Weierstraß–Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Adaptive and spatially adaptive testing of a nonparametric hypothesis

Vladimir G. Spokoiny

submitted: 17th April 1996

Weierstrass Institute
for Applied Analysis
and Stochastics
Mohrenstraße 39
D – 10117 Berlin
Germany

Preprint No. 234
Berlin 1996

1991 Mathematics Subject Classification. 62G10; Secondary 62G20.
Key words and phrases. Signal detection, minimax hypothesis testing, nonparametric alternative, error probabilities, adaptive test.

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

Fax: + 49 30 2044975
e–mail (X.400): c=de;a=d400-gw;p=WIAS-BERLIN;s=preprint
e–mail (Internet): preprint@wias-berlin.de

ABSTRACT. The present paper continues studying the problem of nonparametric hypothesis testing started in Lepski and Spokoiny, 1995 and Spokoiny, 1995. Let a function f be observed with noise. A null simple hypothesis $f \equiv f_0$ is tested against a composite alternative of the form $\|f - f_0\|_r \geq \varrho$. Additionally it is assumed that the underlying function f possesses some smoothness properties, namely, that f belongs to some Besov (or Sobolev) ball $B_{s,p,q}(M) = \{f : \|f\|_{B_{s,p,q}} \leq M\}$. The aim is to evaluate the fastest rate of decay of the radius ϱ to zero as the noise level tends to zero (or, equivalently, as the number of observations tends to infinity) for which testing with prescribed error probabilities is still possible. The ealier results show that the answer depends heavily on the smoothness parameters s, p, q, M. Below we consider the problem of adaptive (assumption free) testing if these parameters are unknown. A test ϕ^* is proposed which is near minimax and adaptive at the same time. Compared with the optimal (minimax) rate, this test has a performance which is worse within a $\log\log$-factor that is inessential but unavoidable payment for adaptation.

1. Introduction

Let a function f be observed with noise on the interval $[0, 1]$. More precisely, we observe a process $X(t)$ for $t \in [0, 1]$ obeying the stochastic equation

$$dX(t) = f(t)dt + n^{-1/2}dW(t), \qquad 0 \leq t \leq 1. \tag{1.1}$$

The factor $n^{-1/2}$ for the noise level is prompt by analogy between this "ideal" statistical model and more realistic models such as the regression model, distribution or spectral density model etc. where n has meaning of the number of observations.

In what follows we consider the problem of testing the null hypothesis about function f. Namely, it is assumed that the function f is completely specified under the null hypothesis. To be more definitive, we assume the null of the form $H_0 : f \equiv 0$. Such a problem is classical in statistical inference, see e.g. Mann and Wald, 1942 or Lehmann, 1959, and it can be treated as the signal detection problem when one has to decide by observations $X(t)$ whether "a signal" f presents or not. Note that the whole statement of the testing problem includes also description of the alternative set. Our aim is to propose such a test which is powerful against as large set H_1 of alternatives as possible.

We would like to consider below alternative sets of the form $H_1(\varrho) : \|f\|_r \geq \varrho$ where $r \geq 1$ and $\varrho > 0$. Unfortunately, such sets of alternatives are too large. It was shown in Burnashev, 1979; Ibragimov and Khasminskii, 1977; Ingster, 1982 that for any $\varrho > 0$ any test has the trivial power on $H_1(\varrho)$ (in the minimax sense). Further progress in this direction was made under the extra assumption that the function f is smooth in the sense that it belongs to some ball in Sobolev, Hölder of Besov space. Hence we look at the alternative of the form

$$H_1: \qquad \|f\|_r \geq \varrho, \quad \|f\|_{B_{s,p,q}} \leq M.$$

Here $\|\cdot\|_{B_{s,p,q}}$ is the Besov norm with parameters s, p, q, see e.g. Triebel, 1992 or Daubechies, 1992. For the case of an integer s and $p = q$, one may apply Sobolev's type seminorm $\|f\|_{H_{s,p}} = \left(\int |f^{(s)}(t)|^p dt\right)^{1/p}$ instead of the mentioned Besov norm. For this case, the parameter s could be viewed as the number of derivatives of the function f integrated in L_p-norm.

The problem is studied in the asymptotic set-up when n tends to infinity or, equivalently, the noise level tends to zero. It is natural to expect that for large n the radius ϱ can be taken small. Therefore, one may ask about the optimal (fastest) rate of decay of $\varrho = \varrho(n)$ to zero as $n \to \infty$ for which a nontrivial testing is still possible. It turned out that the optimal rate depends critically on the smoothness degree s, the norm power r in which we measure the distance between the null and the alternative set and the norm power p in which we measure smoothness properties of the function f.

The reader is referred to Ingster, 1993; Lepski and Spokoiny, 1995 and Spokoiny, 1995 for more detailed historical background. We mention here only a few results closely related to the problem under consideration.

The first results concerning the problem of minimax nonparametric hypothesis testing were obtained for the case when the distance between the null and the alternative set and smoothness properties of the function f are measured in the same norm L_2 (i.e. $r = p = 2$), see Ingster, 1982; Ingster, 1986; Ermakov, 1990. In this case, the optimal rate is $\varrho(n) = n^{-\frac{2s}{4s+1}}$. Note that this rate differs from the rate $\psi(n) = n^{-\frac{s}{2s+1}}$ of estimation of the function f under the same smoothness assumptions (more precisely, the rate of testing is better and it corresponds to the rate of estimation with smoothness parameter $2s$). Next, Ingster, 1986; Ingster, 1993 evaluated the rate of testing if $p = r$ and arbitrary $r \ge 1$. It appeared to be $\varrho(n) = n^{-\frac{s}{2s-1/r+1}}$ if $r \ge 2$ and again $\varrho(n) = n^{-\frac{2s}{4s+1}}$ if $1 \le r < 2$. Such a behavior differs essentially from that of for the estimation problem where, say, for any $p \ge r$, the rate of estimation is the same, namely $n^{-\frac{s}{2s+1}}$. It was shown also in Ingster, 1986 that for the case with $p \ge 2$ and $r \le 2$, the rate of testing is the same as for $p = r = 2$.

Note, however, that the case of $p < r$ was not considered in that papers. But the recent progress in the estimation theory, see Nemirovski, 1985; Donoho and Johnstone, 1995; Donoho et al., 1995; Lepski et al., 1994 stressed importance, both for theory and practice, of considering the situation with $p < r$ corresponding to functions f from (1.1) with inhomogeneous smoothness properties.

The hypothesis testing problem for the case with $p < r = 2$ was studied in Lepski and Spokoiny, 1995. The corresponding rate was found to be of the form $\varrho(n) = n^{-\frac{2s-1/p+1/r}{2(2s-1/p+1)}}$. Note, see Donoho and Johnstone, 1995, that, for the estimation problem, the corresponding rate does not depend on p in the domain $sp > 1$.

A variety of different tests was proposed in the papers cited above. But among all these test procedures, the crucial problem for their practical applications is dependence of their structure on the smoothness parameters s, p, q, M which are typically unknown a'priori.

The problem of adaptive (assumption free) testing for the same nonparametric set-up was considered in Spokoiny, 1995 for the case of L_2-distance ($r = 2$). It was shown that adaptive testing is impossible without loss of efficiency. The optimal adaptive rate was also evaluated there. It occured to be worse than the minimax rate within a $\log\log n$ factor.

Below we continue studying this problem for the case of an arbitrary integral norm L_r. We are aiming, firstly, to evaluate the optimal rate of testing and secondly, to propose an adaptive test which achieves this rate (probably with extra $\log\log$-factor). The results of the paper show that rates of testing differ essentially in three domains described by relations between smoothness parameters s, p, M and the norm power r. Informally one may classify these three domains as follows: the first domain corresponds to function of high regularity with homogeneous smoothness properties; the next domain, on the contrary, corresponds to functions with very poor (inhomogeneous) smoothness properties, particularly, to functions with jumps; finally, the last case can be placed between the extreme two: functions from this domain have again inhomogeneous smoothness properties but of moderate degree.

These three domains can be described also in term of the corresponding least favorable priors. For the first case, the hardest for detection function (signal) is uniformly small and, particularly, signal-to-noise ratio tends to zero. For the second (opposite) case, the hardest for detection signal is zero almost everywhere except one peak at a random place with the altitude which exceeds the noise level within some log-factor. In the last case, the corresponding signal contains polynomial number of peaks with the altitude of the noise level.

Then we propose three kinds of test procedure and show that for each domain, one may show one of these three which is critical for providing the optimal rate of testing. The first procedure corresponds to testing in L_2-norm and it achieves the minimax rate of testing in the first domain. (Note that both r and p may differ from 2 there.) The corresponding rate is $n^{-\frac{2s}{4s+1}}$ and the test is very powerful (in the sense of the rate of testing) but it "works" only for very regular function.

The second procedure corresponds to testing in sup-norm. One could say that this method "works" for any alternative (of course, some minimal smoothness is required) but it leads to relatively poor rate coinciding with the rate of estimation for the same value of smoothness parameters.

The last test procedure, in analogy with the estimation theory, might be called "spatially adaptive". It is based on thresholded empirical wavelet coefficients.

