The Annals of Statistics
1998, Vol. 26, No. 6, 2420–2469

POINTWISE AND SUP-NORM SHARP ADAPTIVE ESTIMATION OF FUNCTIONS ON THE SOBOLEV CLASSES

BY A. B. TSYBAKOV

Université Paris VI and Institute for Problems of Information Transmission, Moscow

The problem of nonparametric function estimation in the Gaussian white noise model is considered. It is assumed that the unknown function belongs to one of the Sobolev classes, with an unknown regularity parameter. Asymptotically exact adaptive estimators of functions are proposed on the scale of Sobolev classes, with respect to pointwise and sup-norm risks. It is shown that, unlike the case of L_2-risk, a loss of efficiency under adaptation is inevitable here. Bounds on the value of the loss of efficiency are obtained.

1. Introduction. The problem of minimax adaptive estimation of a nonparametric function f from noisy data has been studied in a number of papers [see, for example, Efroimovich and Pinsker (1984), Lepski (1990), Golubev and Nussbaum (1992), Donoho, Johnstone, Kerkyacharian and Picard (1995), Härdle, Kerkyacharian, Picard and Tsybakov (1998) and the references cited therein]. These papers deal with adaptation to unknown smoothness of f. It is assumed that f belongs to a smoothness class $\mathscr{F}_\beta$ (usually, Hölder, Sobolev or Besov classes) where β is the unknown smoothness, that is, the number of derivatives of f that are bounded in a certain sense. The aim is to find an estimator f^* of f, independent of β and such that f^* attains asymptotically optimal behavior (in a minimax sense) on all the classes $\mathscr{F}_\beta$, for $\beta \in B$, where B is a given set.

Several questions arise in this context. For the first approximation, the asymptotically optimal behavior can be considered in terms of rates of convergence. In most of the cases it is well known that, for a fixed β, there exists an estimator, depending on β and achieving optimal (minimax) rate of convergence on $\mathscr{F}_\beta$ [Ibragimov and Hasminskii (1981), Stone (1980, 1982)]. The question is whether one can find an estimator f^*, independent of β and attaining this rate uniformly over $\beta \in B$. Such an estimator is called *optimal rate adaptive*. In typical cases of Hölder, Sobolev or Besov scales or classes $\{\mathscr{F}_\beta,\ \beta \in B\}$ and the L_p risks, the answer to this question is positive [Lepski (1991, 1992a, b), Donoho, Johnstone, Kerkyacharian and Picard (1995), Lepski, Mammen and Spokoiny (1997), Goldenshluger and Nemirovskii (1997),

Received November 1995; revised August 1998.
AMS 1991 *subject classifications*. 62G05, 62G20.
Key words and phrases. Gaussian white noise, adaptive nonparametric estimation, Sobolev class, loss of efficiency under adaptation, minimax risk, exact constants.

Juditsky (1997)]. As shown in these papers, optimal rate adaptive estimators can be constructed starting from kernel, spline, piecewise-polynomial and wavelet estimators.

If optimal rate adaptive estimators exist, the next question is how much one loses in the asymptotic constant when using an adaptive estimator. In other words, what is the *loss of efficiency under adaptation*? One can define the loss of efficiency under adaptation as the maximal (over $\beta \in B$) ratio of the risk of the best adaptive estimator to the minimax (nonadaptive) risk. In their pioneering paper, Efroimovich and Pinsker (1984) show that there is no loss of efficiency in the "Pinsker case," where $\{\mathscr{F}_\beta\}$ is the scale of L_2-Sobolev classes and the risk is measured in L_2 as well. This means that the above ratio is asymptotically 1. On the other hand, if $\{\mathscr{F}_\beta\}$ is the scale of Hölder classes with smoothness $0 < \beta \le 1$ and the risk is measured in sup-norm (the "Korostelev case"), Lepski (1992b) shows that, in general, the loss of efficiency does exist; that is, the above ratio is asymptotically strictly greater than 1.

Except for these two cases, the exact asymptotics of the loss of efficiency is not investigated. The problem is very difficult, since one should dispose of the exact asymptotics of both, minimax and "minimax adaptive" risks. However, even the exact asymptotics of minimax risk is known only in special cases [Pinsker (1980), Korostelev (1993)].

The study of the loss of efficiency makes sense only in situations where optimal rate adaptive estimators exist. However, this is not always the case. As shown by Lepski (1990) and Brown and Low (1996), there are no optimal rate adaptive estimators on the scale of Hölder classes $\{\mathscr{F}_\beta\}$ when the pointwise risk is used. They prove that in this situation the best adaptive estimators can only achieve the rate that is slower than the optimal one in a logarithmic factor. The "adaptive," logarithmically slower rates are inherent for this problem. The estimators f^* independent of β and achieving these new rates are called simply *rate adaptive* (to make a distinction from *optimal rate adaptive* ones). The next natural step is to find the best among all rate adaptive estimators in the sense of exact risk asymptotics. For the scale of Hölder classes $\{\mathscr{F}_\beta\}$ and under the pointwise risk this was done by Lepski and Spokoiny (1997).

The purpose of the present paper is to analyze the exact asymptotics of minimax adaptive risks on the scale of L_2 Sobolev classes in the following two cases:

(1) Estimation under the sup-norm risk.
(2) Pointwise estimation.

The results are obtained in the Gaussian white noise model.

In case (1) optimal rate adaptive estimators exist. We find the best among them in the sense of exact "adaptive" risk asymptotics. By comparing to the asymptotics of usual minimax risk, we show that for this model the loss of efficiency under adaptation does occur. We give upper and lower bounds on the loss of efficiency. The exact asymptotics of this quantity, however, remains unknown, since it is related to the unknown exact asymptotics of usual minimax risks over Sobolev classes in sup-norm. To bound the loss of efficiency we find the asymptotics of minimax sup-norm risks for linear estimators. However, the question of whether the exact asymptotics of general minimax risk is attained on linear estimators remains open.

For case (2) we first prove that optimal rate adaptive estimators do not exist and find the adaptive rate of convergence. This result is analogous to Lepski (1990) and Brown and Low (1996), but we consider the Sobolev (rather than Hölder) scale of classes. Next, we find the best (in the sense of exact asymptotics) among rate adaptive estimators on the Sobolev scale of classes. This is a Sobolev scale counterpart of the Hölder scale result by Lepski and Spokoiny (1997). Their adaptation procedure works in the range of smoothness $0 < \beta \leq 2$ and can be written explicitly in the range $0 < \beta \leq 1$. This limitation is related to a nonnestedness property of Hölder classes and to the fact that the explicit solution to optimal recovery problems for Hölder classes, in general, is not known [cf. Donoho (1994)]. For the Sobolev classes considered here, the situation turns out to be more favorable. As shown below, the explicit construction of asymptotically exact adaptive estimators for the Sobolev scale is possible in a wide range of values β, which is not the case for the Hölder scale.

2. The model and definitions. Consider the stochastic process $Y(t)$ on $[0, 1]$ satisfying the stochastic differential equation

$$dY(t) = f(t)\,dt + \varepsilon\,dW(t), \tag{2.1}$$

where $W(t)$ is the standard Wiener process, $f(t)$ is an unknown function in $L_2[0, 1]$, and $0 < \varepsilon < 1$ is a small number. The problem is to estimate the function f, given a sample path of the process $\{Y(t),\ 0 \leq t \leq 1\}$. The model (2.1) is called the Gaussian white noise model [Ibragimov and Hasminskii (1981)].

Assume that f is a smooth function belonging to a Sobolev class on $[0, 1]$. The degree of smoothness is characterized by a positive parameter β. If β is an integer, the Sobolev class is usually defined as the set of all β times differentiable functions f on $[0, 1]$ such that

$$\int_0^1 \left(f^{(\beta)}(t)\right)^2 dt \leq L^2, \tag{2.2}$$

where $L > 0$ and $f^{(\beta)}$ is the βth derivative of f. This definition is extended in a standard way to noninteger values β, at the expense of imposing the periodicity constraint on f and its derivatives of order less than β. Under such a constraint, the definition of the Sobolev class is given in terms of Fourier coefficients of f. It is introduced below and used throughout the paper.

Let $\{\varphi_k(t),\ k = 0, 1, \dots\}$ be the orthonormal trigonometric basis on $[0, 1]$,

$$\varphi_0(t) \equiv 1, \qquad \varphi_{2l-1}(t) = \sqrt{2}\,\sin(2\pi l t),$$
$$\varphi_{2l}(t) = \sqrt{2}\,\cos(2\pi l t),\ l = 1, 2, \dots .$$

Define the Fourier coefficients of f,

$$\theta_k = \int_0^1 f(t)\varphi_k(t)\,dt$$

and introduce the following class of functions on $[0,1]$:

$$W_\beta = \left\{ f(x) = \sum_{k=0}^{\infty} \theta_k \varphi_k(x) : \sum_{k=0}^{\infty} a_k^2(\beta)\theta_k^2 \le Q_\beta \right\},$$

where $Q_\beta = L^2/\pi^{2\beta}$, $\beta > 1/2$ and

$$(2.3) \quad a_0(\beta) = 0, \qquad a_k(\beta) = \begin{cases} (k+1)^\beta, & \text{for } k \text{ odd}, \\ k^\beta, & \text{for } k \text{ even}, \end{cases} \qquad k = 1, 2, \dots .$$

The functions $f \in W_\beta$ satisfy (2.2) and the periodicity constraints mentioned above. In the sequel the name *Sobolev class* will be used for W_β. It is assumed everywhere that $\beta > 1/2$. This condition guarantees that the functions $f \in W_\beta$ are continuous.

Let T_ε be an estimator of f based on the observations $\{Y(t),\ 0 \le t \le 1\}$. The estimation error of T_ε is defined by its maximal risk

$$\mathscr{R}_{\varepsilon,\beta}(T_\varepsilon, \psi_\beta) = \sup_{f \in W_\beta} E_f\left(\psi_\beta^{-p}(\varepsilon) d^p(T_\varepsilon, f)\right),$$

where $p > 0$ is a fixed number, E_f denotes the expectation with respect to the distribution P_f of observations satisfying (2.1), $d(\cdot,\cdot)$ is a given distance and $\psi_\beta(\varepsilon)$, $0 < \varepsilon < 1$, is a normalizing factor, depending on ε and β, such that $\psi_\beta(\varepsilon) > 0$ and $\psi_\beta(\varepsilon) \to 0$, as $\varepsilon \to 0$, for every β.

If the smoothness β is known, then, for various distances d, it is possible to construct estimators $\hat{f}_{\varepsilon,\beta}$ (depending on β) that achieve the *optimal rate of convergence* (ORC) in the following sense: there exists $\psi_\beta^*(\varepsilon)$ (called ORC on W_β for the distance d), such that

$$(2.4) \qquad c \le \liminf_{\varepsilon \to 0} \inf_{T_\varepsilon} \mathscr{R}_{\varepsilon,\beta}(T_\varepsilon, \psi_\beta^*) \le \limsup_{\varepsilon \to 0} \mathscr{R}_{\varepsilon,\beta}\left(\hat{f}_{\varepsilon,\beta}, \psi_\beta^*\right) \le C,$$

with positive constants c and C [Ibragimov and Hasminskii (1981, 1984), Stone (1980, 1982), Birgé (1983), Nemirovskii (1985)]. Here and later $\inf_{T_\varepsilon}$ denotes the infimum over all estimators. The ORC is not uniquely defined, and we denote $\{\psi_\beta^*(\cdot)\}$ the set of all ORC $\psi_\beta^*(\cdot)$ for fixed β.

A harder problem is to find an estimator $\hat{f}_{\varepsilon,\beta}$ which satisfies the exact asymptotic equality

$$(2.5) \qquad \lim_{\varepsilon \to 0} \inf_{T_\varepsilon} \mathscr{R}_{\varepsilon,\beta}(T_\varepsilon, \psi_\beta^*) = \lim_{\varepsilon \to 0} \mathscr{R}_{\varepsilon,\beta}\left(\hat{f}_{\varepsilon,\beta}, \psi_\beta^*\right).$$

The estimator $f_{\varepsilon,\beta}^*$ satisfying (2.5) is called *asymptotically efficient on* W_β *for the distance* d. Asymptotically efficient estimators on W_β are known only for the situation where d is the L_2-distance [Pinsker (1980)]. [The result of Pinsker (1980) covers the case $p = 2$; for its extension to all $p > 0$ see Tsybakov (1997).]

If β is unknown, the following adaptive set-up can be used. Assume that, instead of β, the statistician knows only a set of possible smoothness values $B \subset (1/2, \infty)$, such that $\beta \in B$. Then it is desirable to find an estimator f_ε^* which is independent of β and attains the ORC for any $\beta \in B$.

The estimator f_ε^* is called *optimal rate adaptive on the scale of classes* $\{W_\beta, \beta \in B\}$ *for the distance* d, if for $\Psi_\beta(\varepsilon) = \psi_\beta^*(\varepsilon)$, one has

$$\limsup_{\varepsilon \to 0} \sup_{\beta \in B} \mathscr{R}_{\varepsilon, \beta}(f_\varepsilon^*, \Psi_\beta) \le C, \tag{2.6}$$

where $\psi_\beta^*(\varepsilon)$ is any ORC on W_β for the distance d, and $C > 0$ is a constant. As compared to (2.4), the uniformity over $\beta \in B$ is required here.

Optimal rate adaptive estimates on the Sobolev scale $\{W_\beta\}$ are known for different L_p distances d [Efroimovich and Pinsker (1984), Lepski (1991), Kneip (1994), Donoho and Johnstone (1995), Lepski, Mammen and Spokoiny (1997), Goldenshluger and Nemirovskii (1997), Juditsky (1997)]. Using the methods of these papers, it is easy to construct optimal rate adaptive estimates on $\{W_\beta, \beta \in B\}$ for the sup-norm distance

$$d_\infty(f, g) = \|f - g\|_\infty = \sup_{x \in [0,1]} |f(x) - g(x)|,$$

under rather general assumptions on B. An ORC on W_β for the sup-norm distance is $\psi_\beta^*(\varepsilon) = C(\beta)(\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta}$ where $C(\beta)$ is a positive constant depending on β.

In this paper we propose asymptotically the best estimator among the variety of optimal rate adaptive ones, that is, the estimator that guarantees the smallest value of the constant C in (2.6). This property is expressed by the condition

$$\lim_{\varepsilon \to 0} \inf_{T_\varepsilon} \sup_{\beta \in B} \mathscr{R}_{\varepsilon, \beta}(T_\varepsilon, \Psi_\beta) = \lim_{\varepsilon \to 0} \sup_{\beta \in B} \mathscr{R}_{\varepsilon, \beta}(f_\varepsilon^*, \Psi_\beta). \tag{2.7}$$

DEFINITION 1. An optimal rate adaptive estimator f_ε^* is called *asymptotically exact adaptive* (AEA) *on the scale of classes* $\{W_\beta, \beta \in B\}$ *for the distance* d if it satisfies (2.7) where $\psi_\beta(\varepsilon) \in \{\Psi_\beta^*(\cdot)\}$ for every fixed β.

If d is the $L_2[0,1]$-distance, the AEA estimators on the Sobolev scale $\{W_\beta, \beta \in B\}$ are known [Efroimovich and Pinsker (1984), Golubev (1990), Golubev and Nussbaum (1992)]. For other distances d this problem is not solved. In this paper we give its solution for the case where $d = d_\infty$ (the sup-norm distance). We also evaluate the loss of efficiency under adaptation defined as follows.

For an arbitrary estimator T_ε, denote

$$Q_{\varepsilon, \beta}(T_\varepsilon) = \sup_{f \in W_\beta} \left[E_f(\|T_\varepsilon - f\|_\infty^p) \right]^{1/p}.$$

DEFINITION 2. The value

$$\mathrm{LEF}(\varepsilon, p) = \inf_{f_\varepsilon^*} \sup_{\beta \in B} \frac{Q_{\varepsilon, \beta}(f_\varepsilon^*)}{\inf_{T_\varepsilon} Q_{\varepsilon, \beta}(T_\varepsilon)},$$

where $\inf_{f_\varepsilon^*}$ and $\inf_{T_\varepsilon}$ denote the infimum over all estimators is called *loss of efficiency under adaptation on the scale of classes* $\{W_\beta,\ \beta \in B\}$, *for the distance* d_∞.

The denominator in Definition 2 is the usual minimax risk for fixed β, and $\inf_{T_\varepsilon}$ is also the infimum over all estimators. But since in the denominator β is fixed, T_ε can depend on β, while f_ε^* cannot. The value $\mathrm{LEF}(\varepsilon, p)$ measures the relative loss of efficiency when using the "best adaptive" estimator f_ε^* instead of the best estimator with known β. Under the assumptions used below, the value $\mathrm{LEF}(\varepsilon, p)$ is well defined for ε small enough. The bounds on $\mathrm{LEF}(\varepsilon, p)$ are given in Section 3.

The second problem considered in this paper is adaptive estimation at a fixed point $x_0 \in (0, 1)$; we take $d = d_0$, where d_0 is the pointwise distance

$$d_0(f, g) = |f(x_0) - g(x_0)|.$$

If β is fixed, the ORC on W_β for the pointwise distance d_0 is $\psi_\beta^*(\varepsilon) = \varepsilon^{(2\beta-1)/2\beta}$ [Donoho and Low (1992)]. However, as shown below, in the case $d = d_0$, the condition (2.6) cannot be satisfied with $\Psi_\beta(\varepsilon) = \psi_\beta^*(\varepsilon) = \varepsilon^{(2\beta-1)/2\beta}$, and thus optimal rate adaptive estimators do not exist. Furthermore, if $d = d_0$, the relation (2.6) holds with the rate $\Psi_\beta(\varepsilon) = (\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta} \gg \psi_\beta^*(\varepsilon)$. Our aim is to study the exact asymptotics of the risk in the spirit of Definition 1, with $\Psi_\beta(\varepsilon) = (\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta}$. However, first we have to answer the question whether the rate $\Psi_\beta(\varepsilon) = (\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta}$ gives a proper normalization (i.e., whether it cannot be improved). This question is more delicate than for the case $d = d_\infty$: in fact, for $d = d_0$ the rate $\Psi_\beta(\varepsilon) = (\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta}$ is no longer an ORC. At first glance, to show the optimality of this rate, it suffices to complete (2.6) by the analogous lower bound

$$\liminf_{\varepsilon \to 0} \inf_{T_\varepsilon} \sup_{\beta \in B} \mathscr{R}_{\varepsilon, \beta}(T_\varepsilon, \Psi_\beta) \geq c, \tag{2.8}$$

where $c > 0$. However, it turns out that this is not enough, since (2.6) and (2.8) can be simultaneously satisfied with quite different normalizing factors $\Psi_\beta(\varepsilon)$. In fact, consider the following example. Let $d = d_0$ and let B contain only two values: $B = \{\beta', \beta''\}$ such that $\beta' < \beta''$. Consider the normalizing factors $\Psi_{\beta,1}(\varepsilon) \equiv \varepsilon^{(2\beta'-1)/2\beta'}$, $\forall\, \beta \in B$ and $\Psi_{\beta',2}(\varepsilon) = (\varepsilon^2 \log(1/\varepsilon))^{(2\beta'-1)/4\beta'}$, $\Psi_{\beta'',2}(\varepsilon) = \varepsilon^{(2\beta''-1)/2\beta''}$. Note that $\Psi_{\beta,1}(\varepsilon)$ is the ORC for the worst smoothness β'. It is easy to show that there exists an estimator f_ε^* (for example, a kernel estimator with bandwidth $\sim \varepsilon^{1/\beta'}$), such that

$$\limsup_{\varepsilon \to 0} \mathscr{R}_{\varepsilon, \beta'}(f_\varepsilon^*, \Psi_{\beta',1}) < \infty.$$

This implies, using the inclusion $W_{\beta'} \supset W_{\beta''}$ and the equality $\Psi_{\beta',1} = \Psi_{\beta'',1}$, that

$$\limsup_{\varepsilon\to 0} \mathscr{R}_{\varepsilon,\beta''}\big(f_\varepsilon^*, \Psi_{\beta'',1}\big) < \infty.$$

Hence, (2.6) holds with $\Psi_\beta = \Psi_{\beta,1}$. Next, as $\Psi_{\beta',1}$ is an ORC on $W_{\beta'}$, we get

$$\liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B} \mathscr{R}_{\varepsilon,\beta}\big(T_\varepsilon, \Psi_{\beta,1}\big) \ge \liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \mathscr{R}_{\varepsilon,\beta'}\big(T_\varepsilon, \Psi_{\beta',1}\big) \ge c > 0.$$

We conclude that both (2.6) and (2.8) are satisfied with $\Psi_\beta = \Psi_{\beta,1}$. On the other hand, as follows from the proof of Theorem 4 below, (2.6) and (2.8) hold with $\Psi_\beta = \Psi_{\beta,2}$ as well. The nonuniqueness effect present in this example suggests that (2.6) and (2.8) are not sufficient to define the correct normalizing factor Ψ_β. One should rather use the following definition [Lepski (1996)].

DEFINITION 3. The normalizing factor $\Psi_\beta(\varepsilon)$ is called *adaptive rate of convergence* (ARC) *on the scale of classes* $\{W_\beta,\ \beta \in B\}$ *for the distance* d if:

1°. condition (2.6) holds for some estimator f_ε^*;
2°. the rate of convergence $S_\beta(\varepsilon) > 0$ satisfies for some estimator f_ε^{**} the analog of (2.6):

$$\limsup_{\varepsilon\to 0} \sup_{\beta\in B} \mathscr{R}_{\varepsilon,\beta}\big(f_\varepsilon^{**}, S_\beta\big) \le C, \tag{2.9}$$

and the condition

$$\exists\, \beta' \in B : S_{\beta'}(\varepsilon)/\Psi_{\beta'}(\varepsilon) \to 0 \quad \text{as } \varepsilon \to 0, \tag{2.10}$$

then there exists $\beta'' \in B$ such that

$$\big[S_{\beta'}(\varepsilon)/\Psi_{\beta'}(\varepsilon)\big]\big[S_{\beta''}(\varepsilon)/\Psi_{\beta''}(\varepsilon)\big] \to \infty \quad \text{as } \varepsilon \to 0. \tag{2.11}$$

The estimator f_ε^* satisfying (2.6), where $\Psi_\beta(\varepsilon)$ is the ARC, is called *rate adaptive*.

Definition 3 allows ruling out the effect described in the above example. In words, Definition 3 states that the ARC $\Psi_\beta(\varepsilon)$ is such that any improvement of this rate at a point $\beta' \in B$ is possible only at the expense of much greater loss at another point $\beta'' \in B$. In fact, relation (2.11) states that not only does $S_{\beta''}(\varepsilon)/\Psi_{\beta''}(\varepsilon)$ converge to ∞, as $\varepsilon \to 0$, but it also converges to ∞ faster than $S_{\beta'}(\varepsilon)/\Psi_{\beta'}(\varepsilon)$ goes to 0.

REMARK 1. Similarly to ORC, the ARC is not unique. It is defined up to a bounded positive factor: if $\Psi_\beta(\varepsilon)$ is an ARC, then $a_{\varepsilon,\beta}\Psi_\beta(\varepsilon)$ is also an ARC, for any $a_{\varepsilon,\beta}$ such that $c' \le a_{\varepsilon,\beta} \le c''$, $\forall\, \varepsilon, \beta$, where $0 < c' < c'' < \infty$.

REMARK 2. If optimal rate adaptive estimators exist, they are also rate adaptive, and the ARC $\Psi_\beta(\varepsilon)$ coincides with an ORC $\psi_\beta^*(\varepsilon)$. In fact, in this case it is not possible to find S_β satisfying (2.9) and (2.10) simultaneously, and only condition 1° of Definition 3 is active.

