Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Local adaptivity to inhomogeneous smoothness. 1. Resolution level

Oleg V. Lepskii[1,3], Vladimir G. Spokoiny[2,3]

submitted: 29th March 1994

[1] Institute for System Analysis
Prospekt 60–Let Oktjabrja, 9
Moscow, 117312
Russia

[2] Institute for Information
Transmission Problems
Ermolovoy 19
Moscow, 101447
Russia

[3] Institute of Applied Analysis
and Stochastics
Mohrenstraße 39
D – 10117 Berlin
Germany

Preprint No. 92
Berlin 1994

1991 Mathematics Subject Classification. Primary 62G07; Secondary 62G20.

Edited by
Institut für Angewandte Analysis und Stochastik (IAAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2004975
e–mail (X.400): c=de;a=d400;p=iaas-berlin;s=preprint
e–mail (Internet): preprint@iaas-berlin.d400.de

LOCAL ADAPTIVITY TO INHOMOGENEOUS SMOOTHNESS.1. RESOLUTION LEVEL

LEPSKII, O.V. AND SPOKOINY, V.G.

ABSTRACT. The problem of nonparametric estimation of functions of inhomogeneous smoothness is considered.
The goal is to define the notion of local smoothness of a function $f(\cdot)$, to evaluate the optimal rate of convergence of estimators (depending on this local smoothness) and to construct an asymptotically efficient locally adaptive estimator.
We treat local (or δ-local) smoothness properties of a function $f(\cdot)$ at a point t as the corresponding characteristics of this function on the interval $[t-\delta, t+\delta]$. The value δ measures the "locality" of our procedure. The smaller this value is taken the more precise is our resolution analysis. But this value can not be taken arbitrary small since we should ce able to restore local smoothness properties of a function from the noisy data.
The main result of the paper describes just the maximal rate of convergence of this parameter δ to zero as the noise level ε goes to zero. We call this value the *resolution level.* The value of this level strongly depends on the upper considered smoothness β^* what we wish to attain. If κ_ε^* is the bandwidth corresponding to this smoothness β^* then the resolution level δ_ε^* can not be chosen less (in order) than κ_ε^*.
In particular, this yields that it is impossible to improve at the same time the accuracy of our procedure (which is measured by the upper smoothness β^*) and its local adaptive properties. If we improve the accuracy of estimation at subintervals where a function is of high smoothness then we will have a low accuracy in a larger vicinity near a point with small smoothness.
The main results claim that if the parameter of locality δ is taken less (in order) than the resolution level, then the corresponding risk is (asymptotically) infinite. After that we construct estimators with a finite asymptotic risk for the case of δ coinciding with the resolution level.

1. INTRODUCTION

We consider the problem of estimation of functions with inhomogeneous smoothness properties. We suppose the simplest nonparametric model "signal + noise"

$$dX(t) = f(t)\,dt + \varepsilon\,dW(t).$$

The standard nonparametric minimax approach to the estimation problem (see, e.g., Ibragimov and Khasminskii, 1981) is based on the assumption that a function to be estimated from noisy data X belongs to some functional class, e.g. Hölder, Sobolev, Besov, etc. The corresponding rate of convergence of efficient (in asymptotic sense) estimators is determined by the parameters of this class. Under this point of view one can speak on the global smoothness properties and global rate of convergence of estimators.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 62G07; Secondary 62G20.

Besov, etc. The corresponding rate of convergence of efficient (in asymptotic sense) estimators is determined by the parameters of this class. Under this point of view one can speak on the global smoothness properties and global rate of convergence of estimators. But this approach has the evident disadvantage. If the observed function possesses inhomogeneous smoothness characteristics, for instance, one part is more smooth than another, then the global rate will be defined by the worsest part and the corresponding accuracy for the smooth part will be less than if we estimate this function only within this part.

The natural idea is arised to split the whole interval of observation onto the parts with different smoothness properties and to estimate the function separately on each part with the corresponding smoothing parameters. But under this approach we need in some oracle who tells us how to divide the interval of observations and what are the corresponding smoothing parameters.

To bypass this problem one can assume that the smoothness properties vary slowly and, therefore, that the smoothness properties are homogeneous within each subinterval with a small length δ. After that the estimation problem can be considered separately for each subinterval with the adaptive choice of the smoothing parameters. Of course, the smaller is the length δ of each subinterval in our partition, the more accurate is our local resolution analysis. But again the question: How small the length of each interval can be taken to be able to perform locally adaptive estimation?

The presented paper is devoted to the complete investigation of this problem.

We define local smoothness characteristic of a function at a given point as smoothness characteristics of this function being restricted on a small interval (of length δ) around this point.

We show that if the level ε of the noise is small (goes to zero) then the locality parameter δ also can be taken small (goes to zero). Moreover, we describe the maximal rate of convergence δ_ε^* of this parameter to zero for which we are still able to perform locally efficient estimation. We call this value the *resolution level.* The value of this level strongly depends on the upper considered smoothness β^* what we wish to attain. If κ_ε^* is the bandwidth corresponding to this smoothness β^* then the resolution level δ_ε^* can not be chosen less (in order) than κ_ε^*.

In particular, this yields that it is impossible to improve at the same time the accuracy of our procedure (which is measured by the upper smoothness β^*) and its local adaptive properties. If we improve the accuracy of estimation at subintervals where a function is of high smoothness then we will have a low accuracy in a larger vicinity near a point with bad smoothness characteristics.

Notice also another principal feature of the approach proposed. For the classical minimax nonparametric situation the optimal accuracy (rate of convergence) of estimators is described by one number which depends on the smoothing parameters of the functional class containig $f(\cdot)$ (see Ibragimov and Khasminskii, 1980, Stone, 1980, Bretagnolle and Huber, 1976). But in the inhomogeneous situation, since the smoothness properties of the function $f(\cdot)$ varies from point to point, the corresponding rate of accuracy of estimators also should vary from point to point. Therefore, we treat the risk of an estimator $\tilde{f}_\varepsilon$ as the sum (or the integral) of pointwise risk and the normalizing factor is taken corresponding to the pointwise smoothness characteristics. Such a construction allows to combine minimax approach (the function $f(\cdot)$ is assumed to be in some very large functional class of functions with inhomogeneous smoothness properties) with more detailed pointwise description of the properties of estimators since the rate of an estimator $\tilde{f}_\varepsilon(\cdot)$ at a point $t \in [0,1]$ depends on the function $f(\cdot)$ and on t (through the corresponding smoothness characterists of f at t). From this point of view, an estimator $\tilde{f}_\varepsilon$ can be

called efficient (or rate efficient) if the corresponding risk is finite on the considered large functional class.
Note also that speaking about the construction of efficient estimators, the crucial point is an adaptive choice of the local smoothness parameters (since one could not expect some oracle who informs on the local smoothness properties of a function to be estimated).
We study the properties of the following natural construction of locally adaptive estimators. We split the whole interval onto subintervals of length equals to the resolution level, and then we construct independently an adaptive estimator within each subinterval. Of course, the natural idea is to apply locally one of standard (global) adaptive procedures. We don't give a survey of existing minimax adaptive procedures and note only the closed to the considered framework parers of Efroimovich and Pinsker(1984), Golubev(1987), Lepskii(1990,1991,1992), Poljak and Tsybakov(1990). We apply the procedure from Lepskii(1991) because only this one is universal and the remainings are valid only for the squared losses. We believe that in the case of estimation in L_2 other procedures can be applied as well.
We believe also that the nonlinear wavelet procedure introduced into statistics by Donoho and Johnstone(1992), Kerkicharjan and Picard(1993) posesses some locally adaptive properties but the relating consideratios beyond the scope of the present paper.
The paper is organized as follows. In Section 2 we give the main definitions and results, in Section 3 we prove the results and in Section 4 we discuss the possible directions to develop the results obtained.

2. Definition and the main results

2.1. Local Smoothness Characteristics. We start with the notion of a local smoothness characteristic of a function $f(\cdot)$. We define local smoothness in the Hölder sense but other ones can be taken as well, for instance, based on Sobolev norm. We choose the approach based on Hölder norm since we think it better corresponding to intuitive feeling of smoothness properties.
Let β, L be some positive, $m = \lfloor\beta\rfloor$ (i.e. $0 < \beta - m \le 1$), $\alpha = \beta - m$, and let I be some subinterval of $[0,1]$. We say that $f \in \Sigma(\beta, L, I)$ if $\left|f^{(m)}(s) - f^{(m)}(t)\right| \le |s-t|^{\alpha}$, $\forall\, s,t \in I$.
Our definition of a local smoothness characteristic of a function $f(\cdot)$ involves three parameters: β^*, L, δ. The value β^* can be treated as the upper considered smoothness (the maximal number of derivatives) for a given function. The number L plays the role of a Liptschitz constant, and for us only important that $0 < L < \infty$. At last, the value δ measures "locality" of our procedure. We understand local smoothness properties of a function $f(\cdot)$ at a point t as properties of this function considered on the interval $[t-\delta, t+\delta]$.

Definition 1. *Given β^*, L, δ and a function $f(\cdot)$ define a function $\beta_f(\cdot,\delta)$ with*

$$\beta_f(t,\delta) = \sup\left\{\beta \le \beta^* : f \in \Sigma\left(\beta, L, [t-\delta, t+\delta]\right)\right\}. \tag{2.1}$$

Of course, to be correctly defined the corresponding set in (2.1) should be nonempty. Hence we consider below only functions with this property.

Definition 2. *Given β^*, L, δ we say that a function $f(\cdot)$ belongs to the class $\Sigma(\sigma)$ if for any $t \in [0,1]$ one has $f \in \Sigma\left(\beta, L, [t-\delta, t+\delta]\right)$ with some $\delta > 0$.*

To illustrate this notion we give some examples.

