t_0) = O\left(n^{-\frac{u(r-1/p)}{2(r-1/p)+1}} \right) \quad \textit{as} \quad n \to \infty.$$

Proof of Theorem S1.1. It is easily seen that the risk is bounded by

$$\begin{aligned} R^u_n(\hat{f}, B^r_{p,q}(A), t_0) \quad \le \quad & \Bigg[\sum_{k \in K_{L-1}(t_0)} 2^{L/2} [\mathbb{E}(\hat{\theta}_k - \theta_k)^u]^{1/u} ||\phi||_\infty \\ + & \sum_{k \in K_{L-1}(t_0)} 2^{L/2} |\tilde{\theta}_k - \theta_k| \; ||\phi||_\infty \\ + & \sum_{j=L}^{J-1} \sum_{k \in K_j(t_0)} 2^{j/2} [\mathbb{E}(\hat{\theta}_{jk} - \theta_{jk})^u]^{1/u} ||\psi||_\infty \\ + & \sum_{j=L}^{J-1} \sum_{k \in K_j(t_0)} 2^{j/2} |\tilde{\theta}_{jk} - \theta_{jk}| \; ||\psi||_\infty \\ + & \sum_{j=J}^{\infty} \sum_{k \in K_j(t_0)} 2^{j/2} |\tilde{\theta}_{jk}| \; ||\psi||_\infty \Bigg]^u \\ = \quad & [Q_{11} + Q_{12} + Q_{21} + Q_{22} + Q_3]^u. \end{aligned} \tag{7.3}$$

Term $Q_{11} + Q_{12}$ in (7.3) is bounded by

$$\begin{aligned} & C \sum_{k \in K_{L-1}} 2^{L/2} [\mathbb{V}(\hat{\theta}_k)]^{1/2} + C \sum_{k \in K_{L-1}} 2^{L/2} |\tilde{\theta}_k - \theta_k| \le C n^{-1/2} \sigma_{L-1} + C n^{-r} \\ = \; & O\left(n^{-1/2}\right) + o\left(n^{-(r-1/p)}\right) = o\left(n^{-\frac{(r-1/p)}{2(r-1/p)+1}}\right), \end{aligned}$$

due to Lemma A.4 in Bochkina and Sapatinas (2006) and the fact that $\mathbb{V}(\hat{\theta}_k) = O(n^{-1})$.

On the other hand, term Q_3 in (7.3) is bounded by

$$
\begin{aligned}
C\sum_{j=J}^{\infty}\sum_{k\in K_j(t_0)} 2^{j/2}|\tilde{\theta}_{jk}| &\le C\sum_{j=J}^{\infty} 2^{j/2}2^{-j(r-1/p+1/2)} = O\left(2^{-J(r-1/p)}\right)\\
&= O\left(n^{-(r-1/p)}\right) = o\left(n^{-\frac{(r-1/p)}{2(r-1/p)+1}}\right),
\end{aligned}
$$

due to Lemma A.4 in Bochkina and Sapatinas (2006) which also implies that term Q_{22} in (7.3) is dominated by $C\,n^{-(r-1/p)}(\log n)^{\mathbb{I}(p=\infty)}$.

Thus, we have that

$$
\left[R_n^u(\hat{f}, B^r_{p,q}(A), t_0)\right]^{1/u} \leqslant Cn^{-\frac{(r-1/p)}{2(r-1/p)+1}} + C\sum_{j=L}^{J-1}\sum_{k\in K_j(t_0)} 2^{j/2}\left[\mathbb{E}|\theta_{jk}-\hat{\theta}_{jk}|^u\right]^{1/u}. \tag{7.4}
$$

Consider separately the sums for low and high resolution levels, using Lemma A.3 in Bochkina and Sapatinas (2006) to bound the summand in the first sum and Lemma S5.6 to bound the summand in the second sum:

$$
\begin{aligned}
[R_n^u(\hat{f}, B^r_{pq}(A), t_0)]^{1/u} &\leqslant Cn^{-(r-1/p)/(2(r-1/p)+1)} + C\sum_{j=L}^{j_1} 2^{j/2}t_{jn}\\
&+ C\sum_{j=j_1+1}^{J-1} 2^{j/2}\sum_{k\in K_j(t_0)} [\mathbb{E}|d_{jk}-\theta_{jk}|^u\mathbb{I}(|d_{jk}|>t_{jn})]^{1/u}.
\end{aligned}
$$

The first sum is easy to calculate since, by definition, t_{jn} is independent of j for $j \leqslant j_1$, and it is equal to $Cn^{-\frac{(r-1/p)}{2(r-1/p)+1}}$. Since for $j > j_1$ $t_{jn}\sqrt{n} \to \infty$, by Lemma A.4 in Bochkina and Sapatinas (2006) and Lemma S5.7, the second sum is bounded from above by

$$
\begin{aligned}
&Cn^{-1/2}\sum_{j=j_1+1}^{J-1} 2^{j/2}j^{(u+1)/2u}2^{-j} + C\sum_{j=j_1+1}^{J-1}\sum_{k\in K_j(t_0)} 2^{j/2}|\theta_{jk}|\\
\leqslant\ &Cn^{-1/2}[\log n]^{(u+1)/2u}2^{-j_1/2} + C\sum_{j=j_1+1}^{J-1} 2^{-j(r-1/p)} = O\big(n^{-\frac{(r-1/p)}{2(r-1/p)+1}}\big).
\end{aligned}
$$

Thus, for any $t_0 \in (0,1)$, $R_n^u(\hat{f}, B^r_{pq}(A), t_0) \leqslant Cn^{-\frac{u(r-1/p)}{2(r-1/p)+1}}$ as $n \to \infty$. This completes the proof of Theorem S1.1.

Thus, the proof for the upper bound of Theorem 1 is completed and, hence, Theorem 1 is proved.

We now prove Theorem 2 (adaptive case). In order to do that, we need some preliminary results.

Theorem S1.2. Take (r, p, q) such that $r > \frac{1}{p}$, $1 \leqslant p, q, \leqslant \infty$, and a sequence B_n such that $B_n \to \infty$, $n/\log(B_n) \to \infty$ as $n \to \infty$. Let $\tilde{f}$ be an estimator of f based on observations from the standard nonparametric regression model (2.1). If $f_0 \in B^r_{p,q}(A')$ with $0 < A' < A$ satisfying $\limsup \left[n^{\frac{u(r-1/p)}{2(r-1/p)+1}} B_n \, \mathbb{E}|\tilde{f}(t_0) - f_0(t_0)|^u \right] < \infty$, then

$$\liminf \left[\left(n/\log(B_n)\right)^{\frac{u(r-1/p)}{2(r-1/p)+1}} \sup_{f \in B^r_{p,q}(A)} \mathbb{E}|\tilde{f}(t_0) - f(t_0)|^u \right] > 0.$$

Proof of Theorem S1.2. Let X be a random variable having either distribution $\mathbb{P}_{\theta_0}$ with density f_{θ_0} or distribution $\mathbb{P}_{\theta_1}$ with density f_{θ_1}, with respect to some dominating measure. For any estimator $\delta = \delta(X)$ of $\theta \in \{\theta_0, \theta_1\}$, its l^u-risk ($1 \leq u < \infty$) is defined by $R^u(\delta, \theta) = \mathbb{E}|\delta(X) - \theta|^u$. Denote by $\kappa(x) = f_{\theta_1(x)}/f_{\theta_0}(x)$ the ratio of the two densities. ($\kappa(x) = \infty$ for some x is possible, with the obvious interpretation $\kappa(x) f_{\theta_0}(x) = f_{\theta_1}(x)$.)