In Theorem 4 we find the ARC on $\{W_\beta,\ \beta \in B\}$ for the pointwise distance d_0 under certain assumptions on B. Except for the largest value β in B, the ARC has the form $\Psi_\beta(\varepsilon) \sim (\varepsilon^2 \log(1/\varepsilon))^{(2\beta-1)/4\beta}$ (up to a factor depending on β). Using this $\Psi_\beta(\varepsilon)$ in the definition of the risk, we construct the AEA estimator for the pointwise distance in the sense defined in Section 4.

REMARK 3. The fact that $\{W_\beta,\ \beta \in B\}$ is the Sobolev scale of classes is not crucial for the definitions given above. One can use them in the general situation meaning that $\{W_\beta,\ \beta \in B\}$ is an arbitrary scale of classes.

REMARK 4. The presentation here is restricted to the periodic Sobolev classes. This is done to simplify the technicalities. The results of the paper can be extended, under appropriate assumptions, to the general Sobolev classes. In this case the trigonometric basis should be replaced by a special orthonormal basis $\{\varphi_k\}$ guaranteeing the equivalence of (2.2) to the ellipsoid W_β where the coefficients $a_k(\beta)$ satisfy (2.3) only asymptotically, as $k \to \infty$ [see, e.g., Oudshoorn (1996)].

REMARK 5. We work with the Gaussian white noise model (2.1) but the results of the paper can be extended to other statistical models as well. An extension to nonparametric regression with regular deterministic design can be obtained along the same lines. An extension to nonparametric density estimation is given by Butucea (1998). Simulation study in Butucea (1998) shows that, for density estimation, the exact adaptive procedure analogous to the one proposed below works well on the data.

To finish this section, we make some remarks on the construction of adaptive estimators. The largest group of adaptive methods proposed in the literature is based on the empirical L_2-risks estimators [the idea goes back to Mallows' C_p and Akaike's criteria or Stein's unbiased risk estimator, and in a general form it was recently developed by Kneip (1994), Birgé and Massart (1997), Barron, Birgé and Massart (1995), where one can find other references]. The wavelet thresholding adaptation procedure of Donoho and Johnstone (1995) and Donoho, Johnstone, Kerkyacharian and Picard (1995) may be also interpreted along these lines. A different idea of adaptive estimation, unrelated to the L_2 structure and based on implicit bias-variance comparison schemes, is due to Lepski (1990, 1991, 1992a, b). This idea is used below. We show that, both for the sup-norm and pointwise distance, the AEA estimators can be constructed as spline-type estimators, "adaptively" modified following the scheme of Lepski.

3. Exact adaptation in sup-norm and loss of efficiency effect. In this section we present the AEA estimators for the sup-norm and evaluate the loss of efficiency under adaptation.

First, introduce the assumptions on the set B. Let $\beta_1 > 1/2$ and β_ε^* be the real numbers such that β_1 is fixed and $\beta_\varepsilon^* > \beta_1$ depends on ε. In this paper, B is the discrete set

$$B = \{\beta_1, \beta_2, \dots, \beta_s\}$$

where $1/2 < \beta_1 < \cdots < \beta_s = \beta_\varepsilon^*$ and $s > 1$ is an integer. In general, s and β_i, $i > 1$, can depend on ε, but we skip this dependence in the notation. Assume that $\beta_{i+1} - \beta_i \ge \Delta$, $i = 1, \dots, s-1$, where $\Delta = \Delta_\varepsilon > 0$. The value Δ_ε can be either finite and bounded away from 0 (for example, an interesting case is $\beta_i = i$, $\Delta_\varepsilon = 1$) or $\Delta_\varepsilon \to 0$, as $\varepsilon \to 0$.

Assume the following conditions on β_ε^* and Δ_ε:

$$\tag{3.1} \begin{gathered} \lim_{\varepsilon \to 0} \beta_\varepsilon^* = \infty, \qquad \limsup_{\varepsilon \to 0} \Delta_\varepsilon < \infty, \\ \nu_\varepsilon = \frac{\Delta_\varepsilon \log(1/\varepsilon)}{(\beta_\varepsilon^*)^2 \log\log(1/\varepsilon)} \to \infty \quad \text{as } \varepsilon \to 0. \end{gathered}$$

The assumption $\limsup_{\varepsilon \to 0} \Delta_\varepsilon < \infty$ is introduced for notational convenience only. Moreover, we assume in the proofs that $\Delta_\varepsilon < 1$. This is done without loss of generality, since Δ_ε is a lower bound on the differences $\beta_{i+1} - \beta_i$.

Define the positive numbers b_β and v_β by

$$b_\beta^2 = \frac{1}{\pi} \int_0^\infty \frac{t^{2\beta}}{(1 + t^{2\beta})^2}\, dt = \frac{1}{2\pi\beta} \mathscr{B}\big(1 + (2\beta)^{-1}, 1 - (2\beta)^{-1}\big),$$

$$v_\beta^2 = \frac{1}{\pi} \int_0^\infty \frac{1}{(1 + t^{2\beta})^2}\, dt = \frac{1}{2\pi\beta} \mathscr{B}\big((2\beta)^{-1}, 2 - (2\beta)^{-1}\big),$$

where $\mathscr{B}(x, y)$ denotes the beta-function and $\beta > 1/2$.

Let j be an index taking the values $0, 1$ and ∞. The values $j = 0$ and $j = \infty$ correspond to estimation in d_0 and d_∞ distances, respectively. The value $j = 1$ will appear in the context of linear minimax risks. Define

$$\alpha_j = \begin{cases} 2, & \text{if } j = 1, \\ p, & \text{if } j = 0, \\ p + 2, & \text{if } j = \infty, \end{cases}$$

$$\kappa_{\beta, j} = \left(\frac{\alpha_j}{L^2 \beta (2\beta - 1)} \right)^{1/2\beta};$$

$$c(\beta, j) = b_\beta (2\beta)^{(2\beta+1)/4\beta} \left(\frac{2\alpha_j}{2\beta - 1} \right)^{(2\beta-1)/4\beta} L^{1/2\beta};$$

$$h_{\beta, j} = \kappa_{\beta, j} \left(\varepsilon^2 \log \frac{1}{\varepsilon} \right)^{1/2\beta} \quad \text{if } \beta \ne \beta_\varepsilon^* \text{ or } j \ne 0; \qquad h_{\beta_\varepsilon^*, 0} = \varepsilon^{1/\beta_\varepsilon^*};$$

$$\psi_{\beta, j} = c(\beta, j) \left(\varepsilon^2 \log \frac{1}{\varepsilon} \right)^{(2\beta-1)/4\beta};$$

$$r_{\beta, j} = v_\beta^2 \varepsilon^2 h_{\beta, j}^{-1}; \qquad \eta_j(\beta) = \left(r_{\beta, j} \alpha_j \frac{1}{\beta} \log \frac{1}{\varepsilon} \right)^{1/2}, \qquad j = 0, 1, \infty.$$

Introduce the estimator

$$f_{\varepsilon,\beta}(x) = \sum_{k=0}^{N} \frac{\hat{\theta}_k}{1 + a_k^2(\beta)(\pi h)^{2\beta}} \varphi_k(x), \tag{3.2}$$

where $h > 0$, $\beta > 1/2$, $N = N_\varepsilon \to \infty$, as $\varepsilon \to 0$, and $\hat{\theta}_k$ are the empirical Fourier coefficients

$$\hat{\theta}_k = \int_0^1 \varphi_k(t)\, dY(t) = \theta_k + \varepsilon \xi_k, \qquad k = 0, 1, \dots,$$

where ξ_k are i.i.d. standard normal random variables.

Assume everywhere in the following that N is the minimal even number satisfying

$$N \geq \varepsilon^{-2/\min(1,\beta_1 - 1/2)}. \tag{3.3}$$

If β is an integer, $\beta \geq 2$, the estimator (3.2) approximates, as $N \to \infty$, the usual spline estimator of degree β, with the smoothing parameter $\lambda_\beta = h^{2\beta}$. For example, if $\beta = 2$, one obtains the cubic spline.

Denote $f_{\varepsilon,\beta,j}$ the estimator (3.2) with $h = h_{\beta,j}$, respectively, ($j = 0, 1, \infty$). Clearly, $f_{\varepsilon,\beta,j}$ is a linear estimator w.r.t. $\{\hat{\theta}_k\}$.

Consider the adaptive version of the estimator (3.2) constructed following the scheme of Lepski. Namely, set

$$f^*_{\varepsilon,j}(x) = f_{\varepsilon,\hat{\beta}_j,j}(x), \qquad j = 0, \infty, \tag{3.4}$$

where

$$\hat{\beta}_j = \max\left\{ \beta \in B\colon \max_{\substack{\beta' \in B,\\ \beta' \leq \beta}} d_j\big(f_{\varepsilon,\beta,j}, f_{\varepsilon,\beta',j}\big)/\eta_j(\beta') \leq 1 \right\}, \qquad j = 0, \infty.$$

In words, we use the estimator (3.2) with $h = h_{\hat{\beta}_\infty,\infty}$, $\beta = \hat{\beta}_\infty$, for estimation in the sup-norm distance and with $h = h_{\hat{\beta}_0,0}$, $\beta = \hat{\beta}_0$, for estimation at a fixed point $x_0 \in (0, 1)$. The estimator (3.4) is nonlinear.

THEOREM 1. *Assume* (3.1). *Then the estimator* $f^*_{\varepsilon,\infty}$ *defined by* (3.4) *with* $j = \infty$ *is* AEA *on the scale of Sobolev classes* $\{W_\beta,\ \beta \in B\}$ *for the sup-norm distance* d_∞. *Moreover, the normalizing factor* $\psi_{\beta,\infty}$ *is such that*

$$\lim_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B} \sup_{f\in W_\beta} E_f\big(\psi_{\beta,\infty}^{-p} \|T_\varepsilon - f\|_\infty^p\big) = \lim_{\varepsilon\to 0} \sup_{\beta\in B} \sup_{f\in W_\beta} E_f\big(\psi_{\beta,\infty}^{-p} \|f^*_{\varepsilon,\infty} - f\|_\infty^p\big) = 1,$$

for any $p > 0$.

To prove Theorem 1 we show the upper bound on the risk

$$\limsup_{\varepsilon\to 0} \sup_{\beta\in B} \sup_{f\in W_\beta} E_f\big(\psi_{\beta,\infty}^{-p} \|f^*_{\varepsilon,\infty} - f\|_\infty^p\big) \leq 1 \tag{3.5}$$

(Section 6), and the corresponding lower bound

$$\liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B} \sup_{f\in W_\beta} E_f\big(\psi_{\beta,\infty}^{-p} \|T_\varepsilon - f\|_\infty^p\big) \geq 1 \tag{3.6}$$

(Section 7).

Note that the normalizing factor $\psi_{\beta,\infty}$ in Theorem 1 corresponds to Ψ_β in terms of Definition 1. It is interesting that $\psi_{\beta,\infty}$ depends on p (in fact, $\alpha_\infty = p+2$). The dependence is such that a loss of efficiency under adaptation does occur. To explain this, we study asymptotics of the loss of efficiency $\mathrm{LEF}(\varepsilon, p)$ as $\varepsilon \to 0$. Consider first the asymptotics of the minimax risk $\inf_{T_\varepsilon} Q_{\varepsilon,\beta}(T_\varepsilon)$.

THEOREM 2. *Assume* (3.1), *and let* $p > 0$. *Then for any* $0 < \delta < 1$ *there exists* $\tilde{\delta}_\varepsilon > 0$, *such that* $\lim_{\varepsilon\to 0} \tilde{\delta}_\varepsilon = 0$, *and*

$$\sup_{f\in W_\beta} E_f\left(\|f_{\varepsilon,\beta,1} - f\|_\infty^p\right) \le (1+\delta)^p\left(1+\tilde{\delta}_\varepsilon\right)\psi_{\beta,1}^p \qquad \forall\, \beta \in B, \tag{3.7}$$

$$\inf_{T_\varepsilon}\sup_{f\in W_\beta} E_f\left(\|T_\varepsilon - f\|_\infty^p\right) \ge (1-\delta)^p\left(1-\tilde{\delta}_\varepsilon\right)\left(\psi_{\beta,1}/2\right)^p, \qquad \forall\, \beta \in B. \tag{3.8}$$

Proofs of (3.7) and (3.8) are given in Sections 6 and 7, respectively. Theorem 2 implies that, for fixed β, the linear estimator $f_{\varepsilon,\beta,1}$ is "to within a factor 1/2" asymptotically efficient w.r.t. the risk $Q_{\varepsilon,\beta}(\cdot)$.

Using (3.8), one can bound $\mathrm{LEF}(\varepsilon, p)$ as follows:

$$\begin{aligned}\mathrm{LEF}(\varepsilon, p) &\le 2(1-\delta)^{-1}\left(1-\tilde{\delta}_\varepsilon\right)^{-1/p} \sup_{\beta\in B}\frac{Q_{\varepsilon,\beta}(f^*_{\varepsilon,\infty})}{\psi_{\beta,1}}\\ &\le 2(1-\delta)^{-1}\left(1-\tilde{\delta}_\varepsilon\right)^{-1/p}\\ &\quad\times \sup_{\beta\in B}\left(\mathscr{R}_{\varepsilon,\beta}\left(f^*_{\varepsilon,\infty}, \psi_{\beta,\infty}\right)\right)^{1/p} \sup_{\beta\in B}\frac{\psi_{\beta,\infty}}{\psi_{\beta,1}}\end{aligned} \tag{3.9}$$

if ε is small enough to have $\tilde{\delta}_\varepsilon < 1$. Now

$$\sup_{\beta\in B}\frac{\psi_{\beta,\infty}}{\psi_{\beta,1}} = \sup_{\beta\in B}\left(\frac{p+2}{2}\right)^{(2\beta-1)/4\beta} = \left(\frac{p+2}{2}\right)^{1/2-1/4\beta^*_\varepsilon}. \tag{3.10}$$

Applying (3.9), (3.10), Theorem 1 and the fact that $\delta > 0$ can be chosen arbitrarily small, one gets

$$\limsup_{\varepsilon\to 0}\mathrm{LEF}(\varepsilon, p) \le [2(p+2)]^{1/2}. \tag{3.11}$$

On the other hand, (3.7) entails

$$\begin{aligned}\mathrm{LEF}(\varepsilon, p) &\ge \inf_{f^*_\varepsilon}\sup_{\beta\in B}\frac{Q_{\varepsilon,\beta}(f^*_\varepsilon)}{Q_{\varepsilon,\beta}(f_{\varepsilon,\beta,1})}\\ &\ge (1+\delta)^{-1}\left(1+\tilde{\delta}_\varepsilon\right)^{-1/p}\\ &\quad\times \inf_{f^*_\varepsilon}\sup_{\beta\in B}\left(\mathscr{R}_{\varepsilon,\beta}\left(f^*_\varepsilon, \psi_{\beta,\infty}\right)\right)^{1/p}\left(\frac{\psi_{\beta,\infty}}{\psi_{\beta,1}}\right).\end{aligned} \tag{3.12}$$

Clearly,

$$\frac{\psi_{\beta,\infty}}{\psi_{\beta,1}} = \left(\frac{p+2}{2}\right)^{(2\beta-1)/4\beta} \geq \left(\frac{p+2}{2}\right)^{1/2-1/4\beta_1}, \qquad \beta \in B.$$

Using this, (3.12), (3.6) and the fact that $\delta > 0$ can be chosen arbitrarily small, we find

$$\liminf_{\varepsilon\to 0} \mathrm{LEF}(\varepsilon, p) \geq \left(\frac{p+2}{2}\right)^{1/2-1/4\beta_1}. \tag{3.13}$$

THEOREM 3. *Assume* (3.1) *and let* $p > 0$. *Then*

$$\left(\frac{p+2}{2}\right)^{1/2-1/4\beta_1} + o(1) \leq LEF(\varepsilon, p) \leq [2(p+2)]^{1/2} + o(1), \tag{3.14}$$

as $\varepsilon \to 0$. *If, in addition,*

$$\max\{\beta \in B\colon \beta \leq t\beta_\varepsilon^*\} \to \infty \qquad \forall\, 0 < t \leq 1, \tag{3.15}$$

then

$$[(p+2)/2]^{1/2} + o(1) \leq LEF(\varepsilon, p) \leq [2(p+2)]^{1/2} + o(1), \tag{3.16}$$

as $\varepsilon \to 0$.

For the proof of Theorem 3 note that (3.14) is an immediate consequence of (3.11) and (3.13). The left-hand inequality in (3.16) follows from (3.12) if we show that

$$\liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B} \left[\mathscr{R}_{\varepsilon,\beta}(T_\varepsilon, \psi_{\beta,\infty})(\psi_{\beta,\infty}/\psi_{\beta,1})^p\right] \geq [(p+2)/2]^{p/2}. \tag{3.17}$$

Proof of (3.17) under assumption (3.15) is given in Section 7. Note that (3.15) is not a restrictive assumption. It claims that the sets B were not too sparse. For example, (3.15) rules out the case $B = \{\beta_1, \beta_\varepsilon^*\}$.

With the squared risk ($p = 2$), (3.16) yields

$$\sqrt{2} + o(1) \leq \mathrm{LEF}(\varepsilon, 2) \leq 2\sqrt{2} + o(1).$$

Remark that, for any $p > 0$,

$$\liminf_{\varepsilon\to 0} \mathrm{LEF}(\varepsilon, p) > 1.$$

In other words, the loss of efficiency under adaptation in sup-norm does occur. The right-hand inequality in (3.14) and (3.16) gives an upper bound on this loss. The question about the exact asymptotics of $\mathrm{LEF}(\varepsilon, p)$ remains open. It would be answered if one knew the exact asymptotics of the minimax risk $\inf_{T_\varepsilon} Q_{\varepsilon,\beta}(T_\varepsilon)$.

4. Exact pointwise adaptation. In this section we show that the ARC for the pointwise distance on the Sobolev classes is worse than the ORC by a logarithmic factor, and we find the AEA estimator in this setup.

Denote $B_- = \{\beta \in B\colon \beta < \beta_\varepsilon^*\}$.

THEOREM 4. *Assume* (3.1) *and let* $p > 0$. *Then optimal rate adaptive estimators on the scale of Sobolev classes* $\{W_\beta,\ \beta \in B\}$ *for the pointwise distance* d_0 *do not exist. An ARC on this scale of classes for the distance* d_0 *is*

$$\Psi_\beta(\varepsilon) = \begin{cases} \psi_{\beta,0}, & \text{if } \beta \in B_-, \\ \varepsilon^{(2\beta_\varepsilon^* - 1)/2\beta_\varepsilon^*}, & \text{if } \beta = \beta_\varepsilon^*. \end{cases}$$

Theorem 4 shows that the value $\beta = \beta_\varepsilon^*$ is an outlier in terms of the ARC. The ARC for $\beta = \beta_\varepsilon^*$ is faster than for $\beta \in B_-$: it does not contain a logarithmic factor and equals ORC. This is a kind of boundary effect. To define the pointwise AEA estimator, we exclude from consideration the boundary value $\beta = \beta_\varepsilon^*$ (i.e., consider the set B_- in the place of B). The next theorem shows that the estimator $f_{\varepsilon,0}^*$ has the AEA property on this smaller set.

THEOREM 5. *Assume* (3.1) *and* (3.15). *Then the estimator* $f_{\varepsilon,0}^*$ *defined by* (3.4) *with* $j = 0$ *is rate adaptive on the scale of Sobolev classes* $\{W_\beta,\ \beta \in B\}$ *for the distance* d_0 *at a fixed point* $x_0 \in (0,1)$. *Moreover, the normalizing factor* $\psi_{\beta,0}$ *is such that*

$$\begin{aligned} &\lim_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B_-} \sup_{f\in W_\beta} E_f\Big(\psi_{\beta,0}^{-p} |T_\varepsilon(x_0) - f(x_0)|^p\Big) \\ &\quad = \lim_{\varepsilon\to 0} \sup_{\beta\in B_-} \sup_{f\in W_\beta} E_f\Big(\psi_{\beta,0}^{-p} |f_{\varepsilon,0}^*(x_0) - f(x_0)|^p\Big) = 1 \end{aligned}$$

for any $p > 0$.

Note that, similarly to the sup-norm case, the exact normalizing factor $\psi_{\beta,0}$ depends on the power p of the loss function.

Proofs of Theorems 4 and 5 are organized as follows. In Section 6 we prove that under the assumption (3.1) the following upper bounds are valid:

$$\limsup_{\varepsilon\to 0} \sup_{\beta\in B} \sup_{f\in W_\beta} E_f\Big(\psi_{\beta,0}^{-p} |f_{\beta,0}^*(x_0) - f(x_0)|^p\Big) \le 1 \tag{4.1}$$

and

$$\limsup_{\varepsilon\to 0} \sup_{f\in W_{\beta_\varepsilon^*}} E_f\Big(\varepsilon^{-p(2\beta_\varepsilon^* - 1)/2\beta_\varepsilon^*} |f_{\varepsilon,0}^*(x_0) - f(x_0)|^p\Big) < \infty. \tag{4.2}$$

Section 7 contains the proof of the following lower bound (under the assumptions of Theorem 5):

$$\liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{\beta\in B_-} \sup_{f\in W_\beta} E_f\Big(\psi_{\beta,0}^{-p} |T_\varepsilon(x_0) - f(x_0)|^p\Big) \ge 1. \tag{4.3}$$

The relations (4.1)–(4.3) and Theorem 4 entail Theorem 5. Proof of Theorem 4 is given in Section 7.

5. Lemmas. In the following, c_l, $l = 1, 2, \dots,$ are positive constants, that depend only on p, L and β_1. Introduce the notation

$$h^*_\varepsilon = \max_j h_{\beta^*_\varepsilon, j}, \qquad h_{\min,\varepsilon} = \min_j h_{\beta_1, j},$$

$$b_j(\beta) = Lb_\beta h_{\beta,j}^{\beta-1/2}, \qquad j = 0, 1, \infty,$$

$$Z_{\beta,j}(t) = \sum_{k=0}^{N} \frac{\varepsilon \xi_k}{1 + a_k^2(\beta)(\pi h_{\beta,j})^{2\beta}} \varphi_k(t),$$

$$\tilde{Z}_{\beta,j} = \sup_{t \in [0,1]} |Z_{\beta,j}(t)|, \; j = 1, \infty, \qquad \tilde{Z}_{\beta,0} = |Z_{\beta,0}(x_0)|,$$

$$\tilde{\beta}(\beta, \beta') = \begin{cases} \beta, & \text{if } \beta' > \beta/2, \\ \beta' + (\beta_1 + 1/2)/2, & \text{if } \beta' \le \beta/2, \end{cases} \qquad \forall \beta, \beta' \in B, \; \beta' \le \beta.$$

Note that $\beta' \le \tilde{\beta}(\beta, \beta') \le \beta$, and $\tilde{\beta}(\beta, \beta') < 2\beta'$.

In the following, it is supposed w.l.o.g. that $0 < \Delta_\varepsilon < 1$ and $\varepsilon > 0$ is so small that

$$\log(1/\varepsilon) \ge \beta^*_\varepsilon \ge e, \tag{5.1}$$

and

$$h_{\min,\varepsilon} N_\varepsilon \ge 1. \tag{5.2}$$

Denote

$$\tilde{\nu}_\varepsilon = \frac{\Delta_\varepsilon \log(1/\varepsilon)}{(\beta^*_\varepsilon)^2}.$$

The definition of β^*_ε and Δ_ε yields that $\beta^*_\varepsilon/\Delta_\varepsilon \ge 1$. This and (3.1), (5.1) imply

$$\frac{1}{\beta^*_\varepsilon} \log \frac{1}{\varepsilon} \ge \tilde{\nu}_\varepsilon \ge \nu_\varepsilon \log \beta^*_\varepsilon \ge \nu_\varepsilon \to \infty \quad \text{as } \varepsilon \to 0.$$

LEMMA 1. *There exist positive constants $\kappa_{\min}$, $\kappa_{\max}$, $Q_{\max}$, $b_{\max}$, $v_{\max}$, and $v_{\min}$ depending only on L and β_1, such that*

$$\max_{\beta \in B} \kappa_{\beta,j}^{\beta-1/2} \le \kappa_{\max}, \qquad \max_{\beta \in B} \kappa_{\beta,j} \le \kappa_{\max},$$

$$\min_{\beta \in B} \kappa_{\beta,j} \ge \kappa_{\min}, \; j = 0, 1, \infty,$$

$$\max_{\beta \in B} Q_\beta \le Q_{\max}, \qquad \max_{\beta \in B} b_\beta \le b_{\max},$$

$$\max_{\beta \in B} v_\beta \le v_{\max}, \qquad \min_{\beta \in B} v_\beta \ge v_{\min}.$$

Moreover,

$$b_\beta = (2\beta - 1)^{-1/2} v_\beta. \tag{5.3}$$

Proof of this lemma is an easy consequence of the properties of the beta-function.