Example 1. Consider the function $f(t) = |\frac{1}{2} - t|$. This function is locally linear everywhere except the point $\frac{1}{2}$ and for $\beta^* \geq 2,\ L \geq 2$ one has

$$\beta_f(t,\delta) = \begin{cases} \beta^*, & t \notin [\frac{1}{2}-\delta, \frac{1}{2}+\delta] \\ 1, & t \in [\frac{1}{2}-\delta, \frac{1}{2}+\delta] \end{cases} .$$

Example 2. Let now $f(t) = C\left|\frac{1}{2} - t\right|$ with a large $C > 0$. Then as above we have $\beta_f(t,\delta) = \beta^*$ for $t \notin [\frac{1}{2}-\delta, \frac{1}{2}+\delta]$. But for $t \in [\frac{1}{2}-\delta, \frac{1}{2}+\delta]$ one has

$$\beta_f(t,\delta) = 1 - \frac{\ln 2C/L}{\ln(2\delta)^{-1}}$$

if this value is positive. Otherwise (if $\frac{2C}{L} > \frac{1}{2\delta}$) this function does not lie in $\Sigma(\sigma)$. Therefore, $\beta_f(t,\delta)$ is well defined for large C only if δ is small enough.

Example 3. Let $n \geq 1$ and $f(t) = |\frac{1}{2n} - (t - \frac{k-1}{n})|$ for $t \in [\frac{k-1}{n}, \frac{k}{n}],\ k = 1, \ldots, n$. If $\delta < \frac{1}{2n}$, then $\beta_f(t,\delta) = \beta^*$ for $t \notin [\frac{k}{n} - \delta, \frac{k}{n} + \delta]$ and $\beta_f(t,\delta) = 1$ otherwise.

We use below some simple properties of the function $\beta_f(t,\delta)$ and the classes $\Sigma(\delta)$.

Lemma 1. *The following statements hold:*

(1) *If a function $f(\cdot)$ is constant on the interval $I = [c_I - 2\delta, c_I + 2\delta]$, then*

$$\beta_f(t,\delta) = \beta^*, \qquad t \in [c_I - \delta, c_I + \delta] .$$

(2) *If $\delta' \leq \delta$, then $\beta_f(t,\delta') \geq \beta_f(t,\delta)$;*

(3) *If $\delta' \leq \delta$, then $\Sigma(\delta') \supseteq \Sigma(\delta)$.*

Proof. Obvious. □

2.2. Statistical Model and Estimation Problem. We consider the simplest non-parametric model "signal + white noise":

$$dX(t) = f(t)dt + \varepsilon dW(t) \tag{2.2}$$

where an unknown function $f(\cdot)$ is to be estimated for $t \in [0,1]$, ε is a small parameter (the error level) and $W = (W(t))$ is a standard Wiener process.

Moreover, not to deal with the boundary effect, we suppose that the process $X = (X(t))$ is observed on the larger interval $[-\delta_0, 1+\delta_0]$. Such assumptions allow us to simplify the exposition and to emphasize in more clear form the main ideas and results. Of course, more realistic models with discrete nongaussian errors can be considered as well, but we concentrate in this paper on the simplest case of the model (2.2).

Now we define the estimation problem. We distinguish below between two cases: estimation in L_p-norm for $1 \leq p < \infty$ and in sup-norm.

First some notations. Let I be some subinterval of $[0,1]$. Denote for a function $f(\cdot)$ on $[0,1]$

$$\begin{aligned} \|f(\cdot)\|_{p,I} &= \left[\frac{1}{|I|}\int_I |f(t)|^p\, dt\right]^{1/p}, \qquad 1 \leq p < \infty; \\ \|f(\cdot)\|_{\infty,I} &= \sup_{t\in I} |f(t)| . \end{aligned}$$

We suppress the index I for $I = [0,1]$.

Given ε and $\beta \geq 0$, put

$$\varphi_\varepsilon(\beta) = \begin{cases} \varepsilon^{\frac{2\beta}{2\beta+1}}, & 1 \leq p < \infty \\ \left(\varepsilon\sqrt{\ln\frac{1}{\varepsilon}}\right)^{\frac{2\beta}{2\beta+1}}, & p = \infty \end{cases} .$$

We use also the notation $\varphi_\varepsilon^* = \varphi_\varepsilon(\beta^*)$.

It is well known that $\varphi_\varepsilon(\beta)$ is the minimax rate of convergence of estimators of the function $f(\cdot)$ by the observations from (2.2) on the H ölder class $\Sigma(\beta, L)$ (or some other functional classes type of Sobolev, Besov etc.) with the smoothness parameter β. But we don't suppose $f(\cdot)$ to possess some global smoothness properties described by a unique β. In the contrary, we assume this parameter to vary from point to point. Therefore, the following definition of the risk of estimator seems to be natural.

Definition 3. *Let $\tilde{f}_\varepsilon(\cdot)$ be some estimator of $f(\cdot)$. Given $\delta > 0$, we define the risk (or δ-local risk) of this estimator for $p < \infty$ as*

$$R_\varepsilon^{(p)}\left(\tilde{f}_\varepsilon(\cdot), \delta\right) = \sup_{f \in \Sigma(\delta)} E_f \left\| \frac{\tilde{f}_\varepsilon(\cdot) - f(\cdot)}{\varphi_\varepsilon\left(\beta_f(\cdot, \delta)\right)} \right\|_p^p \tag{2.3}$$

and for $p = \infty$ as

$$R_\varepsilon^{(\infty)}\left(\tilde{f}_\varepsilon(\cdot), \delta\right) = \sup_{f \in \Sigma(\delta)} E_f \left\| \frac{\tilde{f}_\varepsilon(\cdot) - f(\cdot)}{\varphi_\varepsilon\left(\beta_f(\cdot, \delta)\right)} \right\|_\infty .$$

2.3. Main Results. We give separately the main results for the cases $p < \infty$ and $p = \infty$.

Theorem 1. ($p < \infty$). *Let $1 \leq p < \infty$ and*

$$\delta_\varepsilon^* = \varepsilon^{\frac{2}{2\beta^*+1}}.$$

For each $a \in (0,1]$, any sequence $\delta_\varepsilon > 0$ with $\delta_\varepsilon \leq a\delta_\varepsilon^*$ and any estimators $\tilde{f}_\varepsilon(\cdot)$

$$\varliminf_{\varepsilon \to 0} R_\varepsilon^{(p)}\left(\tilde{f}_\varepsilon(\cdot), \delta_\varepsilon\right) \geq \frac{C_1}{a^{p/2}}$$

Theorem 2. ($p = \infty$). *Let*

$$\delta_\varepsilon^* = \left(\varepsilon\sqrt{\ln\frac{1}{\varepsilon}}\right)^{\frac{2}{2\beta^*+1}}.$$

For each $a \in (0,1]$, any sequence $\delta_\varepsilon > 0$ with $\delta_\varepsilon \leq a\delta_\varepsilon^*$ and any estimators $\tilde{f}_\varepsilon(\cdot)$

$$\varliminf_{\varepsilon \to 0} R_\varepsilon^{(\infty)}\left(\tilde{f}_\varepsilon(\cdot), \delta_\varepsilon\right) \geq \frac{C_1'}{\sqrt{a}}$$

Here and below by C with index or not we denote some absolute constants depending possibly on β^*, L but not on ε, a.

Remark. The results of Theorems 1 and 2 claim that the locality parameter δ can not be taken less (in order) than δ_ε^*. We call this value *the resolution level.*

The next two results claim existence of estimators $\hat{f}_\varepsilon(\cdot)$ with a finite asymptotic risk if the parameter of locality δ_ε is taken not less (in order) than the resolution level δ_ε^*.

Theorem 3. ($p < \infty$). *Given $a \in (0,1]$ and a sequence δ_ε with $\delta_\varepsilon \geq a\delta_\varepsilon^* = a\varepsilon^{2/(2\beta^*+1)}$ there exist such estimators $\hat{f}_\varepsilon(\cdot)$ that*

$$\varlimsup_{\varepsilon \to 0} R_\varepsilon^{(p)}\left(\hat{f}_\varepsilon(\cdot), \delta_\varepsilon\right) \leq \frac{C_2}{a^{p/2}}.$$

Theorem 4. ($p = \infty$). *Given* $a \in (0,1]$ *and a sequence* δ_ε *with* $\delta_\varepsilon \ge a\delta_\varepsilon^* = a\left(\varepsilon\sqrt{\ln 1/\varepsilon}\right)^{2/(2\beta^*+1)}$ *there exist such estimators* $\hat f_\varepsilon(\cdot)$ *that*

$$\lim_{\varepsilon\to 0} R_\varepsilon^{(\infty)}\left(\hat f_\varepsilon(\cdot), \delta_\varepsilon\right) \le \frac{C_2'}{\sqrt{a}}.$$

2.4. Construction of locally adaptive estimators. Now we define the structure of the upper estimators $\hat f_\varepsilon(\cdot)$ from Theorems 3 and 4. First denote

$$\kappa_\varepsilon(\beta) = \begin{cases} \frac{a}{2}\varepsilon^{\frac{2}{2\beta+1}} & 1 \le p < \infty \\ \frac{a}{2}\left(\varepsilon\sqrt{\ln\frac{1}{\varepsilon}}\right)^{\frac{2}{2\beta+1}} & p = \infty \end{cases}.$$

We will write κ_ε^* as well as $\kappa_\varepsilon(\beta^*)$.
Now introduce the family of kernel estimators $\{f_\varepsilon(\cdot,\beta),\ 0 < \beta \le \beta^*\}$ with

$$f_\varepsilon(t,\beta) = \frac{1}{\kappa_\varepsilon(\beta)}\int K\left(\frac{u-t}{\kappa_\varepsilon(\beta)}\right) dX_\varepsilon(u)$$

where for simplicity a kernel $K(\cdot)$ is taken universal for all β and we suppose K to satisfy the following properties:

(1) K is supported on $[-1,1]$;
(2) $\int K^2(s)\,ds < \infty$;
(3) $\int K(s)\,ds = 1$;
(4) for each $1 \le m \le \lfloor\beta\rfloor$

$$\int K(s)\,s^m\,ds = 0.$$

Remark. Typically the bandwidth $\kappa_\varepsilon(\beta)$ and the rate sequence $\varphi_\varepsilon(\beta)$ are chosen to satisfy the balance relation

$$\varphi_\varepsilon(\beta) = |\kappa_\varepsilon(\beta)|^\beta, \qquad \varphi_\varepsilon(\beta)\sqrt{\kappa_\varepsilon(\beta)} = \begin{cases} \varepsilon, & 1 \le p < \infty \\ \varepsilon\sqrt{\ln\frac{1}{\varepsilon}}, & p = \infty \end{cases}$$

but through our definition one has

$$\varphi_\varepsilon(\beta) = \left|\frac{2}{a}\kappa_\varepsilon(\beta)\right|^\beta$$

$$\varphi_\varepsilon(\beta)\sqrt{\kappa_\varepsilon(\beta)} = \begin{cases} \sqrt{\frac{a}{2}}\varepsilon, & 1 \le p < \infty \\ \sqrt{\frac{a}{2}}\left(\varepsilon\sqrt{\ln\frac{1}{\varepsilon}}\right), & p = \infty \end{cases}$$