In the case when no information about smoothness properties of the function f is available a'priori, our recommendation is extremely simple and natural: let apply all these procedures simultaneously. We show that a proper combination of the indicated above tests can be taken to provide prescribed error probabilities. Of course, the sensitivity of such a test still depends on the smoothness properties of the underlying function but the proposed goodness-of-fit procedure is "assumption free".

Resuming, let us highlight one more important benefit of considering arbitrary $p, r \geq 1$. This enables us to observe how quantitative change of smoothness parameters leads to qualitative change of the structure of least favorable priors and the corresponding optimal test structure.

The paper is organized as follows. In the next section we formulate the problem and state the results. The proposed test procedure is discussed in Section 3. The proofs are postponed to Section 4 and Appendix.

2. Model and hypothesis testing problem

Assume we are given by the data $X(t)$, $0 \leq t \leq 1$, obeying the following stochastic differential equation

$$dX(t) = f(t)dt + n^{-1/2}dW(t), \qquad 0 \leq t \leq 1,\, X(0) = 0. \tag{2.1}$$

The function $f(\cdot)$ is unknown and the following statistical problem is considered: to test the null hypothesis H_0 that the function f is identically zero,

$H_0: \qquad f \equiv 0$.

We wish to test this hypothesis against as large class of alternatives as possible. That is why we do not assume any special (parametric) structure for the alternative set. This leads to considering a nonparametric alternative set. Following to Ingster, 1982; Ingster, 1993; Lepski and Spokoiny, 1995 we assume only that the function f obeys some smoothness conditions. More precisely, the function f is supposed to lie in some Besov ball $B_{s,p,q}(M)$,

$$B_{s,p,q}(M) = \{f : \|f\|_{B_{s,p,q}} \leq M\}.$$

The definition of the Besov norm $\|\cdot\|_{B_{s,p,q}}$ can be found, e.g., in Triebel, 1992. For the discussion of this notion in statistical context, see Donoho and Johnstone, 1995 or Donoho et al., 1995. Note that the definition of a Besov space can be done also in terms of the wavelet decomposition, see the property ISO2 below.

To be able to test the null against the alternative, we assume also that the alternative set is separated away from the null in some integral L_r-norm where the number $r \geq 1$ is specified. Hence we arrive at the following alternative

$H_1: \qquad \mathcal{F}_{\sigma,r}(\varrho) = \{f \in B_{s,p,q}(M) : \|f\|_r \geq \varrho\}$.

Now we define the hypothesis testing problem. A (non-randomized) test ϕ is a measurable function of the observation $X(\cdot)$ with two values $\{0,1\}$. As usual, the event $\{\phi = 0\}$ is treated as accepting the null hypothesis, and $\{\phi = 1\}$ means that the null is rejected. To simplify the exposition, we do not consider randomized tests. All the results can be extended on the case of randomized tests in a standard way, see e.g. Lehmann, 1959 or Ingster, 1993.

Let $\mathbf{P}_0$ be the distribution of the process $X(\cdot)$ under the null i.e. if we observe the only noise, and let $\mathbf{P}_f$ means the distribution of the process X under f due to (2.1), $\mathbf{P}_f = \mathcal{L}(X|f)$.

The quality of any test ϕ is measured by the corresponding error probabilities of the first and the second kinds. For the case under consideration with a simple hypothesis, the error probability of the first kind is

$$\alpha(\phi) = \mathbf{P}_0(\phi = 1).$$

If f is a point from the alternative set, $f \in \mathcal{F}_{\sigma,r}(\varrho)$, then the error probability of the second kind at f is defined as usual by $\beta(f) = \mathbf{P}_f(\phi = 0)$. The value $1 - \beta(f)$ is called the *power* of the test ϕ at f.

We consider further the minimax set-up which leads to the following criteria

$$\beta_{\sigma,r}(\phi;\varrho) = \sup_{f \in \mathcal{F}_{\sigma,r}(\varrho)} \mathbf{P}_f(\phi = 0). \tag{2.2}$$

2.1. Minimax rate of testing

Below we are focusing on the asymptotic hypothesis testing problem as the noise level tends to zero ($n \to \infty$). We are interested in evaluating the optimal (fastest) rate of decay of the radius ϱ as a function of n to zero as $n \to \infty$ for which testing with prescribed error probabilities is still possible. The following definition of the minimax rate $\varrho(n)$ was proposed in Ingster, 1993.

Definition 2.1. A sequence $\varrho(n)$ is called the *minimax rate of testing* if $\varrho(n) \to 0$ as $n \to \infty$ and the following two conditions hold:

(*i*) For any sequence $\varrho'(n)$ such that

$$\varrho'(n)/\varrho(n) = o_n(1)$$

one has

$$\inf_{\phi_n} [\mathbf{P}_0(\phi_n) + \beta_{\sigma,r}(\phi_n; \varrho'(n))] = 1 - o_n(1).$$

(*ii*) For any $\alpha, \beta > 0$ there exist a constant $C > 0$ and tests ϕ_n^* such that

$$\begin{aligned} \mathbf{P}_0(\phi_n^*) &\leq \alpha + o_n(1) \\ \beta_{\sigma,r}(\phi_n^*; C\varrho(n)) &\leq \beta + o_n(1). \end{aligned}$$

Here and below we denote by $o_n(1)$ any sequence tending to zero as $n \to \infty$.

Remark 2.1. The first condition of the above definition means that testing with the rate faster than $\varrho(n)$ is impossible i.e. if the distance between the null and the alternative set is less in order than $\varrho(n)$, then any test has asymptotically trivial power in the sense that the sum of the error probabilities of the first and second kinds is close to 1. The second condition means roughly that, on the contrary, if the distance is of the order $\varrho(n)$ then testing could be done with prescribed error probabilities.

It turns out that the rate $\varrho(n)$ depends on the smoothness parameters $\sigma = (s, p, q, M)$ and the norm r in which we measure the distance between the null and the alternative.

We present below without proof the result about the rate $\varrho(n)$. This result is partially (for particular values of p, r) proved by different authors, see the remark after the statement of the result.

It is useful to introduce the folowing three numbers

$$\gamma^{(1)} = \frac{1}{4s+1}, \tag{2.3}$$

$$\gamma^{(2)} = \frac{1-1/r}{2s+1-1/p}, \tag{2.4}$$

$$\gamma^{(3)} = \frac{1-2/r}{2s+1-2/p}. \tag{2.5}$$

Denote also

$$\Delta_{2,3} = \gamma^{(2)} - \gamma^{(3)} = \frac{2sp-(r-p)}{pr(2s+1-1/p)(2s+1-2/p)}, \tag{2.6}$$

$$\Delta_{1,2} = \gamma^{(1)} - \gamma^{(2)} = \frac{p-r-2sp(r-2)}{pr(4s+1)(2s+1-1/p)}. \tag{2.7}$$

The following technical statements are straightforward.

Lemma 2.1. *Let $sp > 1$. If $\gamma^{(1)} > \gamma^{(2)}$, then also $\gamma^{(2)} > \gamma^{(3)}$. On the contrary, $\gamma^{(2)} < \gamma^{(3)}$ implies $\gamma^{(1)} < \gamma^{(2)}$.*

The proof is left to the reader.

Now we are ready to formulate the result. We consider only the domain $sp > 1$ (this constrains comes from the approximation theory) and we exclude the points with $\Delta_{1,2} = 0$ or $\Delta_{2,3} = 0$. The latter special cases are more involved and they require special consideration.

Theorem 2.1. *Given r and $\sigma = (s,p,q,M)$ with $sp > 1$, let*

$$\varrho_{\sigma,r}(n) = \begin{cases} \varrho^{(1)}_{\sigma,r}(n) = M^{\gamma^{(1)}} n^{-\frac{1-\gamma^{(1)}}{2}} & \text{if } \gamma^{(1)} > \gamma^{(2)}, \\ \varrho^{(2)}_{\sigma,r}(n) = M^{\gamma^{(2)}} n^{-\frac{1-\gamma^{(2)}}{2}} & \text{if } \gamma^{(2)} > \gamma^{(1)} \text{ and } \gamma^{(2)} > \gamma^{(3)}, \\ \varrho^{(3)}_{\sigma,r}(n) = M^{\gamma^{(3)}} (n/\log n)^{-\frac{1-\gamma^{(3)}}{2}} & \text{if } \gamma^{(2)} < \gamma^{(3)}. \end{cases} \tag{2.8}$$

If $\varrho'(n)/\varrho_{\sigma,r}(n) = o_n(1)$, *then*

$$\inf_{\phi_n}[\mathbf{P}_0(\phi_n) + \beta_{\sigma,r}(\phi_n; \varrho'(n))] = 1 - o_n(1).$$

Remark 2.2. Such kind of results is called typically "a lower bound of the rate of convergence". It means that testing with a faster rate than $\varrho_{\sigma,r}(n)$ is impossible.

We do not state a precise "upper bound" result i.e. we do not present a test which achives exactly the rate shown in (2.8). Instead we propose later on a test which achives (and even adaptively) this lower bounds up to extra $\log\log$-term. But the following conjecture can be formulated: let $\sigma = (s, p, q, M)$ and r be given with $sp > 1$ and $\gamma^{(1)} \neq \gamma^{(2)}$, $\gamma^{(2)} \neq \gamma^{(3)}$. Then $\varrho_{\sigma,r}(n)$ is just the minimax rate of testing. This conjecture is proved for some particular cases, see the remarks below.