Without loss of generality, suppose everywhere in the following that $\varepsilon > 0$ is small enough to satisfy

$$\kappa_{\max}(1 + 1/\kappa_{\min})\exp(-\nu_{\varepsilon}/2) \le 1, \qquad \tilde{\nu}_{\varepsilon} \ge \frac{1}{\beta_1}\log\log\frac{1}{\varepsilon} \tag{5.4}$$

and

$$\delta_{\varepsilon 0} = \frac{2\max\{\kappa_{\max}, 1\}}{v_{\min}^2}\exp\left(-\frac{1}{2\beta_{\varepsilon}^*}\log\frac{1}{\varepsilon}\right) < 1/2. \tag{5.5}$$

LEMMA 2. *As* $\varepsilon \to 0$,

$$\left(\frac{1}{\beta_{\varepsilon}^*}\log\frac{1}{\varepsilon}\right)^{1/2}\frac{1}{\log\ \beta_{\varepsilon}^*} \to \infty, \qquad \left(\frac{1}{\beta_{\varepsilon}^*}\log\frac{1}{\varepsilon}\right)^{1/2}\frac{1}{\log(1/\Delta_{\varepsilon})} \to \infty \tag{5.6}$$

and

$$\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{\Delta_{\varepsilon}/(\beta_{\varepsilon}^*)^2} \le \exp(-\tilde{\nu}_{\varepsilon}). \tag{5.7}$$

The proof follows easily from (3.1) and (5.1).

LEMMA 3. *For* $j = 0, 1, \infty$ *and any* $\beta \in B$, *except for the combination* $j = 0$, $\beta = \beta_{\varepsilon}^*$, *we have*

$$\psi_{\beta, j} = b_j(\beta) + \eta_j(\beta) = L b_{\beta} h_{\beta, j}^{\beta - 1/2} + \left(r_{\beta, j}\alpha_j \frac{1}{\beta}\log\frac{1}{\varepsilon}\right)^{1/2}, \tag{5.8}$$

$$\eta_j(\beta) = b_j(\beta)(2\beta - 1) = L(2\beta - 1)^{1/2} v_{\beta} h_{\beta, j}^{\beta - 1/2}. \tag{5.9}$$

The proof is given by direct calculations using the definitions from Section 3 and (5.3).

LEMMA 4. *Let* $\beta' < \beta$ *and* $\beta', \beta \in B$. *Then*, *for* $\varepsilon > 0$ *small enough*,

$$h_{\beta', j}/h_{\beta, j} \le c_1 \exp(-\tilde{\nu}_{\varepsilon}/2) \le 1, \qquad j = 0, 1, \infty, \tag{5.10}$$

and

$$\begin{aligned} h_{\varepsilon}^* &\le \max\{\kappa_{\max}, 1\} \exp\left(-\frac{1}{2\beta_{\varepsilon}^*}\log\frac{1}{\varepsilon}\right) \\ &\le \max\{\kappa_{\max}, 1\} \exp\left(-\frac{\tilde{\nu}_{\varepsilon}}{2}\right) \le 1. \end{aligned} \tag{5.11}$$

In particular,

$$h_{\beta', j} \le h_{\beta, j} \le h_{\varepsilon}^* \le 1, \qquad j = 0, 1, \infty. \tag{5.12}$$

The proof is straightforward, in view of Lemma 1, (5.4) and (5.7).

The following five lemmas are proved in the Appendix. Lemma 5 yields an evaluation of the bias terms. Lemmas 6 and 7 provide bounds for the stochastic terms of estimation error in sup-norm and at a fixed point. Lemma 8 shows that it is very improbable that the "estimated smoothness" $\hat{\beta}_j$ is strictly smaller than the "true smoothness" β. Lemma 9 gives uniform over β bounds on the risk of the linear estimator $f_{\varepsilon,\beta,\infty}$.

LEMMA 5. *Let β', $\beta \in B$. If $\beta' < \beta$, then*

$$\sup_{f \in W_\beta} \left\| E_f(f_{\varepsilon,\beta',j}) - f \right\|_\infty \le c_2 h_{\beta',j}^{\tilde{\beta}-1/2}, \tag{5.13}$$

where $\tilde{\beta} = \tilde{\beta}(\beta, \beta')$, and if $\beta' = \beta$ then

$$\sup_{f \in W_\beta} \left\| E_f(f_{\varepsilon,\beta,j}) - f \right\|_\infty \le b_j(\beta)\left(1 + c_3 \exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right), \tag{5.14}$$

$$j = 0, 1, \infty.$$

LEMMA 6. *There exists $c_4 > 0$ such that*

$$P\left(\tilde{Z}_{\beta,j} \ge u\right) \le \frac{c_4}{h_{\beta,j}} \exp\left(-\frac{u^2}{2r_{\beta,j}}(1 - \delta_{\varepsilon 0})\right),$$

$$E\left(\tilde{Z}_{\beta,j}^p I\left\{\tilde{Z}_{\beta,j} \ge u\right\}\right) \le \frac{c_4}{h_{\beta,j}} r_{\beta,j}^{p/2} \exp\left(-\frac{u^2}{2r_{\beta,j}}(1 - \delta_{\varepsilon 0})\right),$$

for $j = 1, \infty$, and any $u > 0$, $p > 0$, $\beta \in B$, where $I\{\cdot\}$ denotes the indicator function.

LEMMA 7. *There exists $c_5 > 0$ such that for any $x_0 \in (0,1)$,*

$$P\left(\tilde{Z}_{\beta,0} \ge u\right) \le c_5 \exp\left(-\frac{u^2}{2r_{\beta,0}}(1 - \delta_{\varepsilon 0})\right),$$

$$E\left(\tilde{Z}_{\beta,0}^p I\left\{\tilde{Z}_{\beta,0} \ge u\right\}\right) \le c_5 r_{\beta,0}^{p/2} \exp\left(-\frac{u^2}{2r_{\beta,0}}(1 - \delta_{\varepsilon 0})\right)$$

$\forall\ u > 0$, $p > 0$, $\beta \in B$.

LEMMA 8. *Let $\gamma, \beta \in B$, $\gamma < \beta$. There exists $c_6 > 0$ such that*

$$\sup_{f \in W_\beta} P_f\left(\hat{\beta}_j = \gamma\right) \le c_6(\beta_\varepsilon^*/\Delta_\varepsilon)\varepsilon^{p/2\gamma}, \qquad j = 0, \infty.$$

LEMMA 9. *For any $\delta > 0$ there exists $\varepsilon_0 > 0$ that depends only on δ, such that*

$$\sup_{f \in W_\beta} P_f\left\{\psi_{\beta,\infty}^{-1}\|f_{\varepsilon,\beta,\infty} - f\|_\infty \ge 1 + \delta\right\} \le c_7 \exp\left(-\frac{p}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right),$$

$$0 < \varepsilon < \varepsilon_0,\ \beta \in B.$$

Also

$$\sup_{\varepsilon<\varepsilon_0}\sup_{\beta\in B}\sup_{f\in W_\beta} E_f\left(\psi_{\beta,\infty}^{-p}\|f_{\varepsilon,\beta,\infty}-f\|_\infty^p\right)\le c_8(p) \qquad \forall\, p>0.$$

6. Proofs of the upper bounds. This section is devoted to the proof of (3.5), (3.7), (4.1) and (4.2). It is assumed throughout that $\varepsilon>0$ is small enough to satisfy the conditions (5.1), (5.2), (5.4), (5.5), (5.10). It is also assumed w.l.o.g. that $0<\Delta_\varepsilon<1$, so that (5.6) makes sense.

In the following k_l, $l=1,2,\dots$ are positive constants that depend only on p, L and β_1. Denote

$$R^*_{\varepsilon,\beta,j}=\sup_{f\in W_\beta} E_f\left(\psi_{\beta,j}^{-p}d_j^p(f^*_{\varepsilon,j},f)\right), \qquad j=0,\infty.$$

Clearly,

$$R^*_{\varepsilon,\beta,j}\le R^-_{\varepsilon,\beta,j}+R^+_{\varepsilon,\beta,j},$$

where

$$R^-_{\varepsilon,\beta,j}=\sup_{f\in W_\beta}E_f\left(\psi_{\beta,j}^{-p}d_j^p(f^*_{\varepsilon,j},f)I\left\{\hat\beta_j<\beta\right\}\right),$$

$$R^+_{\varepsilon,\beta,j}=\sup_{f\in W_\beta}E_f\left(\psi_{\beta,j}^{-p}d_j^p(f^*_{\varepsilon,j},f)I\left\{\hat\beta_j\ge\beta\right\}\right), \qquad j=0,\infty.$$

Hence, to prove (3.5) and (4.1) it suffices to show

$$\limsup_{\varepsilon\to 0}\sup_{\beta\in B}R^-_{\varepsilon,\beta,j}=0, \qquad j=0,\infty \tag{6.1}$$

and

$$\limsup_{\varepsilon\to 0}\sup_{\beta\in B}R^+_{\varepsilon,\beta,j}\le 1, \qquad j=0,\infty. \tag{6.2}$$

PROOF OF (6.1). For any $\gamma\in B$, such that $\gamma\le\beta$, denote

$$\tau_j(\gamma)=r_{\gamma,j}^{1/2}\left(\left(\left[p\left(\frac{1}{\gamma}-\frac{1}{\beta}\right)+(\alpha_j-p)\frac{1}{\gamma}\right]\log\frac{1}{\varepsilon}\right)^{1/2}+\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{1/4}\right),$$
$$j=0,\infty.$$

Using (5.13), the fact that $h_{\gamma,j}\le 1$ (see Lemma 4) and the inequality $\gamma\le\tilde\beta(\beta,\gamma)$, we get for any $\gamma\in B$, such that $\gamma<\beta$,

$$\begin{aligned} d_j(f_{\varepsilon,\gamma,j},f)&\le\left\|E_f(f_{\varepsilon,\gamma,j})-f\right\|_\infty+\tilde Z_{\gamma,j}\\ &\le c_2h_{\gamma,j}^{\tilde\beta-1/2}+\tilde Z_{\gamma,j}\le c_2h_{\gamma,j}^{\gamma-1/2}+\tilde Z_{\gamma,j},\end{aligned}$$

where $\tilde\beta=\tilde\beta(\beta,\gamma)$. This and the definition of $R^-_{\varepsilon,\beta,j}$ entail

$$R^-_{\varepsilon,\beta,j}\le\sum_{\substack{\gamma\in B,\\ \gamma<\beta}}\sup_{f\in W_\beta}E_f\left(\psi_{\beta,j}^{-p}d_j^p(f_{\varepsilon,\gamma,j},f)I\left\{\hat\beta_j=\gamma\right\}\right)\le\rho_{1,j}+\rho_{2,j}, \tag{6.3}$$
$$j=0,\infty,$$

where

$$\rho_{1,j} = \rho_{1,j}(\beta) = \sum_{\substack{\gamma\in B,\\ \gamma<\beta}} \sup_{f\in W_\beta} P_f\big(\hat{\beta}_j = \gamma\big)\psi_{\beta,j}^{-p}\big(c_2 h_{\gamma,j}^{\gamma-1/2} + \tau_j(\gamma)\big)^p,$$

$$\rho_{2,j} = \rho_{2,j}(\beta) = \sum_{\substack{\gamma\in B,\\ \gamma<\beta}} \psi_{\beta,j}^{-p} E\Big(\big(c_2 h_{\gamma,j}^{\gamma-1/2} + \tilde{Z}_{\gamma,j}\big)^p I\big\{\tilde{Z}_{\gamma,j} \ge \tau_j(\gamma)\big\}\Big).$$

To prove (6.1) it suffices to show the relations

$$\sup_{\beta\in B} \rho_{l,j} = o(1) \quad \text{as } \varepsilon \to 0,\ l = 1,2,\ j = 0,\infty. \tag{6.4}$$

PROOF OF (6.4) FOR $l = 1$. Using the inequalities $\alpha_j \ge p$, $j = 0,\infty$, $\gamma < \beta_\varepsilon^*$, we get, for $\gamma < \beta$,

$$\tau_j(\gamma)/\eta_j(\gamma) \le 2 + p^{-1/2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{-1/4} \le 2 + p^{-1/2}. \tag{6.5}$$

It follows from (5.9) that

$$h_{\gamma,j}^{\gamma-1/2}/\eta_j(\gamma) \le \left[L(2\beta_1 - 1)^{1/2} v_{\min}\right]^{-1}. \tag{6.6}$$

Next, (5.8) and Lemma 1 imply

$$\begin{aligned}\frac{\eta_j(\gamma)}{\psi_{\beta,j}} &\le \frac{\eta_j(\gamma)}{\eta_j(\beta)} = \left(\frac{\beta}{\gamma}\,\frac{v_\gamma^2 h_{\beta,j}}{v_\beta^2 h_{\gamma,j}}\right)^{1/2} \le \left(\frac{\beta_\varepsilon^*}{\beta_1}\right)\frac{v_{\max}}{v_{\min}}\left(\frac{h_{\beta,j}}{h_{\gamma,j}}\right)^{1/2}\\ &\le k_1\,\beta_\varepsilon^*\left(\varepsilon^2\log\frac{1}{\varepsilon}\right)^{1/4\beta-1/4\gamma} \le k_1\,\beta_\varepsilon^*\varepsilon^{1/2\beta-1/2\gamma},\end{aligned} \tag{6.7}$$

for any $\gamma, \beta \in B$, such that $\gamma < \beta$, except for the combination $j = 0$, $\beta = \beta_\varepsilon^*$. It is easy to see that for $j = 0$, $\beta = \beta_\varepsilon^*$ the result of (6.7) remains valid.

Now, Lemma 8, (6.5)–(6.7) and (5.6) yield

$$\begin{aligned}\rho_{1,j} &\le k_2\,\mathrm{card}(B)(\beta_\varepsilon^*)^{p+1}\Delta_\varepsilon^{-1}\varepsilon^{p/2\beta} \le 2k_2(\beta_\varepsilon^*)^{p+2}\Delta_\varepsilon^{-2}\varepsilon^{p/2\beta}\\ &\le 2k_2\exp\left((p+2)\log\beta_\varepsilon^* + 2\log\left(\frac{1}{\Delta_\varepsilon}\right) - p\frac{1}{2\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\\ &= o(1) \quad \text{as } \varepsilon \to 0,\end{aligned} \tag{6.8}$$

for $j = 0, \infty$. This completes the proof of (6.4) for $l = 1$. □

PROOF OF (6.4) FOR $l = 2$ AND $j = \infty$. It follows from Lemma 6 and (6.6) that

$$\begin{aligned}\rho_{2,\infty} &\le k_3 \sum_{\substack{\gamma\in B,\\ \gamma<\beta}} \psi_{\beta,\infty}^{-p}\Big(\eta_\infty^p(\gamma) P\big(\tilde{Z}_{\gamma,\infty} \ge \tau_\infty(\gamma)\big) + E\Big(\tilde{Z}_{\gamma,\infty}^p I\big\{\tilde{Z}_{\gamma,\infty} \ge \tau_\infty(\gamma)\big\}\Big)\Big)\\ &\le k_3 c_5 \sum_{\substack{\gamma\in B,\\ \gamma<\beta}} \psi_{\beta,\infty}^{-p}\big(\eta_\infty^p(\gamma) + r_{\gamma,\infty}^{p/2}\big)h_{\gamma,\infty}^{-1}\exp\left(-\frac{1}{2r_{\gamma,\infty}}\tau_\infty^2(\gamma)(1-\delta_{\varepsilon 0})\right).\end{aligned} \tag{6.9}$$

Now,

$$r_{\gamma,j}^{1/2} \le \left(\alpha_j \frac{1}{\beta_\varepsilon^*} \log\frac{1}{\varepsilon}\right)^{-1/2} \eta_j(\gamma) \le \alpha_j^{-1/2}\eta_j(\gamma), \qquad j = 0, \infty, \tag{6.10}$$

and $h_{\gamma,\infty}^{-1} \le \kappa_{\min}^{-1}\varepsilon^{-1/\gamma}$, since, by (5.1), ε is supposed to satisfy $\log(1/\varepsilon) \ge \beta_\varepsilon^* \ge e$. This and (6.7) imply

$$\psi_{\beta,\infty}^{-p}\big(\eta_\infty^p(\gamma) + r_{\gamma,\infty}^{p/2}\big)h_{\gamma,\infty}^{-1} \le k_4(\beta_\varepsilon^*)^p \varepsilon^{p/2\beta-(p+2)/2\gamma}. \tag{6.11}$$

On the other hand,

$$\begin{aligned} &\frac{1}{2r_{\gamma,\infty}}\tau_\infty^2(\gamma)(1-\delta_{\varepsilon 0}) \\ &\quad\ge \left(\left[\frac{p+2}{2\gamma} - \frac{p}{2\beta}\right]\log\frac{1}{\varepsilon} + \frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2}\right)(1-\delta_{\varepsilon 0}) \\ &\quad\ge \left[\frac{p+2}{2\gamma} - \frac{p}{2\beta}\right]\log\frac{1}{\varepsilon} + \frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2} - k_5\delta_{\varepsilon 0}\log\frac{1}{\varepsilon}. \end{aligned} \tag{6.12}$$

Note that $\delta_{\varepsilon 0}\log(1/\varepsilon) = o(1)$, as $\varepsilon \to 0$, in view of (3.1) and the definition of $\delta_{\varepsilon 0}$. Using this and substituting (6.11) and (6.12) into (6.9), one obtains

$$\begin{aligned} \rho_{2,\infty} &\le k_6 \operatorname{card}(B)(\beta_\varepsilon^*)^p \exp\left(-\frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2}\right) \\ &\le 2k_6(\beta_\varepsilon^*)^{p+1}\Delta_\varepsilon^{-1}\exp\left(-\frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2}\right) = o(1) \quad \text{as } \varepsilon \to 0, \end{aligned}$$

where (5.6) was applied. This completes the proof of (6.4) for $l = 2$ and $j = \infty$.

Proof of (6.4) for $l = 2$ and $j = 0$. Using Lemma 7 and (6.6), we obtain, as in (6.9),

$$\rho_{2,0} \le k_3 c_5 \sum_{\substack{\gamma\in B,\\ \gamma<\beta}} \psi_{\beta,0}^{-p}\big(\eta_0^p(\gamma) + r_{\gamma,0}^{p/2}\big)\exp\left(-\frac{1}{2r_{\gamma,0}}\tau_0^2(\gamma)(1-\delta_{\varepsilon 0})\right). \tag{6.13}$$

Now, (6.7) and (6.10) imply

$$\psi_{\beta,0}^{-p}\big(\eta_0^p(\gamma) + r_{\gamma,0}^{p/2}\big) \le k_7(\beta_\varepsilon^*)^p\varepsilon^{p/2\beta-p/2\gamma}, \tag{6.14}$$

and, as in (6.12), one gets

$$\begin{aligned} \frac{1}{2r_{\gamma,0}}\tau_0^2(\gamma)(1-\delta_{\varepsilon 0}) &\ge \left[\frac{p}{2\gamma} - \frac{p}{2\beta}\right]\log\frac{1}{\varepsilon} + \frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2} \\ &\qquad - k_5\delta_{\varepsilon 0}\log\frac{1}{\varepsilon}. \end{aligned} \tag{6.15}$$

Combining (6.13)–(6.15), observing that $\delta_{\varepsilon 0}\log(1/\varepsilon) = o(1)$ as $\varepsilon \to 0$ and using (5.6), one comes to

$$\rho_{2,0} \le k_8(\beta_\varepsilon^*)^{p+1}\Delta_\varepsilon^{-1}\exp\left(-\frac{1}{2}\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{1/2}\right) = o(1) \quad \text{as } \varepsilon \to 0. \quad \square \tag{6.16}$$

PROOF OF (6.2) FOR $j = \infty$. Fix a number $\delta > 0$. Consider the random event

$$\mathfrak{A}_f(\beta,\gamma) = \left\{\psi_{\beta,\infty}^{-1}\|f_{\varepsilon,\gamma,\infty} - f\|_\infty \ge 1 + \delta\right\}.$$

Then

$$\begin{aligned}
R^+_{\varepsilon,\beta,\infty} &\le \sup_{f\in W_\beta}\sum_{\substack{\gamma\in B,\\ \gamma\ge\beta}} E_f\left(\psi_{\beta,\infty}^{-p}\|f_{\varepsilon,\gamma,\infty} - f\|_\infty^p I\left\{\hat\beta_\infty = \gamma\right\}\right)\\
&\le (1+\delta)^p \sup_{f\in W_\beta} P_f\left\{\hat\beta_\infty \ge \beta\right\}\\
&\qquad + \sum_{\substack{\gamma\in B,\\ \gamma\ge\beta}}\sup_{f\in W_\beta} E_f\left(\psi_{\beta,\infty}^{-p}\|f_{\varepsilon,\gamma,\infty} - f\|_\infty^p I\left\{\mathfrak{A}_f(\beta,\gamma)\cap\left\{\hat\beta_\infty = \gamma\right\}\right\}\right) \qquad (6.17)\\
&\le (1+\delta)^p\\
&\qquad + \sum_{\substack{\gamma\in B,\\ \gamma\ge\beta}}\left(\sup_{f\in W_\beta}\left(E_f\left(\psi_{\beta,\infty}^{-2p}\|f_{\varepsilon,\gamma,\infty} - f\|_\infty^{2p}\right)\right)^{1/2}\sup_{f\in W_\beta}\rho_f^{1/2}(\beta,\gamma)\right),
\end{aligned}$$

where $\rho_f(\beta,\gamma) = P_f\{\mathfrak{A}_f(\beta,\gamma)\cap\{\hat\beta_\infty = \gamma\}\}$.

Note that if $\hat\beta_\infty = \gamma \ge \beta$, then, by definition of $\hat\beta_\infty$,

$$\|f_{\varepsilon,\gamma,\infty} - f_{\varepsilon,\beta,\infty}\| \le \eta_\infty(\beta). \tag{6.18}$$

Thus

$$\|f_{\varepsilon,\gamma,\infty} - f\|_\infty \le \eta_\infty(\beta) + \|f_{\varepsilon,\beta,\infty} - f\|_\infty, \tag{6.19}$$

and, in view of Lemma 9, for $0 < \varepsilon < \varepsilon_0$,

$$\begin{aligned}
E_f\left(\psi_{\beta,\infty}^{-2p}\|f_{\varepsilon,\gamma,\infty} - f\|_\infty^{2p}\right) &\le 2^{2p-1}\left(\left(\eta_\infty(\beta)/\psi_{\beta,\infty}\right)^{2p} + c_8(2p)\right)\\
&\le 2^{2p-1}(1 + c_8(2p)).
\end{aligned}$$

Substitution of this inequality into (6.17) gives

$$R^+_{\varepsilon,\beta,\infty} \le (1+\delta)^p + k_9\sum_{\substack{\gamma\in B,\\ \gamma\ge\beta}}\sup_{f\in W_\beta}\rho_f^{1/2}(\beta,\gamma). \tag{6.20}$$

It remains to estimate $\rho_f(\beta,\gamma)$. By Lemma 9,

$$\sup_{f\in W_\beta}\rho_f(\beta,\beta) \le \sup_{f\in W_\beta}P_f\left\{\mathfrak{A}_f(\beta,\beta)\right\} \le c_7\exp\left(-\frac{p}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right), \qquad 0 < \varepsilon < \varepsilon_0,\ \beta \in B. \tag{6.21}$$

Consider now $\rho_f(\beta,\gamma)$, $\gamma>\beta$. Since $E_f(f_{\varepsilon,\gamma,\infty}(t))$ is a continuous function of t, there exists a (nonrandom) point $t^*\in[0,1]$, such that

$$\left|E_f\left(f_{\varepsilon,\gamma,\infty}(t^*)\right)-E_f\left(f_{\varepsilon,\beta,\infty}(t^*)\right)\right|=\left\|E_f\left(f_{\varepsilon,\gamma,\infty}\right)-E_f\left(f_{\varepsilon,\beta,\infty}\right)\right\|_\infty \tag{6.22}$$

(clearly, t^* depends on $f,\varepsilon,\gamma,\beta$).