For $1 \leq u < \infty$, denote by u^* the value satisfying $1/u + 1/u^* = 1$ (i.e., u and u^* are conjugate numbers). Let $I_{u^*} := I_{u^*}(\theta_0, \theta_1) = \left[\mathbb{E}_{\theta_0}(\kappa(X))^{u^*}\right]^{1/u^*}$, with obvious change for $u^* = \infty$ (this is a measure of distance between the two distributions $\mathbb{P}_{\theta_0}$ and $\mathbb{P}_{\theta_1}$).

Take $f_n(t) = \gamma_n^{-1} g(\beta_n(t - t_0)) + f_0(t)$, $t \in [0, 1]$, where

1) g is a compactly supported and monotonically decreasing function satisfying (i) $g \in B^r_{p,q}(A - A')$, (ii) $g(x) \geqslant 0$ for $x \in (0, 1]$, $g(0) > 0$, and (iii) $\|g\|_2^2 > 0$ (such a function is easy to construct, either directly or by using wavelets); denote by $b = [u c_u \|g\|_2^2]^{-1} > 0$ where $c_u = 1/[2(u-1)]$ for $u > 1$ and $c_u = 1$ for $u = 1$;

2) $\gamma_n = (n/(b\log(B_n)))^{\frac{(r-1/p)}{2(r-1/p)+1}}$ and $\beta_n = (n/(b\log(B_n)))^{\frac{1}{2(r-1/p)+1}}$.

Note that 2) above implies that $\gamma_n^2 \beta_n = n/(b\log(B_n))$ and $\gamma_n^{-1} \beta_n^{r-1/p} = 1$. Then, in view of Lemma 1 in Cai (2003), $f_n \in B^r_{p,q}(A)$.

Write $\mathbb{P}_0^n$ and $\mathbb{P}_1^n$ for the probability measures on $\mathbb{R}^n$ generated from the standard nonparametric regression model (2.1) with $f = f_0$ and $f = f_n$, respectively. Since σ^2 is assumed known, we take $\sigma^2 = 1$ without loss of generality. Then a sufficient statistic for the family of probability measures $\{\mathbb{P}_0^n, \mathbb{P}_1^n\}$ is given by the log likelihood ratio $T_n = \log(d\mathbb{P}_1^n/d\mathbb{P}_0^n)$ (see, e.g., Brown and Low (1996b)). Set $\rho_n' = \sum_{i=1}^n \left\{ \frac{g^2(\beta_n(t_{ni} - t_0))}{\gamma_n^2} \right\}$, where $t_{ni} := i/n$, $i = 0, 1, \ldots, n$. Then,

$$T_n \sim \begin{cases} N(-\rho_n'/2,\ \rho_n') & \text{under} \quad \mathbb{P}_0^n \\ N(\ \ \rho_n'/2,\ \rho_n') & \text{under} \quad \mathbb{P}_1^n, \end{cases}$$

and $\kappa(x) = e^x$. Note that, for large n and $r > 1/p + 1/2$, $\rho_n' \approx \rho_n := n \int_0^1 \frac{g^2(\beta_n(t-t_0))}{\gamma_n^2} dt$, where $\rho_n = n||f_n - f_0||_2^2 = n\gamma_n^{-2}\beta_n^{-1}||g||_2^2$ (by following the arguments of Section 5 in Brown and Low (1996b)). Now, define $\bar{g}$ by

$$\bar{g}(\beta_n(t - t_0)) = g(\beta_n(t_{ni} - t_0)), \quad t_{n,i-1} < t \le t_{ni},\ i = 1, 2, \dots, n,$$

with $\bar{g}(0) = g(0)$. Then, $\rho_n' = n \int_0^1 \frac{\bar{g}^2(\beta_n(t-t_0))}{\gamma_n^2} dt$.

Following the arguments in the proof of Theorem 4 in Cai (2003), we can rewrite $I_{u^*} = e^{\rho_n'(u^*-1)/2} = e^{\frac{\rho_n'}{2(u-1)}}$ for $1 < u^* < \infty$, and for $u^* = \infty$, $\tilde{I}_\infty = ||\kappa(x)\mathbb{I}(x \le \rho_n')||_\infty = e^{\rho_n'}$. We can unify the two cases by writing $I_{u^*} = e^{n\gamma_n^{-2}\beta_n^{-1}/(ub)} e^{\mu_n c_u}$, where $\mu_n = \rho_n' - \rho_n$. Substituting the values of β_n and γ_n in the definition of g, we get that $I_{u^*} = B_n^{1/u} e^{\mu_n c_u}$ since $n\gamma_n^{-2}\beta_n^{-1}/b = \log(B_n)$.

Since μ_n does not necessarily goes to zero when $1/p < r < 1/p + 1/2$, we use the assumption that g is non-negative and decreasing to obtain the inequality that $\mu_n \leqslant 0$ for $r > 1/p$, which is sufficient for our purpose. Indeed, since g is non-negative and decreasing, for $t \in [t_{n,i-1}, t_{ni}]$, we get

$$\bar{g}^2(\beta_n(t - t_0)) = g^2(\beta_n(t_{ni} - t_0)) \leqslant g^2(\beta_n(t - t_0)),$$

implying that $\rho_n' \leqslant \rho_n$ (i.e., $\mu_n \leqslant 0$) and thus that $I_{u^*} = e^{\rho_n' c_u} \leqslant e^{\rho_n c_u} = B_n^{1/u}$.

Let now $\delta_n = \tilde{f}_n(t_0)$, $\theta_0 = f_0(t_0)$ and $\theta_1 = f_n(t_0)$. Note that for sufficiently large n,

$$\Delta = |f_n(t_0) - f_0(t_0)| = g(0)\gamma_n^{-1}.$$

For some $C > 0$ and large enough n, Lemma 2(i) in Cai (2003) states that

$$R^u(\delta_n, \theta_1) \geqslant \Delta^u(1 - u\varepsilon_u I_{u^*}/\Delta),$$

where $R^u(\delta_n, \theta_0) \leqslant \varepsilon_u^u$; in our case, it is given that $\varepsilon_u = C^{1/u} n^{-\frac{(r-1/p)}{2(r-1/p)+1}} B_n^{-1/u}$. Substituting the value of ε_u and the upper bound for I_{u^*}, we get

$$\begin{aligned} R^u(\delta_n, \theta_1) = \mathbb{E}|\tilde{f}_n(t_0) - f_n(t_0)|^u \geq & \left(\frac{g(0)}{\gamma_n}\right)^u \left\{1 - u\, C^{1/u} n^{\frac{(r-1/p)}{2(r-1/p)+1}} B_n^{-1/u} B_n^{1/u}(1 + o(1))\gamma_n (g(0))^{-1}\right\} \\ =& [b\, g(0)]^{\frac{u(r-1/p)}{2(r-1/p)+1}} \left(\frac{\log B_n}{n}\right)^{\frac{u(r-1/p)}{2(r-1/p)+1}} \times \\ & \times \left\{1 - u\, C^{1/u}(b \log(B_n))^{\frac{-(r-1/p)}{2(r-1/p)+1}}(1 + o(1))(g(0))^{-1}\right\} \\ =& [b\, g(0)]^{\frac{u(r-1/p)}{2(r-1/p)+1}} \left(\frac{\log B_n}{n}\right)^{\frac{u(r-1/p)}{2(r-1/p)+1}} (1 + o(1)). \end{aligned}$$

This completes the proof of Theorem S1.2.

We are now ready to prove Theorem 2 (adaptive case).