Remark 2.3. In view of Lemma 2.1, one may say that up to log term $\varrho_{\sigma,r}(n) \asymp M^{\gamma} n^{-(1-\gamma)/2}$ where $\gamma = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$.

Remark 2.4. The rate $\varrho^{(1)}(n) = M^{\frac{1}{4s+1}} n^{-\frac{2s}{4s+1}}$ appeared first in Ingster, 1982 for the case $p = r = 2$ and was extended to the case $p \geq 2$, $r \leq 2$ in Ingster, 1986. The expression for $\varrho^{(2)}(n)$ in the particular case of $p = r > 2$ was found in Ingster, 1986; Ingster, 1993 and for the case $r = 2$, $p < r$ in Lepski and Spokoiny, 1995.

Remark 2.5. To compare the results on the rate of testing with the similar results for the estimation problem, we recall first the expression for the rate $\psi_{\sigma,r}(n)$ of estimation of the function f over the Besov ball $B_{s,p,q}(M)$ for the global error measure in L_r-norm, see Nemirovski, 1985; Donoho et al., 1995; Lepski et al., 1994:

$$\psi_{\sigma,r}(n) = \begin{cases} M^{\gamma^{(0)}} n^{-\frac{1-\gamma^{(0)}}{2}} & \text{if } sp > (r-p)/2, \\ M^{\gamma^{(3)}} (n/\log n)^{-\frac{1-\gamma^{(3)}}{2}} & \text{if } sp < (r-p)/2. \end{cases} \tag{2.9}$$

Here $\gamma^{(0)} = 1/(2s+1)$ and $\gamma^{(3)}$ is from (2.5).

First we observe that expression (2.8) for the rate of testing $\varrho_{\sigma,r}(n)$ differs essentially in three different domains whereas there are only two cases for the estimation problem. But the last case and the corresponding rate $\varrho^{(3)}(n)$ are common for both problems. The related domain (with $sp < (r-p)/2$) corresponds to functions with very inhomogeneous smoothness characteristics, see Donoho et al., 1995.

Coincideness of the rates for the testing and estimation problems leads to the following conclusion: the test of the form $\|f_n^*\|_r > C\psi_{\sigma,r}(n)$ is rate minimax where f_n^* is any rate optimal estimator and C is a proper constant. Note that this conclusion is not true for the remaining cases where $sp > (r-p)/2$ and where the rate of tesing is better than the rate of estimation (since $\gamma^{(0)} > \gamma^{(1)}$ and $\gamma^{(0)} > \gamma^{(2)}$).

Now we turn to the problem of adaptive testing.

2.2. Adaptive testing

Let the smoothness parameters $\sigma = (s, p, q, M)$ be unknown. We assume only that these parameters belong to some prescibed set $\mathcal{T}$. We say that the set $\mathcal{T}$ is *nontrivial* if there are numbers p, M and $s_{\min} < s_{\max}$ such that $\sigma = (s, p, q, M) \in \mathcal{T}$ for any $s_{\min} \le s \le s_{\max}$.

The problem of adaptive testing for a nontrivial set $\mathcal{T}$ was considered for $r = 2$ in Spokoiny, 1995. It was shown that an adaptive testing without loss of efficiency (i.e. with the same rate $\varrho_{\sigma,r}(n)$) is impossible. Moreover, the optimal "adaptive" rate was evaluated, it appeared to be (for $r = p = 2$) $\tilde{\varrho}_{\sigma,r}(n) = \varrho_{\sigma,r}(n/\sqrt{\log\log n})$. and an adaptive test ϕ_n^* was constructed based on the wavelet decomposition. The stricture of this test depends inessentially on the range of adaptation $\mathcal{T}$ but not on the particular value σ from this domain.

Below we present a slightly different test based again on wavelet decomposition and show that it "works" for any norm L_r and for any compact range of adaptation $\mathcal{T}$. The structure of this test is described in the next section. Now we state the result describing its asymptotic performance.

Denote for $\sigma = (s, p, q, M)$

$$
\begin{aligned}
h_\sigma^{(1)} &= \left(M^2 n\right)^{-\frac{2}{4s+1}}, &(2.10)\\
h_\sigma^{(2)} &= \left(M^2 n\right)^{-\frac{1}{2s+1-1/p}}, &(2.11)\\
h_\sigma^{(3)} &= \left(M^2 n/\log n\right)^{-\frac{1}{2s+1-2/p}}. &(2.12)
\end{aligned}
$$

Theorem 2.2. *Let a set $\mathcal{T}$ be such that its projection $\mathcal{T}_{s,p} = \Pi_{s,p}\mathcal{T}$ on the plane (s,p) is a compact set non-intersecting with the curves $\gamma^{(1)} = \gamma^{(2)}$ and $\gamma^{(2)} = \gamma^{(3)}$, see (2.6), (2.7). Particularly, one has*

$$
\inf_{\sigma\in\mathcal{T}} |\gamma_{\sigma,r}^{(2)} - \gamma_{\sigma,r}^{(3)}| > 0,
$$
$$
\inf_{\sigma\in\mathcal{T}} |\gamma_{\sigma,r}^{(2)} - \gamma_{\sigma,r}^{(1)}| > 0,
$$

whehe $\gamma^{(i)} = \gamma_{\sigma,r}^{(i)}$ is due to (2.3) - (2.5), $i = 1,2,3$. Let also there be a constant $\delta > 0$ such that

$$
n^{-\delta/2} \le h_\sigma^{(i)} \le n^{-(1-\delta)/2}, \qquad i = 1,2,3,
$$

for any $\sigma \in \mathcal{T}$. Then there is a constant $C = C(\mathcal{T}_{s,p}, r)$ and tests ϕ_n^* (see the next section) such that

$$
\mathbf{P}_0(\phi_n^* = 0) = o_n(1) \tag{2.13}
$$

and uniformly in $\sigma \in \mathcal{T}$ and $f \in B_{s,p,q}(M)$ with $\|f\|_r \ge C\tilde{\varrho}_{\sigma,r}(n)$ one has

$$
\mathbf{P}_f(\phi_n^* = 0) = o_n(1)
$$

where

$$\tilde{\varrho}_{\sigma,r}(n) = \begin{cases} M^{\gamma^{(1)}} n^{-\frac{1-\gamma^{(1)}}{2}} (\log\log n)^{-\frac{1-\gamma^{(1)}}{4}} & \text{if } \gamma^{(1)} > \gamma^{(2)}, \\ M^{\gamma^{(2)}} n^{-\frac{1-\gamma^{(2)}}{2}} (\log\log n)^{\frac{1}{2r}-\frac{\gamma^{(2)}}{2p}} & \text{if } \gamma^{(2)} > \gamma^{(1)} \text{ and } \gamma^{(2)} > \gamma^{(3)}, \\ M^{\gamma^{(3)}} (n/\log n)^{-\frac{1-\gamma^{(3)}}{2}} & \text{if } \gamma^{(3)} > \gamma^{(2)}. \end{cases} \tag{2.14}$$

Remark 2.6. Easy to see that the condition of the theorem related to $h_\sigma^{(1)}$, $h_\sigma^{(2)}$ or $h_\sigma^{(3)}$ is fulfilled if the parameter M belongs to some compact set that is $M \in [M_{\min}, M_{\max}]$ for all $\sigma \in \mathcal{T}$. But the presented condition is much weaker and it allows parameter M to be close to zero or infinity for s inside the interval $[s_{\min}, s_{\max}]$.

Remark 2.7. We see that $\tilde{\varrho}_{\sigma,r}(n) = \varrho_{\sigma,r}(n) = \varrho_{\sigma,r}^{(3)}(n)$ if $sp < (r-p)/2$. This means that the test ϕ_n^* is adaptive and minimax simultaneously in this domain.

The possibility of adaptive testing for this case could be explained also by the fact that the rates of testing and estimation coincide and that an adaptive estimation is possible, see Juditsky, 1995, Donoho et al., 1995, Lepski et al., 1994.

Remark 2.8. The factor $\log\log n$ in the domain $\Delta_{2,3} > 0$ can be viewed as "payment for adaptation". It was shown in Spokoiny, 1995 for $r = p = 2$ that this factor can be neither removed nor improved.

3. Test procedure

In this section we describe the structure of the adaptive test ϕ_n^* declared in Theorem 2.2.

The whole construction uses only two external parameters $s_{\max}$ and δ where $s_{\max}$ has the meaning of the upper considered value of the smoothness degree s (recall that s belongs to a compact set on the real line). The knowledge of the norm L_r in which we measure the distance between the null and the alternative set, is also important. The construction of the test makes heavy use of the wavelet decomposition. We proceed as in Spokoiny, 1995.

3.1. Wavelet transform

Assume we are given an orthonormal basis of compactly supported wavelets of $L_2[0,1]$. One may use the construction from Meyer, 1990 or Cohen et al., 1993b. Let $\phi_{j,k}, \psi_{j,k}$ be a system of compactly supported orthogonal wavelets ($\operatorname{supp}\phi \subseteq [-0, A]$ and $\operatorname{supp}\psi \subseteq [-0, A]$). We suppose that ϕ and $\psi \in C^m$, where m is the maximal integer smaller than $s_{\max}$. This implies (cf. Daubechies, 1992[Ch.7]) that $\psi(x)$ has at least m vanishing moments.