Let

$$h_\varepsilon = \left(\ln\frac{1}{\varepsilon}\right)^{-1}$$

and B_ε be the h_ε-net of the interval $[0,\beta^*]$:

$$B_\varepsilon = \{\beta \ge 0 : \beta = \beta^* - kh_\varepsilon\ ,\ k = 0,1,2,\dots\}.$$

By $B_\varepsilon(\beta)$ for any $\beta \in B_\varepsilon$ we denote the subset of B_ε containing all $\beta' \le \beta$,

$$B_\varepsilon(\beta) = \{\beta' \in B_\varepsilon : \beta' \le \beta\}.$$

Let now $\mathcal{I}_\varepsilon$ be the partition of the interval $[0,1]$ onto subintervals of length $\delta_\varepsilon/2$. We construct our estimators $\hat{f}_\varepsilon(\cdot)$ in the following way. For each $I \in \mathcal{I}_\varepsilon$ we determine adaptively the corresponding smoothness characteristic $\hat{\beta}_\varepsilon(I)$ and then take $\hat{f}_\varepsilon(\cdot)$ coinciding on I with the kernel estimator $f_\varepsilon(\cdot,\beta)$ for $\beta = \hat{\beta}_\varepsilon(I)$,

$$\hat{f}_\varepsilon(t) = \sum_{I \in \mathcal{I}_\varepsilon} f_\varepsilon\left(t, \hat{\beta}_\varepsilon(I)\right) 1(t \in I).$$

To choose $\beta_\varepsilon(I)$ we apply the procedure from Lepskii(91):

$$\hat{\beta}_\varepsilon(I) = \sup\left\{\beta \in B_e : \|f_\varepsilon(\cdot,\beta) - f_\varepsilon(\cdot,\beta')\|_{p,I} \le C_a \varphi_\varepsilon(\beta') \quad \forall \beta' \in B_\varepsilon(\beta)\right\}$$

where

$$C_a = \begin{cases} 2L + \frac{2(1+C_p)^{1/p}}{\sqrt{a/2}}, & 1 \le p < \infty \\ 2L + \frac{8}{\sqrt{a/2}}, & p = \infty \end{cases},$$

$$C_p = \|K\|^p E|\xi|^p, \qquad \xi \sim N(0,1)$$

with

$$\|K\|^2 = \int K^2(s)\,ds$$

3. Proof of Theorems

3.1. Proof of Theorem 1. First define a function $g(\cdot)$ on $[-\frac{1}{2}, \frac{1}{2}]$ as follows

$$g(t) = \begin{cases} 1, & |t| \le 1/4 \\ 4\left(\frac{1}{2} - |t|\right), & 1/4 \le |t| \le 1/2 \end{cases}.$$

This function is obviously continuous and piecewise linear, and $g\left(\pm\frac{1}{2}\right) = 0$.
Below we refer to the following Bayes problem. Let the observation model $\mathcal{E}(\sigma)$ be described by the stochastic equation

$$dX(t) = \pm g(t)\,dt + \sigma\,dW(t), \qquad t \in \left[-\frac{1}{2}, \frac{1}{2}\right]$$

and the parameter set consists of two functions $\pm g(\cdot)$. Define the loss function l for an estimator $\tilde{g}(\cdot)$ at the point $\pm g(\cdot)$ as follows

$$l(\tilde{g} - \pm g) = \int_{-1/8}^{1/8} |\tilde{g}(t) - \pm g(t)|^p\,dt = \int_{-1/8}^{1/8} |\tilde{g}(t) \mp 1|^p\,dt.$$

Denote by $b(\sigma)$ the solution of the Bayes problem for this model and the uniform prior π (i.e. $\pi(\pm g) = 1/2$):

$$b(\sigma) = \inf_{\tilde{g}} \frac{1}{2}\left[E_g l(\tilde{g} - g) + E_{-g} l(\tilde{g} + g)\right].$$

Obviously $b(\sigma) > 0$ for each $\sigma > 0$.
Now we turn to the proof of the theorem. Define for each $\varepsilon > 0$ the value β_ε in such a way that

$$\varepsilon^{\frac{2}{2\beta_\varepsilon+1}} = 8\delta_\varepsilon \tag{3.1}$$

Later we write also Δ_ε instead of $8\delta_\varepsilon$ and φ_ε instead of $\varphi_\varepsilon(\beta_\varepsilon) = \varepsilon^{\frac{2\beta_\varepsilon}{2\beta_\varepsilon+1}}$. Obviously

$$\begin{aligned} \varphi_\varepsilon &= |\Delta_\varepsilon|^{\beta_\varepsilon}, \\ \varphi_\varepsilon\sqrt{\Delta_\varepsilon} &= \varepsilon \end{aligned} \tag{3.2}$$

Let now $\mathcal{I}_\varepsilon$ be the partition of the interval $[0,1]$ onto the subintervals of length Δ_ε. Without loss of generality we suppose $card\,\mathcal{I}_\varepsilon = 1/\Delta_\varepsilon$.
For each $I \in \mathcal{I}_\varepsilon$ we construct the function $g_I(\cdot)$ on I with

$$g_I(t) = |I|^{\beta_\varepsilon} g\left(\frac{t-c_I}{|I|}\right), \qquad t \in I, \tag{3.3}$$

where $|I| = 8\delta_\varepsilon = \Delta_\varepsilon$ and c_I is the center of I.
Now we introduce the set G_ε of functions f_θ on $[0,1]$ of the form

$$f_\theta(t) = \sum_{I\in\mathcal{I}_\varepsilon} \theta_I\, g_I(t)\, 1(t \in I)$$

where $\theta = (\theta_I)_{I\in\mathcal{I}_\varepsilon}$ is a vector with coordinates $\theta_I = \pm 1$.
Evidently these functions f_θ belong to the set $\Sigma(\delta)$ for any small $\delta > 0$ and we show below that even being restricted on this finite set G_ε the risk of arbitrary estimators $\tilde{f}_\varepsilon$ is greater than $Ca^{-p/2}$.
First we notice that by the construction each function f_θ from G_ε is constant on the intervals $[c_I - 2\delta_\varepsilon, c_I + 2\delta_\varepsilon]$, $I \in \mathcal{I}_\varepsilon$. Through Lemma 1 this yields

$$\beta_{f_\theta}(t,\delta_\varepsilon) = \beta^*, \qquad t \in [c_I - \delta_\varepsilon, c_I + \delta_\varepsilon], \qquad I \in \mathcal{I}_\varepsilon \tag{3.4}$$

Now put for any estimators $\tilde{f}_\varepsilon$

$$l\left(\tilde{f}_\varepsilon - f_\theta\right) = \sum_{I\in\mathcal{I}_\varepsilon} \int_{c_I-\delta_\varepsilon}^{c_I+\delta_\varepsilon} \left|\tilde{f}_\varepsilon(t) - f_\theta(t)\right|^p dt \tag{3.5}$$

and let

$$\mathcal{R}_\varepsilon = \inf_{\tilde{f}_\varepsilon} \sup_{f_\theta} E_{f_\theta} l\left(\tilde{f}_\varepsilon - f_\theta\right).$$

From (3.4) follows that

$$R_\varepsilon^{(p)}\left(\tilde{f}_\varepsilon, \delta_\varepsilon\right) \ge |\varphi_\varepsilon^*|^{-p} \mathcal{R}_\varepsilon$$

and it suffices to state that

$$\mathcal{R}_\varepsilon \ge |\varphi_\varepsilon^*|^p \frac{C_1}{a^{p/2}}.$$

Let Π_ε be the uniform measure on the finite set G_ε. Of course, Π_ε can be represented as the direct product of measures π_I ,

$$\Pi_\varepsilon = \prod_{I\in\mathcal{I}_\varepsilon} \pi_I, \tag{3.6}$$

where π_I is for each $I \in \mathcal{I}_\varepsilon$ the uniform measure on the two-points set $\{\pm g_I(\cdot)\}$ of functions on I.
Denote by $\mathcal{R}_\varepsilon(\Pi_\varepsilon)$ the Bayes risk for this prior Π_ε and the loss function from (3.5),

$$\mathcal{R}_\varepsilon(\Pi_\varepsilon) = \inf_{\tilde{f}_\varepsilon} \frac{1}{card\, G_\varepsilon} \sum_{f_\theta\in G_\varepsilon} E_{f_\theta} l\left(\tilde{f}_\varepsilon - f_\theta\right).$$

Now we decompose the whole Bayes risk $\mathcal{R}_\varepsilon(\Pi_\varepsilon)$ on the sum of Bayes risks for the submodels corresponding to each subinterval $I \in \mathcal{I}_\varepsilon$. This can be done because of the direct product structure of the original model (2.2) and the priors Π_ε as well as the additive structure of losses (3.5). Namely, let E_I be the submodel corresponding to the interval I i.e. E_I describes the observations $(X_\varepsilon(t),\, t \in I)$. Denote

$$r_\varepsilon(\pi_I) = \inf_{\tilde{f}_{\varepsilon,I}} \frac{1}{2}\left[E_{g_I} l_I\left(\tilde{f}_{\varepsilon,I} - g_I\right) + E_{-g_I} l_I\left(\tilde{f}_{\varepsilon,I} + g_I\right)\right]$$

where

$$l_I\left(\tilde{f}_{\varepsilon,I} - \pm g_I\right) = \int_{c_I-\delta_\varepsilon}^{c_I+\delta_\varepsilon} \left|\tilde{f}_{\varepsilon,I}(t) - \pm g_I(t)\right|^p dt$$

and $\tilde{f}_{\varepsilon,I}$ is an estimator of the signal $\pm g_I$ for the model $\mathcal{E}_I$ (i.e. by the observations $X_\varepsilon(t)$, $t \in I$). Obviously these observations are sufficient for this subproblem and one can not improve the risk of observations taking into account the observations beyond I.