If $\hat{\beta}_\infty=\gamma>\beta$, then, in view of (6.18) and (6.22),

$$\left\|E_f\left(f_{\varepsilon,\gamma,\infty}\right)-E_f\left(f_{\varepsilon,\beta,\infty}\right)\right\|_\infty\le\eta_\infty(\beta)+|\xi^*|,$$

where $\xi^*=Z_{\beta,\infty}(t^*)-Z_{\gamma,\infty}(t^*)$. Therefore, if $\hat{\beta}_\infty=\gamma>\beta$,

$$\begin{aligned}\|f_{\varepsilon,\gamma,\infty}-f\|_\infty&\le\left\|E_f\left(f_{\varepsilon,\gamma,\infty}\right)-E_f\left(f_{\varepsilon,\beta,\infty}\right)\right\|_\infty+\tilde{Z}_{\gamma,\infty}+\left\|E_f\left(f_{\varepsilon,\beta,\infty}\right)-f\right\|_\infty\\&\le\eta_\infty(\beta)+|\xi^*|+\tilde{Z}_{\gamma,\infty}+b_\infty(\beta)\left(1+c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right)\\&\le\psi_{\beta,\infty}\left(1+c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right)+|\xi^*|+\tilde{Z}_{\gamma,\infty},\end{aligned}$$

where (5.14) of Lemma 5 and (5.8) of Lemma 3 were used. This entails that, if $\gamma>\beta$,

$$\begin{aligned}\rho_f(\beta,\gamma)&\le P_f\left\{\psi_{\beta,\infty}^{-1}\left(|\xi^*|+\tilde{Z}_{\gamma,\infty}\right)\ge\delta-c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right\}\\&\le P\left(\tilde{Z}_{\gamma,\infty}\ge\frac{\delta\psi_{\beta,\infty}}{3}\right)+P_f\left(|\xi^*|\ge\frac{\delta\psi_{\beta,\infty}}{3}\right),\end{aligned}\tag{6.23}$$

if ε is so small that $c_3\exp(-(1/4\beta_\varepsilon^*)\log(1/\varepsilon))<\delta/3$. By Lemma 6,

$$P\left(\tilde{Z}_{\gamma,\infty}\ge\delta\psi_{\beta,\infty}/3\right)\le c_4h_{\beta,\infty}^{-1}\exp\left(-(\delta^2/18)\psi_{\beta,\infty}^2r_{\gamma,\infty}^{-1}(1-\delta_{\varepsilon0})\right).\tag{6.24}$$

By (5.8), (5.10) and the fact that $\delta_{\varepsilon0}<1/2$,

$$\begin{aligned}\psi_{\beta,\infty}^2r_{\gamma,\infty}^{-1}(1-\delta_{\varepsilon0})&\ge\frac{\eta_\infty^2(\beta)r_{\gamma,\infty}^{-1}}{2}=\frac{r_{\beta,\infty}}{r_{\gamma,\infty}}\frac{p+2}{2}\frac{1}{\beta}\log\frac{1}{\varepsilon}\\&\ge\left(\frac{v_{\min}}{v_{\max}}\right)^2\left[\frac{p+2}{2}\right]\left(\frac{h_{\gamma,\infty}}{h_{\beta,\infty}}\right)\frac{1}{\beta}\log\frac{1}{\varepsilon}\ge k_{10}\exp\left(\frac{\tilde{\nu}_\varepsilon}{2}\right)\frac{1}{\beta}\log\frac{1}{\varepsilon}.\end{aligned}$$

On the other hand, $h_{\beta,\infty}^{-1}\le\kappa_{\min}^{-1}\varepsilon^{-1/\beta}$, since ε is supposed to satisfy $\log(1/\varepsilon)\ge e$. This and (6.24) imply

$$\begin{aligned}P\left(\tilde{Z}_{\gamma,\infty}\ge\frac{\delta\psi_{\beta,\infty}}{3}\right)&\le k_{11}\exp\left(\frac{1}{\beta}\log\frac{1}{\varepsilon}\left(1-k_{10}\left(\frac{\delta^2}{18}\right)\exp\left(\frac{\tilde{\nu}_\varepsilon}{2}\right)\right)\right)\\&\le k_{11}\exp\left(-k_{12}\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\exp\left(\frac{\tilde{\nu}_\varepsilon}{2}\right)\right)=o(1)\text{ as }\varepsilon\to0.\end{aligned}\tag{6.25}$$

Now, the random variable ξ^* is Gaussian with mean 0 and with the variance $\mathrm{Var}(\xi^*)$, satisfying

$$\begin{aligned}\mathrm{Var}(\xi^*) &\le 2\big[\mathrm{Var}\big(Z_{\beta,\infty}(t^*)\big) + \mathrm{Var}\big(Z_{\gamma,\infty}(t^*)\big)\big] = 2\big(r(0,h_{\beta,\infty}) + r(0,h_{\gamma,\infty})\big)\\ &\le 2(r_{\beta,\infty} + r_{\gamma,\infty})(1+\delta_{\varepsilon 0}) \le k_{13}\varepsilon^2 h_{\beta,\infty}^{-1},\end{aligned}$$

where we used Lemma 1, (A.13) and the inequalities $\delta_{\varepsilon 0} \le 1/2$, $h_{\beta,\infty} \le h_{\gamma,\infty}$, $\beta < \gamma$ (see Lemma 4). Hence, for ε small enough,

$$\begin{aligned}P_f\left(|\xi^*| \ge \frac{\delta\psi_{\beta,\infty}}{3}\right) &\le P_f\left(|\xi^*| \ge \frac{\delta\eta_\infty(\beta)}{3}\right)\\ &\le 2\exp\left(-\left(\frac{\delta^2}{18}\right)\frac{\eta_\infty^2(\beta)}{\mathrm{Var}(\xi^*)}\right) \le 2\exp\left(-k_{14}\frac{1}{\beta}\log\frac{1}{\varepsilon}\right) \qquad (6.26)\\ &\le 2\exp\left(-k_{14}\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right) = o(1) \quad \text{as } \varepsilon \to 0.\end{aligned}$$

It follows from (6.23), (6.25) and (6.26) that

$$\sup_{f\in W_\beta} \rho_f(\beta,\gamma) \le k_{15}\exp\left(-\frac{1}{k_{15}\,\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right), \qquad \gamma > \beta,\ \gamma \in B.$$

This and (6.20), (6.21) yield

$$\begin{aligned}R^+_{\varepsilon,\beta,\infty} &\le (1+\delta)^p + k_{16}\,\mathrm{card}(B)\exp\left(-\frac{1}{k_{16}\,\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\\ &= (1+\delta)^p + o(1) \quad \text{as } \varepsilon \to 0.\end{aligned}$$

Since $\delta > 0$ can be chosen arbitrarily small, this completes the proof of (6.2) with $j = \infty$. □

PROOF OF (6.2) FOR $j = 0$. Note first that

$$\begin{aligned}R^+_{\varepsilon,\beta_\varepsilon^*,0} &= \sup_{f\in W_{\beta_\varepsilon^*}} E_f\Big(\psi^{-p}_{\beta_\varepsilon^*,0}|f^*_{\varepsilon,0}(x_0) - f(x_0)|^p I\big\{\hat\beta_0 = \beta_\varepsilon^*\big\}\Big)\\ &= o(1)R^+_\varepsilon, \text{ as } \varepsilon \to 0,\end{aligned}$$

where R^+_ε is defined in the proof of (4.2) below. Since $R^+_\varepsilon = O(1)$, as $\varepsilon \to 0$ [see the proof of (4.2)], we have $R^+_{\varepsilon,\beta_\varepsilon^*,0} = o(1)$, and it suffices to prove a weaker version of (6.2) where $\sup_{\beta\in B}$ is replaced by $\sup_{\beta\in B_-}$. Therefore, in the rest of the proof we assume $\beta \in B_-$.

If $\hat\beta_0 = \gamma \ge \beta$, then, by definition of $\hat\beta_0$,

$$\big|f^*_{\varepsilon,0}(x_0) - f_{\varepsilon,\beta,0}(x_0)\big| = \big|f_{\varepsilon,\gamma,0}(x_0) - f_{\varepsilon,\beta,0}(x_0)\big| \le \eta_0(\beta),$$

and thus, for $\hat{\beta}_0 \geq \beta$ and $f \in W_\beta$, using (5.14), we get

$$
\begin{aligned}
|f^*_{\varepsilon,0}(x_0) - f(x_0)| &\leq \left|f_{\varepsilon,\beta,0}(x_0) - f(x_0)\right| + \eta_0(\beta) \\
&\leq \eta_0(\beta) + b_0(\beta)\left(1 + c_3 \exp\left(-\frac{1}{4\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)\right) + \tilde{Z}_{\beta,0} \\
&\leq \left(\psi_{\beta,0} + \tilde{Z}_{\beta,0}\right)(1 + o(1)),
\end{aligned}
$$

where $o(1) \to 0$ as $\varepsilon \to 0$ uniformly in $\beta \in B$. Hence, for any $\beta \in \mathrm{B}_-$,

$$
\begin{aligned}
R^+_{\varepsilon,\beta,0} &\leq \psi^{-p}_{\beta,0} E\left(\left(\psi_{\beta,0} + \tilde{Z}_{\beta,0}\right)^p\right)(1 + o(1)) \\
&\leq \left[\psi^{-p}_{\beta,0}\left(\psi_{\beta,0} + \tau_0(\beta)\right)^p + \psi^{-p}_{\beta,0} E\left(\left(\psi_{\beta,0} + \tilde{Z}_{\beta,0}\right)^p I\left\{\tilde{Z}_{\beta,0} \geq \tau_0(\beta)\right\}\right)\right] \\
&\quad \times (1 + o(1)) \\
&\leq \left[\left(1 + \left(\tau_0(\beta)/\psi_{\beta,0}\right)\right)^p\right. \\
&\quad \left. + 2^{p-1} c_5 \left(1 + \left(r^{1/2}_{\beta,0}/\psi_{\beta,0}\right)^p\right) \exp\left(-\left(\tau^2_0(\beta)/2r_{\beta,0}\right)(1 - \delta_{\varepsilon 0})\right)\right] \\
&\quad \times (1 + o(1)),
\end{aligned}
\tag{6.27}
$$

where Lemma 7 was used. Now, for any $\beta \in B_-$,

$$
\begin{aligned}
\frac{r^{1/2}_{\beta,0}}{\psi_{\beta,0}} &\leq \frac{r^{1/2}_{\beta,0}}{\eta_0(\beta)} \leq p^{-1/2}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{-1/2}, \\
\frac{\tau_0(\beta)}{\psi_{\beta,0}} &= r^{1/2}_{\beta,0}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{1/4} \psi^{-1}_{\beta,0} \leq p^{-1/2}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{-1/4}
\end{aligned}
\tag{6.28}
$$

and

$$
\begin{aligned}
&\left(\frac{\tau^2_0(\beta)}{2r_{\beta,0}}\right)(1 - \delta_{\varepsilon 0}) \\
&\quad = \frac{1}{2}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{1/2}(1 - \delta_{\varepsilon 0}) \geq \frac{1}{4}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{1/2}.
\end{aligned}
\tag{6.29}
$$

The substitution of (6.28) and (6.29) into (6.27) gives, for any $\beta \in B_-$,

$$
\begin{aligned}
R^+_{\varepsilon,\beta,0} \leq &\left[\left(1 + p^{-1/2}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{-1/4}\right)^p\right. \\
&\left. + k_{17}\exp\left(-\frac{1}{4}\left(\frac{1}{\beta^*_\varepsilon}\log\frac{1}{\varepsilon}\right)^{1/2}\right)\right](1 + o(1)).
\end{aligned}
$$

This entails (6.2) for $j = 0$. $\square$

Proof of (3.7). Using (5.14) with $j = 1$, we get, for any $\beta \in B$, $\delta > 0$,

$$\sup_{f \in W_\beta} E_f\Big(\|f_{\varepsilon,\beta,1} - f\|_\infty^p\Big)$$
$$\le \sup_{f \in W_\beta} E_f\Big(\Big(\tilde{Z}_{\beta,1} + \|E_f(f_{\varepsilon,\beta,1}) - f\|_\infty\Big)^p\Big)$$
$$\le \sup_{f \in W_\beta} E_f\Big(\Big(\tilde{Z}_{\beta,1} + \|E_f(f_{\varepsilon,\beta,1}) - f\|_\infty\Big)^p I\Big\{\tilde{Z}_{\beta,1} < (1+\delta)\eta_1(\beta)\Big\}\Big)$$
$$+ \sup_{f \in W_\beta} E_f\Big(\Big(\tilde{Z}_{\beta,1} + \|E_f(f_{\varepsilon,\beta,1}) - f\|_\infty\Big)^p I\Big\{\tilde{Z}_{\beta,1} \ge (1+\delta)\eta_1(\beta)\Big\}\Big) \tag{6.30}$$
$$\le (1+\delta)^p(\eta_1(\beta) + b_1(\beta))^p\left(1 + c_3 \exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right)^p$$
$$+ 2^{p-1}\Bigg(E\Big(\tilde{Z}_{\beta,1}^p I\Big\{\tilde{Z}_{\beta,1} \ge (1+\delta)\eta_1(\beta)\Big\}\Big)$$
$$+ b_1^p(\beta)\left(1 + c_3 \exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right)^p P\Big\{\tilde{Z}_{\beta,1} \ge (1+\delta)\eta_1(\beta)\Big\}\Bigg).$$

By Lemma 6,

$$P\Big(\tilde{Z}_{\beta,1} \ge (1+\delta)\eta_1(\beta)\Big) \le \frac{c_4}{h_{\beta,1}}\exp\left(-(1+\delta)^2(1-\delta_{\varepsilon 0})\frac{1}{\beta}\log\frac{1}{\varepsilon}\right) \tag{6.31}$$

and

$$E\Big(\tilde{Z}_{\beta,1}^p I\Big\{\tilde{Z}_{\beta,1} \ge (1+\delta)\eta_1(\beta)\Big\}\Big)$$
$$\le \frac{c_4}{h_{\beta,1}} r_{\beta,1}^{p/2}\exp\left(-(1+\delta)^2(1-\delta_{\varepsilon 0})\frac{1}{\beta}\log\frac{1}{\varepsilon}\right) \tag{6.32}$$
$$\le \eta_1^p(\beta)\left(\frac{2}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)^{-p/2}\frac{c_4}{h_{\beta,1}}\exp\left(-(1+\delta)^2(1-\delta_{\varepsilon 0})\frac{1}{\beta}\log\frac{1}{\varepsilon}\right).$$

Now, in view of Lemma 1, $h_{\beta,1} \ge \kappa_{\min}(\varepsilon^2 \log(1/\varepsilon))^{1/2\beta}$, and thus

$$\frac{c_4}{h_{\beta,1}}\exp\left(-(1+\delta)^2(1-\delta_{\varepsilon 0})\frac{1}{\beta}\log\frac{1}{\varepsilon}\right)$$
$$\le c_4\kappa_{\min}^{-1}\exp\left(-\Big[(1+\delta)^2(1-\delta_{\varepsilon 0}) - 1\Big]\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon} - \frac{1}{2\beta_\varepsilon^*}\log\log\frac{1}{\varepsilon}\right) \tag{6.33}$$

for all $\beta \in B$ and ε small enough. Combining (6.30)–(6.33) and using (5.8) and the fact that $\delta_{\varepsilon 0} \to 0, (1/\beta_\varepsilon^*)\log(1/\varepsilon) \to \infty$, one obtains (3.7). □

PROOF OF (4.2). Now let $\beta = \beta_\varepsilon^*$, $j = 0$. We have

$$\sup_{f \in W_{\beta_\varepsilon^*}} E_f\Big(\varepsilon^{-p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}\big|f_{\varepsilon,0}^*(x_0) - f(x_0)\big|^p\Big) \le R_\varepsilon^- + R_\varepsilon^+, \tag{6.34}$$

where

$$R_\varepsilon^- = \sup_{f \in W_{\beta_\varepsilon^*}} E_f\Big(\varepsilon^{-p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}\big|f_{\varepsilon,0}^*(x_0) - f(x_0)\big|^p I\big\{\hat\beta_0 \in B_-\big\}\Big),$$

$$R_\varepsilon^+ = \sup_{f \in W_{\beta_\varepsilon^*}} E_f\Big(\varepsilon^{-p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}\big|f_{\varepsilon,0}^*(x_0) - f(x_0)\big|^p I\big\{\hat\beta_0 = \beta_\varepsilon^*\big\}\Big).$$

As in (6.3),

$$\begin{aligned} R_\varepsilon^- &\le \varepsilon^{-p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*} \sum_{\gamma \in B_-} \sup_{f \in W_{\beta_\varepsilon^*}} E_f\Big(\big|f_{\varepsilon,\gamma,0}(x_0) - f(x_0)\big|^p I\big\{\hat\beta_0 = \gamma\big\}\Big) \\ &\le \varepsilon^{-p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}\psi_{\beta_\varepsilon^*,0}^p\big(\rho_{1,0}(\beta_\varepsilon^*) + \rho_{2,0}(\beta_\varepsilon^*)\big) \\ &\le c^p(\beta_\varepsilon^*,0)\Big(\log\frac{1}{\varepsilon}\Big)^{p(2\beta_\varepsilon^*-1)/4\beta_\varepsilon^*}\big(\rho_{1,0}(\beta_\varepsilon^*) + \rho_{2,0}(\beta_\varepsilon^*)\big). \end{aligned}$$

This, together with the definition of $c(\beta,0)$ and (6.8), (6.16), (5.6) implies

$$\limsup_{\varepsilon \to 0} R_\varepsilon^- = 0. \tag{6.35}$$

Next, if $\hat\beta_0 = \beta_\varepsilon^*$ and $f \in W_{\beta_\varepsilon^*}$, using (5.14) we get

$$\begin{aligned} \big|f_{\varepsilon,0}^*(x_0) - f(x_0)\big| &= \big|f_{\varepsilon,\beta_\varepsilon^*,0}(x_0) - f(x_0)\big| \\ &\le b_0(\beta_\varepsilon^*)\Big(1 + c_3\exp\Big(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\Big)\Big) + \tilde Z_{\beta_\varepsilon^*,0} \\ &\le Lb_{\max}\varepsilon^{(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}(1 + o(1)) + \tilde Z_{\beta_\varepsilon^*,0}. \end{aligned} \tag{6.36}$$

Furthermore, using Lemma 7 with $u = (\varepsilon^2 h_{\beta_\varepsilon^*,0}^{-1})^{1/2}$, one obtains

$$\begin{aligned} E\big(\tilde Z_{\beta_\varepsilon^*,0}^p\big) &\le \big(\varepsilon^2 h_{\beta_\varepsilon^*,0}^{-1}\big)^{p/2} \\ &\quad + c_5 r_{\beta_\varepsilon^*,0}^{p/2}\exp\Big(-\frac{1}{2r_{\beta_\varepsilon^*,0}}\varepsilon^2 h_{\beta_\varepsilon^*,0}^{-1}(1-\delta_{\varepsilon 0})\Big) \\ &\le \varepsilon^{p(2\beta_\varepsilon^*-1)/2\beta_\varepsilon^*}\big[1 + c_5 v_{\max}^p\big]. \end{aligned} \tag{6.37}$$

It follows from (6.36) and (6.37) that

$$\limsup_{\varepsilon \to 0} R_\varepsilon^+ < \infty.$$

Combining this result with (6.34) and (6.35) one gets (4.2). □

7. Proofs of the lower bounds. This section is devoted to the proofs of the lower bounds (3.6), (3.8), (4.3), (3.17) and Theorem 4. Let us start with some auxiliary facts.

Define the Fourier transform of a function $f(x) \in L_1(-\infty, \infty)$ as

$$\hat{f}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x) e^{-ix\omega}\, dx,$$

and denote $f^{(\beta)}$ the Weyl derivative of order $\beta > 1/2$ of a function f, that is,

$$f^{(\beta)}(x) = \int_{-\infty}^{\infty} \hat{f}(\omega)(i\omega)^{\beta} e^{ix\omega}\, d\omega,$$

if the last integral exists.

Denote, for every $\beta > 1/2$,

$$K_{0,\beta}(x) = \int_{-\infty}^{\infty} \frac{e^{itx}}{1+|t|^{2\beta}}\, dt = 2\int_0^{\infty} \frac{\cos(tx)}{1+t^{2\beta}}\, dt.$$

Note that the Fourier transform of $K_{0,\beta}$ is $\hat{K}_{0,\beta}(\omega) = 1/(1+|\omega|^{2\beta})$. By Plancherel's formula, the L_2-norm of $K_{0,\beta}$ equals

$$\|K_{0,\beta}\|_2 = \left(2\pi\int_{-\infty}^{\infty} \frac{1}{(1+|t|^{2\beta})^2}\, dt\right)^{1/2} = 2\pi v_\beta. \tag{7.1}$$

Similarly, the L_2-norm of the Weyl derivative $K_{0,\beta}^{(\beta)}$ equals

$$\left\|K_{0,\beta}^{(\beta)}\right\|_2 = \left(2\pi\int_{-\infty}^{\infty} \frac{|t|^{2\beta}}{(1+|t|^{2\beta})^2}\, dt\right)^{1/2} = 2\pi b_\beta. \tag{7.2}$$

Define the function

$$K_\beta(x) = s_1 K_{0,\beta}(s_2 x),$$

where $s_1 = (2\pi v_\beta)^{-1}(2\beta-1)^{1/4\beta}$, $s_2 = (2\beta-1)^{1/2\beta}$. For $j = 0, 1, \infty$, $\beta \in B$, denote

$$\bar{\kappa}_{\beta,j} = \left(\frac{\alpha_j}{L^2\beta}\right)^{1/2\beta}, \qquad \bar{h}_{\beta,j} = \bar{\kappa}_{\beta,j}\left(\varepsilon^2 \log(1/\varepsilon)\right)^{1/2\beta}.$$

Lemma 10. *We have*

$$\|K_\beta\|_2 = 1, \qquad \left\|K_\beta^{(\beta)}\right\|_2 = 1, \qquad K_\beta(0) = 2\beta(2\beta-1)^{-(2\beta-1)/4\beta} b_\beta. \tag{7.3}$$

Furthermore, for any $\varepsilon \in (0,1)$, $\delta \in (0,1)$ and $\beta \in B$, there exist a number $D = D(\varepsilon, \beta, \delta) > 0$ and a function $\overline{K}_\beta(\cdot)$ such that:

(i) $\operatorname{supp} \overline{K}_\beta = (-D, D)$;
(ii) $|\hat{\overline{K}}_\beta(\omega)| \le (1-\delta/2)|\hat{K}_\beta(\omega)|$, $\forall \omega \in (-\infty, \infty)$;
(iii) $\|\overline{K}_\beta\|_2 \le 1-\delta/2$;
(iv) $\overline{K}_\beta(0) \ge K_\beta(0)(1-\delta)$;
(v) $\sup_{\beta\in B} (D(\varepsilon,\beta,\delta)\bar{h}_{\beta,j}) = o(1)$ *as* $\varepsilon \to 0$, $j = 0, 1, \infty$,

where $\hat{\overline{K}}_\beta(\omega)$ and $\hat{K}_\beta(\omega)$ are the Fourier transforms of $\overline{K}_\beta$ and K_β respectively.