Let j_0 be such that $2^{j_0} > A+1$. It has been shown in Cohen et al., 1993b and Cohen et al., 1993a that an orthogonal wavelet basis on $[0,1]$ can be constructed by retaining

$\psi_{j,k}$ and $\phi_{j,k}$ as the interiour wavelets and scaling functions and adding adapted edge wavelets and scaling functions. These edge elements are tailored so that the total number is exactly 2^j at resolution j. For the sake of simplicity we use the same notation for the edge corrected and original functions. This construction provides an unconditional basis for the $B_{s,p,q}[0,1]$ space for $s > m$, $sp > 1$.

It is useful to use for $\phi_{j_0,k}$ also the notation ψ_k, $k = 1, \ldots, 2^{j_0}$. Denote also by $\mathcal{J}$ the set of resolution levels for the considered wavelet basis,

$$\mathcal{J} = \{j \geq j_0\}$$

and let $\mathcal{I}_j$ be the index set for jth level,

$$\begin{aligned}
\mathcal{I}_{j_0} &= \{k : k = 1, \ldots, 2^{j_0}\} \bigcup \{(j_0, k) : k = 1, \ldots, 2^{j_0}\}, \\
\mathcal{I}_j &= \{(j,k) : k = 1, \ldots, 2^j\}.
\end{aligned}$$

One has obviously for the number of elements $N_j = \#(\mathcal{I}_j)$ in jth level

$$N_j = \#(\mathcal{I}_j) = \begin{cases} 2^{j_0+1} & j = j_0, \\ 2^j & j > j_0. \end{cases}$$

By $\mathcal{I}$ we denote the global index set for the considered basis, $\mathcal{I} = \{\mathcal{I}_j, \, j \in \mathcal{J}\}$. Now the wavelet decomposition of a function f can be represented in the form

$$f(t) = \sum_{I \in \mathcal{I}} \theta_I \psi_I(t) = \sum_{j \in \mathcal{J}} \sum_{I \in \mathcal{I}_j} \theta_I \psi_I(t)$$

where θ_I is Ith wavelet coefficient,

$$\theta_I = \int_0^1 f(t) \psi_I(t) dt, \qquad I \in \mathcal{I}.$$

Let now X_I, $I \in \mathcal{I}$ be empirical wavelet coefficients for the model (2.1),

$$X_I = \int_0^1 \psi_I(t) dX(t).$$

The model equation (2.1) yields

$$X_I = \theta_I + n^{-1/2} \int_0^1 \psi_I(t) dW(t)$$

and the original functional model (2.1) is translated into the sequence space model

$$X_I = \theta_I + n^{-1/2} \xi_I, \qquad I \in \mathcal{I}, \tag{3.1}$$

where $\xi_I = \int \psi_I dW$ are standard normal and independent for different I.

Given j, denote by D_j the projector on the function subspace corresponding to jth level,

$$D_j f(t) = \sum_{I \in \mathcal{I}_j} \theta_I \psi_I(t)$$

and by E_j the projector on the first levels till j,

$$E_j f(t) = \sum_{j'=j_0}^{j} \sum_{I \in \mathcal{I}_{j'}} \theta_I \psi_I(t).$$

The wavelet transform is justified by the following (isometric) properties, cf. Triebel, 1992[p.240]:

(ISO1) for any function $f \in L_2[0,1]$

$$\|f\|^2 = \|\theta\|^2 := \sum_{\mathcal{I}} \theta_I^2, \tag{3.2}$$

(ISO2) there are two constants C_1 and C_2 such that

$$C_1 \|f\|_{B_{s,p,q}} \le \|\theta\|_{b_{s,p,q}} \le C_2 \|f\|_{B_{s,p,q}},$$

where

$$\|\theta\|_{b_{s,p,q}} = \begin{cases} \left\{ \sum_{j \ge j_0} \left[2^{j(s+\frac{1}{2}-\frac{1}{p})} \left(\sum_{\mathcal{I}_j} |\theta_I|^p \right)^{1/p} \right]^q \right\}^{1/q}, & q < \infty, \\ \sup_{j \ge j_0} \left\{ 2^{j(s+\frac{1}{2}-\frac{1}{p})} \left(\sum_{\mathcal{I}_j} |\theta_I|^p \right)^{1/p} \right\}, & q = \infty. \end{cases} \tag{3.3}$$

(ISO3) For any $r \ge 1$ there exists a constant C depending only on r and the considered basis $\{\psi_I(t), I \in \mathcal{I}\}$ such that for any $j \ge j_0$ and any f

$$\|D_j f\|_r \le C 2^{j(\frac{1}{2}-\frac{1}{r})} \left(\sum_{I \in \mathcal{I}_j} |\theta_I|^r \right)^{1/r}.$$

For the technical reason, we suppose that the initial level j_0 of the considered wavelet basis $\{\psi_I(t), I \in \mathcal{I}\}$ is chosen depending on n in such a way that $j_0 \to \infty$ as $n \to \infty$ and

$$2^{j_0} > n^{\delta/2}. \tag{3.4}$$

Here δ is from the theorem conditions.

3.2. Test procedure. Preliminaries

The presented test is essentially levelwise. This means that all the considered statistics are based on empirical wavelet coefficients within a certain wavelet level. In the other words, we test each wavelet level independently and the resulting test rejects the null if at least one level test does.

Note that a test procedure presented in Spokoiny, 1995 used "global" test statistics depending on coefficients from a number of levels.

Let the empirical wavelet coefficients X_I be defined by (3.1). Denote also

$$Y_I = n^{1/2} X_I = n^{1/2}\theta_I + \xi_I.$$

Given $j \le j_0$, set

$$S_j = \sum_{I \in \mathcal{I}_j} Y_I^2 = n \sum_{I \in \mathcal{I}_j} X_I^2.$$

For $r > 2$ set also

$$S_{j,r} = \sum_{I \in \mathcal{I}_j} |Y_I|^r = n^{r/2} \sum_{I \in \mathcal{I}_j} |X_I|^r.$$

It was shown in Ingster, 1982; Ingster, 1986; Ingster, 1993 that a rate optimal test for the case with $p = r$ can be based on the sum $T_{j(n)} = S_{j_0} + \ldots + S_{j(n)}$ where the index $j(n)$ is to be chosen depending on s, p. More precisely, determine the level $j(n)$ for $p = r = 2$ by

$$2^{j(n)} \approx n^{\frac{4}{4s+1}}.$$

Then the rate optimal test could be taken in the form

$$\phi_{n,2} = \mathbf{1}\left(\eta_{j(n)} > \chi_\alpha\right)$$

where $\eta_j = \frac{T_j - \mathbf{E}_0 T_j}{\sqrt{\mathbf{D}_0 T_j}}$, $\mathbf{E}_0 \eta$ and $\mathbf{D}_0 \eta$ mean the expectation and the variance of a random variable ζ under the measure $\mathbf{P}_0$ and χ_α is $(1-\alpha)$-quantile of the standard normal low.

For the case with $r = p > 2$, see Ingster, 1993, a rate optimal test can be constructed in the similar way with $S_{j,r}$ in place of S_j,

$$\phi_{n,r} = \mathbf{1}\left(\eta_{j,r} > \chi_\alpha\right).$$

Here $\eta_{j,r} = \frac{T_{j,r} - \mathbf{E}_0 T_{j,r}}{\sqrt{\mathbf{D}_0 T_{j,r}}}$, $T_{j,r} = S_{j_0,r} + \ldots + S_{j,r}$ and the level $j(n)$ is to be taken by

$$2^{j(n)} \approx n^{\frac{2}{2s+1-1/p}}.$$

Note that the proper choice of the level $j(n)$ is crucial here and above for $r = 2$.

In the case of adaptive testing, when s and p are unknown, such a test could not be implemented. It was proposed in Spokoiny, 1995 to use for testing the supremum value of η_j or $\eta_{j,r}$ over all feasible levels j. Below we proceed in the similar way but the test is based directly on S_j (resp. $S_{j,r}$) in place of T_j (resp. $T_{j,r}$).

As it was pointed out in Spokoiny, 1995, tests based on S_j or $S_{j,r}$ allow to test with the optimal rate for the case with $p \ge r$. If the norm L_r for the distance is stronger than the norm L_p in which we measure smoothness properties of f, then such tests do not provide the optimal rate. The situation here is similar to that of in the estimation problem when linear methods are rate optimal only for $r \le p$. For the case $p < r$, similarly to the estimation problem, see Donoho and Johnstone, 1995 or Donoho et al., 1995, we involve the idea of thresholding of empirical wavelet coefficients.

3.2.1. An adaptive test

The test proposed below can be viewed as combination of four different tests each of them is in its turn composite and operates levelwise. The first test corresponds to testing in L_2-norm, the second one (to be apply only for $r > 2$) tests in L_r norm, and the third one does the same but for thresholded empirical wavelet coefficients. Finally, the last subtest analyses each empirical wavelet coefficient separately and it corresponds to testing in sup-norm.