Proof of Theorem 2

[*Lower bound*] Consider two Besov classes $B^{r_i}_{p_i,q_i}(A_i)$ with $r_i > 1/p_i$ for $i = 1, 2$. Let $0 < r_2 - 1/p_2 < r_1 - 1/p_1$. Applying Theorem S1.2, it immediately follows that if an estimator $\hat{f}_n$ satisfies, as $n \to \infty$,

$$\limsup \left[n^l \sup_{f \in B^{r_1}_{p_1,q_1}(A_1)} \mathbb{E}|\hat{f}_n(t_0) - f(t_0)|^u \right] < \infty$$

for some $l > u(r_2 - 1/p_2)/(1 + 2(r_2 - 1/p_2)$, then

$$\liminf \left[(n/\log n)^{\frac{u(r_2-1/p_2)}{2(r_2-1/p_2)+1}} \sup_{f \in B^{r_2}_{p_2,q_2}(A_2)} \mathbb{E}|\hat{f}_n(t_0) - f(t_0)|^u \right] > 0.$$

This completes the proof for the lower bound of Theorem 2.

[*Upper bound*] The validity of the upper bound follows from the following theorem.

Theorem S1.3. *Assume (S1), (S2) and (S3),* $1 \le u < \infty$*, and* $f \in B^r_{pq}(A)$ *with* $1 \le p, q \le \infty$*,* $A > 0$ *and* $1/p < r < s$*. Let* φ_j *be the pdf of* $N(0, \sigma_j^2/2)$*,* $0 < \underline{\sigma} \leqslant \sigma_j \leqslant \bar{\sigma} < \infty$*, and let* $\hat{f}$ *be a hard thresholding wavelet estimator with threshold* $t_{jn} = \sigma_j n^{-1/2}(uj \log 2)^{1/2}$ *for* $j = L, L+1, \dots, J-1$*. Then, for any* $t_0 \in (0, 1)$*,*

$$R^u_n(\hat{f}, B^r_{pq}(A), t_0) = O\left(\left(\frac{n}{\log n}\right)^{-\frac{u(r-1/p)}{2(r-1/p)+1}} \right) \quad as \quad n \to \infty.$$

Proof of Theorem S1.3. Let j_2 be such that $2^{j_2} = \left(\frac{n}{\log n}\right)^{\frac{1}{2\nu+1}}$. From the proof of Theorem S1.1 it follows that we need to find an upper bound on (7.4) which we consider separately for low and high resolution levels:

$$\sum_{j=L}^{j_2} \sum_{k \in K_j(t_0)} 2^{j/2}[\mathbb{E}(\hat{\theta}_{jk} - \theta_{jk})^u]^{1/u} + \sum_{j=j_2+1}^{J-1} \sum_{k \in K_j(t_0)} 2^{j/2}[\mathbb{E}(\hat{\theta}_{jk} - \theta_{jk})^u]^{1/u} = R_1 + R_2.$$

By Lemma A.3 in Bochkina and Sapatinas (2006), the first sum is bounded from above by

$$\begin{aligned} & C\sum_{j=L}^{j_2} 2^{j/2} t_{jn} + O(n^{-1/2}2^{j_2/2}) = Cn^{-1/2}\sum_{j=L}^{j_2} 2^{j/2}\sigma_j j^{1/2} + O(n^{-1/2}2^{j_2/2}) \\ \leqslant \; & Cn^{-1/2}[\log_2 n]^{1/2}2^{j_2/2} + O(n^{-1/2}2^{j_2/2}) \leqslant C\left(\frac{n}{\log n}\right)^{-\nu/(2\nu+1)}. \end{aligned}$$

In the proof of Theorem S1.1, we showed that if $t_{jn}n^{1/2} \to \infty$ as $n \to \infty$,

$$E|d_{jk} - \theta_{jk}|^u I(|d_{jk}| > t_{jn}) \leqslant C\sigma_j^u n^{-u/2}|t_{jn}\sqrt{n}/\sigma_j|^{u+1}\exp\{-(t_{jn}\sqrt{n}/\sigma_j)^2/2\},$$

which holds here since for $j > j_2$, $t_{jn} n^{1/2} \geqslant C j_2^{1/2} = C \log_2\left(\frac{n}{\log n}\right) \to \infty$ as $n \to \infty$.

Note that since $j_2 < j_1$ and $t_{jn} n^{1/2} \to \infty$ for $j > j_2$ as $n \to \infty$, we can apply Lemma S5.7 to obtain an upper bound on the second sum:

$$\begin{aligned}
& C \sum_{j=j_2+1}^{J-1} 2^{j/2} \sigma_j n^{-1/2} |t_{jn}\sqrt{n}/\sigma_j|^{1+1/u} \exp\{-(t_{jn}\sqrt{n}/\sigma_j)^2/2u\} \\
\leqslant\ & C n^{-1/2} \sum_{j=j_2+1}^{J-1} 2^{j/2} j^{1/2+1/2u} \exp\{-j \log 2/2\} \leqslant C n^{-1/2} (\log n)^{1/2+1/2u} \sum_{j=j_2+1}^{J-1} 1 \\
=\ & C \left(\frac{n}{\log n}\right)^{-1/2} (\log n)^{1+1/2u} = o\left(\frac{n}{\log n}\right)^{-\nu/(2\nu+1)}.
\end{aligned}$$

This completes the proof of Theorem S1.3.

Thus, the proof for the upper bound of Theorem 2 is completed and, hence, Theorem 2 is proved.

S2. Proof of theorems for non-adaptive Bayes factor wavelet estimators

Set $\varepsilon_1 = \frac{1}{2(r-1/p)+1}$.

Proof of Theorem 3. Following the proof of Theorem S1.1, we only need to consider (7.4):

$$R = \sum_{j=L}^{J-1} 2^{j/2} \sum_{k \in K_j(t_0)} \left[\mathbb{E}|\theta_{jk} - \hat{\theta}_{jk}|^u\right]^{1/u} = S_{low} + S_{high},$$

where S_{low} represents the sum over indices $L \leqslant j \leqslant j_1$ and S_{high} represents the sum over $j_1 < j \leqslant J-1$. Note that for the low resolution levels

$$\nu_j/\sqrt{n} = C 2^{m_1 j} n^{-1/2} = C 2^{-m_1(j_1-j)} n^{m_1\varepsilon_1 - 1/2} \to 0 \quad \text{as} \quad n \to \infty, \tag{7.5}$$

since $m_1\varepsilon_1 - 1/2 = \frac{m_1 - (r-1/p+1/2)}{2(r-1/p)+1} \leqslant 0$. Similarly, for the high resolution levels,

$$\nu_j/\sqrt{n} = C 2^{m_2 j} n^{-1/2} = C 2^{m_2(j-j_1)} n^{m_2\varepsilon_1 - 1/2} \to \infty \quad \text{as} \quad n \to \infty, \tag{7.6}$$

since $m_2\varepsilon_1 - 1/2 = \frac{m_2 - (r-1/p+1/2)}{2(r-1/p)+1} \geqslant 0$.

Low resolution levels, $L \leqslant j \leqslant j_1$.

We use Lemma A.3 in Bochkina and Sapatinas (2006) to bound S_{low} from above:

$$S_{low} \;=\; \sum_{j=L}^{j_1} 2^{j/2} \sum_{k \in K_j(t_0)} \min(t_{jn}, |\theta_{jk}|) + O(2^{j_1/2} n^{-1/2}),$$

since $\kappa_{u,j} = \int_{-\infty}^{+\infty} |x|^u \varphi_j(x) dx = c\sigma_j^u \int_{-\infty}^{+\infty} |z|^u e^{-|z|^\beta} dz < C\,\overline{\sigma}^u < \infty$ (uniformly).