Lemma 2. *The following assertions are satisfied:*

(1)

$$\mathcal{R}_\varepsilon(\Pi_\varepsilon) = \sum_{I\in\mathcal{I}_\varepsilon} r_\varepsilon(\pi_I).$$

(2)

$$r_\varepsilon(\pi_I) = |I|^{\beta_\varepsilon p+1} b\left(\varepsilon^{-1}|I|^{\beta_\varepsilon+1/2}\right) = \Delta_\varepsilon^{\beta_\varepsilon p+1} b(1).$$

Proof. The first assertion is based on (3.5), (3.6) and the structure of the model (2.2). The second assertion is obtained by simple renormalization arguments, the definition (3.3) and the equalities $|I| = \Delta_\varepsilon$ and $\Delta_\varepsilon^{\beta_\varepsilon+1/2} = \varepsilon$ (see M.Low, 1991). We omit the details. □

Therefore,

$$\mathcal{R}_\varepsilon \geq \mathcal{R}_\varepsilon(\Pi_\varepsilon) = \sum_{I\in\mathcal{I}_\varepsilon} r_\varepsilon(\pi_I) = card\,\mathcal{I}_\varepsilon\, \Delta_\varepsilon^{\beta_\varepsilon p+1} b(1) = b(1)\left(\Delta_\varepsilon^{\beta_\varepsilon}\right)^p = b(1)\left(\varphi_\varepsilon\right)^p.$$

But from (3.1) and (3.2)

$$\frac{\varphi_\varepsilon}{\varphi_\varepsilon^*} = \sqrt{\frac{\delta_\varepsilon^*}{\Delta_\varepsilon}} = \sqrt{\frac{\delta_\varepsilon^*}{8\delta_\varepsilon}} \geq \sqrt{\frac{1}{8a}}.$$

Finally

$$\mathcal{R}_\varepsilon \geq (\varphi_\varepsilon^*)^p b(1) 8^{-p/2} \frac{1}{a^{p/2}}$$

and the theorem follows.

3.2. Proof of Theorem 2. First notice that without loss of generality on can suppose

$$\delta_\varepsilon = a\delta_\varepsilon^*.$$

Define now the values Δ_ε, β_ε, φ_ε by the equalities

$$\begin{aligned} \Delta_\varepsilon &= \left(\varepsilon\sqrt{\ln\tfrac{1}{\varepsilon}}\right)^{\frac{2}{2\beta_\varepsilon+1}} = 8\delta_\varepsilon, \\ \varphi_\varepsilon &= \Delta_\varepsilon^{\beta_\varepsilon} = \left(\varepsilon\sqrt{\ln\tfrac{1}{\varepsilon}}\right)^{\frac{2\beta_\varepsilon}{2\beta_\varepsilon+1}}. \end{aligned} \tag{3.7}$$

We will use later the relation

$$\varphi_\varepsilon\sqrt{\Delta_\varepsilon} = \varepsilon\sqrt{\ln\frac{1}{\varepsilon}}.$$

Let again $\mathcal{I}_\varepsilon$ be the partition of $[0,1]$ onto intervals I of length Δ_ε and again we suppose

$$card\,\mathcal{I}_\varepsilon = \Delta_\varepsilon^{-1}.$$

Let the function $g(\cdot)$ be defined on $\left[-\frac{1}{2},\frac{1}{2}\right]$ as before in the proof of Theorem 1. Put for each $I \in \mathcal{I}_\varepsilon$

$$g_I(t) = g_0\,|I|^{\beta_\varepsilon} g\left(\frac{t-c_I}{|I|}\right), \qquad t \in I,$$

c_I being the center of I, $|I| = \Delta_\varepsilon$, its length, and a constant g_0 will be chosen later. Let also the function $f_I(\cdot)$ on $[0,1]$ coincide with g_I on I and vanishe outside I.

Introduce the finite parameter set G_ε of function on $[0,1]$ with

$$G_\varepsilon = \{f_0, f_I,\ I \in \mathcal{I}_\varepsilon\}$$

where $f_0 \equiv 0$.
By Π_ε we denote the prior on G_ε with

$$\begin{aligned} \Pi_\varepsilon(f_0) &= \frac{1}{2}, \\ \Pi_\varepsilon(f_I) &= \frac{1}{2}\Delta_\varepsilon, \quad I \in \mathcal{I}_\varepsilon. \end{aligned}$$

Now we reduce again the whole problem to the subproblem corresponding to the finite subset G_ε. First note that as before one has for each $f \in G_\varepsilon$

$$\beta_f(t, \delta_\varepsilon) = \beta^*, \qquad t \in [c_I - \delta_\varepsilon, c_I + \delta_\varepsilon].$$

Define the loss of an estimator $\tilde{f}_\varepsilon$ at a point $f \in G_\varepsilon$ as follows

$$\left(\tilde{f}_\varepsilon - f\right) = \sup_{I \in \mathcal{I}_\varepsilon} \sup_{t \in [c_I - \delta_\varepsilon, c_I + \delta_\varepsilon]} \left|\tilde{f}_\varepsilon - f\right|.$$

Similarly to that in the proof of Theorem 1

$$R_\varepsilon^{(\infty)}\left(\tilde{f}_\varepsilon, \delta_\varepsilon\right) \ge |\varphi_\varepsilon^*|^{-1} \sup_{f \in G_\varepsilon} E_f\, l\left(\tilde{f}_\varepsilon - f\right) \ge |\varphi_\varepsilon^*|^{-1} \sum_{f \in G_\varepsilon} \Pi_\varepsilon(f)\, E_f\, l\left(\tilde{f}_\varepsilon - f\right).$$

Now notice that through (3.7)

$$\frac{\varphi_\varepsilon}{\varphi_\varepsilon^*} = \sqrt{\frac{\delta_\varepsilon^*}{\delta_\varepsilon}} = \frac{1}{\sqrt{8a}}$$

and it remains to verify that

$$R_\varepsilon'\left(\tilde{f}_\varepsilon\right) = \varphi_\varepsilon^{-1}\left[\frac{1}{2}E_0\, l\left(\tilde{f}_\varepsilon\right) + \frac{\Delta_\varepsilon}{2} \sum_{I \in \mathcal{I}_\varepsilon} E_f\, l\left(\tilde{f}_\varepsilon - f\right)\right] \ge C > 0,$$

where we write E_0 instead of E_{f_0} , and the observations X from (2.2) coincide with εW under the measure P_0.
Denote

$$Z_{\varepsilon,I} = \frac{dP_{f_I}}{dP_0}(X).$$

By Girsanov's formulae

$$Z_{\varepsilon,I} = \exp\left\{\varepsilon^{-1}\int_I g_I(t)\, dW(t) - \frac{\varepsilon^{-2}}{2}\int_I g_I^2(t)\, dt\right\}.$$

But

$$\int_I g_I^2(t)\, dt = g_0^2\, |I|^{2\beta_\varepsilon+1} \int_{-1/2}^{1/2} g^2(t)\, dt = \sigma^2 \Delta_\varepsilon^{2\beta_\varepsilon+1} = \sigma^2 \varepsilon^2 \ln \varepsilon^{-1} \tag{3.8}$$

with

$$\sigma^2 = g_0^2 \int_{-1/2}^{1/2} g^2(t)\, dt$$

and hence

$$Z_{\varepsilon,I} = \exp\left\{\sigma t_\varepsilon \zeta_{\varepsilon,I} - \frac{1}{2}\sigma^2 t_\varepsilon^2\right\} \tag{3.9}$$

with $t_\varepsilon = \sqrt{\ln \varepsilon^{-1}}$ and

$$\zeta_{\varepsilon,I} = \frac{1}{\sigma t_\varepsilon \varepsilon} \int_I g_I(t)\, dW(t) \sim N(0,1). \tag{3.10}$$

Moreover, $\zeta_{\varepsilon,I}$ are independent for different $I \in \mathcal{I}_\varepsilon$. Put

$$\begin{aligned} Z_\varepsilon &= \frac{1}{card\, \mathcal{I}_\varepsilon} \sum_{I\in\mathcal{I}_\varepsilon} Z_{\varepsilon,I} = \Delta_\varepsilon \sum_{I\in\mathcal{I}_\varepsilon} Z_{\varepsilon,I}\,, \\ A_\varepsilon &= \left\{ l\left(\tilde{f}_\varepsilon\right) \geq \frac{1}{2} g_0 \varphi_\varepsilon \right\}, \end{aligned}$$

and let $\bar{A}_\varepsilon$ be the complement of A_ε. Since

$$l\left(f_I\right) = g_0 \left|I\right|^{\beta_\varepsilon} = g_0 \varphi_\varepsilon$$

then under $\bar{A}_\varepsilon$ we have

$$l\left(\tilde{f}_\varepsilon - f\right) \geq \frac{1}{2} g_0 \varphi_\varepsilon.$$

Hence we obtain through (3.8)

$$\begin{aligned} R'_\varepsilon\left(\tilde{f}_\varepsilon\right) &\geq \frac{1}{2}\varphi_\varepsilon^{-1} E_0 \left[l\left(\tilde{f}_\varepsilon\right) + \Delta_\varepsilon \sum_{I\in\mathcal{I}_\varepsilon} Z_{\varepsilon,I}\, l\left(\tilde{f}_\varepsilon - f\right)\right] \geq \\ &\geq \frac{1}{2} E_0 \left[1\left(A_\varepsilon\right) + Z_\varepsilon 1\left(\bar{A}_\varepsilon\right)\right]. \end{aligned}$$

But, if $Z_\varepsilon \geq C$, then

$$1\left(A_\varepsilon\right) + Z_\varepsilon 1\left(\bar{A}_\varepsilon\right) \geq 1\left(A_\varepsilon\right) + c\, 1\left(\bar{A}_\varepsilon\right) = c + (1-c)\, 1\left(A_\varepsilon\right) \geq c$$

and thus

$$R'_\varepsilon\left(\tilde{f}_\varepsilon\right) \geq \frac{1}{2} c\, P_0\left(Z_\varepsilon \geq C\right).$$