To begin with, we restrict the number of considered wavelet levels. Let j_0 be taken due to (3.4). Define also the highest considered level $j_{\max}$ in such a way that

$$n^{1-\delta/2} < 2^{j_{\max}} < n$$

where δ is shown in the theorem conditions. Denote

$$\mathcal{J}^n = \{j : j_0 \le j \le j_{\max}\}.$$
$$\mathcal{I}^n = \bigcup_{j \in \mathcal{J}^n} \mathcal{I}_j.$$

Note that

$$m_n = \#(\mathcal{J}^n) \le \log_2 n \le 2 \log n \tag{3.5}$$

and

$$\#(\mathcal{I}^n) \le 2n. \tag{3.6}$$

For each $j \in \mathcal{J}^n$, state

$$S_j = \sum_{I \in \mathcal{I}_j} Y_I^2, \tag{3.7}$$
$$\zeta_j = \frac{S_j - \mathbf{E}_0 S_j}{\sqrt{\mathbf{D}_0 S_j}}. \tag{3.8}$$

The first test, which we introduce, is based on ζ_j for $j \in \mathcal{J}^n$,

$$\phi_{n,2} = \mathbf{1}\left\{\sup_{j \in \mathcal{J}^n} \zeta_j > 2\sqrt{\log\log n}\right\}. \tag{3.9}$$

Remark 3.1. Now we are in a position to explain the nature of the $\log\log$-factor entering in the adaptive rate of testing. Obviously, S_j and thus ζ_j are independent for different j. This follows directly from the definition and from independence of empirical wavelet coefficients. Easy to observe also that each ζ_j is under the null asymptotically (as $n \to \infty$) standard normal. Hence our test statistic in (3.9) is supremum of m_n independent asymptotically Gaussian random variables and its distribution is degenerate around $\sqrt{2\log m_n} \approx \sqrt{2\log\log n}$. This explains the choice of the testing level in (3.9).

Similarly we define the test $\phi_{n,r}$ for $r > 2$. Set

$$S_{j,r} = \sum_{I\in\mathcal{I}_j} |Y_I|^r, \tag{3.10}$$

$$\zeta_{j,r} = \frac{S_{j,r} - \mathbf{E}_0 S_{j,r}}{\sqrt{\mathbf{D}_0 S_{j,r}}}. \tag{3.11}$$

$$\phi_{n,r} = \mathbf{1}\left\{\sup_{j\in\mathcal{J}^n} \zeta_{j,r} > 2\sqrt{\log\log n}\right\}. \tag{3.12}$$

The similar test with thresholding is defined as follows. Denote

$$\mathbf{r} = \max\{r, 2\}$$

and, given $\lambda > 0$, $j \in \mathcal{J}^n$ set

$$S_{j,\mathbf{r}}(\lambda) = \sum_{I\in\mathcal{I}_j} |Y_I|^{\mathbf{r}}\mathbf{1}(|Y_I| > \lambda) = n^{\mathbf{r}/2}\sum_{I\in\mathcal{I}_j} |X_I|^{\mathbf{r}}\mathbf{1}(|X_I| > \lambda/\sqrt{n}), \tag{3.13}$$

$$\zeta_{j,\mathbf{r}}(\lambda) = \frac{S_{j,\mathbf{r}}(\lambda) - \mathbf{E}_0 S_{j,\mathbf{r}}(\lambda)}{\sqrt{\mathbf{D}_0 S_{j,\mathbf{r}}(\lambda)}}. \tag{3.14}$$

Define also the set Λ_j by

$$\Lambda_j = \{\lambda = \sqrt{2\mathbf{r}k\log 2}, k = 1, 2, \ldots, \lambda \le \sqrt{0.5\log N_j}\}$$

where recall $N_j = \#(\mathcal{I}_j) = 2^j$ for $j > j_0$, $N_{j_0} = 2^{j_0+1}$. Now state

$$\phi_{n,r,t} = \mathbf{1}\left\{\sup_{j\in\mathcal{J}^n}\sup_{\lambda\in\Lambda_j} \zeta_{j,\mathbf{r}}(\lambda) > 3\sqrt{\log\log n}\right\}. \tag{3.15}$$

Finally set

$$\zeta_{j,\infty} = \sup_{I\in\mathcal{I}_j} |Y_I|,$$

$$\phi_{n,\infty} = \mathbf{1}\left\{\sup_{j\in\mathcal{J}^n} \zeta_{j,\infty} > 2\sqrt{\log N_j}\right\}.$$

The whole test ϕ_n^* rejects the null if at least one of these four does:

$$\phi_n^* = \max\{\phi_{n,2}, \phi_{n,r}, \phi_{n,r,t}, \phi_{n,\infty}\} \tag{3.16}$$

where $\phi_{n,r} = 0$ for $r \le 2$.

Note that the values $\mathbf{E}_0 S_j, \mathbf{D}_0 S_j$, $\mathbf{E}_0 S_{j,r}, \mathbf{D}_0 S_{j,r}$ and even $\mathbf{E}_0 S_{j,r}(\lambda), \mathbf{D}_0 S_{j,r}(\lambda)$ are easy to calculate, see Lemmas 4.2 - 4.4 below, that simplifies the implementation of the test ϕ_n^*.

4. Proofs

In this section we prove Theorem 2.2. The proof can be informally split into three parts. First we study the behavior of the test statistics ζ_j, $\zeta_{j,r}$, $\zeta_{j,\mathrm{r}}(\lambda), \zeta_{j,\infty}$ under the null and show that $\mathbf{P}_0(\phi_n^* = 1) = o_n(1)$. Next we analyse what sort of information about the function f can be extracted from the statistical fact that $\phi_n^* = 0$. Finally we show that this information and smoothness properties of the function f enable us to prove the desired assertion (2.14).

4.1. Error probability of the first kind

In the next lemmas we collect some information about the distrubution of all introduced statistics under the null i.e. if $f \equiv 0$ and hence $\theta_I = 0$, $I \in \mathcal{I}$.

Denote

$$t_n = \sqrt{\log\log n}.$$

Lemma 4.1. *For each $j \in \mathcal{J}^n$ and $I \in \mathcal{I}_j$*

$$\mathbf{P}_0(|Y_I| > 2\sqrt{\log N_j}) \le 2N_j^{-2}$$

and

$$\mathbf{P}_0\left(\zeta_{j,\infty} > 2\sqrt{\log N_j}\right) \le 2N_j^{-1}.$$

Lemma 4.2. *The following assertions hold uniformly in $j \in \mathcal{J}^n$: for n large enough*

(i) $\mathbf{E}_0 S_j = N_j$;

(ii) $\mathbf{D}_0 S_j = 2N_j = d_2^2 N_j$ *with* $d_2 = \sqrt{2}$.

(iii) $\mathbf{P}_0\left(S_j - \mathbf{E}_0 S_j > 2d_2\sqrt{N_j}t_n\right) \le 2\exp\{-2t_n^2\} = 2(\log n)^{-2}$.

The similar assertions are valid for for $S_{j,r}$.

Lemma 4.3. *Let $r > 2$. For n large enough, uniformly in $j \in \mathcal{J}^n$*

(i) $\mathbf{E}_0 S_{j,r} = b_r N_j$ *with* $b_r = \mathbf{E}|\xi|^r$, $\xi \sim \mathcal{N}(0,1)$,

$$b_r = (2\pi)^{-1/2}\int |x|^r e^{-x^2/2}dx.$$

(ii) $\mathbf{D}_0 S_{j,r} = d_r^2 N_j$ *with* $d_r = \mathbf{D}|\xi|^r = b_{2r} - b_r^2$.

(iii) $\mathbf{P}_0\left(S_{j,r} - \mathbf{E}_0 S_{j,r} > 2d_r\sqrt{N_j}t_n\right) \le 2\exp\{-2t_n^2\} = 2(\log n)^{-2}$.

Lemma 4.4. *For n large enough, uniformly in $j \in \mathcal{J}^n$ and $\lambda \in \Lambda_j$*

(i) $\mathbf{E}_0 S_{j,\mathrm{r}}(\lambda) = b_{\mathrm{r}}(\lambda)N_j$ *with* $b_{\mathrm{r}} = \mathbf{E}|\xi|^{\mathrm{r}}\mathbf{1}(|\xi| > \lambda)$, $\xi \sim \mathcal{N}(0,1)$,

$$b_{\mathrm{r}}(\lambda) = 2(2\pi)^{-1/2}\int_\lambda^\infty x^{\mathrm{r}} e^{-x^2/2}dx;$$

(ii) $\quad \mathbf{D}_0 S_{j,\mathrm{r}}(\lambda) = d_{\mathrm{r}}^2(\lambda) N_j$ *with*

$$d_{\mathrm{r}}^2 = \mathbf{D}|\xi|^{\mathrm{r}} \mathbf{1}(|\xi| > \lambda) = b_{2\mathrm{r}}(\lambda) - b_{\mathrm{r}}^2(\lambda);$$

(iii) $\quad \mathbf{P}_0 \left(S_{j,\mathrm{r}}(\lambda) - \mathbf{E}_0 S_{j,\mathrm{r}}(\lambda) > 3 d_{\mathrm{r}}(\lambda) \sqrt{N_j} t_n \right) \le 2 \exp\{-9t_n^2/2\} = 2(\log n)^{-9/2};$