Since $\nu_j/\sqrt{n} \to 0$ due to (7.5) and prior h satisfies the assumptions of Lemma S5.5, we can apply Lemma S5.5 together with the upper bound for the threshold under power-exponential errors (Lemma 4(i) in Pensky and Sapatinas (2007)) to obtain the following bound:

$$\begin{aligned} S_{low} &\leqslant \sum_{j=L}^{j_1} 2^{j/2} t_{jn} + O(2^{j_1/2} n^{-1/2}) \leqslant n^{-1/2} \sum_{j=L}^{j_1} 2^{j/2} \left[\log\left(\beta_{jn}\sqrt{n}\nu_j^{-1}\right)\right]^{1/\beta} + O(2^{j_1/2} n^{-1/2}) \\ &\leqslant n^{-1/2} \sum_{j=L}^{j_1} 2^{j/2} \left[\log\left(\beta_{jn}\sqrt{n}\nu_j^{-1}\right)\right]^{1/\beta} + O(2^{j_1/2} n^{-1/2}) = O\left(n^{-\frac{r-1/p}{2(r-1/p)+1}}\right), \end{aligned}$$

since $\beta_{jn}\sqrt{n}\nu_j^{-1} = C2^{j(a_1-m_1)} n^{b_1+1/2} \leqslant Cn^{b_1+1/2+\varepsilon_1(a_1-m_1)_+} \leqslant C$ under the assumptions of the theorem. Thus, for the low level sum the rate is optimal.

High resolution levels, $j_1 + 1 \leqslant j \leqslant J-1$.

We need to show that the following sum is bounded by $Cn^{-\nu/(2\nu+1)}$:

$$S_{high} = \sum_{j=j_1+1}^{J-1} 2^{j/2} \sum_{k \in K_j(t_0)} \left[\mathbb{E}|\theta_{jk} - \hat{\theta}_{jk}|^u\right]^{1/u}.$$

By Lemma S5.6,

$$S_{high} \leqslant 2 \sum_{j=j_1+1}^{J-1} 2^{j/2} \sum_{k \in K_j(t_0)} [\mathbb{E}|\theta_{jk} - d_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn})]^{1/u} + 2 \sum_{j=j_1+1}^{J-1} 2^{j/2} \sum_{k \in K_j(t_0)} |\theta_{jk}|.$$

Due to Lemma A.4 in Bochkina and Sapatinas (2006), the second sum is bounded by

$$C \sum_{j=j_1+1}^{J-1} 2^{j/2} 2^{-j(r-1/p+1/2)} = C \sum_{j=j_1+1}^{J-1} 2^{-j(r-1/p)} = O(n^{-\nu/(2\nu+1)}) \quad \text{as} \quad n \to \infty.$$

Note also that $\nu_j/\sqrt{n} \to \infty$ due to (7.6). Now we consider the first term.

Consider separately 2 cases: $\beta \leqslant 1$ and $\beta > 1$.

1. $\beta \leqslant 1$**.** In this case $\lambda_\varphi = 0$ and, if $\beta_{jn} > 1$, by Lemma 2(ii) in Pensky (2006),

$$\zeta_{jn} = 1 + o(n\nu_j^{-1}) < \beta_{jn} \quad \forall x,$$

implying that $\mathbb{I}(|d_{jk}| > t_{jk}) = \mathbb{I}(\zeta_{jn}(d_{jk}) > \beta_{jn}) = 0$ and thus the first sum is zero.

Note that $\beta_{jn} > 1$ since $\beta_{jn} = C2^{a_2 j} n^{b_2} \geqslant Cn^{b_2+\varepsilon_1(a_2)_+} \to \infty$ since $b_2 + \varepsilon_1(a_2)_+ > 0$.

2. $\beta > 1$**.** In this case, we bound the summands in the first sum using Lemma S5.7 and Lemma S5.3:

$$\begin{aligned} \mathbb{E}|\theta_{jk} - d_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn}) &\leqslant Cn^{-u/2}[t_{jn}\sqrt{n}/\sigma_j]^{u+1} e^{-(t_{jn}\sqrt{n}/\sigma_j)^\beta} \\ &= Cn^{-u/2}[\log \beta_{jn}]^{(u+1)/\beta} \beta_{jn}^{-1}, \end{aligned}$$

since $\beta_{jn} \to \infty$ for $j > j_1$ due to $b_2 + \varepsilon_1 (a_2)_+ > 0$, and thus $t_{jn} n^{1/2} \to \infty$.

Thus, the second sum in this case is bounded by

$$\begin{aligned} & C \sum_{j=j_1+1}^{J-1} 2^{j/2} n^{-1/2} [\log \beta_{jn}]^{(u+1)/u\beta} \beta_{jn}^{-1/u} \\ = \;& C n^{-1/2} \sum_{j=j_1+1}^{J-1} 2^{j/2} [b_2 \log n + a_2 j \log 2]^{(u+1)/u\beta} n^{-b_2/u} 2^{-a_2 j/u} \\ \leqslant \;& C n^{-1/2 - b_2/u} [(b_2 + \varepsilon_1 (a_2)_+) \log n]^{(u+1)/u\beta} \sum_{j=j_1+1}^{J-1} 2^{j(1/2 - a_2/u)}. \end{aligned}$$

The last sum is equal to

$$\begin{cases} C n^{-(b_2+a_2)/u} [\log n]^{(u+1)/u\beta}, & 1/2 - a_2/u > 0, \\ C n^{-1/2 - b_2/u} [\log n]^{1+(u+1)/u\beta}, & 1/2 - a_2/u = 0, \\ C n^{-(1-\varepsilon_1)/2} n^{-(b_2+\varepsilon_1 a_2)/u} [\log n]^{(u+1)/u\beta}, & 1/2 - a_2/u < 0. \end{cases}$$

Thus, this sum converges to zero at a rate not slower than the optimal if

$$\begin{cases} b_2 + a_2 > u(1-\varepsilon_1)/2 & \text{if} \quad a_2 \leqslant u/2, \\ b_2 + \varepsilon_1 a_2 > 0 & \text{if} \quad a_2 > u/2, \end{cases}$$

which can be rewritten as

$$\begin{cases} b_2 + \varepsilon_1 a_2 > (u/2 - a_2)(1-\varepsilon_1) & \text{if} \quad a_2 \leqslant u/2, \\ b_2 + \varepsilon_1 a_2 > 0 & \text{if} \quad a_2 > u/2, \end{cases}$$

or $b_2 + \varepsilon_1 a_2 > (u/2 - a_2)_+ (1-\varepsilon_1)$. It is easy to check that this condition implies $b_2 + \varepsilon_1 (a_2)_+ > 0$ required earlier. This completes the proof of Theorem 3.

Proof of Theorem 4. Following the proof of Theorem S1.1, we only need to consider (7.4) which we consider separately for low and high resolution levels.

Low resolution levels, $L \leqslant j \leqslant j_1$.

Since $m_1 \leqslant r - 1/p + 1/2$, $\nu_j/\sqrt{n} \to 0$, following the proof for the low levels of Theorem 3, the sum for the low levels is bounded by the optimal rate plus $\sum_{j=L}^{j_1} 2^j t_{jn}$ which achieves the optimal rate if $\beta_{jn}\sqrt{n}/\nu_j \leqslant C$ for the considered j (Lemma S5.5), i.e. given

$$\beta_{jn}\sqrt{n}/\nu_j = n^{1/2+b_1} 2^{(a_1-m_1)j} \leqslant \begin{cases} C n^{1/2+b_1} 2^{(a_1-m_1)j_1} & if \;\; a_1 - m_1 > 0, \\ C n^{1/2+b_1} j_1 & if \;\; a_1 - m_1 = 0 \\ C n^{1/2+b_1} & if \;\; a_1 - m_1 < 0 \end{cases}$$

or, equivalently, if

$$\begin{cases} 1/2 + b_1 + \varepsilon_1(a_1 - m_1) \leqslant 0 & if \;\; a_1 - m_1 > 0, \\ 1/2 + b_1 < 0 & if \;\; a_1 - m_1 = 0 \\ 1/2 + b_1 \leqslant 0 & if \;\; a_1 - m_1 < 0. \end{cases}$$

High resolution levels, $j_1 + 1 \leqslant j \leqslant J - 1$.