Now the assertion of the theorem follows immediately from the fact that

$$Z_\varepsilon \to 1$$

under the measure P_0. Recall that Z_ε is the average of i.i.d. random variable $Z_{\varepsilon,I}$ defined by (3.9) and (3.10), and it suffices to verify that

$$\Delta_\varepsilon^2 E_0 Z_{\varepsilon,I}^2 \to 0, \qquad \varepsilon \to 0.$$

But

$$E_0 Z_{\varepsilon,I}^2 = e^{-\sigma^2 t_\varepsilon^2} E_0 e^{2\sigma t_\varepsilon \zeta_{\varepsilon,I}} = e^{\sigma^2 t_\varepsilon^2} = e^{\sigma^2 \ln \varepsilon^{-1}} = \varepsilon^{-\sigma^2}$$

and

$$\Delta_\varepsilon^2 E_0 Z_{\varepsilon,I}^2 = (8a)^2\, \varepsilon^{\frac{4}{2\beta^*+1} - \sigma^2} \to 0, \qquad \varepsilon \to 0$$

for $\sigma^2 < \frac{4}{2\beta^*+1}$. Finally, if the constant g_0 was chosen properly, then the result follows. □

Before to prove Theorems 3, 4 we recall some simple technical properties of the kernel estimators $f_\varepsilon\left(\cdot,\beta\right)$. First denote

$$K_\varepsilon(\beta) * f(t) = \frac{1}{\kappa_\varepsilon(\beta)} \int K\left(\frac{u-t}{\kappa_\varepsilon(\beta)}\right) f(u)\, du$$

and

$$\xi_\varepsilon\left(t,\beta\right) = \frac{\sqrt{\kappa_\varepsilon(\beta)}}{\varepsilon}\left[f_\varepsilon(t,\beta) - E_f\, f_\varepsilon\left(t,\beta\right)\right].$$

Lemma 3. *The following statements hold:*

(1)

$$E_f\, f_\varepsilon\,(t,\beta) = K_\varepsilon(\beta) * f(t);$$

(2) *For any subinterval I of $[0,1]$, any $\beta > 0$, and any function $f(\cdot)$ such that $f(\cdot) \in \Sigma\,(\beta, L, I_\delta)$ with $\delta = \kappa_\varepsilon(\beta)$ one has*

$$\|f(\cdot) - K_\varepsilon\,(\beta') * f(\cdot)\|_{p,I} \le L\,|\kappa_\varepsilon\,(\beta')|^{\beta'}, \qquad \forall \beta' \le \beta;$$

(3) *The distribution of the process $\xi_\varepsilon\,(\cdot,\beta) = (\xi_\varepsilon\,(t,\beta),\ t \in [0,1])$ under the measure P_f does not depend on f and coincides with the distribution of the process $\left(\xi\left(\frac{t}{\kappa_\varepsilon(\beta)}\right),\ t \in [0,1]\right)$ where*

$$\xi(s) = \int K(u-s)\,dW(u),$$

$(W(u),\ u \ge 0)$ being a standard Wiener process.

(4) *Let $p \in [1,\infty)$ and let I be any subinterval of $[0,1]$. Then*

$$E\,\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^p = E\,|\xi(0)|^p = C_p$$

where $C_p = [\int K^2(s)\,ds]^{p/2}\, E\,|N(0,1)|^p$;

$$E\,\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^{2p} \le C_{2p}$$

for some constant $C_{2p} \le C_p^2$;
Moreover, given $r \ge 1$, there exists such a constant $C(r)$ that

$$E\,\left|\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^p - C_p\right|^{r+1} \le \frac{C(r)}{|T_\varepsilon(\beta)|^r}$$

with

$$T_\varepsilon(\beta) = \frac{|I|}{\kappa_\varepsilon(\beta)}.$$

(5) *Let $p = \infty$. Then*

$$E\ \|\xi_\varepsilon(\cdot,\beta)\|_\infty \le C\sqrt{\ln\frac{1}{\kappa_\varepsilon(\beta)}},$$

$$E\ \|\xi_\varepsilon(\cdot,\beta)\|_\infty^2 \le C\ \ln\frac{1}{\kappa_\varepsilon(\beta)}$$

and

$$P\left(\|\xi_\varepsilon(\cdot,\beta)\|_\infty \le u\,\|K\|\right) \le C\frac{1}{\kappa_\varepsilon(\beta)}\exp\left\{-\frac{u^2}{2}\right\}$$

with $\|K\|^2 = \int K^2(s)\,ds$.

Proof. These assertions describe well known properties of the kernel estimator (see ??). The proof is standard and straightforward and we omit it. □

3.3. Proof of Theorem 3. We use essentially the additive structure of the loss function (2.3). Namely,

$$\begin{aligned} R_\varepsilon^{(p)}\left(\hat f_\varepsilon(\cdot),\delta_\varepsilon\right) &= \sup_{f\in\Sigma} E_f \left\| \frac{\hat f_\varepsilon(\cdot)-f(\cdot)}{\varphi_\varepsilon\left(\beta_f(\cdot,\delta_\varepsilon)\right)} \right\|_p^p = \\ &= \sup_{f\in\Sigma} E_f \frac{1}{card\,\mathcal{I}_\varepsilon} \sum_{I\in\mathcal{I}_\varepsilon} \left\| \frac{f_\varepsilon(\cdot,\hat\beta_\varepsilon(I))-f(\cdot)}{\varphi_\varepsilon(\beta_f(\cdot,\delta_\varepsilon))} \right\|_{p,I}^p \le \\ &\le \frac{1}{card\,\mathcal{I}_\varepsilon} \sum_{I\in\mathcal{I}_\varepsilon} \sup_{f\in\Sigma} E_f \left\| \frac{f_\varepsilon(\cdot,\hat\beta_\varepsilon(I))-f(\cdot)}{\varphi_\varepsilon(\beta_f(\cdot,\delta_\varepsilon))} \right\|_{p,I}^p . \end{aligned} \tag{3.11}$$

Hence it is enough to state that each summand in this sum is bounded by a universal constant C_2.
Let us fix some $I\in\mathcal{I}_\varepsilon$ and some $f\in\Sigma$, and put

$$\begin{aligned} \hat\beta &= \hat\beta_\varepsilon(I), \\ \beta_f &= \sup\left\{\beta\le\beta^* : f(\cdot)\in\Sigma(\beta,L,\tilde I)\right\}, \end{aligned}$$

$\tilde I = I_{\kappa_\varepsilon^*}$ being the κ_ε^*-neighborhood of I. Because $|I|=\delta_\varepsilon/2$ and $\kappa_\varepsilon^*\le\delta_\varepsilon/2$, for each $t\in I$ the interval $[t-\delta_\varepsilon,t+\delta_\varepsilon]$ contains $\tilde I$. This obviously implies $\beta_f(t,\delta_\varepsilon)\le\beta_f$ and hence permits to reduce through (3.11) the assertion of the theorem to the following statement:

$$\sup_{f\in\Sigma} E_f \left\| \frac{f_\varepsilon(\cdot,\hat\beta)-f(\cdot)}{\varphi_\varepsilon(\beta_f)} \right\|_{p,I}^p \le C. \tag{3.12}$$

We arrive to the typical problem of adaptive estimation. The function $f(\cdot)$ to be estimated is supposed (being restricted on the interval $\tilde I$) to belong to some smoothness class $\Sigma(\beta,L)$ with unknown β , and we claim possibility of adaptive estimation. But we cannot refer to the famous result on adaptive estimation, for example, Lepskii(91) since the length of the interval I is not fixed and goes to zero with the rate of bandwidth corresponding to the upper smoothness β^*.
Before to prove (3.12) we notice that one can assume β_f to take its value in B_ε. Otherwise we can replace β_f by the closest from below $\beta_f'\in B_\varepsilon$ (i.e. $0\le\beta_f$ - $\beta_f'\le h_\varepsilon$) and easy to check that $1\le\varphi_\varepsilon\left(\beta_f'\right)/\varphi_\varepsilon(\beta_f)\le C$, where, in particular, $C=e^2$ can be taken.
To prove (3.12) we decompose this expression onto two parts relating to the events $\left\{\hat\beta\ge\beta_f\right\}$ and $\left\{\hat\beta<\beta_f\right\}$:

$$E_f \left\| \frac{f_\varepsilon(\cdot,\hat\beta)-f(\cdot)}{\varphi_\varepsilon(\beta_f)} \right\|_{p,I}^p = R_\varepsilon^+(f)+R_\varepsilon^-(f)$$

with

$$R_\varepsilon^+(f) = E_f \left\| \frac{f_\varepsilon(\cdot,\hat\beta)-f(\cdot)}{\varphi_\varepsilon(\beta_f)} \right\|_{p,I}^p 1\left(\hat\beta\ge\beta_f\right)$$

and

$$R_\varepsilon^-(f) = E_f \left\| \frac{f_\varepsilon(\cdot,\hat\beta)-f(\cdot)}{\varphi_\varepsilon(\beta_f)} \right\|_{p,I}^p 1\left(\hat\beta<\beta_f\right).$$

We start with $R_\varepsilon^+(f)$. The definition of $\hat\beta$ implies

$$\left\{\hat\beta\ge\beta_f\right\} \subseteq \left\{ \left\| f_\varepsilon(\cdot,\hat\beta)-f_\varepsilon(\cdot,\beta_f)\right\|_{p,I} \le C_a\varphi_\varepsilon(\beta_f)\right\}.$$