Proof of Lemma 10 is given in the Appendix.

For $j = 0, 1, \infty,\ \beta \in B$, denote

$$G_{\beta, j}(x) = L\bar{h}_{\beta, j}^{\beta - 1/2} \bar{K}_\beta\big(x/\bar{h}_{\beta, j}\big),$$

where $\bar{K}_\beta$ satisfies the conditions (i) to (iv) of Lemma 10 for a $\delta \in (0, 1/2)$. In particular, $\operatorname{supp} G_{\beta, j} = (-D\bar{h}_{\beta, j}, D\bar{h}_{\beta, j})$.

PROOF OF (3.6). Assume that $0 < \delta < \min\{1, 2p\}/(p + 2)$. Set $M_\beta = [1/(2D\bar{h}_{\beta, \infty})] - 1$, where $D = D(\varepsilon, \beta, \delta)$, and let $\varepsilon > 0$ be so small that $M_\beta \geq 1$ (cf. (v) of Lemma 10). Then the interval $[0, 1]$ contains M_β disjoint subintervals of length $2D\bar{h}_{\beta, \infty}$.

Define the family of functions $\mathscr{F}_\beta = \{f_{k\beta}(\cdot),\ k = 0, 1, \dots, M_\beta\}$ on $[0, 1]$, where

$$f_{0\beta}(x) \equiv 0, \qquad f_{k\beta}(x) = G_{\beta, \infty}(x - x_{k, \infty}),$$

with $x_{k, \infty} = (2k - 1)D\bar{h}_{\beta, \infty}$, $k = 1, \dots, M_\beta$. Using Lemma 10, it is easy to see that

$$\operatorname{supp} f_{k\beta} = \Big((2k - 2)D\bar{h}_{\beta, \infty}, 2kD\bar{h}_{\beta, \infty}\Big) \subseteq (0, 1),$$

for ε small enough, and

$$|\hat{f}_{k\beta}(\omega)| \leq \left(1 - \frac{\delta}{2}\right) L\bar{h}_{\beta, \infty}^{\beta + 1/2} |\hat{K}_\beta\big(\omega \bar{h}_{\beta, \infty}\big)|, \qquad k = 1, \dots, M_\beta, \forall \omega \in (-\infty, \infty), \tag{7.4}$$

$$\begin{aligned} \|f_{k\beta}\|_2^2 &= L^2 \bar{h}_{\beta, \infty}^{2\beta} \|\bar{K}_\beta\|_2^2 \\ &\leq (1 - \delta/2)^2 (\alpha_\infty/\beta) \varepsilon^2 \log(1/\varepsilon),\ k = 1, \dots, M_\beta. \end{aligned} \tag{7.5}$$

It follows from Proposition 4 in the Appendix that $f_{k\beta} \in W_\beta$, $k = 0, 1, \dots, M_\beta$, for ε small enough, and thus $\mathscr{F}_\beta \subset W_\beta$.

Fix $\beta' \in B$, $\beta' \neq \beta^*_\varepsilon$. Consider the binary vectors $\vartheta_k = (\delta_{1k}, \dots, \delta_{Mk})$, $k = 0, 1, \dots, M$, where $M = M_{\beta'}$, and δ_{ik} is the Kronecker delta. For an arbitrary estimator T_ε,

$$\|T_\varepsilon - f_{i\beta'}\|_\infty \geq \max_{1 \leq i \leq M} \big|T_\varepsilon(x_{i, \infty}) - f_{i\beta'}(x_{i, \infty})\big| = G_{\beta', \infty}(0)\, d\big(\hat{\vartheta}, \vartheta_i\big)/(1 - \delta)$$

where $\hat{\vartheta} = (\hat{\vartheta}_1, \dots, \hat{\vartheta}_M)$, $\hat{\vartheta}_i = T_\varepsilon(x_{i, \infty})/G_{\beta', \infty}(0)$, $i = 1, \dots, M$, and the distance

$$d(u, v) = (1 - \delta) \max_{1 \leq i \leq M} |u_i - v_i|$$

for any two M-vectors $u = (u_1, \dots, u_M)$, $v = (v_1, \dots, v_M)$. Taking this into account and denoting $E_i = E_{f_{i\beta'}}$, one obtains

$$\begin{aligned} &\sup_{\beta \in B} \sup_{f \in W_\beta} E_f\big(\psi_{\beta, \infty}^{-p} \|T_\varepsilon - f\|_\infty^p\big) \\ &\quad \geq \max\Big\{E_0\big[\psi_{\beta^*_\varepsilon, \infty}^{-p} \|T_\varepsilon\|_\infty^p\big],\ \max_{1 \leq i \leq M} E_i\big[\psi_{\beta', \infty}^{-p} \|T_\varepsilon - f_{i\beta'}\|_\infty^p\big]\Big\} \\ &\quad \geq \max\Big\{E_0\Big[\big(qd\big(\hat{\vartheta}, \vartheta_0\big)\big)^p\Big],\ \max_{1 \leq i \leq M} E_i\big[d^p\big(\hat{\vartheta}, \vartheta_i\big)\big]\Big\}, \end{aligned} \tag{7.6}$$

where

$$q = G_{\beta',\infty}(0)\psi^{-1}_{\beta^*_\varepsilon,\infty}/(1-\delta),$$

and the following consequence of Lemma 10 was used:

$$G_{\beta,j}(0) = L\bar{h}_{\beta,j}^{\beta-1/2}\bar{K}_\beta(0) \ge (1-\delta)L\bar{h}_{\beta,j}^{\beta-1/2}K_\beta(0) = (1-\delta)\psi_{\beta,j} \tag{7.7}$$

(valid for all $\beta \in B$, $j = 0, 1, \infty$, except for the combination $j = 0$, $\beta = \beta^*_\varepsilon$). To estimate from below the last expression in (7.6), use Theorem 6 from the Appendix. Let us check the conditions (A.1) and (A.2) of Theorem 6 with $\Theta = \{\vartheta_0, \vartheta_1, \dots, \vartheta_M\}$ and $d\{\cdot,\cdot\}$ defined above, $w(u) = u^p$, $\alpha = \delta$ and $\tau = \varepsilon^\gamma$, where $\gamma = [p - \delta(p+2)/2]/2\beta'$. It suffices to check (A.2), since (A.1) is straightforward in view of the definition of ϑ_i and $d(\cdot,\cdot)$. For the rest of the proof, $P_i = P_{f_{i\beta'}}$. Clearly,

$$Q\left(\frac{dP_0}{dQ} \ge \tau\right) = M^{-1}\sum_{i=1}^M \tilde{p}_i, \tag{7.8}$$

where $\tilde{p}_i = P_i(M^{-1}\sum_{k=1}^M (dP_k/dP_0) \le 1/\tau)$. Fix i and estimate $\tilde{p}_i$ from below. Standard results on the absolute continuity of Gaussian measures [see, e.g., Ibragimov and Hasminskii (1981), Appendix 2] imply that under P_i,

$$\frac{dP_k}{dP_0} = \exp\left(\varepsilon^{-1}\sigma\zeta_k + \mu_k\varepsilon^{-2}\sigma^2/2\right), \qquad k = 1, \dots, M,$$

where ζ_k are i.i.d. $\mathcal{N}(0,1)$ random variables, $\sigma^2 = \|f_{i\beta'}\|_2^2$ and

$$\mu_k = \begin{cases} 1, & \text{if } k = i, \\ -1, & \text{if } k \ne i. \end{cases}$$

Hence, using the independence of ζ_i and $\{\zeta_k,\ k \ne i\}$, one obtains

$$\begin{aligned} \tilde{p}_i &= P\left(M^{-1}\sum_{k=1}^M \exp\left(\varepsilon^{-1}\sigma\zeta_k + \mu_k\varepsilon^{-2}\sigma^2/2\right) \le 1/\tau\right) \\ &\ge P\left(M^{-1}\sum_{\substack{k=1,\\ k\ne i}}^M \exp\left(\varepsilon^{-1}\sigma\zeta_k - \varepsilon^{-2}\sigma^2/2\right) \le 1/2\tau\right) \\ &\quad \times P\left(M^{-1}\exp\left(\varepsilon^{-1}\sigma\zeta_i + \varepsilon^{-2}\sigma^2/2\right) \le 1/2\tau\right). \end{aligned} \tag{7.9}$$

By Chebyshev's inequality,

$$\begin{aligned} &P\left(M^{-1}\sum_{\substack{k=1,\\ k\ne i}}^M \exp\left(\varepsilon^{-1}\sigma\zeta_k - \varepsilon^{-2}\sigma^2/2\right) \le 1/2\tau\right) \\ &\quad \ge 1 - 2\tau\frac{M-1}{M}E\left(\exp\left(\varepsilon^{-1}\sigma\zeta_k - \varepsilon^{-2}\sigma^2/2\right)\right) \\ &\quad = 1 - 2\tau\frac{M-1}{M} = 1 + o(1) \quad \text{as } \varepsilon \to 0, \end{aligned} \tag{7.10}$$

since $\tau \to 0$, as $\varepsilon \to 0$. It follows from (7.5) and the definition of $M = M_{\beta'}$ that

$$\sigma^2 \le \left(1 - \frac{\delta}{2}\right)^2 (p+2)\frac{1}{\beta'}\varepsilon^2 \log\frac{1}{\varepsilon},$$
$$\log M = \frac{1}{\beta'}\log\frac{1}{\varepsilon}(1 + o(1)), \qquad \varepsilon \to 0. \tag{7.11}$$

(to get the last equality, use (3.1) and the definition of $D(\varepsilon, \beta, \delta)$ in the proof of Lemma 10; here $o(1) \to 0$ uniformly in $\beta' \in B$).

Now,

$$\begin{aligned} &P\big(M^{-1}\exp\big(\varepsilon^{-1}\sigma\zeta_i + \varepsilon^{-2}\sigma^2/2\big) \le 1/2\tau\big)\\ &\quad = P\big(-\zeta_i \ge (\varepsilon/\sigma)\big[\log(2\tau) - \log M + \varepsilon^{-2}\sigma^2/2\big]\big)\\ &\quad = 1 - \Phi(\lambda_\varepsilon), \end{aligned} \tag{7.12}$$

where $\Phi(\cdot)$ is the standard normal c.d.f., and, in view of (7.11),

$$\begin{aligned} \lambda_\varepsilon &= (\varepsilon/\sigma)\left[\left(-\gamma - \frac{1}{\beta'}\right)\log\frac{1}{\varepsilon}(1 + o(1)) + \frac{\varepsilon^{-2}\sigma^2}{2}\right]\\ &\le \left(\frac{\varepsilon}{\sigma}\right)\left[-\frac{\delta}{4}\left(1 - \frac{\delta}{2}\right)(p+2)\frac{1}{\beta'}\log\frac{1}{\varepsilon}(1 + o(1))\right]\\ &\le -\frac{\delta}{4}\left[(p+2)\frac{1}{\beta'}\right]^{1/2}\left(\log\frac{1}{\varepsilon}\right)^{1/2}(1 + o(1)). \end{aligned}$$

Choose $\beta' = \beta_1$. Then, $\lambda_\varepsilon \to -\infty$, as $\varepsilon \to 0$. This, together with (7.12) implies

$$P\big(M^{-1}\exp\big(\varepsilon^{-1}\sigma\zeta_i + \varepsilon^{-2}\sigma^2/2\big) \le 1/2\tau\big) \ge 1 - \delta/2$$

for ε small enough. Combining this inequality with (7.9) and (7.10), one obtains that $\tilde{p}_i \ge 1 - \delta$ for ε small enough, uniformly over i, which, together with (7.8) entails (A.2) for $\alpha = \delta$. Therefore, for ε small enough, it is possible to apply (A.3). It follows from (A.3) and (7.6) that

$$\inf_{T_\varepsilon}\sup_{\beta\in B}\sup_{f\in W_\beta} E_f\big(\psi_{\beta,\infty}^{-p}\|T_\varepsilon - f\|_\infty^p\big) \ge \frac{(1-\delta)\tau(1-2\delta)^p(q\delta)^p}{(1-2\delta)^p + (q\delta)^p\tau}, \tag{7.13}$$

for ε small enough. It follows from (7.7) and from the definitions of τ and q that $\tau q^p = \varepsilon^\gamma q^p \ge \varepsilon^\gamma(\psi_{\beta_1,\infty}/\psi_{\beta_\varepsilon^*,\infty})^p \to \infty$, as $\varepsilon \to 0$. To get (3.6), it suffices to take the lower limits of both sides of (7.13), as $\varepsilon \to 0$, and to note that $\delta > 0$ can be chosen arbitrarily small. $\square$

PROOF OF (4.3). Let $\beta' \in B_-$ and $\beta'' \in B_-$ be such that $\beta'' > \beta'$. Consider the functions $f_0(x) \equiv 0$ and $f_1(x) = G_{\beta',0}(x - x_0)(1 - \beta'/\beta'')^{1/2}$.

Clearly, $\operatorname{supp} f_1 = (x_0 - D\bar{h}_{\beta',0}, x_0 + D\bar{h}_{\beta',0}) \subset (0,1)$ for ε small enough, and

$$\left|\hat{f}_1(\omega)\right| \le \left(1 - \frac{\delta}{2}\right) L\bar{h}_{\beta',0}^{\beta'+1/2}\left|\hat{K}_{\beta'}\left(\omega\bar{h}_{\beta',0}\right)\right|, \forall\omega \in (-\infty,\infty),$$

$$\|f_1\|_2^2 \le \left(1 - \frac{\delta}{2}\right)^2 \alpha_0\left(\frac{1}{\beta'} - \frac{1}{\beta''}\right)\varepsilon^2 \log\frac{1}{\varepsilon} \tag{7.14}$$

$$= \left(1 - \frac{\delta}{2}\right)^2 p\left(\frac{1}{\beta'} - \frac{1}{\beta''}\right)\varepsilon^2 \log\frac{1}{\varepsilon}$$

[cf. (7.4) and (7.5)]. Following the lines of Proposition 4 in the Appendix it is easy to show that $f_1 \in W_{\beta'}$ for ε small enough.

Let $\Theta = \{\vartheta_0, \vartheta_1\}$, where $\vartheta_0 = 0$, $\vartheta_1 = 1$. For an arbitrary estimator T_ε and $i = 0, 1$, we have

$$|T_\varepsilon(x_0) - f_i(x_0)| = G_{\beta',0}(0)\left(1 - \frac{\beta'}{\beta''}\right)^{1/2}\frac{d(\hat{\vartheta}, \vartheta_i)}{1-\delta},$$

where $\hat{\vartheta} = (T_\varepsilon(x_0)/G_{\beta',0}(0))(1 - \beta'/\beta'')^{-1/2}$, and $d(u,v) = (1-\delta)|u - v|$, $u, v \in \mathbb{R}$.

Denoting $E_i = E_{f_i}$ and using (7.7) with $\beta = \beta'$, $j = 0$, we get

$$\sup_{\beta\in B_-}\sup_{f\in W_\beta} E_f\left(\psi_{\beta,0}^{-p}|T_\varepsilon(x_0) - f(x_0)|^p\right)$$

$$\ge \max\left\{E_0\left[\psi_{\beta'',0}^{-p}|T_\varepsilon(x_0)|^p\right], E_1\left[\psi_{\beta',0}^{-p}|T_\varepsilon(x_0) - f_1(x_0)|^p\right]\right\} \tag{7.15}$$

$$\ge \left(1 - \frac{\beta'}{\beta''}\right)^{p/2}\max\left\{E_0\left[\left(qd(\hat{\vartheta}, \vartheta_0)\right)^p\right], E_1\left[d^p(\hat{\vartheta}, \vartheta_1)\right]\right\},$$

where

$$q = \frac{G_{\beta',0}(0)\psi_{\beta'',0}^{-1}}{1-\delta} = \psi_{\beta',0}\psi_{\beta'',0}^{-1}\left(1 - \frac{\beta'}{\beta''}\right)^{1/2}.$$

The last expression in (7.15) is estimated from below by use of (A.3) of Theorem 6 in the Appendix with $w(u) = u^p$, $M = 1$ and $\Theta = \{\vartheta_0, \vartheta_1\}$ defined above. To apply (A.3), it suffices to check (A.2) with $\alpha = \delta$, $\tau = \varepsilon^{\gamma'}$, $\gamma' = p(1 - \delta/2)(1/2)(1/\beta' - 1/\beta'')$, $Q = P_1$ (in the rest of the proof, $P_i = P_{f_i}$, $i = 0, 1$).

We have

$$P_1\left(\frac{dP_0}{dP_1} \ge \tau\right) = P\left(\exp\left(\varepsilon^{-1}\bar{\sigma}\zeta - \frac{\varepsilon^{-2}\bar{\sigma}^2}{2}\right) \ge \tau\right)$$

$$= P\left(\zeta \ge \frac{\varepsilon}{\bar{\sigma}}\left[\log\tau + \frac{\varepsilon^{-2}\bar{\sigma}^2}{2}\right]\right) = 1 - \Phi(\bar{\lambda}_\varepsilon),$$

where $\zeta \sim \mathcal{N}(0,1)$, $\bar\sigma^2 = \|f_1\|_2^2$, and, in view of (7.14),

$$\begin{aligned}\bar\lambda_\varepsilon &= \frac{\varepsilon}{\bar\sigma}\left[(-\gamma')\log\frac{1}{\varepsilon} + \frac{\varepsilon^{-2}\bar\sigma^2}{2}\right]\\ &\le \frac{\varepsilon}{\bar\sigma}\left[-\frac{\delta}{4}\left(1-\frac{\delta}{2}\right)p\left(\frac{1}{\beta'}-\frac{1}{\beta''}\right)\log\frac{1}{\varepsilon}\right]\\ &\le -\frac{\delta}{4}\left(p\left(\frac{1}{\beta'}-\frac{1}{\beta''}\right)\right)^{1/2}\left(\log\frac{1}{\varepsilon}\right)^{1/2}.\end{aligned}$$

Choose $\beta' = \beta_1$, $\beta'' = \max\{\beta \in B\colon \beta < \beta_\varepsilon^*/2\}$. Then $\beta'' \to \infty$, $\bar\lambda_\varepsilon \to -\infty$, as $\varepsilon \to 0$, and thus $P_1(dP_0/dP_1 \ge \tau) \ge 1-\delta$ for ε small enough. Therefore, by Theorem 6 from the Appendix, for ε small enough, one can use (A.3) to evaluate the last expression in (7.15). This results in

$$\inf_{T_\varepsilon}\sup_{\beta\in B_-}\sup_{f\in W_\beta} E_f\left(\psi_{\beta,0}^{-p}|T_\varepsilon(x_0)-f(x_0)|^p\right) \ge \frac{(1-\delta)\tau(1-2\delta)^p(q\delta)^p}{(1-2\delta)^p+(q\delta)^p\tau}$$

for ε small enough. Here $\tau q^p = \varepsilon^{\gamma'}q^p \ge \varepsilon^{\gamma'}\ (\psi_{\beta',0}/\psi_{\beta'',0})^p(1+o(1)) \to \infty$, as $\varepsilon \to 0$. We finish as in the proof of (3.6). □

PROOF OF (3.8). Use Theorem 6(ii) from the Appendix. Fix $0 < \delta < 1/2$ and $\beta \in B$. Set $M = [1/(2D\bar h_{\beta,1})] - 1$ where $D = D(\varepsilon, \beta, \delta)$, and let $\varepsilon > 0$ be so small that $M \ge 1$. Define the functions on $[0,1]$ as follows

$$\tilde f_{0\beta}(x) \equiv 0, \qquad \tilde f_{k\beta}(x) = G_{\beta,1}(x - x_{k,1}),$$

with $x_{k,1} = (2k-1)D\bar h_{\beta,1}$, $k = 1,\dots,M$. Similarly to (7.4), (7.5) and to the argument after these inequalities, one gets that $\tilde f_{k\beta} \in W_\beta$, $k = 0,1,\dots,M$, and

$$\|\tilde f_{k\beta}\|_2^2 \le (1-\delta/2)^2(2/\beta)\varepsilon^2\log(1/\varepsilon), \qquad k = 1,\dots,M, \tag{7.16}$$

and that for an arbitrary estimator T_ε,

$$\|T_\varepsilon - \tilde f_{i\beta}\|_\infty \ge G_{\beta,1}(0)\,d(\hat\vartheta,\vartheta_i)/(1-\delta), \tag{7.17}$$

where ϑ_i and $d(\cdot,\cdot)$ are as in the proof of (3.6), and $\hat\vartheta = (\hat\vartheta_1,\dots,\hat\vartheta_M)$, $\hat\vartheta_k = T_\varepsilon(x_{k,1})/G_{\beta,1}(0)$, $k = 1,\dots,M$. Now, in view of (7.17) and (7.7),

$$\sup_{f\in W_\beta} E_f\left(\psi_{\beta,1}^{-p}\|T_\varepsilon - f\|_\infty^p\right) \ge \max_{0\le i\le M} E_i\left[d^p(\hat\vartheta,\vartheta_i)\right]. \tag{7.18}$$

Here and later E_i and P_i, respectively, are used as a brief notation for E_f and P_f in case $f = \tilde f_{i\beta}$.