Since φ_j is a heavy tailed density (3.3), by Lemma 2 (ii) in Pensky (2006),

$$\zeta_{jn}(x) = \frac{I_j(x)}{\sqrt{n}\varphi_j(\sqrt{n}x)} = 1 + o(1) < \beta_{jn},$$

since $\beta_{jn} = C2^{a_2 j} n^{b_2} \geqslant Cn^{b_2 + \varepsilon_1 (a_2)_+} \to \infty$ as $n \to \infty$ due to $b_2 + \varepsilon_1(a_2)_+ > 0$. Thus, $\mathbb{I}(\zeta_{jn}(|d_{jk}|) > \beta_{jn}) = 0$ and the second sum is zero. This completes the proof of Theorem 4.

S3. Proof of theorems for adaptive Bayes factor wavelet estimators

To prove adaptive error rates, we use another division of the resolution levels, with the critical level j_2 defined by

$$j_2 = \frac{1}{2(r - 1/p) + 1} \log_2 \left(\frac{n}{\log n} \right). \tag{7.7}$$

Note that this "adaptive" critical level is smaller than then "non-adaptive" critical level j_1.

Lemma S3.1. *Assume (S1), (S2) and (S3),* $1 \leq u < \infty$, *and* $f \in B^r_{pq}(A)$ *with* $1 \leq p, q \leq \infty$, $A > 0$ *and* $1/p < r < s$. *Let* φ_j *be the pdf of* $N(0, \sigma_j^2/2)$, $0 < \underline{\sigma} \leqslant \sigma_j \leqslant \bar{\sigma} < \infty$, *and let* $\hat{f}_{BF}$ *is the corresponding Bayes Factor estimator. We assume that* h *is such that* $\zeta_{jn}(x)$ *increases for* $x > 0$. *Denote* $f_{jn} = \max(1, \nu_j/\sqrt{n})$, $v_{jn} = \beta_{jn} f_{jn} \sqrt{n}/\nu_j \to \infty$ *as* $n \to \infty$. *Then, for any* $t_0 \in (0, 1)$,

$$\begin{aligned} [R^u_n(\hat{f}_{BF}, B^r_{pq}(A), t_0)]^{1/u} \leqslant & C \left(\frac{n}{\log n} \right)^{-\frac{(r-1/p)}{(2(r-1/p)+1)}} \\ & + Cn^{-1/2} \sum_{j=L}^{j_2} 2^{j/2} f_{jn} [\log(C_{\varphi,h} v_{jn})]^{1/2} \\ & + Cn^{-1/2} \sum_{j=j_2+1}^{J-1} 2^{j/2} v_{jn}^{-1/u} [\log v_{jn}]^{(u+1)/2u}. \end{aligned} \tag{7.8}$$

Proof of Lemma S3.1. Following the proof of Theorem S1.1, we only need to consider (7.4), and using Lemma A.3 in Bochkina and Sapatinas (2006) and Lemma S5.6, we have that

$$\begin{aligned} [R^u_n(\hat{f}_{BF}, B^r_{pq}(A), t_0)]^{1/u} & \leqslant Cn^{-\nu/(2\nu+1)} + C \sum_{j=L}^{j_2} 2^{j/2} t_{jn} \\ & + C \sum_{j=j_2+1}^{J-1} 2^{j/2} \sum_{k \in K_j(t_0)} [\mathbb{E}|\theta_{jk} - d_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn})]^{1/u}. \end{aligned}$$

By Lemma S5.4, the first sum is bounded by

$$
\begin{aligned}
& C\sum_{j=L}^{j_2} 2^{j/2} t_{jn} \leqslant C\sum_{j=L}^{j_2} 2^{j/2}\sigma_j n^{-1/2}\left(1+\frac{\nu_j^2}{n}\right)^{1/2}\left[\log(C_{\varphi h}\beta_{jn}\sqrt{1+n/\nu_j^2})\right]^{1/2} \\
\leqslant\ & Cn^{-1/2}\sum_{j=L}^{j_2} 2^{j/2}\sigma_j f_{jn}\left[\log(C_{\varphi h}\beta_{jn} f_{jn}\sqrt{n}/\nu_j)\right]^{1/2}.
\end{aligned}
$$

By Lemma S5.7 and Lemma S5.2, the second sum is bounded by

$$
\begin{aligned}
& Cn^{-1/2}\sum_{j=j_2+1}^{J-1} 2^{j/2}[t_{jn}\sigma_j/\sqrt{n}]^{1+1/u}\exp\{-[t_{jn}\sigma_j/\sqrt{n}]^2/u\} \\
\leqslant\ & Cn^{-1/2}\sum_{j=j_2+1}^{J-1} 2^{j/2}[\log(\beta_{jn}\max(1,\sqrt{n}/\nu_j))]^{(u+1)/2u}[\beta_{jn}\max(1,\sqrt{n}/\nu_j)]^{-1/u} \\
\leqslant\ & Cn^{-1/2}\sum_{j=j_2+1}^{J-1} 2^{j/2}[\log(\beta_{jn} f_{jn}\sqrt{n}/\nu_j)]^{(u+1)/2u}[\beta_{jn} f_{jn}\sqrt{n}/\nu_j]^{-1/u}.
\end{aligned}
$$

Thus, Lemma S3.1 is proved.

If we bound the logarithmic term in the sums from above, we obtain the following corollary.

Corollary S3.1. *Assume that ν_j and β_{jn} are such that for all $j = L, L+1, \ldots, J-1$, $\nu_j \leqslant c\sqrt{n}$ and $\beta_{jn}\sqrt{n}/\nu_j \to \infty$ as $n \to \infty$ such that $\log(\beta_{jn}\sqrt{n}/\nu_j) \leqslant B\log n$ for some c, $B > 0$. Then, under assumptions of Lemma 1, for any $t_0 \in (0,1)$,*

$$
[R_n^u(\hat{f}_{BF}, B_{pq}^r(A), t_0)]^{1/u} \leqslant C\left(\frac{n}{\log n}\right)^{-\frac{r-1/p}{2(r-1/p)+1}} + C\left(\frac{n}{\log n}\right)^{-1/2-1/2u}\sum_{j=j_2+1}^{J-1} 2^{j/2}(\beta_{jn}/\nu_j)^{-1/u}. \tag{7.9}
$$

We are now ready to prove Theorems 5 and 6.

Proof of Theorem 5. By substituting assumption (3) of the theorem into Corollary S3.1, we obtain the following bound:

$$
[R_n^u(\hat{f}_{BF}, B_{pq}^r(A), t_0)]^{1/u} \leqslant C\left(\frac{n}{\log n}\right)^{-\frac{r-1/p}{2(r-1/p)+1}} + \left(\frac{n}{\log n}\right)^{-1/2-1/2u}\sum_{j=j_2+1}^{J-1} 2^{j/2} n^{-b/u} 2^{-aj/u}.
$$

The latter sum is bounded by

$$
\begin{aligned}
& C\left(\frac{n}{\log n}\right)^{-1/2}(\log n)^{1/2u}\begin{cases} n^{-b/u-1/2u}n^{\varepsilon_1(1/2-a/u)}, & 1/2-a/u<0\\ n^{-b/u-1/2u}\log n, & 1/2-a/u=0\\ n^{-b/u-1/2u}n^{(1/2-a/u)}, & 1/2-a/u>0\end{cases}\\
= \;& C\left(\frac{n}{\log n}\right)^{-1/2}(\log n)^{1/2u}\begin{cases} n^{-b/u-1/2u+\varepsilon_1(1/2-a/u)}, & a>u/2\\ n^{-b/u-1/2u}\log n, & a=u/2\\ n^{-b/u-1/2u+1/2-a/u}, & a<u/2\end{cases}\\
\leqslant \;& C\left(\frac{n}{\log n}\right)^{-1/2}(\log n)^{1/2u+1}n^{-b/u-1/2u+(u/2-a)_+/u}\\
\leqslant \;& C\left(\frac{n}{\log n}\right)^{-1/2}(\log n)^{1/2u+1}\leqslant C\left(\frac{n}{\log n}\right)^{-\frac{r-1/p}{2(r-1/p)+1}},
\end{aligned}
$$

due to assumption $b+1/2-(u/2-a)_+\geqslant 0$. Thus, Theorem 5 is proved.