Hence under the event $\left\{\hat{\beta} \geq \beta_f\right\}$ we obtain through (i) and (ii) of Lemma 3

$$\begin{aligned}
R_\varepsilon^+(f) &\leq E_f\left[C_a + \left\|\frac{f_\varepsilon(\cdot,\beta_f) - f(\cdot)}{\varphi_\varepsilon(\beta_f)}\right\|_{p,I}\right]^p \leq \\
&\leq E_f\left[C_a + \left\|\frac{f(\cdot) - K_\varepsilon(\beta_f) * f(\cdot)}{\varphi_\varepsilon(\beta_f)}\right\|_{p,I} + \right. \\
&\left. + \frac{\varepsilon}{\varphi_\varepsilon(\beta_f)\sqrt{\kappa_\varepsilon(\beta_f)}} \|\xi_\varepsilon(\cdot,\beta_f)\|_{p,I}\right]^p \leq \\
&\leq 3^p\left[C_a^p + \left(L\frac{|\kappa_\varepsilon(\beta_f)|^{\beta_f}}{\varphi_\varepsilon(\beta_f)}\right)^p + \left(\frac{a}{2}\right)^{-p/2} E_f\|\xi_\varepsilon(\cdot,\beta_f)\|_{p,I}^p\right] \leq \\
&\leq \frac{C}{a^{p/2}}.
\end{aligned}$$

Here we used the equalities

$$\begin{aligned}
|\kappa_\varepsilon(\beta_f)|^{\beta_f} &= \left(\frac{a}{2}\right)^\beta \varepsilon^{\frac{2\beta}{2\beta+1}} = \left(\frac{a}{2}\right)^\beta \varphi_\varepsilon(\beta) \leq \varphi_\varepsilon(\beta), \\
\varphi_\varepsilon(\beta_f)\sqrt{\kappa_\varepsilon(\beta_f)} &= \varepsilon\sqrt{\frac{a}{2}}.
\end{aligned} \tag{3.13}$$

Now we estimate $R_\varepsilon^-(f)$. One has

$$\begin{aligned}
R_\varepsilon^-(f) &= \sum_{\beta\in B_\varepsilon(\beta_f)} |\varphi_\varepsilon(\beta_f)|_f^{-p} E\left\|f_\varepsilon(\cdot,\hat{\beta}) - f(\cdot)\right\|_{p,I}^p 1\left(\hat{\beta} = \beta\right) \leq \\
&\leq \sum_{\beta\in B_\varepsilon(\beta_f)} |\varphi_\varepsilon(\beta_f)|^{-p}\left[E_f\left\|f_\varepsilon(\cdot,\hat{\beta}) - f(\cdot)\right\|_{p,I}^{2p} P_f\left(\hat{\beta} = \beta\right)\right]^{1/2}
\end{aligned}$$

. For each β as before we get (using also (iv) of Lemma 3)

$$E_f\left\|f_\varepsilon(\cdot,\hat{\beta}) - f(\cdot)\right\|_{p,I}^{2p} \leq \left(\frac{C}{a^{p/2}}\right)^2$$

and thus

$$R_\varepsilon^-(f) \leq \frac{C}{a^{p/2}} \sum_{\beta\in B_\varepsilon(\beta_f)} \left|\frac{\varphi_\varepsilon(\beta)}{\varphi_\varepsilon(\beta_f)}\right|^p \sqrt{P_f\left(\hat{\beta} = \beta\right)}. \tag{3.14}$$

Let us fix for a moment some $\beta \in B_\varepsilon(\beta_f)$ and denote

$$k = \frac{\beta^* - \beta}{h_\varepsilon}.$$

Then we have for $T_\varepsilon(\beta)$ defined in Lemma 3

$$\begin{aligned}
T_\varepsilon(\beta) &= \frac{\kappa_\varepsilon(\beta)}{\delta_\varepsilon} \geq \frac{\kappa_\varepsilon(\beta)}{\kappa_\varepsilon(\beta^*)} = \varepsilon^{\frac{2}{2\beta+1} - \frac{2}{2\beta^*+1}} = \varepsilon^{\frac{-4kh_\varepsilon}{(2\beta+1)(2\beta^*+1)}} = \\
&= \exp\left\{\frac{4k}{(2\beta+1)(2\beta^*+1)}\right\}
\end{aligned} \tag{3.15}$$

(since $\varepsilon^{-h_\varepsilon} = \left(\frac{1}{\varepsilon}\right)^{\left(\ln\frac{1}{\varepsilon}\right)^{-1}} = e$).

Let now $\beta' \in B_\varepsilon(\beta)$ (i.e. $\beta' \le \beta,\ \beta' \in B_\varepsilon$) and

$$\begin{aligned} k' &= \frac{\beta^* - \beta'}{h_\varepsilon}, \\ Z_\varepsilon(\beta', \beta) &= \frac{\varphi_\varepsilon(\beta')}{\varphi_\varepsilon(\beta)}. \end{aligned}$$

Similarly to (3.15) we obtain

$$Z_\varepsilon(\beta', \beta) = \varepsilon^{\frac{2\beta'}{2\beta'+1} - \frac{2\beta}{2\beta+1}} = \exp\left\{ \frac{2(k'-k)}{(2\beta'+1)(2\beta+1)} \right\}. \tag{3.16}$$

Again straightforwardly from the definition of $\hat\beta$

$$\left\{\hat\beta = \beta - h_\varepsilon\right\} \subseteq \bigcup_{\beta' \in B_\varepsilon(\beta)} \left\{ \|f_\varepsilon(\cdot, \beta') - f_\varepsilon(\cdot, \beta)\|_{p,I} \ge C_a \varphi_\varepsilon(\beta') \right\}. \tag{3.17}$$

Applying (i) - (iii) of Lemma 3 and (3.13) we derive

$$\begin{aligned} &\|f_\varepsilon(\cdot,\beta') - f_\varepsilon(\cdot,\beta)\|_{p,I} \le \\ &\le \ \|f_\varepsilon(\cdot,\beta') - E_f\, f_\varepsilon(\cdot,\beta')\|_{p,I} + \|f_\varepsilon(\cdot,\beta) - E_f\, f_\varepsilon(\cdot,\beta)\|_{p,I} + \\ &\quad + \|f(\cdot) - K_\varepsilon(\beta') * f(\cdot)\|_{p,I} + \|f(\cdot) - K_\varepsilon(\beta) * f(\cdot)\|_{p,I} \le \\ &\le \ \frac{\varepsilon}{\sqrt{\kappa_\varepsilon(\beta')}} \|\xi_\varepsilon(\cdot,\beta')\|_{p,I} + \frac{\varepsilon}{\sqrt{\kappa_\varepsilon(\beta)}} \|\xi_\varepsilon(\cdot,\beta)\|_{p,I} + \\ &\quad + L\left[|\kappa_\varepsilon(\beta')|^{\beta'} + |\kappa_\varepsilon(\beta)|^{\beta}\right] \le \\ &\le \ \frac{1}{\sqrt{a/2}} \left[\varphi_\varepsilon(\beta')\, \|\xi_\varepsilon(\cdot,\beta')\|_{p,I} + \varphi_\varepsilon(\beta)\, \|\xi_\varepsilon(\cdot,\beta)\|_{p,I}\right] + \\ &\quad + L\left[\varphi_\varepsilon(\beta') + \varphi_\varepsilon(\beta)\right]. \end{aligned} \tag{3.18}$$

Recall that $C_a = 2L + \frac{2}{\sqrt{a/2}}(1 + C_p)^{1/p}$ and by definition $\varphi_\varepsilon(\beta') \le \varphi_\varepsilon(\beta)$ for $\beta' \le \beta$. Hence (3.17) and (3.18) readily imply

$$\begin{aligned} P_f\left(\hat\beta = \beta - h_\varepsilon\right) &\le \sum_{\beta' \in B_\varepsilon(\beta)} P\left(\|\xi_\varepsilon(\cdot,\beta')\|_{p,I}^p \ge 1 + C_p\right) + \\ &+ \sum_{\beta' \in B_\varepsilon(\beta)} P\left(\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^p \ge (1 + C_p)\, Z_\varepsilon^p(\beta',\beta)\right). \end{aligned} \tag{3.19}$$

Through (iv) of Lemma 3 and (3.15) for each $r \ge 1$

$$\begin{aligned} &\sum_{\beta' \in B_\varepsilon(\beta)} P\left(\|\xi_\varepsilon(\cdot,\beta')\|_{p,I}^p - C_p \ge 1\right) \le \\ &\le \ \sum_{\beta' \in B_\varepsilon(\beta)} E\left|\|\xi_\varepsilon(\cdot,\beta')\|_{p,I}^p - C_p\right|^{r+1} \le \\ &\le \ \sum_{\beta' \in B_\varepsilon(\beta)} \frac{C(r)}{|T_\varepsilon(\beta')|^r} \le \\ &\le \ \sum_{k'=k}^{\infty} C(r) \exp\left\{\frac{-4k'r}{(2\beta'+1)(2\beta^*+1)}\right\} \le \\ &\le \ C \exp\left\{\frac{-4kr}{(2\beta+1)(2\beta^*+1)}\right\}. \end{aligned} \tag{3.20}$$

Similarly through (3.16)

$$
\begin{aligned}
&\sum_{\beta'\in B_\varepsilon(\beta)} P\left(\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^p \geq (1+C_p)\, Z_\varepsilon^p(\beta',\beta)\right) \leq \\
&\leq \sum_{\beta'\in B_\varepsilon(\beta)} P\left(\|\xi_\varepsilon(\cdot,\beta)\|_{p,I}^p - C_p \geq Z_\varepsilon^p(\beta',\beta)\right) \leq \\
&\leq \sum_{\beta'\in B_\varepsilon(\beta)} \frac{C(r)}{|T_\varepsilon(\beta')|^r\, |Z_\varepsilon^p(\beta',\beta)|^{r+1}} \leq \\
&\leq \sum_{k'=k}^{\infty} C(r)\exp\left\{\frac{-4kr}{(2\beta+1)(2\beta^*+1)} - \frac{-2(k'-k)(r+1)}{(2\beta'+1)(2\beta+1)}\right\} \leq \\
&\leq C\exp\left\{\frac{-4kr}{(2\beta+1)(2\beta^*+1)}\right\}. \qquad (3.21)
\end{aligned}
$$

Combining (3.19) - (3.21) we conclude

$$
P_f\left(\hat\beta = \beta - h_\varepsilon\right) \leq C\exp\left\{\frac{-4kr}{(2\beta+1)(2\beta^*+1)}\right\} \leq C\exp\{-2kp\}
$$

if $r \geq \frac{1}{2}p\,(2\beta^*+1)^2$.