To apply Theorem 6(ii) from the Appendix, it is enough to show that (A.4) holds with $\tau = \varepsilon^{(1-\delta/2)/\beta}$, $\alpha = \delta/2$. (in fact, (A.1) is obvious). We have

$$P_k\left(\frac{dP_0}{dP_k} \ge \tau\right) = P\left(\exp\left(\varepsilon^{-1}\sigma\zeta_k - \frac{\varepsilon^{-2}\sigma^2}{2}\right) \ge \tau\right), \qquad k = 1,\dots,M,$$

where $\zeta_k \sim \mathcal{N}(0,1)$ and $\sigma^2 = \|\tilde{f}_{k\beta}\|_2^2$. This and (7.16) entail

$$P_k\left(\frac{dP_0}{dP_k} \geq \tau\right) \geq 1 - \Phi\left(\tilde{\lambda}_\varepsilon\right),$$

where $\Phi(\cdot)$ is the standard normal c.d.f., and, for any $\beta \in B$,

$$\begin{aligned}\tilde{\lambda}_\varepsilon &= \frac{\varepsilon}{\sigma}\left[-\left(1-\frac{\delta}{2}\right)\frac{1}{\beta}\log\frac{1}{\varepsilon} + \frac{1}{\beta}\left(1-\frac{\delta}{2}\right)^2\log\frac{1}{\varepsilon}\right]\\ &= -\delta\left[\frac{1}{8\beta}\log\frac{1}{\varepsilon}\right]^{1/2} \leq -\delta\left[\frac{1}{8\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right]^{1/2} \to -\infty \quad \text{as } \varepsilon \to 0.\end{aligned}$$

Hence, $P_k(dP_0/dP_k \geq \tau) \geq 1 - \Phi(-\delta[(1/8\beta_\varepsilon^*)\log(1/\varepsilon)]^{1/2})$, which yields (A.4). Moreover, there exist constants $c > 0$, $\tilde{c} > 0$, independent of β, such that for all $\beta \in B$,

$$\begin{aligned}\tau M &\geq c\varepsilon^{(1-\delta/2)/\beta}\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{-1/2\beta} D(\varepsilon,\beta,\delta)^{-1}\\ &\geq \tilde{c}\exp\left(\frac{\delta\beta_1}{2}\log\frac{1}{\varepsilon} - \frac{1}{2\beta_\varepsilon^*}\log\log\frac{1}{\varepsilon} - \frac{\beta_\varepsilon^*}{\tilde{c}}\right) \to \infty \quad \text{as } \varepsilon \to 0\end{aligned}$$

[here we use (3.1) and the expression for $D(\varepsilon,\beta,\delta)$ given in the proof of Lemma 10]. Thus, one can apply (A.5) which takes the form

$$\begin{aligned}\inf_{\hat{\vartheta}} \max_{0\leq i\leq M} E_i\left[d^p\left(\hat{\vartheta},\vartheta_i\right)\right] &\geq \left(\frac{1-\delta}{2}\right)^p \left(1 - \Phi\left(-\delta\left[\frac{1}{8\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right]^{1/2}\right)\right)\frac{\tau M}{1+\tau M}\\ &\geq \left(\frac{1-\delta}{2}\right)^p\left(1-\tilde{\delta}_\varepsilon\right),\end{aligned}$$

where $\tilde{\delta}_\varepsilon > 0$ does not depend on β and $\tilde{\delta}_\varepsilon \to 0$, as $\varepsilon \to 0$. To complete the proof of (3.8) it remains to combine this inequality with (7.18). □

PROOF OF (3.17). We act as in the proof of (3.6), with some modifications. The definitions and notation of the proof of (3.6) are used as well. For an arbitrary estimator T_ε, and for any $\beta' \in B$, $\beta' \neq \beta_\varepsilon^*$, we get

$$\begin{aligned}&\sup_{\beta\in B}\sup_{f\in W_\beta} E_f\left(\psi_{\beta,\infty}^{-p}\|T_\varepsilon - f\|_\infty^p\right)\left(\psi_{\beta,\infty}/\psi_{\beta,1}\right)^p\\ &\quad\geq \left(\psi_{\beta',\infty}/\psi_{\beta',1}\right)^p \max\Big\{E_0\left[\left(a\psi_{\beta_\varepsilon^*,\infty}^{-1}\|T_\varepsilon\|_\infty\right)^p\right],\\ &\qquad\qquad \max_{1\leq i\leq M} E_i\left[\psi_{\beta',\infty}^{-p}\|T_\varepsilon - f_{i\beta'}\|_\infty^p\right]\Big\}\\ &\quad\geq \left(\psi_{\beta',\infty}/\psi_{\beta',1}\right)^p \max\left\{E_0\left[\left(q'd\left(\hat{\vartheta},\vartheta_0\right)\right)^p\right], \max_{1\leq i\leq M} E_i\left[d^p\left(\hat{\vartheta},\vartheta_i\right)\right]\right\},\end{aligned} \tag{7.19}$$

where

$$a = \left(\psi_{\beta',1}/\psi_{\beta',\infty}\right)\left(\psi_{\beta_\varepsilon^*,\infty}/\psi_{\beta_\varepsilon^*,1}\right), \qquad q' = aq,$$

and q' is defined in the proof of (3.6). Put

$$\beta' = \max\{\beta \in B: \beta \leq \delta\beta_\varepsilon^*/2\}.$$

Condition (3.15) ensures that $\beta' \to \infty$, as $\varepsilon \to 0$. Let us first show that

$$\liminf_{\varepsilon \to 0} \max\left\{E_0\left[\left(q'd\left(\hat{\vartheta}, \vartheta_0\right)\right)^p\right], \max_{1 \leq i \leq M} E_i\left[d^p\left(\hat{\vartheta}, \vartheta_i\right)\right]\right\} \geq 1. \tag{7.20}$$

To show (7.20) it suffices to check the relations [cf. the proof of (3.6)]

$$\frac{\delta}{4}\left[(p+2)\frac{1}{\beta'}\right]^{1/2}\left(\log\frac{1}{\varepsilon}\right)^{1/2} \to \infty, \tag{7.21}$$

$$\tau = \varepsilon^\gamma = o(1) \quad \text{where } \gamma = [p - \delta(p+2)/2]/2\beta', \tag{7.22}$$

$$\tau(q')^p \to \infty \tag{7.23}$$

as $\varepsilon \to 0$. Note that (7.21) and (7.22) are straightforward in view of the definition of β' and of the relation $(1/\beta_\varepsilon^*)\log(1/\varepsilon) \to \infty$. In view of (7.7) we get $q \geq \psi_{\beta',\infty}/\psi_{\beta_\varepsilon^*,\infty}$. Also, $a = ((p+2)/2)^{(1/\beta' - 1/\beta_\varepsilon^*)/4} \geq 1$, and $\tau(q')^p = \varepsilon^\gamma(aq)^p \geq \varepsilon^\gamma a^p(\psi_{\beta',\infty}/\psi_{\beta_\varepsilon^*,\infty})^p$. Thus, to show (7.23) it suffices to prove

$$\varepsilon^\gamma\left(\psi_{\beta',\infty}/\psi_{\beta_\varepsilon^*,\infty}\right)^p \to \infty.$$

Using Lemma 1, it is easy to show that

$$\liminf_{\beta \to \infty} c(\beta, \infty)\beta^{1/2} > 0, \qquad \limsup_{\beta \to \infty} c(\beta, \infty)\beta^{1/2} < \infty.$$

As $\beta' \to \infty$ and $\beta_\varepsilon^* \to \infty$, we conclude that, for ε small enough,

$$\left[c(\beta', \infty)/c(\beta_\varepsilon^*, \infty)\right]^p \geq c',$$

where $c' > 0$ is an absolute constant. It follows that for ε small enough

$$\begin{aligned}
\varepsilon^\gamma\left(\frac{\psi_{\beta',\infty}}{\psi_{\beta_\varepsilon^*,\infty}}\right)^p &= \varepsilon^\gamma\left[\frac{c(\beta', \infty)}{c({\beta^*}_\varepsilon, \infty)}\right]^p\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{p(1/\beta_\varepsilon^* - 1/\beta')/4} \\
&\geq c'\varepsilon^{-1/\beta_\varepsilon^*}\left(\log\frac{1}{\varepsilon}\right)^{-p/2} \\
&= c' \exp\left(\frac{1}{\beta_\varepsilon^*}\log\frac{1}{\varepsilon} - \frac{p}{2}\log\log\frac{1}{\varepsilon}\right) \to \infty,
\end{aligned}$$

as $\varepsilon \to 0$. Here we used the inequalities $1/2 < \beta' \leq \delta\beta_\varepsilon^*/2$. Thus, (7.23) follows, and consequently (7.20) is satisfied. Finally, as $\beta' \to \infty$, we get $(\psi_{\beta',\infty}/\psi_{\beta',1})^p = ((p+2)/2)^{p(1/2 - 1/4\beta')} \to ((p+2)/2)^{p/2}$, as $\varepsilon \to 0$. This, together with (7.19) and (7.20), gives (3.17). □

PROOF OF THEOREM 4. Since 1° of Definition 3 with the rate

$$\Psi_\beta(\varepsilon) = \begin{cases} \psi_{\beta,0}, & \text{if } \beta \in B_-, \\ \varepsilon^{(2\beta_\varepsilon^* - 1)/2\beta_\varepsilon^*} & \text{if } \beta = \beta_\varepsilon^* \end{cases}$$

follows from (4.1) and (4.2), we need only to prove 2° of Definition 3 with this $\Psi_\beta(\varepsilon)$.

Let $S_\beta(\varepsilon)$ be such that for some estimator f_ε^{**} we have (2.9) and (2.10). First, note that β' in (2.10) must satisfy $\beta' \neq \beta_\varepsilon^*$ (or, equivalently, $\beta \in B_-$), since $\Psi_{\beta_\varepsilon^*}(\varepsilon)$ coincides with the ORC and hence cannot be improved. Let $\beta^* = \beta_\varepsilon^*$ for brevity. Then, in view of (2.9),

$$\limsup_{\varepsilon\to 0} \max\Big\{ \sup_{f\in W_{\beta'}} E_f\big(S_{\beta'}^{-p}(\varepsilon) d_0^p(f_\varepsilon^{**}, f)\big), \tag{7.24}$$
$$\sup_{f\in W_{\beta^*}} E_f\big(S_{\beta^*}^{-p}(\varepsilon) d_0^p(f_\varepsilon^{**}, f)\big)\Big\} \le C$$

and, since $\beta' \neq \beta^*$, (2.10) reads as

$$\frac{S_{\beta'}(\varepsilon)}{c(\beta', 0)}\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{(2\beta'-1)/4\beta'} \to 0 \quad \text{as } \varepsilon \to 0. \tag{7.25}$$

Note first that using Theorem 6(ii) from the Appendix, with $M = 2$, it is straightforward to prove the lower bound for fixed $\beta = \beta'$,

$$\liminf_{\varepsilon\to 0} \inf_{T_\varepsilon} \sup_{f\in W_{\beta'}} E_f\big(\varepsilon^{-p(2\beta'-1)/2\beta'} d_0^p(T_\varepsilon, f)\big) > 0$$

[in fact, $\varepsilon^{(2\beta'-1)/2\beta'}$ is the optimal pointwise rate of convergence on $W_{\beta'}$, cf. Donoho and Low (1992)]. This and (7.24) entail that there exists a constant $\bar{c} > 0$ such that

$$S_{\beta'}(\varepsilon) \ge \bar{c}\varepsilon^{(2\beta'-1)/2\beta'} \tag{7.26}$$

for ε small enough. Set

$$r = 1 - \delta/9\beta'$$

where $0 < \delta < 1$. There are two possibilities:

(i) $$\liminf_{\varepsilon\to 0} S_{\beta^*}(\varepsilon)/\varepsilon^r = \infty$$

or

(ii) $$\exists\, \overline{C} > 0 \quad \text{such that } \liminf_{\varepsilon\to 0} S_{\beta^*}(\varepsilon)/\varepsilon^r \le \overline{C},$$

If (i) holds, then in view of (7.26),

$$\frac{S_{\beta'}(\varepsilon)}{\Psi_{\beta'}(\varepsilon)}\frac{S_{\beta^*}(\varepsilon)}{\Psi_{\beta^*}(\varepsilon)}$$
$$\ge \frac{\bar{c}}{c(\beta', 0)}\frac{S_{\beta^*}(\varepsilon)}{\varepsilon^r}\left(\log\frac{1}{\varepsilon}\right)^{(1-2\beta')/4\beta'} \varepsilon^{-\delta/9\beta' + 1/2\beta^*} \to \infty,$$

as $\varepsilon \to 0$, which yields (2.11), with $\beta'' = \beta^*$, and thus the theorem follows. To end the proof we show that (ii) is not possible if (2.9) and (2.10) hold. We will come to a contradiction with (7.24). In fact, if (ii) is true, there is a sequence $\varepsilon_n \to 0$ such that

$$\limsup_{n\to\infty} S_{\beta^*_{\varepsilon_n}}(\varepsilon_n)/\varepsilon_n^r \le \overline{C}. \tag{7.27}$$

It follows from (7.24) and (7.27) that

$$\begin{aligned} C &\ge \limsup_{n\to\infty} \inf_{T_{\varepsilon_n}} \max\Big\{ a_n \sup_{f\in W_{\beta'}} E_f\big(\psi^{-p}_{\beta',0}(\varepsilon_n) d_0^p(T_{\varepsilon_n}, f)\big), \\ &\qquad \big(\varepsilon_n^{r'-r}/\overline{C}\big)^p \sup_{f\in W_{\beta^*_{\varepsilon_n}}} E_f\big(\varepsilon_n^{-r'p} d_0^p(T_{\varepsilon_n}, f)\big)\Big\} \\ &\ge \limsup_{n\to\infty} \min\Big\{a_n, \big(\varepsilon_n^{r'-r}/\overline{C}\big)^p\Big\} \inf_{T_{\varepsilon_n}} \max\Big\{ \sup_{f\in W_{\beta'}} E_f\big(\psi^{-p}_{\beta',0}(\varepsilon_n) d_0^p(T_{\varepsilon_n}, f)\big), \\ &\qquad \sup_{f\in W_{\beta^*_{\varepsilon_n}}} E_f\big(\varepsilon_n^{-r'p} d_0^p(T_{\varepsilon_n}, f)\big)\Big\}, \end{aligned} \tag{7.28}$$

where $a_n = \psi^p_{\beta',0}(\varepsilon_n)/S^{-p}_{\beta'}(\varepsilon_n)$, $\psi_{\beta',0}(\varepsilon_n) = c(\beta', 0)(\varepsilon_n^2 \log(1/\varepsilon_n))^{(2\beta'-1)/4\beta'}$, and $r' = 1 - \delta/8\beta'$. Clearly, $r' - r < 0$. This and (7.25) yield

$$\min\Big\{a_n, \big(\varepsilon_n^{r'-r}/\overline{C}\big)^p\Big\} \to \infty \quad \text{as } n \to \infty. \tag{7.29}$$

Acting as in (7.15) and using the same notation as there, with $\beta'' = \beta^*_{\varepsilon_n}$, we find

$$\begin{aligned} &\max\Big\{ \sup_{f\in W_{\beta'}} E_f\big(\psi^{-p}_{\beta',0}(\varepsilon_n) d_0^p(T_{\varepsilon_n}, f)\big), \sup_{f\in W_{\beta^*_{\varepsilon_n}}} E_f\big(\varepsilon_n^{-r'p} d_0^p(T_{\varepsilon_n}, f)\big)\Big\} \\ &\qquad \ge \left(1 - \frac{\beta'}{\beta^*_{\varepsilon_n}}\right)^{p/2} \max\Big\{E_0\Big[\big(\bar q d(\hat\vartheta, \vartheta_0)\big)^p\Big], E_1\big[d^p(\hat\vartheta, \vartheta_1)\big]\Big\}, \end{aligned} \tag{7.30}$$

where $\bar q = \psi_{\beta',0}(\varepsilon_n)/\varepsilon_n^{r'}$. The right-hand side of (7.30) can be bounded from below as in the proof of (4.3). The modification is that we replace ε there by ε_n, q by $\bar q$ and put $\gamma' = p(1 - \delta/2)(1/2\beta' - 1/2\beta^*_{\varepsilon_n})$. Then

$$\tau \bar q^p = \varepsilon_n^{\gamma'} \bar q^p \ge (c(\beta', 0))^p \left(\log\frac{1}{\varepsilon_n}\right)^{p(2\beta'-1)/4\beta'} \varepsilon_n^{-\delta p/8\beta'} \to \infty \quad \text{as } n \to \infty.$$

Consequently, as in the proof of (4.3), we find

$$\liminf_{n\to\infty} \max\Big\{E_0\Big[\big(\bar q d(\hat\vartheta, \vartheta_0)\big)^p\Big], E_1\big[d^p(\hat\vartheta, \vartheta_1)\big]\Big\} > 0.$$

This inequality together with (7.28)–(7.30) yields a contradiction and proves the theorem. □

APPENDIX

A1. A theorem on lower bounds. Let $w\colon [0,\infty) \to [0,\infty)$ be a monotone nondecreasing function. Let $(\Theta, \mathscr{S})$ be a measurable space of parameters equipped with a pseudo-distance $d(\cdot,\cdot)$ (i.e., $d(\cdot,\cdot)$ satisfies the definition of a distance, except, perhaps, the condition $d(\vartheta, \vartheta') = 0 \Rightarrow \vartheta = \vartheta'$). For an integer $M \geq 1$, consider $M+1$ elements of Θ: $\vartheta_0, \vartheta_1, \dots, \vartheta_M$, and a family of probability measures $\{P_\vartheta, \vartheta \in \Theta\}$ on a measurable space $(\mathscr{X}, \mathscr{A})$. Denote for brevity $P_k = P_{\vartheta_k}$, $k = 0, 1, \dots, M$, and let E_k denote the expectation w.r.t. P_k.

THEOREM 6. *Let the numbers $q > 0$, $\tau > 0$, $0 < \delta < 1/2$, $0 < \alpha < 1$ be fixed, and let*

$$d(\vartheta_i, \vartheta_k) \geq 1 - \delta, \qquad i, k = 0, 1, \dots, M, \; i \neq k. \tag{A.1}$$

(i) *If, in addition, $P_0 \ll Q$ and*

$$Q\left(\frac{dP_0}{dQ} \geq \tau\right) \geq 1 - \alpha, \tag{A.2}$$

where $Q = M^{-1}\sum_{k=1}^{M} P_k$, then

$$\begin{aligned}&\inf_{\hat\vartheta} \max\left\{E_0\left[w\left(qd\left(\hat\vartheta, \vartheta_0\right)\right)\right], \max_{1 \leq k \leq M} E_k\left[w\left(d\left(\hat\vartheta, \vartheta_k\right)\right)\right]\right\} \\ &\qquad \geq \frac{(1-\alpha)\tau w(1-2\delta) w(q\delta)}{w(1-2\delta) + \tau w(q\delta)},\end{aligned} \tag{A.3}$$

whenever $w(1-2\delta) > 0$, $w(q\delta) > 0$, where $\inf_{\hat\vartheta}$ denotes the infimum over all measurable functions $\hat\vartheta\colon \mathscr{X} \to \{\vartheta_0, \vartheta_1, \dots, \vartheta_M\}$.

(ii) *If, in addition to* (A.1), *$P_0 \ll P_k$, $k = 1, \dots, M$, and*

$$P_k\left(\frac{dP_0}{dP_k} \geq \tau\right) \geq 1 - \alpha, \qquad k = 1, \dots, M, \tag{A.4}$$

then

$$\inf_{\hat\vartheta} \max_{0 \leq k \leq M} E_k\left[w\left(d\left(\hat\vartheta, \vartheta_k\right)\right)\right] \geq w\left(\frac{1-\delta}{2}\right)\frac{(1-\alpha)\tau M}{1 + \tau M}. \tag{A.5}$$

PROOF. Prove (A.3) first. By monotonicity of $w(\cdot)$,

$$\begin{aligned}\tilde R &\stackrel{\text{def}}{=} \max\left\{E_0\left[w\left(qd\left(\hat\vartheta, \vartheta_0\right)\right)\right], \max_{1 \leq k \leq M} E_k\left[w\left(d\left(\hat\vartheta, \vartheta_k\right)\right)\right]\right\} \\ &\geq \max\left\{w(q\delta) P_0\left\{d\left(\hat\vartheta, \vartheta_0\right) \geq \delta\right\}, w(1-2\delta) \max_{1 \leq k \leq M} P_k\left\{d\left(\hat\vartheta, \vartheta_k\right) \geq 1 - 2\delta\right\}\right\},\end{aligned}$$

In view of (A.1), if $d(\hat\vartheta, \vartheta_0) < \delta$, then $d(\hat\vartheta, \vartheta_k) \geq 1 - 2\delta$, $k = 1, \dots, M$. Therefore,

$$\begin{aligned}\tilde R &\geq \max\left\{w(q\delta) P_0\{\Omega\}, w(1-2\delta) \max_{1 \leq k \leq M} P_k\{\overline{\Omega}\}\right\} \\ &\geq \max\left\{w(q\delta) P_0\{\Omega\}, w(1-2\delta) Q\{\overline{\Omega}\}\right\},\end{aligned} \tag{A.6}$$

where Ω is the random event $\Omega = \{d(\hat\vartheta, \vartheta_0) \ge \delta\}$ and $\overline{\Omega}$ denotes the complement of Ω. Introduce the random event $A = \{dP_0/dQ \ge \tau\}$. Then, by (A.2),

$$P_0(\Omega) = \int \frac{dP_0}{dQ} I(\Omega)\, dQ \ge \tau Q(\Omega \cap A) \ge \tau(Q(\Omega) - \alpha). \tag{A.7}$$

Substitution of (A.7) into (A.6) gives

$$\begin{aligned}\tilde R &\ge \max\{w(q\delta)\tau(Q(\Omega) - \alpha), w(1-2\delta)(1 - Q(\Omega))\}\\ &\ge \min_{0\le t\le 1} \max\{w(q\delta)\tau(t-\alpha), w(1-2\delta)(1-t)\}\\ &= \frac{(1-\alpha)\tau w(1-2\delta)w(q\delta)}{w(1-2\delta) + \tau w(q\delta)}.\end{aligned}$$

This yields (A.3),

Now, prove (A.5). By monotonicity of $w(\cdot)$,

$$\max_{0\le k\le M} E_k\Big[w\Big(d\big(\hat\vartheta, \vartheta_k\big)\Big)\Big] \ge w\left(\frac{1-\delta}{2}\right)\max\Big\{P_0(\Omega_0),\ \max_{1\le k\le M} P_k(\Omega_k)\Big\}, \tag{A.8}$$

where $\Omega_k = \{d(\hat\vartheta, \vartheta_k) \ge (1-\delta)/2\}$, $k = 0, 1, \dots, M$, are random events. Note that $\overline{\Omega}_i \cap \overline{\Omega}_k = \varnothing$, $i \ne k$, in view of (A.1). Hence, denoting $A_k = \{dP_0/dP_k \ge \tau\}$ and using (A.4), one gets

$$\begin{aligned}P_0(\Omega_0) &\ge P_0\left\{\bigcup_{k=1}^M \overline{\Omega}_k\right\} = \sum_{k=1}^M P_0\big(\overline{\Omega}_k\big)\\ &= \sum_{k=1}^M \int \frac{dP_0}{dP_k} I\big\{\overline{\Omega}_k\big\}\, dP_k \ge \tau \sum_{k=1}^M P_k\big(\overline{\Omega}_k \cap A_k\big)\\ &\ge \tau\left(\sum_{k=1}^M P_k\big(\overline{\Omega}_k\big) - M\alpha\right) \ge \tau M\Big(1 - \max_{1\le k\le M} P_k(\Omega_k) - \alpha\Big).\end{aligned}$$

Together with (A.8) this yields

$$\begin{aligned}\max_{0\le k\le M} E_k\Big[w\Big(d\big(\hat\vartheta, \vartheta_k\big)\Big)\Big] &\ge w\left(\frac{1-\delta}{2}\right)\min_{0\le t\le 1}\max\{\tau M(1-t-\alpha), t\}\\ &= w\left(\frac{1-\delta}{2}\right)\frac{(1-\alpha)\tau M}{1+\tau M}.\end{aligned}$$

□

A2. Proofs of Lemmas 5–10.

PROPOSITION 1. *Let* $g(x) = x^{q_1}(1 + x^{q_2})^{-2}$, *where* $q_2 \ge q_1 > 0$, $q_2 > 1$. *Then, for every* $h > 0$,

$$\sum_{m=1}^\infty g(mh)h \le \int_0^\infty g(x)\, dx + 2h.$$

The proof of Proposition 1 is easy, using the fact that $g(x)$ has two monotone pieces and $\max_{x\geq 0} g(x) \leq 1$.

PROOF OF LEMMA 5. Let us prove that for any $h > 0$ the estimator (3.2) satisfies

$$\text{(A.9)}\qquad \begin{aligned} \sup_{f\in W_\beta} \left\| E_f(f_{\varepsilon,\beta'}) - f\right\|_\infty &\leq Lh^{\tilde{\beta}-1/2}(C + 2h^{1/2}) \\ &\quad + \left(Q_{\max}/(\beta_1 - 1/2)\right)^{1/2}\varepsilon^2, \end{aligned}$$

where $C = b_\beta$ if $\beta' = \beta$, and $C = [\pi^{-1}(1 + (2\beta_1 - 1)^{-1})]^{1/2}$ otherwise. This entails (5.13) and (5.14). In fact, $\tilde{\beta}(\beta,\beta) = \beta$, and $\beta' \leq \tilde{\beta}(\beta,\beta') \leq 2\beta'$. Therefore, since $\varepsilon^2 \log(1/\varepsilon) \leq 1$, we have

$$\begin{aligned} h^{\tilde{\beta}}_{\beta',j} &= \kappa^{\tilde{\beta}}_{\beta',j}\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{\tilde{\beta}/2\beta'} \geq \kappa^{\tilde{\beta}}_{\beta',j}\varepsilon^2 \log\frac{1}{\varepsilon} \geq \left(\frac{\alpha_j}{L^2\beta'(2\beta'-1)}\right)^{\tilde{\beta}/2\beta'}\varepsilon^2 \log\frac{1}{\varepsilon} \\ &\geq \min\left\{1, \alpha_j(2L\beta^*_\varepsilon)^{-2}\right\}\varepsilon^2 \log\frac{1}{\varepsilon} \geq c_9\varepsilon^2\nu_\varepsilon. \end{aligned}$$

except for the case $j = 0$, $\beta = \beta' = \tilde{\beta} = \beta^*_\varepsilon$, where the proof of the similar inequality is straightforward. (Here and later in this proof we write for brevity $\tilde{\beta} = \tilde{\beta}(\beta,\beta')$). By Lemma 4, $h^{1/2}_{\beta,j} \leq (h^*_\varepsilon)^{1/2} \leq \max\{1, \kappa^{1/2}_{\max}\}$ $\exp(-(1/4\beta^*_\varepsilon)\log(1/\varepsilon))$. These remarks, together with (A.9) and the fact that $\nu_\varepsilon \to \infty$, yield (5.13) and (5.14).