Proof of Theorem 6. Following the proof of Theorem S1.1, we only need to consider (7.4), and using Lemma A.3 in Bochkina and Sapatinas (2006) and Lemma S5.6, we have that

$$
\begin{aligned}
[R_n^u(\hat f_{BF}, B_{pq}^r(A), t_0)]^{1/u} \;\leqslant\;& Cn^{-\nu/(2\nu+1)} + C\sum_{j=L}^{j_2} 2^{j/2}t_{jn}\\
&+C\sum_{j=j_2+1}^{J-1} 2^{j/2}\sum_{k\in K_j(t_0)}[\mathbb{E}|\theta_{jk}-d_{jk}|^u\mathbb{I}(|d_{jk}|>t_{jn})]^{1/u}.
\end{aligned}
$$

a) *Low resolution levels.* Since $\nu_j/\sqrt{n}\to 0$ and h is either heavy tailed (3.3) or normal, we can apply Lemma 4(i) in Pensky and Sapatinas (2007) to bound the first sum by

$$
Cn^{-1/2}\sum_{j=L}^{j_2}2^{j/2}[\log(\beta_{jn}\sqrt{n}/\nu_j)]^{1/\beta}\leqslant Cn^{-1/2}(\log n)^{1/2}2^{j_2/2} = C\left(\frac{n}{\log n}\right)^{-(r-1/p)/(2(r-1/p)+1)},
$$

since we assume that $\log(\beta_{jn}\sqrt{n}/\nu_j)\leqslant B[\log n]^{\beta/2}$.

b) *High resolution levels.* By Lemma S5.7, we have

$$
\mathbb{E}|\hat\theta_{jk}-\theta_{jk}|^u\mathbb{I}(|d_{jk}|>t_{jn})\leqslant Cn^{-u/2}[t_{jn}\sqrt{n}/\sigma_j]^{u+1}e^{-[t_{jn}\sqrt{n}/\sigma_j]^\beta},
$$

and by Lemma S5.2, we have $\varphi_j(t_{jn}\sqrt{n})\leqslant\varphi_j(0)(\beta_{jn}\sqrt{n}/\nu_j)^{-1}$ since $\nu_j/\sqrt{n}\to 0$. Hence, $t_{jn}\geqslant C\sigma_j n^{-1/2}[\log(\beta_{jn}\sqrt{n}/\nu_j)]^{1/\beta}$ with $\beta_{jn}\sqrt{n}/\nu_j\to\infty$, implying

$$
\mathbb{E}|\hat\theta_{jk}-\theta_{jk}|^u\mathbb{I}(|d_{jk}|>t_{jn})\leqslant Cn^{-u/2}[\log(\beta_{jn}\sqrt{n}/\nu_j)]^{(u+1)/\beta}[\beta_{jn}\sqrt{n}/\nu_j]^{-1},
$$

and

$$\sum_{j=j_2}^{J-1} \sum_{k \in K_j(t_0)} 2^{j/2} [\mathbb{E}|\hat{\theta}_{jk} - \theta_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn})]^{1/u} \leqslant C n^{-1/2-1/2u} [\log n]^{(u+1)/(2u)} \sum_{j=j_2}^{J-1} 2^{j/2} [\beta_{jn}/\nu_j]^{-1/u}.$$

We have the same sum as in the proof of Theorem 5 but with a different power of the logarithmic factor. Therefore, under the same assumptions as in Theorem 5, we obtain the optimal rate of convergence. Hence, Theorem 6 is proved.

S4. Proof of propositions for a priori Besov regularity

Proof of Proposition 1. According to Theorem 3 in Pensky and Sapatinas (2007), we need to check that

$$\lim_{n\to\infty} \sum_{j=L}^{J-1} [2^{j(r+1/2)} \nu_j^{-1} \beta_{jn}^{-1/p}]^{\min(p,\,q)} < \infty.$$

Denote $\kappa = \min(p,\, q) \in [1, \infty)$, then the sum can be written as

$$\begin{aligned} & C n^{-\kappa b_1/p} \sum_{j=L}^{j_1} 2^{\kappa j (r+1/2-m_1-a_1/p)} + C n^{-\kappa b_2/p} \sum_{j=j_1+1}^{J-1} 2^{\kappa j(r+1/2-m_2-a_2/p)} \\ = \; & C n^{-\kappa b_1/p + \kappa (r+1/2-m_1-a_1/p)_+} [\log n]^{I(r+1/2-1/p-m_1-a_1/p=0)} \\ + \; & C n^{-\kappa b_2/p + \kappa Z(r+1/2-m_2-a_2/p)} [\log n]^{I(r+1/2-1/p-m_2-a_2/p=0)}, \end{aligned}$$

where $Z = 1$ if $r+1/2-1/p-m_2-a_2/p \geqslant 0$ and $Z = \frac{1}{2(r-1/p)+1}$ if $r+1/2-1/p-m_2-a_2/p < 0$. This sum is finite as $n \to \infty$ if $b_1/p - (r+1/2-m_1-a_1/p)_+ \geqslant 0$ and $b_2/p - Z(r+1/2-m_2-a_2/p) \geqslant 0$, the inequalities are strict if $r+1/2-m_1-a_1/p = 0$ or $r+1/2-m_2-a_2/p = 0$ respectively.

Proof of Proposition 2. According to Theorem 3 in Pensky and Sapatinas (2007), we need to check that

$$\lim_{n\to\infty} \sum_{j=L}^{J-1} [2^{j(r+1/2)} \nu_j^{-1} \beta_{jn}^{-1/p}]^{\min(p,\,q)} < \infty.$$

Denote $\kappa = \min(p,\, q) \in [1, \infty)$, then the sum can be written as

$$C n^{-\kappa b/p} \sum_{j=L}^{J-1} 2^{\kappa j(r+1/2-m-(a+m)/p)} = C n^{-\kappa b/p + \kappa(r+1/2-m-(a+m)/p)_+} [\log n]^{I(r+1/2-1/p-m-(a+m)/p=0)}$$

which is finite as $n \to \infty$ if $b/p - (r+1/2-m-(a+m)/p)_+ \geqslant 0$, and the inequality is strict if $r+1/2-m-(a+m)/p = 0$.

S5. Auxiliary lemmas.

Lemma S5.2. *If φ_j and h are symmetric unimodal density functions, then,*

$$\varphi_j(\sqrt{n} t_{jn}) \;\; \leqslant \;\; \beta_{jn}^{-1} \min\left(\varphi_j(0), \frac{h(0)\nu_j}{\sqrt{n}}\right).$$

Proof of Lemma S5.2. Note that the symmetry and unimodality of φ_j implies that $\varphi_j(x) \le \varphi_j(0)$ for any x. Therefore, the equation for the threshold $t_{j,n}$ (see expression above (3.11)) can be rewritten as follows

$$\begin{aligned}\beta_{j,n} = \zeta_{j,n}(t_{j,n}) &= \frac{\int_{-\infty}^{+\infty} \sqrt{n}\,\varphi_j(\sqrt{n}(t_{j,n}-x))\nu_j h(\nu_j x)dx}{\sqrt{n}\,\varphi_j(\sqrt{n}t_{j,n})}\\ &\le \frac{\int_{-\infty}^{+\infty} \sqrt{n}\,\varphi_j(0)\nu_j h(\nu_j x)dx}{\sqrt{n}\,\varphi_j(\sqrt{n}t_{j,n})} = \frac{\varphi_j(0)}{\varphi_j(\sqrt{n}t_{j,n})}.\end{aligned}$$

Similarly, by symmetry and unimodality of h, we have

$$\begin{aligned}\beta_{j,n} = \zeta_{j,n}(t_{j,n}) &= \frac{\int_{-\infty}^{+\infty} \sqrt{n}\,\varphi_j(\sqrt{n}x)\nu_j h(\nu_j (t_{j,n}-x))dx}{\sqrt{n}\,\varphi_j(\sqrt{n}t_{j,n})}\\ &\le \frac{\int_{-\infty}^{+\infty} \sqrt{n}\,\varphi_j(\sqrt{n}x)\nu_j h(0)dx}{\sqrt{n}\,\varphi_j(\sqrt{n}t_{j,n})} = \frac{\nu_j h(0)}{\sqrt{n}\,\varphi_j(\sqrt{n}t_{j,n})}.\end{aligned}$$

Rearranging the terms, we have

$$\varphi_j(\sqrt{n}t_{j,n}) \le \min\left\{\beta_{j,n}^{-1}\varphi_j(0), \beta_{j,n}^{-1}h(0)\nu_j/\sqrt{n}\right\}. \tag{7.10}$$

Thus, Lemma S5.2 is proved.