(iv) *There are such* $C, C' > 0$ *that for any* $\lambda \ge 1$ *and* $r \ge 2$ *one has*

$$C e^{-\lambda^2/4} \le d_r(\lambda) \le C' \lambda^r e^{-\lambda^2/4}.$$

We defer the proof of these lemmas to the Appendix. Using these results, one can easily prove that ϕ_n^* obeys (2.13). In fact, due to Lemma 4.2, (i) and (iii), one has

$$\begin{aligned}
\mathbf{P}_0(\phi_{n,2} = 1) &= \mathbf{P}_0 \left(\sup_{j \in \mathcal{J}^n} \zeta_j > 2t_n^2 \right) \le \\
&\le \sum_{j \in \mathcal{J}^n} \mathbf{P}_0 \left(S_j - \mathbf{E}_0 S_j > 2 d_2 \sqrt{N_j} t_n \right) \le \sum_{j \in \mathcal{J}^n} 2(\log n)^{-2} \le \\
&\le 2\#(\mathcal{J}^n)(\log n)^{-2} = 2 m_n (\log n)^{-2} \to 0, \qquad n \to \infty,
\end{aligned}$$

since $m_n \le 2 \log n$. Similarly one estimates the probabilities $\mathbf{P}_0(\phi_{n,r} = 1)$. Next, to estimate $\mathbf{P}_0(\phi_{n,r,t} = 1)$, we apply Lemma 4.4 in the same manner. One has

$$\begin{aligned}
&\mathbf{P}_0(\phi_{n,r,t} = 1) \le \sum_{j \in \mathcal{J}^n} \sum_{\lambda \in \Lambda_j} P\left(\zeta_{j,\mathrm{r}}(\lambda) > 3 d_{\mathrm{r}}(\lambda) t_n \right) \le \\
&\le \sum_{j \in \mathcal{J}^n} \#(\Lambda_j) \exp\{-9t_n^2/2\} \le C(\log n)^2 (\log n)^{-9/2} \to 0, \qquad n \to \infty.
\end{aligned}$$

The assertion of Lemma 4.1 implies obviously

$$\mathbf{P}_0(\phi_{n,\infty} = 1) \le \sum_{j \in \mathcal{J}^n} N_j^{-1} \le 2^{-\delta j_0 + 2} = o_n(1).$$

and the statement (2.13) follows.

At the next step of the proof, we study which functions f are detectable for the test ϕ_n^*.

4.2. Sensivity of the test ϕ_n^*

Let an arbitrary function f be observed with noise due to (2.1) and let $\{\theta_I, I \in \mathcal{I}\}$ be the corresponding wavelet coefficients. Denote for $j \in \mathcal{J}^n$ and $\lambda \in \Lambda_j$

$$\begin{aligned}
\bar{S}_j &= \sum_{I \in \mathcal{I}_j} |\sqrt{n}\theta_I|^2, \\
\bar{S}_{j,r} &= \sum_{I \in \mathcal{I}_j} |\sqrt{n}\theta_I|^r, \\
\bar{S}_{j,r}(\lambda) &= \sum_{I \in \mathcal{I}_j} |\sqrt{n}\theta_I|^r \mathbf{1}(|\sqrt{n}\theta_I| > \lambda).
\end{aligned}$$

Proposition 4.1. *If for some $j \in \mathcal{J}^n$ and $I \in \mathcal{I}_j$ one has $|\theta_I| > 4n^{-1/2}\sqrt{\log N_j}$, then*

$$\mathbf{P}_f(\phi_{n,\infty} = 0) = o_n(1).$$

Proposition 4.2. *If for some $j \in \mathcal{J}^n$*

$$\bar{S}_j \geq 4d_2\sqrt{N_j}t_n,$$

then

$$\mathbf{P}_f\left(\phi_{n,2} = 0\right) = o_n(1).$$

The similar fact can be stated for $\phi_{n,r}$ and $\phi_{n,r,t}$. For technical reason, we formulate these statements in combination with the result of Proposition 4.1.

Proposition 4.3. *Let $r > 2$. If for some $j \in \mathcal{J}^n$*

$$\bar{S}_{j,r} \geq 4d_r\sqrt{N_j}t_n,$$

and $|\theta_I| \leq 4n^{-1/2}\sqrt{\log N_j}$ for all $I \in \mathcal{I}_j$, then

$$\mathbf{P}_f\left(\phi_{n,r} = 0\right) = o_n(1).$$

Proposition 4.4. *If for some $j \in \mathcal{J}^n$ and $\lambda \in \Lambda_j$*

$$\bar{S}_{j,\mathbf{r}}(\lambda) \geq 4d_{\mathbf{r}}(\lambda)\sqrt{N_j}t_n,$$

and $|\theta_I| \leq 4n^{-1/2}\sqrt{\log N_j}$ for all $I \in \mathcal{I}_j$, then

$$\mathbf{P}_f\left(\phi_{n,r,t} = 0\right) = o_n(1).$$

Proof of these propositions can be found in the Appendix.

The results of Propositions 4.1 – 4.4 can be treated in the following way. Given function f, if $\{\phi_n^* = 0\}$, then the following statements hold true simultaneosly for all $j \in \mathcal{J}^n$ and $\lambda \in \Lambda_j$ with $\mathbf{P}_f$-probability close to 1

$$\sup_{I \in \mathcal{I}_j} |\theta_I| < 4n^{-1/2}\sqrt{\log N_j}, \tag{4.1}$$

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^2 \leq 4n^{-1}d_2\sqrt{N_j}t_n, \tag{4.2}$$

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^{\mathbf{r}} \leq 4n^{-\mathbf{r}/2}d_{\mathbf{r}}\sqrt{N_j}t_n, \tag{4.3}$$

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^{\mathbf{r}}\mathbf{1}(|\theta_I| > \lambda n^{-1/2}) \leq 4n^{-\mathbf{r}/2}d_{\mathbf{r}}(\lambda)\sqrt{N_j}t_n. \tag{4.4}$$

4.3. Rate of testing

To complete the proof of the theorem, it remains to show that the latter statements and smoothness conditions $f \in B_{s,p,q}(M)$ imply the desired assertion (2.14).

The inclusion $f \in B_{s,p,q}(M)$ yields, see ISO2,

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^p \le CM^p 2^{-jp(s+\frac{1}{2}-\frac{1}{p})}. \tag{4.5}$$

Let D_j and E_j be the projector operators associated with the given wavelet transform. By ISO2, see also (4.5), and ISO3, for any function $f \in B_{s,p,q}(M)$

$$\|D_j f\|_p \le C2^{-js}. \tag{4.6}$$

For $r \le p$, one has also

$$\|D_j f\|_r \le \|D_j f\|_p \le C2^{-js} \tag{4.7}$$

and for $r > p$ in view of (4.5) and Hölder inequality

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^r \le \left| \sum_{I \in \mathcal{I}_j} |\theta_I|^p \right|^{r/p} \le CM^r 2^{-jr(s+\frac{1}{2}-\frac{1}{p})}$$

that implies by ISO3

$$\|D_j f\|_r \le C2^{-j(s-\frac{1}{p}+\frac{1}{r})}. \tag{4.8}$$

This statement is well known as the "embedding theorem for Besov classes", see e.g. Triebel (1992).

Note also that, using ISO1, ISO3 and the equality $N_j = 2^j$, $j > j_0$, one can rewrite the inequalities (4.2) and (4.3) in the following form: for each $j \in \mathcal{J}^n$

$$\|D_j f\|_2^2 \le Cn^{-1}2^{j/2}t_n, \tag{4.9}$$

$$\|D_j f\|_{\mathrm{r}}^{\mathrm{r}} \le C2^{j(\mathrm{r}/2-1)}n^{-\mathrm{r}/2}2^{j/2}t_n = C2^{j(\mathrm{r}/2-1/2)}n^{-\mathrm{r}/2}t_n. \tag{4.10}$$

Here and in what follows, C means a constant depending possibly on r, s, p but not on f, j, n and bounded uniformly over the whole range of adaptation $\mathcal{T}$.

Further we show that the above inequalities enable us to prove the result of the theorem. We use the following fact which is a simple consequence of the triangle inequality in the L_r-norm:

$$\begin{aligned} \|f\|_r &= \|f - E_{j_{\max}} f + D_{j_0} f + \ldots + D_{j_{\max}} f\|_r \le \\ &\le \|f - E_{j_{\max}} f\|_r + \sum_{j \in \mathcal{J}^n} \|D_j f\|_r. \end{aligned} \tag{4.11}$$

In view of (4.7) and (4.8) one may conlude that

$$\|f - E_{j_{\max}} f\|_r \le C2^{-j_{\max}s'} \le Cn^{-s'(1-\delta)/2} \ll \varrho_{\sigma,r}(n).$$

with $s' = \min\{s, s - 1/p + 1/r\}$. Therefore, it suffices to show that

$$\sum_{j \in \mathcal{J}^n} \|D_j f\|_r \le C \tilde{\varrho}_{\sigma,r}(n)$$

with $\tilde{\varrho}_{\sigma,r}(n)$ from (2.14).

We are checking this statement separately for different domains of the parameters p, r, s and for each case and each j, we pick one from the indicated above inequalities in an optimal way. We separate between the following cases taking into account the value of $\operatorname{argmax}\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$ and the relation between $p, r, 2$:

$\gamma^{(1)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$:

1. $r \le p$, $r \le 2$;
2. $p < r \le 2$;
3. $2 \le r \le p$;

$\gamma^{(2)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$:

1. $2 \le r \le p$;
2. $p < r$, $r \ge 2$;
3. $p < r \le 2$.