Substituting this inequality in (3.14) and using again (3.16) we arrive to the final calculation

$$
\begin{aligned}
R_\varepsilon^-(f) &\leq \frac{C}{a^{p/2}} \sum_{\beta\in B_\varepsilon(\beta_f)} |Z_\varepsilon(\beta-h_\varepsilon,\beta_f)|^p \sqrt{P_f\left(\hat\beta=\beta-h_\varepsilon\right)} \leq \\
&\leq \frac{C}{a^{p/2}} \sum_{k=k_f}^{\infty} C(r)\exp\left\{\frac{p(k-k_f-1)}{(2\beta+1)(2\beta_f+1)} - kp\right\} \leq \frac{C}{a^{p/2}}
\end{aligned}
$$

and the theorem follows. □

3.4. Proof of Theorem 4. Given $f\in\Sigma$ we have the following representation for the risk of the estimator $\hat f_\varepsilon(\cdot)$ "at the point" $f(\cdot)$:

$$
R_\varepsilon(f) = E_f \sup_{I\in\mathcal{I}_\varepsilon} \left\| \frac{f_\varepsilon(\cdot,\hat\beta_\varepsilon(I)) - f(\cdot)}{\varphi_\varepsilon(\beta_f(\cdot,\delta_\varepsilon))} \right\|_{\infty,I}.
$$

Similarly to above denote for each $I\in\mathcal{I}_\varepsilon$

$$
\beta_f(I) = \sup\{\beta\leq\beta^* : f(\cdot)\in\Sigma(\beta,L,I_\delta)\}
$$

with $\delta=\kappa_\varepsilon(\beta^*)$. Arguing as in the case of $p<\infty$ we infer that the following assertion is sufficient for the whole theorem:

$$
\sup_{f\in\Sigma} E_f \sup_{I\in\mathcal{I}_\varepsilon} \left\| \frac{f_\varepsilon\left(\cdot,\hat\beta_\varepsilon(I)\right) - f(\cdot)}{\varphi_\varepsilon(\beta_f(I))} \right\|_{\infty,I} \leq \frac{C}{\sqrt{a}}.
$$

Fix for a moment some $I\in\mathcal{I}_\varepsilon$ and denote

$$
\begin{aligned}
\hat\beta &= \hat\beta_\varepsilon(I),\\
\beta_f &= \beta_f(I).
\end{aligned}
$$

Further,

$$\begin{aligned}\left\|f_\varepsilon(\cdot,\hat\beta)-f(\cdot)\right\|_{\infty,I} &= \left\|f_\varepsilon(\cdot,\hat\beta)-f(\cdot)\right\|_{\infty,I} 1\left(\hat\beta\ge\beta_f\right)+\\ &+ \left\|f_\varepsilon(\cdot,\hat\beta)-f(\cdot)\right\|_{\infty,I} 1\left(\hat\beta<\beta_f\right).\end{aligned}$$

Using the definition of $\hat\beta$ and (ii) of Lemma 3 we derive

$$\begin{aligned}&\left\|f_\varepsilon(\cdot,\hat\beta)-f(\cdot)\right\|_{\infty,I} 1\left(\hat\beta\ge\beta_f\right)\le\\ &\le C_a\varphi_\varepsilon(\beta_f)+\|f_\varepsilon(\cdot,\beta_f)-f(\cdot)\|_{\infty,I}\le\\ &\le C_a\varphi_\varepsilon(\beta_f)+L\,|\kappa_\varepsilon(\beta_f)|^{\beta_f}+\frac{\varepsilon}{\sqrt{\kappa_\varepsilon(\beta_f)}}\|\xi_\varepsilon(\cdot,\beta_f)\|_{\infty,I}\le\\ &\le \frac{1}{\sqrt{a/2}}\varphi_\varepsilon(\beta_f)\left[C_a+\frac{1}{t_\varepsilon}\|\xi_\varepsilon(\cdot,\beta_f)\|_{\infty,I}\right].\end{aligned}\tag{3.22}$$

(Here we used the notation $t_\varepsilon=\sqrt{\ln\frac{1}{\varepsilon}}$ and the relations (3.1)).
Now consider the loss corresponding to the inverse event $\left\{\hat\beta<\beta_f\right\}$. The definition of $\hat\beta$ yields

$$\left\{\hat\beta<\beta_f\right\}=\bigcup_{\beta\in B_\varepsilon(\beta_f)}\left\{\hat\beta=\beta\right\}\tag{3.23}$$

and

$$\left\{\hat\beta=\beta\right\}\subseteq\bigcup_{\beta'\in B_\varepsilon(\beta)}\left\{\|f_\varepsilon(\cdot,\beta')-f_\varepsilon(\cdot,\beta^+)\|_{\infty,I}>C_a\varphi_\varepsilon(\beta')\right\}\tag{3.24}$$

with $\beta^+=\beta+h_\varepsilon$.
Using the definition $C_a=2L+\frac{8}{\sqrt{a/2}}$ similarly to the case of $p<\infty$ we obtain

$$\left\{\hat\beta=\beta\right\}\subseteq\bigcup_{\beta'\in B_\varepsilon(\beta)}A_{\varepsilon,I}(\beta',\beta)\tag{3.25}$$

with

$$A_{\varepsilon,I}(\beta',\beta)=\left\{\|\xi_\varepsilon(\cdot,\beta')\|_{\infty,I}>4t_\varepsilon\right\}\cup\left\{\|\xi_\varepsilon(\cdot,\beta')\|_{\infty,I}>4t_\varepsilon\,Z_\varepsilon(\beta',\beta^+)\right\}\tag{3.26}$$

where for each $\beta'<\beta$

$$Z_\varepsilon(\beta',\beta)=\frac{\varphi_\varepsilon(\beta')}{\varphi_\varepsilon(\beta)}=\left(\varepsilon\sqrt{\ln\frac{1}{\varepsilon}}\right)^{\frac{2\beta}{2\beta+1}-\frac{2\beta'}{2\beta'+1}}\le\varepsilon^{-1}.\tag{3.27}$$

Now, again by (ii) of Lemma 3 for each $\beta<\beta_f$

$$\varphi_\varepsilon^{-1}(\beta)\,\|f_\varepsilon(\cdot,\beta)-f(\cdot)\|_{\infty,I}\le\zeta_{\varepsilon,I}(\beta)\tag{3.28}$$

with

$$\zeta_{\varepsilon,I}(\beta)=L+\frac{1}{t_\varepsilon\sqrt{a/2}}\|\xi_\varepsilon(\cdot,\beta)\|_{\infty,I},\tag{3.29}$$

and, therefore, through (3.23) - (3.29)

$$\begin{aligned}
&\varphi_\varepsilon^{-1}(\beta_f)\left\|f_\varepsilon(\cdot,\hat\beta)-f(\cdot)\right\|_{\infty,I}1\left(\hat\beta<\beta_f\right)\le\\
&\le\sum_{\beta\in B_\varepsilon(\beta_f)}Z_\varepsilon(\beta,\beta_f)\varphi_\varepsilon^{-1}(\beta)\left\|f_\varepsilon(\cdot,\beta)-f(\cdot)\right\|_{\infty,I}1\left(\hat\beta=\beta\right)\le\\
&\le\sum_{\beta\in B_\varepsilon(\beta_f)}\sum_{\beta'\in B_\varepsilon(\beta)}\varepsilon^{-1}\zeta_{\varepsilon,I}(\beta)\,1\left(A_{\varepsilon,I}(\beta',\beta)\right). \qquad (3.30)
\end{aligned}$$

The next step is to evaluate the risk for the whole interval $[0,1]$. Put

$$R_\varepsilon^-(f)=E_f\sup_{I\in\mathcal{I}_\varepsilon}\varphi_\varepsilon^{-1}(\beta_f(I))\left\|f_\varepsilon\left(\cdot,\hat\beta_\varepsilon(I)\right)-f(\cdot)\right\|_{\infty,I}1\left(\hat\beta_\varepsilon(I)<\beta_f(I)\right).$$

From (3.30) and (3.26)

$$\begin{aligned}
R_\varepsilon^-(f)&\le\varepsilon_f^{-1}E\sup_{I\in\mathcal{I}_\varepsilon}\sum_{\beta\in B_\varepsilon}\sum_{\beta'\in B_\varepsilon(\beta)}\zeta_{\varepsilon,I}(\beta)\,1\left(A_{\varepsilon,I}(\beta',\beta)\right)\le\\
&\le\varepsilon_f^{-1}E\sum_{\beta\in B_\varepsilon}\sum_{\beta'\in B_\varepsilon(\beta)}\zeta_\varepsilon(\beta)\,1\left(A_\varepsilon(\beta',\beta)\right) \qquad (3.31)
\end{aligned}$$

where through (3.29)

$$\zeta_\varepsilon(\beta)=\sup_{I\in\mathcal{I}_\varepsilon}\zeta_{\varepsilon,I}(\beta)=C+\frac{1}{t_\varepsilon\sqrt{a/2}}\left\|\xi_\varepsilon(\cdot,\beta)\right\|_\infty \qquad (3.32)$$

and

$$\begin{aligned}
A_\varepsilon(\beta',\beta)&\le\bigcup_{I\in\mathcal{I}_\varepsilon}A_{\varepsilon,I}(\beta',\beta)\subseteq\\
&\subseteq\left\{\left\|\xi_\varepsilon(\cdot,\beta')\right\|_\infty>4t_\varepsilon\right\}\cup\left\{\left\|\xi_\varepsilon(\cdot,\beta^+)\right\|_\infty>4t_\varepsilon\right\}. \qquad (3.33)
\end{aligned}$$

Using Cauchy - Schwartz inequality we get from (3.31)

$$R_\varepsilon^-(f)\le\varepsilon^{-1}\sum_{\beta\in B_\varepsilon}\sum_{\beta'\in B_\varepsilon(\beta)}\left[E_f\,\zeta_\varepsilon^2(\beta)\,P_f\left(A_\varepsilon(\beta',\beta)\right)\right]^{1/2}. \qquad (3.34)$$