The following lemma is an obvious corollary from Lemma S5.2.

Lemma S5.3. If $\varphi_j(x) = c_\beta \sigma_j^{-1} e^{-|x/\sigma_j|^\beta}$, $\beta > 0$, and h is a symmetric unimodal density, then

$$\sqrt{n}\, t_{j,n} \;\geqslant\; \sigma_j \max\left\{\left[\log\left(\frac{\beta_{j,n}\sqrt{n}}{h(0)\nu_j}\right)\right]^{1/\beta},\ [\log(\beta_{j,n})]^{1/\beta}\right\}.$$

Lemma S5.4. *Take $\varphi_j(x) \sim N(0,\sigma_j^2/2)$, $\varphi_j(x)/h(x) \leqslant C_{\varphi h}$, and let $\zeta_{jn}(x)$ be increasing for $x > 0$. Then,*

$$t_{jn} \leqslant \sigma_j n^{-1/2}\left(1 + \frac{\nu_j^2}{n}\right)^{1/2}\left[\log\left(C_{\varphi h}\beta_{jn}\sqrt{1+\frac{n}{\nu_j^2}}\right)\right]^{1/2}.$$

Proof of Lemma S5.4. Since $\zeta_{jn}(x)$ increases for $x > 0$, $\zeta_{jn}(x) \geqslant \beta_{jn}$ if and only if $x \geqslant t_{jn}$. We can find the following lower bound on $\zeta_{jn}(x)$ using $h(x) \geqslant C_{\varphi h}^{-1}\varphi_j(x)$. More precisely,

$$\begin{aligned}\zeta_{jn}(x) &= [\varphi_j(x\sqrt{n})]^{-1}\nu_j \int_{-\infty}^{\infty} \varphi_j((x-y)\sqrt{n})h(\nu_j y)dy\\ &\geqslant C_{\varphi h}^{-1}[\varphi_j(x\sqrt{n})]^{-1}\nu_j \int_{-\infty}^{\infty} \varphi_j((x-y)\sqrt{n})\varphi_j(\nu_j y)dy\\ &= C_{\varphi h}^{-1}\frac{\nu_j}{\sqrt{\pi}\sigma_j}\int_{-\infty}^{\infty} \exp\{-(n+\nu_j^2)y^2/\sigma_j^2 + 2nxy/\sigma_j^2\}dy\\ &= C_{\varphi h}^{-1}\frac{\nu_j}{\sqrt{\nu_j^2+n}}\exp\left\{\frac{n^2x^2}{(n+\nu_j^2)\sigma_j^2}\right\}.\end{aligned}$$

Take $x > 0$ such that $C_{\varphi h}^{-1} \frac{\nu_j}{\sqrt{\nu_j^2+n}} \exp\left\{\frac{n^2x^2}{(n+\nu_j^2)\sigma_j^2}\right\} = \beta_{jn}$. Then,

$$t_{jn} \leqslant x = \sqrt{n}\sigma_j \left(1 + \frac{\nu_j^2}{n}\right)^{1/2} \left[\log\left(C_{\varphi h}\beta_{jn}\sqrt{1 + \frac{n}{\nu_j^2}}\right)\right]^{1/2}.$$

Thus, Lemma S5.4 is proved.

Lemma S5.5. *Let φ_j and h be symmetric unimodal densities, and φ_j have finite variances σ_j^2, $0 < \underline{\sigma} \le \sigma_j \le \bar{\sigma} < \infty$, h has a bounded second derivative, and $\zeta_{jn}(x)$ increases for $x > 0$. Then, if $\nu_j/\sqrt{n} \to 0$ as $n \to \infty$, $\beta_{jn}\sqrt{n}/\nu_j \leqslant C_1$ if and only if $t_{jn} \leqslant C_0 n^{-1/2}$, for some $C_1, C_0 > 0$ which depend on φ and h but not on ν_j or n.*

Remark S5.1. *If the condition that h has a bounded second derivative is replaced with the conditions that $\varphi_j(x)/h(x) \leqslant C_{\varphi h}$ and φ_j have bounded second derivatives (uniformly in j), then $\beta_{jn}\sqrt{n}/\nu_j \leqslant C_1$ implies that $t_{jn} \leqslant C_0 n^{-1/2}$.*

Proof of Lemma S5.5. Consider the function $\zeta_{jn}(x)$ at point $x_n = \sqrt{n}v$, where v is independent of n. Then,

$$\begin{aligned}
\zeta_{jn}(x_n) &= [\varphi_j(\sqrt{n}x_n)]^{-1}\nu_j \int_{-\infty}^{+\infty} \varphi_j(\sqrt{n}y)h(\nu_j(x_n - y))dy \\
&= [\varphi_j(v)]^{-1}\frac{\nu_j}{\sqrt{n}} \int_{-\infty}^{+\infty} \varphi_j(z)h\left(\frac{\nu_j}{\sqrt{n}}(v - z)\right)dz \\
&= [\varphi_j(v)]^{-1}\frac{\nu_j}{\sqrt{n}} \int_{-\infty}^{+\infty} \varphi_j(z)\left[h(0) + \left(\frac{\nu_j}{\sqrt{n}}\right)^2 (v-z)^2 h''(c(v-z))/2\right]dz \\
&= [\varphi_j(v)]^{-1}\frac{\nu_j}{\sqrt{n}} \left[h(0) + O\left(\frac{\nu_j}{\sqrt{n}}\right)^2 (v^2 + \sigma_j^2).\right]
\end{aligned}$$

Since $\zeta_{jn}(x)$ increases, $t_{jn} \leqslant x_n$ if and only if $\zeta_{jn}(x_n) \geqslant \zeta_{jn}(t_{jn}) = \beta_{jn}$. The last inequality can be written as

$$[\varphi_j(v)]^{-1}\left[h(0) + O\left(\frac{\nu_j}{\sqrt{n}}\right)^2 (v^2 + \sigma_j^2)\right] \geqslant \beta_{jn}\sqrt{n}/\nu_j.$$

As $\nu_j/\sqrt{n} \to 0$, the left hand side tends to $C = h(0)/\varphi_j(v)$ uniformly in n, since σ_j^2 are bounded from above and below implying that $\beta_{jn}\sqrt{n}/\nu_j \leqslant C_0$ if and only if $t_{jn} \leqslant C_1 n^{-1/2}$, where $C_1 = v$ is a constant. Thus, Lemma S5.5 is proved.

We state the obvious lemma below due to its frequent use in the proofs.