$\gamma^{(3)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$:

In the last case, one has automatically $p < r$ and $r > 2$. Using Lemma 2.1, it is not hard to check that any point outside the curves $\gamma^{(1)} = \gamma^{(2)}$ and $\gamma^{(2)} = \gamma^{(3)}$, see (2.3) - (2.5), belongs to one from the indicated above cases.

We start with the simplest case with $p \ge r$ and $r \le 2$ considered first in Ingster, 1982 (for $p = r = 2$) and Ingster, 1986. In this domain one has automatically $\gamma^{(1)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$.

4.3.1. The case of $r \le p, r \le 2$

For this case, it suffices to use only (4.9) and (4.6) i.e. we use the information delivered by the test $\phi_{n,2}$ and the smoothness properties of the function f. This means that the L_2-test $\phi_{n,2}$ provides the optimal rate of testing in this domain.

One has by (4.6)

$$\|D_j f\|_r \le \|D_j f\|_p \le C M 2^{-js}$$

and by (4.9)

$$\|D_j f\|_r \le \|D_j f\|_2 \le C (n^{-1} t_n 2^{j/2})^{1/2} = C n^{-1/2} t_n^{1/2} 2^{j/4}.$$

The following technical statement is helpful.

Lemma 4.5. *Let $\{d_j, j \in \mathcal{J}^n\}$ be a sequence such that for some positive $b_0, b_1, \epsilon_0, \epsilon_1$ one has*

$$d_j \le \min\{b_0 2^{-j\epsilon_0}, b_1 2^{j\epsilon_1}\}.$$

Then

$$\sum_{j\in\mathcal{J}^n} d_j \le C(\epsilon_0,\epsilon_1) b_0 2^{-j^*\epsilon_0} = \frac{2}{1-2^{\epsilon_0\wedge\epsilon_1}} b_0^{\frac{\epsilon_1}{\epsilon_0+\epsilon_1}} b_1^{\frac{\epsilon_0}{\epsilon_0+\epsilon_1}}.$$

Here j^* *is determined by the equality*

$$b_0 2^{-j^*\epsilon_0} = b_1 2^{j^*\epsilon_1} \tag{4.12}$$

and $C(\epsilon_0,\epsilon_1)$ *is uniformly bounded if* ϵ_0 *and* ϵ_1 *are separated away from zero.*

Proof. The proof is obvious. One applies $d_j \le b_0 2^{-j\epsilon_0}$ for $j \le j^*$ and $d_j \le b_1 2^{j\epsilon_1}$ for $j < j^*$. One may also use for $C(\epsilon_0,\epsilon_1)$ an estimate

$$C(\epsilon_0,\epsilon_1) \le \frac{2}{1-2^{-\epsilon_0\wedge\epsilon_1}}$$

and the assertion follows. □

We apply now this lemma with $d_j = \|D_j f\|$, $\epsilon_0 = s$ in (4.7) and $\epsilon_1 = 1/4$ in (4.9)

$$\sum_{j\in\mathcal{J}^n} d_j \le C(\epsilon_0,\epsilon_1) M 2^{-j^*s}$$

where j^* is due to

$$M2^{-j^*s} = n^{-1/2} t_n^{1/2} 2^{j^*/4}. \tag{4.13}$$

By the condition of the theorem, the values $\epsilon_0 = s$ and $\epsilon_1 = 1/4$ are separated away from zero in the whole range of adaptation. This implies

$$\sup_{\sigma\in\mathcal{T}} C(\epsilon_0,\epsilon_1) \le \sup_{\sigma\in\mathcal{T}} (1-2^{-s\wedge\frac{1}{4}})^{-1} \le C < \infty. \tag{4.14}$$

The relation (4.13) is equivalent to

$$2^{j^*} = (M^2 n/t_n)^{\frac{2}{4s+1}}.$$

With this choice

$$\sum_{j\in\mathcal{J}^n} \|D_j f\|_r \le CM2^{-j^*s} = CM^{\frac{1}{4s+1}} (n/t_n)^{-\frac{2s}{4s+1}} = CM^{\gamma^{(1)}} (n/t_n)^{-(1-\gamma^{(1)})/2}$$

and the theorem follows in the considered case.

4.3.2. The case of $p < r \le 2$, $\gamma^{(1)} > \gamma^{(2)}$

We apply again only (4.9) and (4.6) hence L_2-test $\phi_{n,2}$ provides again rate optimality.

The relation $p < r \le 2$ allows to estimate $\|D_j f\|_r$ for each $j \in \mathcal{J}^n$ with the help of Minkovskii's inequality

$$\|D_j f\|_r^r \le \|D_j f\|_2^{2\rho} \|D_j f\|_p^{p(1-\rho)}$$

where $\rho = (p-r)/(p-2)$. By (4.6), (4.9) we get for any $j \in \mathcal{J}^n$

$$\|D_j f\|_r^r \le C\left(n^{-1} t_n 2^{j/2}\right)^{\rho} \left(M2^{-js}\right)^{p(1-\rho)} = CM^{p(1-\rho)} n^{-\rho} t_n^{\rho} 2^{j[\rho/2 - p(1-\rho)]}. \tag{4.15}$$

Denote

$$\epsilon = \frac{\rho/2 - sp(1-\rho)}{r} = \frac{p - r - 2sp(r-2)}{2r(p-2)}. \tag{4.16}$$

and under the case into consideration one has, see (2.7),

$$\operatorname{sign}\epsilon = -\operatorname{sign}(\gamma^{(1)} - \gamma^{(2)}) = -1.$$

This allows to apply Lemma 4.5 with $\epsilon_0 = -\epsilon$ in (4.15) and $\epsilon_1 = 1/4$ in (4.9). One gets

$$\sum_{j\in\mathcal{J}^n} \|D_j f\|_r \le C n^{-1/2} t_n^{1/2} 2^{j^*/4}$$

where j^* is determined by the equation

$$(n^{-1/2} t_n^{1/2} 2^{j^*/4})^r = M^{p(1-\rho)} n^{-\rho}\, t_n^{\rho}\, 2^{j^*[\rho/2 - sp(1-\rho)]}. \tag{4.17}$$

The definition of ρ yields the identity $p(1-\rho) = r - 2\rho$ and thus (4.17) leads again to the selection rule for j^* in the form

$$2^{j^*} = (M^2 n/t_n)^{\frac{2}{4s+1}}.$$

This implies similarly to the above

$$\sum_{j\in\mathcal{J}^n} \|D_j f\|_r \le C M^{\gamma^{(1)}} (n/t_n)^{-(1-\gamma^{(1)})/2}.$$

4.3.3. The case of $2 \le r \le p$, $\gamma^{(1)} > \gamma^{(2)}$

We proceed similarly to the preceding case and again the L_2-test $\phi_{n,2}$ provides the optimal rate.

The Minkovskii inequality allows to get (4.15) but now ϵ from (4.16) is positive (since $p > 2$). We apply Lemma 4.5 with $\epsilon_1 = \epsilon$ in (4.15) and $\epsilon_0 = s$ in (4.7). One obtains

$$\sum_{j\in\mathcal{J}^n} \|D_j f\|_r \le C M 2^{-j^* s}$$

where j^* is determined by the equation

$$(M 2^{-j^* s})^r = M^{p(1-\rho)} n^{-\rho}\, t_n^{\rho}\, 2^{j[\rho/2 - sp(1-\rho)]}.$$

Using again the identity $p(1-\rho) = r - 2\rho$ we get

$$2^{j^*} = (M^2 n/t_n)^{\frac{2}{4s+1}}.$$

This implies similarly to the above

$$\sum_{j\in\mathcal{J}^n} \|D_j f\|_r \le C M^{\gamma^{(1)}} (n/t_n)^{-(1-\gamma^{(1)})/2}.$$

We resume that for all cases where $\gamma^{(1)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$, the L_2-test $\phi_{n,2}$ provides the optimal rate of testing.

4.3.4. The case of $2 \le r \le p$, $\gamma^{(1)} < \gamma^{(2)}$

In this case, we apply (4.9) (more precisely (4.15) which is based on (4.9) and (4.6)) and (4.10). This means that the optimal rate is obtained by a combination of testing in L_2-norm and in L_r-norm.

In the considered case of $\gamma^{(1)} < \gamma^{(2)}$, one has $\epsilon < 0$ in (4.16). Now we apply Lemma 4.5 for (4.15) with $\epsilon_0 = -\epsilon$ and for (4.10) with $\epsilon_1 = 1/2 - 1/(2r)$. We get

$$\sum_{j \in \mathcal{J}^n} \|D_j f\|_r \le C n^{-1/2} t_n^{1/r} 2^{j^* \epsilon_1} \tag{4.18}$$

where j^* is determined by

$$M^{p(1-s)} n^{-\rho} t_n^{\rho} 2^{j^*[\rho/2 - sp(1-\rho)]} = n^{-r/2} t_n 2^{j^*[r/2-1/2)]}$$

or

$$2^{j^*} = (M^2 n)^{\frac{1}{2s-1/p+1}} t_n^{-\frac{2}{p(2s-1/p+1)}}.$$

Substituting in (4.18), we obtain

$$\begin{aligned}
\sum_{j \in \mathcal{J}^n} \|D_j f\|_r &\le C n^{-1/2} t_n^{1/r} 2^{j^*(r-1)/(2r)} = \\
&= C M^{\frac{r-1}{2r(2s-1/p+1)}} n^{-1/2 + \frac{r-1}{2r(2s-1/p+1)}} t_n^{1/r - \frac{r-1}{pr(2s-1/p+1)}} = \\
&= C M^{\gamma^{(2)}} n^{-(1-\gamma^{(2)})/2} t_n^{1/r - \gamma^{(2)}/p}
\end{aligned} \tag{4.19}$$

and the assertion follows.