By (3.32) and (v) of Lemma 3 for ε small enough

$$\begin{aligned}
\left[E_f\,\zeta_\varepsilon^2(\beta)\right]^{1/2}&\le C+\frac{1}{\sqrt{a/2}}\left[\frac{C}{t_\varepsilon^2}\ln\kappa_\varepsilon^{-1}(\beta)\right]^{1/2}\le\\
&\le\frac{C}{\sqrt{a/2}\ln\frac{1}{\varepsilon}}\left[\ln\frac{2}{a}+\frac{2}{2\beta+1}\ln\frac{1}{\varepsilon}\right]\le\frac{C}{\sqrt{a}}.
\end{aligned}$$

Using once more (v) of Lemma 3 we have also

$$\begin{aligned}
P\left(\left\|\xi_\varepsilon(\cdot,\beta')\right\|_\infty>4t_\varepsilon\right)&\le\frac{C}{\kappa_\varepsilon(\beta)}\exp\left\{-\frac{(4t_\varepsilon)^2}{2}\right\}\le\\
&\le C\varepsilon^{-\frac{2}{2\beta+1}}\exp\left\{-8\ln\frac{1}{\varepsilon}\right\}\le C\varepsilon^6. \qquad (3.35)
\end{aligned}$$

Since the cardinality of B_ε is not greater than $\beta^*/h_\varepsilon=\beta^*\ln\frac{1}{\varepsilon}$ we conclude through (3.33) - (3.35) that

$$R_\varepsilon^-(f)\le\left(\beta^*\ln\frac{1}{\varepsilon}\right)^2\varepsilon^{-1}\frac{C}{\sqrt{a}}\left(2\varepsilon^6\right)^{1/2}\to0,\qquad\varepsilon\to0. \qquad (3.36)$$

It remains to estimate the risk $R_\varepsilon^+(f)$ with

$$R_\varepsilon^+(f) = E_f \sup_{I\in\mathcal{I}_\varepsilon} \varphi_\varepsilon^{-1}(\beta_f(I)) \left\| f_\varepsilon\left(\cdot, \hat\beta_\varepsilon(I)\right) - f(\cdot)\right\|_{\infty,I} 1\left(\hat\beta_\varepsilon(I) \geq \beta_f(I)\right).$$

We deduce straightforwardly from (3.22) that it suffices to verify only the following assertion:

$$\sup_{f\in\Sigma} E_f \sup_{I\in\mathcal{I}_\varepsilon} \frac{1}{t_\varepsilon} \|\xi_\varepsilon(\cdot,\beta_f(I))\|_{\infty,I} \leq C. \tag{3.37}$$

Let N_ε be the cardinality of $\mathcal{I}_\varepsilon$. Then

$$N_\varepsilon = \frac{2}{\delta_\varepsilon} \leq \frac{1}{a}\varepsilon^{-\frac{2}{2\beta^*+1}} < \varepsilon^{-2}$$

if ε is small enough.
For each $f \in \Sigma$ by (v) of Lemma 3

$$\begin{aligned} P_f\left(\sup_{I\in\mathcal{I}_\varepsilon} \frac{1}{t_\varepsilon} \|\xi_\varepsilon(\cdot,\beta_f(I))\|_{\infty,I} > u\right) &\leq \\ \leq\ & N_\varepsilon \sup_{I\in\mathcal{I}_\varepsilon} P_f\left(\|\xi_\varepsilon(\cdot,\beta_f(I))\|_\infty > ut_\varepsilon\right) \leq \\ \leq\ & \varepsilon^{-2}\frac{C}{\kappa_\varepsilon(\beta)}\exp\left\{-\frac{u^2}{2}\ln\frac{1}{\varepsilon}\right\} \leq \\ \leq\ & C\varepsilon^{\frac{u^2}{2}-4} \end{aligned}$$

and obviously

$$\int_3^\infty \varepsilon^{\frac{u^2}{2}-4} du \to 0, \qquad \varepsilon \to 0,$$

that implies (3.37). $\square$

References

1. Bretagnolle,J. and Huber,C. Estmation des densites: Risque minimax *Z. Wahrsch. Verw.Gebiete* **47** 119-137
2. Donoho,D.L. and Johnstone,I.M. (1992a) Ideal spatial adaptation by wavelet shrinkage *Technical Report 400* Dep. of Statistics, Stanford University
3. Donoho,D.L. and Johnstone,I.M. (1992b) Minimax estimation via wavelet shrinkage *Technical Report 402* Dep. of Statistics, Stanford University
4. Efroimovich,S. and Pinsker,M.S.(1984) Learning algorithm for nonparametric filtering. *Automation and Remote Control* **11** 1434-1440
5. Golubev,G.K.(1990) Quasilinear estimates of signal in L_2. *Prob. of Info. Transmission* **26** 15-20
6. Härdle,W. and Marron,J.S.(1985) Optimal bandwidth selection in nonparametric regression function estimation *Ann. Statist.* **12** 1466-1481
7. Ibragimov,I.A. and Khasminskii,R.Z. (1980) Estimates of signal, its derivatives, and point of maximum for Gaussian observations *Treory Probab. Appl.* **25** N.4 703-716
8. Ibragimov,I.A. and Khasminskii,R.Z. (1981) *Statistical Estimation. Asymptotic Theory* Springer New York
9. Kerkyacharian,G. and Picard,D.(1993) Density estimation by kernel and wavelet method, optimality in Besov space *Statistics and Probabilty Letters* **18** 327-336
10. Lepskii,O.V.(1990) One problem of adaptive estimation in Gaussian white noise. *Treory Probab. Appl.* **35** N.3 459-470
11. Lepskii,O.V.(1991) Asymptotic minimax adaptive estimation. 1. Upper bounds. *Treory Probab. Appl.* **36** N.4 645-659
12. Lepskii,O.V.(1992) Asymptotic minimax adaptive estimation. 2. Statistical model without optimal adaptation. Adaptive estimators *Treory Probab. Appl.* **37** N.3 468-481

13. Low,M.G.(1992) Renormalizing upper and lower bounds for integrated risk in the white noise model *Ann.Statist.* **21** 577-589
14. Poljak,B.T. and Tsybakov,A.B. Asymptotic optimality of C_p-test for the orthogonal series estimation of regression *Treory Probab. Appl.* **35** N.2 293-306
15. Stone,C.J.(1982) Optimal global rates of convergence for nonparametric regression *Ann. Statist.* **10** N.4 1040-1053

INSTITUTE FOR SYSTEM ANALYSIS, PROSPEKT 60-LET OKTJABRJA, 9, MOSCOW, 117312, RUSSIA

INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS, ERMOLOVOY 19, MOSCOW,101447, RUSSIA, AND INSTITUTE OF APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTR. 39, 10117, BERLIN, GERMANY

Recent publications of the
Institut für Angewandte Analysis und Stochastik

Preprints 1993

63. Joachim Förste: Zur mathematischen Modellierung eines Halbleiterinjektionslasers mit Hilfe der Maxwellschen Gleichungen bei gegebener Stromverteilung.

64. Andreas Juhl: On the functional equations of dynamical theta functions I.

65. Jürgen Borchardt, Ingo Bremer: Zur Analyse großer strukturierter chemischer Reaktionssysteme mit Waveform–Iterationsverfahren.

66. Günther Albinus, Hans–Christoph Kaiser, Joachim Rehberg: On stationary Schrödinger–Poisson equations.

67. Jörg Schmeling, Reinhard Winkler: Typical dimension of the graph of certain functions.

68. Ale Jan Homburg: On the computation of hyperbolic sets and their invariant manifolds.

69. John W. Barrett, Peter Knabner: Finite element approximation of transport of reactive solutes in porous media. Part 2: Error estimates for equilibrium adsorption processes.

70. Herbert Gajewski, Willi Jäger, Alexander Koshelev: About loss of regularity and "blow up" of solutions for quasilinear parabolic systems.

71. Friedrich Grund: Numerical solution of hierarchically structured systems of algebraic–differential equations.

72. Henri Schurz: Mean square stability for discrete linear stochastic systems.

73. Roger Tribe: A travelling wave solution to the Kolmogorov equation with noise.

74. Roger Tribe: The long term behavior of a Stochastic PDE.

75. Annegret Glitzky, Konrad Gröger, Rolf Hünlich: Rothe's method for equations modelling transport of dopants in semiconductors.

76. Wolfgang Dahmen, Bernd Kleemann, Siegfried Prößdorf, Reinhold Schneider: A multiscale method for the double layer potential equation on a polyhedron.

77. Hans–Günter Bothe: Attractors of non invertible maps.

78. Gregori Milstein, Michael Nussbaum: Autoregression approximation of a nonparametric diffusion model.

Preprints 1994

79. Anton Bovier, Véronique Gayrard, Pierre Picco: Gibbs states of the Hopfield model in the regime of perfect memory.

80. Roland Duduchava, Siegfried Prößdorf: On the approximation of singular integral equations by equations with smooth kernels.

81. Klaus Fleischmann, Jean–François Le Gall: A new approach to the single point catalytic super–Brownian motion.

82. Anton Bovier, Jean–Michel Ghez: Remarks on the spectral properties of tight binding and Kronig–Penney models with substitution sequences.

83. Klaus Matthes, Rainer Siegmund–Schultze, Anton Wakolbinger: Recurrence of ancestral lines and offspring trees in time stationary branching populations.

84. Karmeshu, Henri Schurz: Moment evolution of the outflow–rate from nonlinear conceptual reservoirs.

85. Wolfdietrich Müller, Klaus R. Schneider: Feedback stabilization of nonlinear discrete–time systems.

86. Gennadii A. Leonov: A method of constructing of dynamical systems with bounded nonperiodic trajectories.

87. Gennadii A. Leonov: Pendulum with positive and negative dry friction. Continuum of homoclinic orbits.

88. Reiner Lauterbach, Jan A. Sanders: Bifurcation analysis for spherically symmetric systems using invariant theory.

89. Milan Kučera: Stability of bifurcating periodic solutions of differential inequalities in $\mathbb{R}^3$.

90. Peter Knabner, Cornelius J. van Duijn, Sabine Hengst: An analysis of crystal dissolution fronts in flows through porous media Part I: Homogeneous charge distribution.

91. Werner Horn, Philippe Laurençot, Jürgen Sprekels: Global solutions to a Penrose–Fife phase–field model under flux boundary conditions for the inverse temperature.