It remains to prove (A.9). Since $E_f(\hat{\theta}_k) = \theta_k$ and $\|\varphi_k\|_\infty = 2^{1/2}$,

$$\text{(A.10)}\qquad \left\|E_f(f_{\varepsilon,\beta'}) - f\right\|_\infty \leq \Lambda(\beta') + 2^{1/2}\sum_{k=N+1}^{\infty}|\theta_k|,$$

where

$$\Lambda(\beta') = \sup_{t\in[0,1]}\left|\sum_{k=1}^{\infty}\frac{a_k^2(\beta')(\pi h)^{2\beta'}}{1 + a_k^2(\beta')(\pi h)^{2\beta'}}\theta_k\varphi_k(t)\right|.$$

By the Cauchy–Schwarz inequality, and in view of the definition of N,

$$\text{(A.11)}\qquad \begin{aligned} \sum_{k=N+1}^{\infty}|\theta_k| &\leq Q_\beta^{1/2}\left(\sum_{k=N+1}^{\infty}a_k^{-2}(\beta)\right)^{1/2} \leq Q_\beta^{1/2}\left(\frac{N^{1-2\beta_1}}{2\beta_1 - 1}\right)^{1/2} \\ &\leq \left(\frac{Q_{\max}}{\beta_1 - 1/2}\right)^{1/2}\varepsilon^2. \end{aligned}$$

Again, the Cauchy–Schwarz inequality and the fact that $a_k^2(\beta') = a_k(\tilde{\beta})a_k(2\beta' - \tilde{\beta})$, entail

$$\Lambda(\beta') \leq (\pi h)^{\tilde{\beta}-1/2}\left(\sum_{k=0}^{\infty}a_k^2(\tilde{\beta})\theta_k^2\right)^{1/2}\Lambda_1^{1/2}(h),$$

where

$$\Lambda_1(h) = \sup_{t\in[0,1]} \sum_{k=1}^{\infty} s_k \varphi_k^2(t), \qquad s_k = \frac{a_k^2(2\beta' - \tilde{\beta})(\pi h)^{2(2\beta'-\tilde{\beta})}}{\left(1 + a_k^2(\beta')(\pi h)^{2\beta'}\right)^2}\pi h.$$

As $\tilde{\beta} \le \beta$, we find $\sum_{k=0}^{\infty} a_k^2(\tilde{\beta})\theta_k^2 \le Q_\beta$. Thus,

$$\Lambda(\beta') \le Q_\beta^{1/2}(\pi h)^{\tilde{\beta}-1/2}\Lambda_1^{1/2}(h). \tag{A.12}$$

Now, $a_{k-1}(\beta) = a_k(\beta) = k^\beta$, for k even and any $\beta \in B$. Therefore,

$$s_{2m-1} = s_{2m} = \frac{(2\pi hm)^{2(2\beta'-\tilde{\beta})}}{\left(1 + (2\pi hm)^{2\beta'}\right)^2}\pi h, \qquad m = 1, 2, \dots .$$

This and the equality $\varphi_{2m}^2(t) + \varphi_{2m-1}^2(t) = 2$, $\forall\ t \in [0,1]$, $m = 1, 2, \dots,$ yield

$$\sum_{k=1}^{\infty} s_k \varphi_k^2(t) = 2\sum_{m=1}^{\infty} s_{2m} = \sum_{m=1}^{\infty} g(m\bar{h})\bar{h} \qquad \forall\ t \in [0,1],$$

where $\bar{h} = 2\pi h$, and $g(x) = x^{2(2\beta'-\tilde{\beta})}/(1 + x^{2\beta'})^2$. Note that this function g satisfies the conditions of Proposition 1. In fact, $2\beta' \ge 2(2\beta' - \tilde{\beta})$, since $\tilde{\beta} \ge \beta'$, and $2\beta' > 1$ since $\beta' \in B$. Also, $2\beta' - \tilde{\beta} > 0$. The application of Proposition 1 gives

$$\sum_{k=1}^{\infty} s_k \varphi_k^2(t) \le \Lambda_2(\beta', \beta) + 4\pi h,$$

where $\Lambda_2(\beta', \beta) = \int_0^\infty x^{2(2\beta'-\tilde{\beta})}/(1 + x^{2\beta'})^2\,dx$. If $\beta' = \beta$, then $\tilde{\beta} = \beta$, and $\Lambda_2(\beta, \beta) = \pi b_\beta^2$. Hence, in view of (A.12) and the definition of $\Lambda_1(h)$, one gets

$$\begin{aligned}\Lambda(\beta) &\le Q_\beta^{1/2}(\pi h)^{\beta-1/2}\left(\pi b_\beta^2 + 4\pi h\right)^{1/2} = Lh^{\beta-1/2}\left(b_\beta^2 + 4h\right)^{1/2}\\ &\le Lh^{\beta-1/2}\left(b_\beta + 2h^{1/2}\right).\end{aligned}$$

This, together with (A.10), and (A.11), proves the Lemma for the case $\beta' = \beta$.

If $\beta' < \beta$, then $\Lambda_2(\beta', \beta) \le 1 + \int_1^\infty t^{-2\tilde{\beta}}\,dt = 1 + (2\tilde{\beta} - 1)^{-1} \le 1 + (2\beta_1 - 1)^{-1}$, and

$$\begin{aligned}\Lambda(\beta') &\le Q_\beta^{1/2}(\pi h)^{\tilde{\beta}-1/2}\left(1 + (2\beta_1 - 1)^{-1} + 4\pi h\right)^{1/2}\\ &\le Lh^{\tilde{\beta}-1/2}(C + 2h^{1/2}).\end{aligned}$$

□

Proposition 2. *Let $Z(t)$ be a stationary Gaussian random process with mean* 0 *and correlation function $r(t)$, such that $\lambda_2 = -r''(0)$ is finite. Then*

$$P\left(\sup_{t\in[0,1]} |Z(t)| \ge u\right) \le \left(\frac{1}{\pi}(\lambda_2/r(0))^{1/2} + e^{1/2}\right)\exp\left(-\frac{u^2}{2r(0)}\right),$$

for every $u > 0$.

PROOF. The result is a simple corollary of the Rice formula [Leadbetter, Rootzen, Lindgren (1986) Theorem 7.3.2] which states that

$$E(\mathbf{N}_u) = (2\pi)^{-1}\left(\frac{\lambda_2}{r(0)}\right)^{1/2} \exp\left(-\frac{u^2}{2r(0)}\right),$$

where $\mathbf{N}_u$ is the number of upcrossings of the level u by the process $Z(t)$ on $[0,1]$. $\square$

PROOF OF LEMMA 6. The process $Z_{\beta,j}(t)$ is a stationary Gaussian process on $\mathbb{R}$, with mean 0 and correlation function $r(t) = r(t, h_{\beta,j})$, where

$$r(t,h) = \varepsilon^2\left(1 + \sum_{m=1}^{N/2} \frac{2\cos(2\pi mt)}{\left(1+(2\pi mh)^{2\beta}\right)^2}\right) \qquad \forall\, t \in \mathbb{R},\, h > 0.$$

In particular,

$$r(0,h) = \varepsilon^2(1 + \Gamma_1 + \Gamma_2),$$

where

$$\Gamma_1 = \sum_{m=1}^{\infty} \frac{2}{\left(1+(2\pi mh)^{2\beta}\right)^2}$$

and

$$|\Gamma_2| \le \sum_{m=N/2+1}^{\infty} \frac{2}{(2\pi mh)^{4\beta}} \le (\pi h)^{-1}\int_{N/2}^{\infty} t^{-4\beta}\,dt \le \frac{2}{\pi h_{\min,\varepsilon} N} \le 1$$

$$\forall\, h \ge h_{\min,\varepsilon}.$$

Here we used the assumption (5.2). It is easy to see that, for any $h > 0$,

$$h^{-1}v_\beta^2 - 1 \le \Gamma_1 \le h^{-1}v_\beta^2.$$

Hence, $|r(0,h) - \varepsilon^2 h^{-1} v_\beta^2| \le 2\varepsilon^2$, for any $h \ge h_{\min,\varepsilon}$. This entails

$$\text{(A.13)} \qquad \left|\frac{r(0,h_{\beta,j})}{r_{\beta,j}} - 1\right| \le \frac{2\varepsilon^2}{r_{\beta,j}} \le \frac{2h_\varepsilon^*}{v_{\min}^2} \le \delta_{\varepsilon 0}, \qquad j = 0, 1, \infty,$$

by virtue of (5.11) and of the definition (5.5) of $\delta_{\varepsilon 0}$. Simple calculations yield the following estimate of the second derivative of the correlation function

$$\text{(A.14)} \qquad \lambda_2 = |r''(0)| \le c_{10}\varepsilon^2 \sum_{m=1}^{\infty} \frac{(2\pi m)^2}{\left(1+(2\pi m h_{\beta,j})^{2\beta}\right)^2} \le c_{11}\varepsilon^2 h_{\beta,j}^{-3}.$$

The first inequality of the lemma follows immediately from Proposition 2 and (A.13), (A.14). The second inequality follows from the first one. In fact, it suffices to use the relation

$$E\left(\tilde{Z}_{\beta,j}^p I\left\{\tilde{Z}_{\beta,j} \ge u\right\}\right) = p\int_u^{\infty} t^{p-1} P\left(\tilde{Z}_{\beta,j} \ge t\right) dt,$$

and to note that, for any $a > 0$, $p \ge 0$,

$$\int_a^\infty t^p \exp(-t^2/2)\,dt \le C(p)\exp(-a^2/2), \tag{A.15}$$

where $C(p) > 0$ depends only on p. □

PROOF OF LEMMA 7. Note that $Z_{\beta,0}(x_0)$ is a normal random variable with mean 0 and variance $r(0, h_{\beta,0})$. It follows from (A.13) that

$$r(0, h_{\beta,0}) \le r_{\beta,0}(1 + \delta_{\varepsilon 0}) \le r_{\beta,0}/(1 - \delta_{\varepsilon 0}).$$

Now, to prove the lemma it suffices to use (A.15). □

PROOF OF LEMMA 8. By definition of $\hat{\beta}_j$,

$$\sup_{f \in W_\beta} P_f\big(\hat{\beta}_j = \gamma\big) \le \sum_{\substack{\beta' \in B,\\ \beta' \le \gamma}} \sup_{f \in W_\beta} p_j(\beta', \gamma) \le \operatorname{card}(B) \max_{\substack{\beta' \in B\\ \beta' \le \gamma}} \sup_{f \in W_\beta} p_j(\beta', \gamma),$$

where

$$p_j(\beta', \gamma) = P_f\big(d_j(f_{\varepsilon,\gamma',j}, f_{\varepsilon,\beta',j}) > \eta_j(\beta')\big), \qquad j = 0, \infty,$$

$\gamma' = \gamma'(\gamma)$ is the element of B: $\gamma' = \gamma'(\gamma) = \min\{\beta \in B\colon \beta > \gamma\}$ closest from above to γ and $\operatorname{card}(B) \le \beta_\varepsilon^*/\Delta_\varepsilon + 1$. Hence, to prove the Lemma it is sufficient to show the relation

$$\sup_{f \in W_\beta} p_j(\beta', \gamma) \le c_{12}\,\varepsilon^{p/2\gamma}, \qquad j = 0, \infty. \tag{A.16}$$

First, prove the following auxiliary result.

PROPOSITION 3. *Let* $\gamma, \beta, \beta' \in B$, $\beta' \le \gamma < \beta$, $\tilde{\beta} = \tilde{\beta}(\beta, \beta')$, $\gamma' = \gamma'(\gamma)$ *and* $\tilde{\gamma} = \tilde{\beta}(\beta, \gamma')$. *Then*

$$h_{\beta',j}^{\tilde{\beta}-1/2}/\eta_j(\beta') \le \delta_{\varepsilon 1}, \qquad h_{\gamma',j}^{\tilde{\gamma}-1/2}/\eta_j(\beta') \le \delta_{\varepsilon 1}, \qquad j = 0, \infty,$$

where $\delta_{\varepsilon 1} = c_{13}\exp(-\tilde{\nu}_\varepsilon/8)$.

PROOF OF PROPOSITION 3. Since $\tilde{\beta} \ge \beta'$ and since, by Lemma 4, $h_{\beta,j} \le h_{\beta_\varepsilon^*,j}$ $\forall \beta \in B$, $j = 0, \infty$, one obtains

$$h_{\beta',j}^{\tilde{\beta}-1/2} \le \big(h_{\beta_\varepsilon^*,j}\big)^{\tilde{\beta}-\beta'} h_{\beta',j}^{\beta'-1/2}.$$

Here, by (5.11) and Lemma 1,

$$\begin{aligned}
h_{\beta_\varepsilon^*,j}^{\tilde{\beta}-\beta'} &\le \max\Big\{1, \kappa_{\beta_\varepsilon^*,j}^{\tilde{\beta}-\beta'}\Big\}\exp\bigg(-\frac{1}{2\beta_\varepsilon^*}(\tilde{\beta}-\beta')\log\frac{1}{\varepsilon}\bigg)\\
&\le \max\Big\{1, \kappa_{\beta_\varepsilon^*,j}^{\beta_\varepsilon^*-1/2}\Big\}\exp\bigg(-\frac{1}{2\beta_\varepsilon^*}(\tilde{\beta}-\beta')\log\frac{1}{\varepsilon}\bigg)\\
&\le \max\{1, \kappa_{\max}\}\exp\bigg(-\frac{1}{2\beta_\varepsilon^*}(\tilde{\beta}-\beta')\log\frac{1}{\varepsilon}\bigg).
\end{aligned}$$

Hence,

$$h_{\beta',j}^{\tilde{\beta}-1/2} \le \max\{1, \kappa_{\max}\}\exp\left(-\frac{1}{2\beta_\varepsilon^*}(\tilde{\beta}-\beta')\log\frac{1}{\varepsilon}\right)h_{\beta',j}^{\beta'-1/2}.$$

Now, $\beta \ge \beta' + \Delta$ since B is a discrete set with the step at least Δ, and $\beta' < \beta$. This, the inequality $\Delta < 1$ and the definition of $\tilde{\beta}$ yield

$$\tilde{\beta} \ge \beta' + \min\{(\beta_1 + 1/2)/2, \Delta\} \ge \beta' + \min\{1/2, \Delta\} \ge \beta' + \Delta/2.$$

Hence,

$$\frac{1}{2\beta_\varepsilon^*}\log\frac{1}{\varepsilon}(\tilde{\beta}-\beta') \ge \frac{\Delta}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon} \ge \frac{\Delta}{8(\beta_\varepsilon^*)^2}\log\frac{1}{\varepsilon} = \tilde{v}_\varepsilon/8,$$

and

$$h_{\beta',j}^{\tilde{\beta}-1/2} \le \max\{1, \kappa_{\max}\}\exp(-\tilde{\nu}_\varepsilon/8)\,h_{\beta',j}^{\beta'-1/2}. \tag{A.17}$$

Similarly, considering in turn the cases $\gamma' < \beta$ and $\gamma' = \beta (\Rightarrow \tilde{\gamma} = \gamma' = \beta)$ we find

$$h_{\gamma',j}^{\tilde{\gamma}-1/2} \le \max\{1, \kappa_{\max}\exp(-\tilde{\nu}_\varepsilon/8)\}h_{\gamma',j}^{\gamma'-1/2}. \tag{A.18}$$

It follows from (A.17) and (5.9) that

$$\begin{aligned} h_{\beta',j}^{\tilde{\beta}-1/2}/\eta_j(\beta') &\le \max\{1, \kappa_{\max}\}\exp(-\tilde{\nu}_\varepsilon/8)/\left[L(2\beta'-1)^{1/2}v_{\beta'}\right] \\ &\le \left(\max\{1, \kappa_{\max}\}/\left[L(2\beta_1-1)^{1/2}v_{\min}\right]\right)\exp(-\tilde{\nu}_\varepsilon/8). \end{aligned}$$

This shows the first inequality of Proposition 3.

Next, in view of (5.10) and of the fact that $\beta' < \gamma'$,

$$\begin{aligned} \frac{\eta_j(\gamma')}{\eta_j(\beta')} &= \left(\frac{\beta'}{\gamma'}\frac{v_{\gamma'}^2 h_{\beta',j}}{v_{\beta'}^2 h_{\gamma',j}}\right)^{1/2} \le \frac{v_{\max}}{v_{\min}}\left(\frac{h_{\beta',j}}{h_{\gamma',j}}\right)^{1/2} \\ &\le \frac{v_{\max}}{v_{\min}}c_1^{1/2}\exp\left(-\frac{\tilde{\nu}_\varepsilon}{4}\right). \end{aligned} \tag{A.19}$$

Finally, (5.9), (A.18) and (A.19) entail [excluding the special case with $j = 0$ and $\tilde{\gamma} = \gamma' = \beta_\varepsilon^*$ where (5.9) does not apply]:

$$\begin{aligned} h_{\gamma',j}^{\tilde{\gamma}-1/2}/\eta_j(\beta') &\le \left(\left[1 + \kappa_{\max}\exp\left(-\frac{\tilde{\nu}_\varepsilon}{8}\right)\right]/\left[L(2\gamma'-1)^{1/2}v_{\gamma'}\right]\right) \\ &\quad\times \frac{\eta_j(\gamma')}{\eta_j(\beta')} \\ &\le \left(\frac{1+\kappa_{\max}}{L(2\beta_1-1)^{1/2}v_{\min}}\right)\frac{v_{\max}}{v_{\min}}c_1^{1/2}\exp\left(-\frac{\tilde{\nu}_\varepsilon}{4}\right), \end{aligned}$$

which shows the second inequality of Proposition 3 (it is easy to see that in the case $j = 0$, $\tilde{\gamma} = \gamma' = \beta_\varepsilon^*$ the last inequalities also hold, but with different constants). □

PROOF OF (A.16). Recall that $d_\infty(f,g)=\|f-g\|_\infty$ and $d_0(f,g)=|f(x_0)-g(x_0)|$. In view of Lemma 5 and Proposition 3,

$$
\begin{aligned}
d_j(f_{\varepsilon,\gamma',j}, f_{\varepsilon,\beta',j}) &\le \left\|E_f(f_{\varepsilon,\gamma',j})-f\right\|_\infty + \left\|E_f(f_{\varepsilon,\beta',j})-f\right\|_\infty \\
&\quad + \tilde{Z}_{\beta',j} + \tilde{Z}_{\gamma',j} \\
&\le c_2\left(h_{\beta',j}^{\tilde{\beta}-1/2} + h_{\gamma',j}^{\tilde{\gamma}-1/2}\right) + \tilde{Z}_{\beta',j} + \tilde{Z}_{\gamma',j} \\
&\le 2c_2\delta_{\varepsilon 1}\eta_j(\beta') + \tilde{Z}_{\beta',j} + \tilde{Z}_{\gamma',j}, \qquad j=0,\infty.
\end{aligned}
\tag{A.20}
$$

Hence,

$$
p_j(\beta',\gamma) \le p_{1,j} + p_{2,j}, \qquad j=0,\infty, \tag{A.21}
$$

where

$$
p_{1,j} = P\left\{\tilde{Z}_{\beta',j} \ge \eta_j(\beta')(1-3c_2\delta_{\varepsilon 1})\right\}, \qquad p_{2,j} = P\left\{\tilde{Z}_{\gamma',j} \ge c_2\delta_{\varepsilon 1}\eta_j(\beta')\right\}.
$$

Let us evaluate separately $p_{1,j}$ and $p_{2,j}$.

Evaluation of $p_{1,0}$ *and* $p_{1,\infty}$. By Lemma 6

$$
p_{1,\infty} \le c_4 h_{\beta',\infty}^{-1} \exp\left(-\frac{1}{2r_{\beta',\infty}}\eta_\infty^2(\beta')(1-3c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0})\right), \tag{A.22}
$$

and by Lemma 7

$$
p_{1,0} \le c_5 \exp\left(-\frac{1}{2r_{\beta',0}}\eta_0^2(\beta')(1-3c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0})\right). \tag{A.23}
$$

Here

$$
(1-3c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0}) \ge 1-6c_2\delta_{\varepsilon 1}-\delta_{\varepsilon 0} \ge 1-c_{14}\exp(-\tilde{\nu}_\varepsilon/8),
$$

and

$$
\begin{aligned}
&\frac{1}{2r_{\beta',j}}\eta_j^2(\beta')(1-3c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0}) \\
&\quad \ge \frac{1}{2\beta'}\alpha_j \log\frac{1}{\varepsilon} - \frac{1}{2\beta'}\alpha_j c_{14}\log\frac{1}{\varepsilon}\exp(-\tilde{\nu}_\varepsilon/8) \\
&\quad \ge \frac{1}{2\beta'}\alpha_j \log\frac{1}{\varepsilon} - c_{15}\delta_{\varepsilon 2}, \qquad j=0,\infty \quad \text{where } \delta_{\varepsilon 2} = \log\frac{1}{\varepsilon}\exp\left(-\frac{\tilde{\nu}_\varepsilon}{8}\right).
\end{aligned}
$$

Substitute this expression with $j=\infty$ into (A.22) and note that $\alpha_\infty = p+2$. This gives

$$
\begin{aligned}
p_{1,\infty} &\le c_4 h_{\beta',\infty}^{-1}\varepsilon^{1/\beta'}\varepsilon^{p/2\beta'}\exp(c_{15}\delta_{\varepsilon 2}) \le c_4\kappa_{\min}^{-1}\exp(c_{15}\delta_{\varepsilon 2})\varepsilon^{p/2\beta'} \\
&\le c_4\kappa_{\min}^{-1}\exp(c_{15}\delta_{\varepsilon 2})\varepsilon^{p/2\gamma} \le c_{16}\varepsilon^{p/2\gamma},
\end{aligned}
\tag{A.24}
$$

where the last inequality is due to the relation $\delta_{\varepsilon 2} = o(1)$, as $\varepsilon \to 0$, that follows directly from (3.1).