Lemma S5.6. *Let $\hat\theta_{jk}$ be a hard thresholding estimator of θ_{jk} with threshold t_{jn} based on observation d_{jk}, and $1 \leqslant u < \infty$. Then,*

$$\begin{aligned}\mathbb{E}|\hat\theta_{jk} - \theta_{jk}|^u &= |\theta_{jk}|^u \mathbb{I}(|d_{jk}| \leqslant t_{jn}) + \mathbb{E}|d_{jk} - \theta_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn}) \\ &\leqslant |\theta_{jk}|^u + \mathbb{E}|d_{jk} - \theta_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn}).\end{aligned}$$

Lemma S5.7. *If $\varphi_j(x) = C_\beta \sigma_j^{-1} \exp\{-|x/\sigma_j|^\beta\}$ and $t_{jn}\sqrt{n} \to \infty$ for $j > j_1$ as $n \to \infty$, then*

$$[\mathbb{E}|\theta_{jk} - d_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn})]^{1/u} \leqslant \frac{C}{\beta}\left[t_{jn}\sigma_j/\sqrt{n}\right]^{u+1} \exp\{-[t_{jn}\sigma_j/\sqrt{n}]^\beta\}.$$

Proof of Lemma S5.7. It is easily seen that

$$\begin{aligned}\mathbb{E}|d_{jk} - \theta_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn}) &= \frac{\sqrt{n}}{\sigma_j\sqrt{2\pi}} \int_{|x|>t_{jn}} |x - \theta_{jk}|^u \exp\{-n|(x-\theta_{jk})/\sigma_j|^\beta\} dx \\ &= \frac{\sigma_j^u n^{-u/2}}{\sqrt{2\pi}} \int_{|y\sigma_j/\sqrt{n}+\theta_{jk}|>t_{jn}} |y|^u \exp\{-|y|^\beta\} dy.\end{aligned}$$

Since for $B > \max(0, [(u+1)/\beta - 1]^{1/\beta})$,

$$\int_B^\infty |z|^u e^{-|z|^\beta} dz \leqslant \frac{1}{\beta} B^{u+1} e^{-B^\beta}, \tag{7.11}$$

we have the following bound:

$$\begin{aligned}\mathbb{E}|d_{jk} - \theta_{jk}|^u \mathbb{I}(|d_{jk}| > t_{jn}) &\leqslant C\sigma_j^u n^{-u/2} |t_{jn}\sqrt{n}/\sigma_j|^{u+1} \exp\{-(t_{jn}\sqrt{n}/\sigma_j)^\beta\} \\ &= C\sigma_j^u n^{-u/2} (2uj\log 2)^{(u+1)/2} e^{-uj\log 2},\end{aligned}$$

by Lemma S5.6 and, for $j \geq j_1 + 1$,

$$\sqrt{n}|\theta_{jk}| \leq A\sqrt{n} 2^{-j(r-1/p+1/2)} \leq A\sqrt{n} 2^{-j_1(r-1/p+1/2)} = A \tag{7.12}$$

according to Lemma A.4 in Bochkina and Sapatinas (2006).

Now we prove (7.11).

$$\int_B^\infty z^u e^{-z^\beta} dz = [x = 1/z] = \int_0^{1/B} x^{-u-2} e^{-x^{-\beta}} dx$$

Function $x^{-u-2}e^{-x^{-\beta}}$ increases for $x < [\beta/(u+2)]^{1/\beta}$, since then

$$\frac{d}{dx}(x^{-u-2}e^{-x^{-\beta}}) = e^{-x^{-\beta}} x^{-u-3}[-(u+2) + \beta x^{-\beta}] < 0.$$

Thus, for all $x \leqslant 1/B < [\beta/(u+2)]^{1/\beta}$,

$$\int_B^\infty z^u e^{-z^\beta} dz = \int_0^{1/B} x^{-u-2} e^{-x^{-\beta}} dx \leqslant B^{-1} B^{u+2} e^{-B^\beta} = B^{u+1} e^{-B^\beta}.$$

Thus, Lemma S5.7 is proved.

Lemma S5.8. *Let φ and h be Student's t distributions with density (4.2) with ρ and γ degrees of freedom respectively, $0<\gamma<\rho$, $\varphi_j(x)=\sigma_j^{-1}\varphi(x/\sigma_j)$, $0<\underline{\sigma}\le\sigma_j\le\bar{\sigma}<\infty$. Then, if $\nu_j/\sqrt{n}\to 0$ as $n\to\infty$, the threshold $t_{j,n}$ defined by (3.10) asymptotically satisfies*

$$t_{jn}\sqrt{n}/\sigma_j=(1+o(1))\begin{cases} C_1, & \beta_{j,n}<C_0\frac{\nu_j\sigma_j}{\sqrt{n}},\\ \left[C_0^{-1}\beta_{j,n}\frac{\sqrt{n}}{\nu_j\sigma_j}\right]^{1/(\rho+1)}, & C_0\frac{\nu_j\sigma_j}{\sqrt{n}}\leqslant\beta_{j,n}\leqslant C_0\left(\frac{\sqrt{n}}{\nu_j\sigma_j}\right)^{\rho},\\ \left[C_0^{-1}\beta_{j,n}\left(\frac{\nu_j\sigma_j}{\sqrt{n}}\right)^{\gamma}\right]^{1/(\rho-\gamma)}, & \beta_{j,n}>C_0\left(\frac{\sqrt{n}}{\nu_j\sigma_j}\right)^{\rho},\end{cases}$$

for some constants $C_0,C_1>0$ depending only on γ and ρ.

Proof of Lemma S5.8. We can obtain the asymptotic expression for the threshold following the proof of Lemma 4(ii) in Pensky and Sapatinas (2007) under the assumptions of the lemma, that for some constant $C_0>0$, $\zeta_{j,n}(x)=C_0\frac{\nu_j\sigma_j}{\sqrt{n}}F(x)$ with $F(x)=(\gamma+\nu_j^2x^2)^{-(\gamma+1)/2}(\rho+nx^2/\sigma_j^2)^{(\rho+1)/2}(1+o(1))$ (due to Lemma 2 in Pensky and Sapatinas (2007)), and

$$F(x)=\begin{cases} c_\rho c_\gamma^{-1}(1+o(1)), & |x|<\sigma_j/\sqrt{n},\\ c_\gamma^{-1}(\sqrt{n}x/\sigma_j)^{\rho+1}[1+o(1)], & \sigma_j/\sqrt{n}\leqslant|x|\leqslant 1/\nu_j,\\ (\sqrt{n}x/\sigma_j)^{\rho+1}(\nu_jx)^{-\gamma-1}[1+o(1)], & |x|>1/\nu_j,\end{cases}$$

where $c_k=k^{(k+1)/2}$. Since the threshold satisfies $\beta_{j,n}=\zeta_{j,n}(t_{j,n})=C_0\frac{\nu_j\sigma_j}{\sqrt{n}}F(t_{j,n})$, we obtain that

$$t_{jn}\sqrt{n}/\sigma_j=(1+o(1))\begin{cases} C_1, & \frac{\beta_{j,n}\sqrt{n}}{\nu_j\sigma_j}<C_0,\\ \left[C_0^{-1}\frac{\beta_{j,n}\sqrt{n}}{\nu_j\sigma_j}\right]^{1/(\rho+1)}, & C_0\leqslant\frac{\beta_{j,n}\sqrt{n}}{\nu_j\sigma_j}\leqslant C_0\left(\frac{\sqrt{n}}{\nu_j\sigma_j}\right)^{\rho+1},\\ \left[C_0^{-1}\frac{\beta_{j,n}\sqrt{n}}{\nu_j\sigma_j}\left(\frac{\nu_j\sigma_j}{\sqrt{n}}\right)^{\gamma+1}\right]^{1/(\rho-\gamma)}, & \frac{\beta_{j,n}\sqrt{n}}{\nu_j\sigma_j}>C_0\left(\frac{\sqrt{n}}{\nu_j\sigma_j}\right)^{\rho+1},\end{cases}$$

which is the statement of the lemma.