Similarly, if $j = 0$ ($\alpha_0 = p$), then

$$p_{1,0} \le c_5 \varepsilon^{p/2\beta'} \exp(c_{15}\delta_{\varepsilon 2}) \le c_{17}\varepsilon^{p/2\gamma}. \tag{A.25}$$

Evaluation of $p_{2,0}$ *and* $p_{2,\infty}$. By Lemma 6 and Lemma 7

$$\begin{aligned} p_{2,\infty} &\le c_4 h_{\gamma',\infty}^{-1} \exp\left(-\frac{1}{2r_{\gamma',\infty}}\eta_\infty^2(\beta')(c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0})\right),\\ p_{2,0} &\le c_5 \exp\left(-\frac{1}{2r_{\gamma',0}}\eta_0^2(\beta')(c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0})\right). \end{aligned} \tag{A.26}$$

Using the inequality $\delta_{\varepsilon 0} < 1/2$, the definition of $\delta_{\varepsilon 1}$, (5.10) and the inequality $\log\log(1/\varepsilon) \ge 1$, one gets

$$\begin{aligned} &\frac{1}{2r_{\gamma',j}}\eta_j^2(\beta')(c_2\delta_{\varepsilon 1})^2(1-\delta_{\varepsilon 0})\\ &\qquad \ge (c_2 c_{13}/2)^2 \frac{1}{\beta'}\alpha_j \log\frac{1}{\varepsilon}\left(\frac{v_{\beta'}^2 h_{\gamma',j}}{v_{\gamma'}^2 h_{\beta',j}}\right)\exp\left(-\frac{\tilde{\nu}_\varepsilon}{4}\right)\\ &\qquad \ge c_1^{-1}\left(\frac{c_2 c_{13}}{2}\right)^2\left(\frac{v_{\min}}{v_{\max}}\right)^2 \frac{1}{\beta_\varepsilon^*}\alpha_j \log\frac{1}{\varepsilon}\exp\left(\frac{\tilde{\nu}_\varepsilon}{4}\right) \ge A_\varepsilon \log\frac{1}{\varepsilon}, \end{aligned}$$

where $A_\varepsilon = c_{18}(\beta_\varepsilon^*)^{-1}\exp(\tilde{\nu}_\varepsilon/4)$. Substitution of this result into (A.26) gives

$$p_{2,\infty} \le c_4 h_{\gamma',\infty}^{-1}\exp\left(-A_\varepsilon \log\frac{1}{\varepsilon}\right) \le c_4 \kappa_{\min}^{-1}\varepsilon^{-1/\gamma'}\exp\left(-A_\varepsilon\log\frac{1}{\varepsilon}\right)$$

and

$$p_{2,0} \le c_5 \exp\left(-A_\varepsilon \log\frac{1}{\varepsilon}\right).$$

Note that $A_\varepsilon \to \infty$ as $\varepsilon \to 0$, since $\tilde{\nu}_\varepsilon/\log\beta_\varepsilon^* \ge \nu_\varepsilon$. Therefore,

$$p_{2,j} \le c_{19}\varepsilon^{p/2\gamma}, \qquad j = 0, \infty.$$

This, together with (A.21), (A.24) and (A.25), proves (A.16).

PROOF OF LEMMA 9. By use of (5.14) one has

$$\|f_{\varepsilon,\beta,\infty} - f\|_\infty \le b_\infty(\beta)\left(1 + c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right) + \tilde{Z}_{\beta,\infty}. \tag{A.27}$$

Choose $\varepsilon_0 = \varepsilon_0(\delta)$ so that

$$\max\left\{\delta_{\varepsilon 0},\, c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right\} < \delta$$

for all $\varepsilon \in (0, \varepsilon_0)$. In view of (5.8),

$$\psi_{\beta,\infty} u - b_\infty(\beta)\left(1 + c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right) \ge u\eta_\infty(\beta),$$

if $u \geq 1 + \delta$ and $\varepsilon \in (0, \varepsilon_0)$. Using this inequality, (A.27) and Lemma 6, one gets, for $u \geq 1 + \delta$ and $\varepsilon \in (0, \varepsilon_0)$,

$$
\begin{aligned}
&P_f\left\{\psi_{\beta,\infty}^{-1}\|f_{\varepsilon,\beta,\infty} - f\|_\infty \geq u\right\}\\
&\qquad \leq P_f\left\{b_\infty(\beta)\left(1 + c_3\exp\left(-\frac{1}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)\right) + \tilde{Z}_{\beta,\infty} \geq \psi_{\beta,\infty}u\right\}\\
&\qquad \leq P_f\left\{\tilde{Z}_{\beta,\infty} \geq u\eta_\infty(\beta)\right\} \leq c_4 h_{\beta,\infty}^{-1}\exp\left(-\frac{u^2}{2}(p+2)\frac{1}{\beta}\log\frac{1}{\varepsilon}(1-\delta_{\varepsilon 0})\right)\\
&\qquad \leq c_4\kappa_{\min}^{-1}\,\varepsilon^{-1/\beta}\exp\left(-\frac{u^2}{2}(p+2)\frac{1}{\beta}\log\frac{1}{\varepsilon}(1-\delta_{\varepsilon 0})\right)\\
&\qquad \leq c_4\kappa_{\min}^{-1}\exp\left(-\frac{1}{\beta}\log\frac{1}{\varepsilon}\left[\frac{u^2}{2}(p+2)(1-\delta_{\varepsilon 0}) - 1\right]\right).
\end{aligned}
$$

The condition $\delta_{\varepsilon 0} < \min(\delta, 1/2)$ entails

$$
\frac{u^2}{2}(p+2)(1-\delta_{\varepsilon 0}) - 1 \geq (1+\delta)^2(1-\delta_{\varepsilon 0}) - 1 + \frac{pu^2}{4} \geq \frac{pu^2}{4}
$$

for all $\varepsilon \in (0, \varepsilon_0)$ and $u \geq 1 + \delta$. Hence,

$$
\begin{aligned}
P_f\left\{\psi_{\beta,\infty}^{-1}\|f_{\varepsilon,\beta,\infty} - f\|_\infty \geq u\right\} &\leq c_4\kappa_{\min}^{-1}\exp\left(-\frac{pu^2}{4\beta}\log\frac{1}{\varepsilon}\right)\\
&\leq c_7\exp\left(-u^2\frac{p}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right),
\end{aligned}
\tag{A.28}
$$

and the first inequality of Lemma 9 follows.

Using (A.28) we get, for all $\varepsilon \in (0, \varepsilon_0)$,

$$
\begin{aligned}
E_f\left(\psi_{\beta,\infty}^{-p}\|f_{\varepsilon,\beta,\infty} - f\|_\infty^p\right) &\leq (1+\delta)^p + \int_{(1+\delta)^p}^\infty P_f\left\{\psi_{\beta,\infty}^{-1}\|f_{\varepsilon,\beta,\infty} - f\|_\infty \geq t^{1/p}\right\}dt\\
&\leq (1+\delta)^p + c_7\int_{(1+\delta)^p}^\infty \exp\left(-t^{2/p}\,\frac{p}{4\beta_\varepsilon^*}\log\frac{1}{\varepsilon}\right)dt.
\end{aligned}
$$

Since the last integral is of the order $O((\beta_\varepsilon^*/\log(1/\varepsilon))^{p/2}) = o(1)$, as $\varepsilon \to 0$, the lemma is proved. $\square$

Proof of Lemma 10. Using (7.1), (7.2) and (5.3), one gets

$$
\|K_\beta\|_2^2 = s_1^2 s_2^{-1}\|K_{0,\beta}\|_2^2 = 1, \qquad \|K_\beta^{(\beta)}\|_2^2 = s_1^2 s_2^{2\beta-1}\|K_{0,\beta}^{(\beta)}\|_2^2 = 1.
$$

Another application of (5.3) yields

$$
K_\beta(0) = s_1\int_{-\infty}^\infty \frac{1}{1+|t|^{2\beta}}\,dt = 2\pi s_1\left(v_\beta^2 + b_\beta^2\right) = 2\beta(2\beta-1)^{-(2\beta-1)/4\beta}b_\beta.
$$

Thus, (7.3) follows.

Next, consider the function $g_0(t) = I\{|t| \le 1/2\}$, and its m-fold convolution $g_m^* = g_0 * \cdots * g_0$, where $m \ge 1$ is an integer. Put

$$g_\beta(t) = C_{\beta 1}^{-1} g_{4[\beta]+4}^*(t),$$

where $C_{\beta 1} = g_{4[\beta]+4}^*(0)$. Note that

$$\begin{aligned} C_{\beta 1} &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \frac{\sin(\omega/2)}{\omega/2} \right)^{4[\beta]+4} d\omega \\ &\ge \int_{|\omega| \le \pi} \left(\frac{1}{2\pi} \frac{\sin(\omega/2)}{\omega/2} \right)^{4[\beta]+4} d\omega \ge \pi^{-(8[\beta]+7)} \end{aligned}$$

since $(\sin(\omega/2))/(\omega/2) \ge 2/\pi$ for $|\omega| \le \pi$ (with $\sin 0/0 = 1$). In the sequel $C_{\beta i}$, $i = 1, 2, \dots,$ are positive constants depending only on β. The function g_β has the following properties:

$$g_\beta(0) = 1, \tag{A.29}$$

$$\operatorname{supp} g_\beta = (-2([\beta] + 1), 2([\beta] + 1)), \tag{A.30}$$

$$\hat{g}_\beta(\omega) = C_{\beta 1}^{-1} \left(\frac{1}{2\pi} \frac{\sin(\omega/2)}{\omega/2} \right)^{4[\beta]+4}. \tag{A.31}$$

Let $D_0 > 1$ be a fixed number. Set

$$K_{1,\beta}(t) = K_\beta(t) g_\beta(t/D_0).$$

The Fourier transform of $K_{1,\beta}$ is

$$\hat{K}_{1,\beta}(\omega) = \int_{-\infty}^{\infty} \hat{K}_\beta(u) D_0 \hat{g}_\beta(D_0(u - \omega))\, du.$$

Since $\hat{K}_\beta(u) = s_1 s_2^{-1} \hat{K}_{0,\beta}(s_2^{-1} u) = s_1 s_2^{-1}/(1 + |u/s_2|^{2\beta})$ and $\int \hat{g}_\beta(\tau)\, d\tau = g_\beta(0) = 1$, one gets

$$\begin{aligned} &\left| \hat{K}_{1,\beta}(\omega) - \hat{K}_\beta(\omega) \right| \\ &\quad \le \left| s_1 s_2^{-1} \int \hat{g}_\beta(\tau) \left(\frac{1}{1 + |(\tau/D_0 + \omega)/s_2|^{2\beta}} - \frac{1}{1 + |\omega/s_2|^{2\beta}} \right) d\tau \right| \\ &\quad \le \left| \hat{K}_\beta(\omega) \right| \int \left| \hat{g}_\beta(\tau) \right| A(\tau, \omega)\, d\tau, \end{aligned} \tag{A.32}$$

where

$$\begin{aligned} A(\tau, \omega) &= \frac{\left| |\omega|^{2\beta} - |\tau/D_0 + \omega|^{2\beta} \right|}{s_2^{2\beta} + |\tau/D_0 + \omega|^{2\beta}} \\ &\le 2\beta |\tau/D_0| \frac{\max(|\omega|, |\tau/D_0 + \omega|)^{2\beta - 1}}{2\beta - 1 + |\tau/D_0 + \omega|^{2\beta}}. \end{aligned}$$

If $|\omega| < 2|\tau/D_0|$ we have $A(\tau,\omega) \le 3^{2\beta} C_{\beta 2} |\tau/D_0|^{2\beta}$, where $\sup_{\beta \in B} C_{\beta 2} < \infty$. If $|\omega| \ge 2|\tau/D_0|$ one obtains $|\tau/D_0 + \omega| \ge |\omega|/2$, and therefore in this case

$$A(\tau,\omega) \le 2\beta|\tau/D_0| \max_{x>0} \frac{x^{2\beta-1}}{2\beta - 1 + (x/2)^{2\beta}} \le 2^{3\beta} C_{\beta 3} |\tau/D_0|$$

where $\sup_{\beta \in B} C_{\beta 3} < \infty$. Thus, for all τ and ω,

$$A(\tau,\omega) \le 3^{3\beta} C_{\beta 4} |t/D_0| \left(1 + |\tau/D_0|^{2\beta-1}\right), \tag{A.33}$$

where $\sup_{\beta \in B} C_{\beta 4} < \infty$.

Together (A.32) and (A.33) entail

$$\begin{aligned} &\left|\hat{K}_{1,\beta}(\omega) - \hat{K}_\beta(\omega)\right| \\ &\quad \le 3^{3\beta} C_{\beta 4} \left|\hat{K}_\beta(\omega)\right| \int \left|\hat{g}_\beta(\tau)\right| |\tau/D_0| \left(1 + |\tau/D_0|^{2\beta-1}\right) d\tau. \end{aligned} \tag{A.34}$$

Now, due to (A.31), for $s = 1$ or $s = 2\beta$, one gets

$$\int \left|\hat{g}_\beta(\tau)\right| |\tau|^s \, d\tau = C_{\beta 1}^{-1} \left[2^{s+1} \int |u|^s \left(\frac{1}{2\pi} \frac{\sin u}{u} \right)^{4[\beta]+4} du \right].$$

Considering separately integration over $|u| \le \pi/2$ and $|u| > \pi/2$, one easily finds that the last expression in square brackets is bounded uniformly in $\beta \in B$. Since also $C_{\beta 1}^{-1} \le \pi^{8[\beta]+7}$ and $D_0 \ge 1$, (A.34) implies

$$|\hat{K}_{1,\beta}(\omega) - \hat{K}_\beta(\omega)| \le 4^{11\beta} C_{\beta 5} |\hat{K}_\beta(\omega)|/D_0 = C_{\beta 6} |\hat{K}_\beta(\omega)|/D_0 \qquad \forall \omega, \tag{A.35}$$

where $\sup_{\beta \in B} C_{\beta 5} < \infty$, $C_{\beta 6} = 4^{11\beta} C_{\beta 5}$.

Set

$$\begin{aligned} D_0 &= \max\left(1, 4^{11\beta_\varepsilon^*} (2/\delta) \sup_{\beta \in B} C_{\beta 5}\right), \\ D &= D(\varepsilon, \beta, \delta) = 2([\beta] + 1) D_0 \end{aligned}$$

and

$$\overline{K}_\beta(t) = \left(1 + C_{\beta 6}/D_0\right)^{-1} (1 - \delta/2) K_{1,\beta}(t).$$

Let us check that $\overline{K}_\beta$ satisfies (i)–(iv) of Lemma 10 if D_0 (resp., D) is large enough. Using (A.30), one gets

$$\operatorname{supp} \overline{K}_\beta = \operatorname{supp} K_{1,\beta} = \operatorname{supp} g_\beta(\cdot/D_0) = (-D, D).$$

This proves (i) of Lemma 10. Next, (A.35) entails

$$\left|\hat{K}_{1,\beta}(\omega)\right| \le \left(1 + C_{\beta 6}/D_0\right) \left|\hat{K}_\beta(\omega)\right| \tag{A.36}$$

which yields (ii) of Lemma 10. Now, in view of (7.3) and (A.36),

$$1 = \|K_\beta\|_2 \ge \left(1 + C_{\beta 6}/D_0\right)^{-1} \|K_{1,\beta}\|_2 = \|\overline{K}_\beta\|_2 (1 - \delta/2)^{-1},$$

which yields (iii) of Lemma 10. It remains to observe that (iv) of Lemma 10 follows from the relations

$$\overline{K}_\beta(0) = K_\beta(0) \left(1 + C_{\beta 6}/D_0\right)^{-1} (1 - \delta/2) \ge K_\beta(0)(1 - \delta),$$

where we used (A.29) and the inequality $D_0 \geq 2C_{\beta 6}/\delta$. To check (v) of Lemma 10 note that, for any $\beta \in B$, $j = 0, 1, \infty$, and ε small enough, one has

$$D\bar{h}_{\beta,j} \leq D_* \beta_\varepsilon^* 4^{11\beta_\varepsilon^*}\left(\varepsilon^2 \log\tfrac{1}{\varepsilon}\right)^{1/2\beta_\varepsilon^*}$$

where $D_* > 0$ is a constant depending only on L, β_1, p. The last expression tends to 0 as $\varepsilon \to 0$ in view of (3.1).

Proposition 4. *Let $\beta \in B$, $\delta \in (0,1)$, $j \in \{0, 1, \infty\}$ and let $f \in L_1(-\infty, \infty)$ be a function such that $\operatorname{supp} f \subseteq [0,1]$ and the Fourier transform of f satisfies*

$$|\hat{f}(\omega)| \leq (1 - \delta/2) L \bar{h}_{\beta,j}^{\beta+1/2} |\hat{K}_\beta(\omega \bar{h}_{\beta,j})|.$$

Then there exists $\varepsilon' \in (0,1)$ independent of β such that for $0 < \varepsilon < \varepsilon'$ we have $f \in W_\beta$, $\beta \in B$.

Proof. Put for brevity $h = \bar{h}_{\beta,j}$. Since $\operatorname{supp} f \subseteq [0,1]$ we have

$$|\hat{f}(2\pi l)|^2 = \left|\frac{1}{2\pi}\int_0^1 f(x) e^{-i2\pi l x}\, dx\right|^2 = (8\pi^2)^{-1}\left(\theta_{2l-1}^2 + \theta_{2l}^2\right), \qquad l = 1, 2, \ldots,$$

where θ_k are the Fourier coefficients of f, see Section 2. Thus,

$$\begin{aligned}
\sum_{k=0}^\infty a_k^2(\beta)\theta_k^2 &= \sum_{l=1}^\infty (2l)^{2\beta}\left(\theta_{2l-1}^2 + \theta_{2l}^2\right) = 8\pi^2 \sum_{l=1}^\infty (2l)^{2\beta} |\hat{f}(2\pi l)|^2 \\
&\leq 8\pi^2 (1 - \delta/2)^2 L^2 \sum_{l=1}^\infty (2l)^{2\beta} h^{2\beta+1} |\hat{K}_\beta(2\pi l h)|^2 \\
&= A \sum_{l=1}^\infty (lh')^{2\beta} \frac{h'}{\left(1 + (lh')^{2\beta}\right)^2}
\end{aligned}$$

where $A = 4\pi(1 - \delta/2)^2 L^2 \pi^{-2\beta} s_1^2 s_2^{2\beta-1}$ and $h' = 2\pi h / s_2$. Thus, by Proposition 1 and (5.3)

$$\begin{aligned}
\sum_{k=0}^\infty a_k^2(\beta)\theta_k^2 &\leq A\left(\int_0^\infty \frac{t^{2\beta}}{(1 + t^{2\beta})^2}\, dt + 2h'\right) \\
&= A\left(\pi b_\beta^2 + 2h'\right) = (1 - \delta/2)^2 L^2 \pi^{-2\beta} + 2Ah'.
\end{aligned}$$

To finish the proof it remains to note that $Q_\beta = L^2 \pi^{-2\beta}$ and that, uniformly in $\beta \in B$, $j = 0, 1, \infty$,

$$Ah'/Q_\beta \leq Q_* \pi^{2\beta_\varepsilon^*}\left(\varepsilon^2 \log\frac{1}{\varepsilon}\right)^{1/2\beta_\varepsilon^*}$$

where Lemma 1 was used and $Q_* > 0$ is a constant depending only on L, β_1, p. The last expression tends to 0 as $\varepsilon \to 0$ in view of (3.1). □

Acknowledgments. It is a pleasure to acknowledge comments of Oleg Lepski, Yuri Golubev and Jussi Klemelä. The referees, Associate Editor and the Editor, Lawrence D. Brown, also provided helpful comments which are greatly appreciated.

REFERENCES

BARRON, A., BIRGE, L. and MASSART, P. (1995). Risk bounds for model selection via penalization. Technical Report 95.54, Univ. Paris-Sud.

BIRGE, L. (1983). Approximation dans les espaces métriques et théorie de l'estimation. *Z. Wahrsch. Verw. Gebiete* **65** 181–237.

BIRGE, L. and MASSART, P. (1997). From model selection to adaptive estimation. In *Festschrift for Lucien Le Cam* (D. Pollard, E. Torgersen, G. L. Yang, eds.) 55–87. Springer, New York.

BROWN, L. and LOW, M. (1996). A constrained risk inequality with applications to nonparametric estimation. *Ann. Statist.* **24** 2524–2535.

BUTUCEA, C. (1998). Estimation non paramétrique adaptative de la densité de probabilité: vitesses de convergence et constante exacte. Ph.D. thesis, Univ. Paris VI.

DONOHO, D. L. (1994). Asymptotic minimax risk for sup-norm loss: solution via optimal recovery. *Probab. Theory Related Fields* **99** 145–170.

DONOHO, D. L. and JOHNSTONE, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.* **90** 1200–1224.

DONOHO, D. L., JOHNSTONE, I. M., KERKYACHARIAN, G. and PICARD, D. (1995). Wavelet shrinkage asymptopia? (with discussion) *J. Roy. Statist. Soc. Ser. B* **57** 301–369.

DONOHO, D. L. and LOW, M. (1992). Renormalization exponents and optimal pointwise rates of convergence. *Ann. Statist.* **20** 944–970.

EFROIMOVICH, S. YU. and PINSKER, M. S. (1984). Learning algorithm for nonparametric filtering. *Automat. Remote Control* **11** 1434–1440.

GOLDENSHLUGER, A. and NEMIROVSKII, A. (1997). On spatial adaptive estimation of nonparametric regression. *Math. Methods Statist.* **6** 135–170.

GOLUBEV, G. K. (1990). Quasi-linear estimates of signals in L_2. *Problems Inform. Transmission* **26** 15–20.

GOLUBEV, G. K. and NUSSBAUM, M. (1992). Adaptive spline estimates for nonparametric regression models. *Theory Probab. Appl.* **37** 521–529.

HÄRDLE, W., KERKYACHARIAN, G., PICARD, D. and TSYBAKOV, A. B. (1998). *Wavelets, Approximation and Statistical Applications. Lecture Notes in Statist.* **129**. Springer, New York.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer, New York. (Originally published in Russian in 1979.)

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1984). Nonparametric estimation of the value of a linear functional in Gaussian white noise. *Theory Probab. Appl.* **29** 18–32.

JUDITSKY, A. (1997). Wavelet estimators: adapting to unknown smoothness. *Math. Methods Statist.* **6** 1–25.

KNEIP, A. (1994). Ordered linear smoothers. *Ann. Statist.* **22** 835–866.

KOROSTELEV, A. P. (1993). Exact asymptotically minimax estimator for nonparametric regression in uniform norm. *Theory Probab. Appl.* **38** 737–743.

LEADBETTER, M. R., LINDGREN, G. and ROOTZEN, H. (1986). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

LEPSKI, O. V. (1990). One problem of adaptive estimation in Gaussian white noise. *Theory Probab. Appl.* **35** 454–466.

LEPSKI, O. V. (1991). Asymptotically minimax adaptive estimation I: upper bounds. Optimal adaptive estimates. *Theory Probab. Appl.* **36** 682–697.

LEPSKI, O. V. (1992a). Asymptotically minimax adaptive estimation II: statistical models without optimal adaptation. Adaptive estimates. *Theory Probab. Appl.* **37** 433–468.

LEPSKI, O. V. (1992b). On problems of adaptive estimation in white Gaussian noise. In *Topics in Nonparametric Estimation.* (R. Z. Khasminskii, ed.) 87–106. Amer. Math. Soc., Providence, RI.

LEPSKI, O. V. (1996). Personal communication.

LEPSKI, O. V., MAMMEN, E. and SPOKOINY, V. G. (1997). Ideal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selection. *Ann. Statist.* **25** 929–947.

LEPSKI, O. V. and SPOKOINY, V. G. (1997). Optimal pointwise adaptive methods in nonparametric estimation. *Ann. Statist.* **25** 2512–2546.

NEMIROVSKII, A. (1985). On nonparametric estimation of smooth regression functions. *Soviet J. Comput. Syst. Sci.* **23** 1–11.

OUDSHOORN, K. (1996). Optimality and adaptivity in nonparametric regression. Ph.D. thesis, Univ. Utrecht.

PINSKER, M. S. (1980). Optimal filtering of square integrable signals in Gaussian white noise. *Problems Inform. Transmission* **16** 120–133.

STONE, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.* **8** 1348–1360.

STONE, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053.

TSYBAKOV, A. B. (1997). Asymptotically efficient signal estimation in L_2 with general loss functions. *Problems Inform. Transmission* **33** 78–88.

UPRES-A CNRS 7055
UNIVERSITÉ PARIS VI
4, PL. JUSSIEU, B.P. 188
75252 PARIS
FRANCE
E-MAIL: tsybakov@ccr.jussieu.fr