Now we consider the case with $\gamma^{(3)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$.

4.3.5. The case of $sp < (r-p)/2$

In this case, the inequality $sp > 1$ implies $p < r$ and $r > 2$. We will apply now (4.6), (4.8) and (4.1). This means that in this domain the test $\phi_{n,\infty}$ provides the optimal rate of testing.

By (4.1) and (4.5), one has for each $j \in \mathcal{J}^n$

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^r \le \left(\frac{\log N_j}{n}\right)^{\frac{r-p}{2}} \sum_{I \in \mathcal{I}_j} |\theta_I|^p \le C \left(\frac{\log N_j}{n}\right)^{\frac{r-p}{2}} M^p 2^{-jp(s-1/p+1/2)}. \tag{4.20}$$

In view of ISO3 this implies

$$\begin{aligned}
&\|D_j f\|_r^r \le C 2^{-j(\frac{r}{2}-1)} \left(\frac{\log N_j}{n}\right)^{\frac{r-p}{2}} M^p 2^{-jp(s-\frac{1}{p}+\frac{1}{2})} = \\
&= C \left(\frac{\log N_j}{n}\right)^{\frac{r-p}{2}} M^p 2^{j(\frac{r-p}{2}-sp)} \le C \left(\frac{\log n}{n}\right)^{\frac{r-p}{2}} M^p 2^{j(\frac{r-p}{2}-sp)}.
\end{aligned}$$

Now we use the latter inequality, (4.8) and Lemma 4.5 with $\epsilon_0 = (r-p-2sp)/(2r)$ and $\epsilon_1 = s - 1/p + 1/r$. We get

$$\sum_{j \in \mathcal{J}^n} \|D_j f\|_r \le CM 2^{-j^*(s-1/p+1/r)}$$

where j^* is due to

$$M^r 2^{-j^* r(s-1/p+1/r)} = \left(\frac{\log n}{n}\right)^{\frac{r-p}{2}} M^p 2^{j(r-p-2sp)/2}$$

or

$$2^{j^*} = \left(\frac{M^2 n}{\log n}\right)^{\frac{1}{2s-2/p+1}}.$$

With this choice, one gets

$$\sum_{j \in \mathcal{J}^n} \|D_j f\|_r \le CM \left(\frac{M^2 n}{\log n}\right)^{\frac{s-1/p+1/r}{2s-2/p+1}} = CM^{\gamma^{(3)}} \left(\frac{n}{\log n}\right)^{-(1-\gamma^{(3)})/2}$$

and the assertion follows.

It remains to consider the cases with $\gamma^{(2)} = \max\{\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}\}$ and $r > p$. The calculations differ slightly for $r \ge 2$ and $r < 2$.

4.3.6. The case of $r > p$, $r \ge 2$, $2sp > r - p$

Here we apply the statements (4.3) and (4.4) based on the tests $\phi_{n,r}$ and $\phi_{n,r,t}$. First we note that by (4.4) for each $j \in \mathcal{J}^n$ and any $\lambda \in \Lambda_j$

$$\begin{aligned} \sum_{I \in \mathcal{I}_j} |\theta_I|^r &= \sum_{I \in \mathcal{I}_j} |\theta_I|^r \mathbf{1}(|\theta_I| > \lambda/\sqrt{n}) + \sum_{I \in \mathcal{I}_j} |\theta_I|^r \mathbf{1}(|\theta_I| \le \lambda/\sqrt{n}) \le \\ &\le 4n^{-r/2} d_r(\lambda) 2^{j/2} t_n + (\lambda/\sqrt{n})^{r-p} \sum_{I \in \mathcal{I}_j} |\theta_I|^p. \end{aligned} \tag{4.21}$$

Since $f \in B_{s,p,q}(M)$, we get using ISO2 and ISO3

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^p \le CM^p 2^{-jp(s+1/2-1/p)}.$$

Let us fix now some $j^* \in \mathcal{J}^n$ and state for $j > j^*$

$$\lambda_j = \sqrt{2r(j-j^*)\log 2}.$$

Then, using again ISO3 and Lemma 4.4,iv, one has for $j > j^*$ and $2r(j-j^*)\log 2 \le 0.5 \log N_j$

$$\begin{aligned} &\|D_j f\|_r^r \le C 2^{j(\frac{r}{2}-1)} n^{-r/2} d_r(\lambda_j) 2^{j/2} t_n + (\lambda/\sqrt{n})^{r-p} M^p 2^{-jp(s+\frac{1}{2}-\frac{1}{p})} \le \\ &\le C 2^{j^*(\frac{r}{2}-\frac{1}{2})} n^{-r/2} t_n (j-j^*)^{r/2} 2^{-(j-j^*)/2} + C(j-j^*)^{-\frac{r-p}{2}} n^{-\frac{r-p}{2}} M^p 2^{-j(sp+\frac{p}{2}-\frac{r}{2})}. \end{aligned}$$

If $2r(j-j^*)\log 2 > 0.5 \log N_j$, then $\|D_j f\|_r$ can be estimated even better (with the only second term) using (4.20).

Easy to check that the last inequality yields

$$\sum_{j=j^*+1}^{j_{\max}} \|D_j\|_r \le C\left[2^{j^*(\frac{r}{2}-\frac{1}{2})} n^{-r/2} t_n\right]^{1/r} + C\left[n^{-\frac{r-p}{2}} M^p 2^{-j^*(sp+\frac{p}{2}-\frac{r}{2})}\right]^{1/r}.$$

Next, for $j \le j^*$ we use (4.10) which gives

$$\sum_{j=j_0}^{j^*} \|D_j\|_r \le C\left[n^{-r/2} t_n 2^{j^*(\frac{r}{2}-\frac{1}{2})}\right]^{1/r}.$$

Let us take level j^* due to

$$n^{-r/2} t_n 2^{j^*(\frac{r}{2}-\frac{1}{2})} = n^{-\frac{r-p}{2}} M^p 2^{-j^*(sp+\frac{p}{2}-\frac{r}{2})}$$

or

$$2^{j^*} = (M^2 n t_n^{-2/p})^{1/(2s+1-1/p)}. \tag{4.22}$$

We get now

$$\sum_{j=j_0}^{j_{\max}} \|D_j\|_r \le C\left[n^{-r/2} t_n 2^{j^*(\frac{r}{2}-\frac{1}{2})}\right]^{1/r} = C M^{\gamma^{(2)}} n^{-(1-\gamma^{(2)})/2} t_n^{1/r-\gamma^{(2)}/p}$$

as required.

4.3.7. The case of $p < r < 2$, $\gamma^{(1)} < \gamma^{(2)}$

We proceed as above but with 2 in place of r (since $r < 2$).

For each $j \in \mathcal{J}^n$ and any $\lambda \in \Lambda_j$

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^r = \sum_{I \in \mathcal{I}_j} |\theta_I|^r \mathbf{1}(|\theta_I| > \lambda/\sqrt{n}) + \sum_{I \in \mathcal{I}_j} |\theta_I|^r \mathbf{1}(|\theta_I| \le \lambda/\sqrt{n}).$$

Let again some $j^* \in \mathcal{J}^n$ be fixed. We set for $j > j^*$

$$\lambda_j = \sqrt{4(j-j^*)\log 2}.$$

Now by (4.4), using then Hölder inequality and Lemma 4.4, iv, one gets

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^r \mathbf{1}(|\theta_I| > \lambda_j/\sqrt{n}) \le \left[\sum_{I \in \mathcal{I}_j} |\theta_I|^2 \mathbf{1}(|\theta_I| > \lambda_j/\sqrt{n})\right]^{r/2} \le$$
$$\le \left[4 n^{-1} t_n d_2(\lambda_j) 2^{j/2}\right]^{r/2} \le C\left[n^{-1} t_n 2^{j/2} \sqrt{j-j^*}\, 2^{-(j-j^*)}\right]^{r/2}.$$

Since $f \in B_{s,p,q}(M)$, we get using ISO2 and ISO3

$$\sum_{I \in \mathcal{I}_j} |\theta_I|^p \le C M^p 2^{-jp(s+1/2-1/p)}.$$

Now we obtain as above

$$\begin{aligned}
\sum_{j=j^*+1}^{j_{\max}} \|D_j\|_r &\le C \sum_{j=j^*+1}^{j_{\max}} \left[2^{j(\frac{r}{2}-1)} \sum_{I\in\mathcal{I}_j} |\theta_I|^r \right]^{1/r} \le \\
&\le C \sum_{j=j^*+1}^{j_{\max}} \left[2^{j^*(\frac{3r}{4}-1)} n^{-r/2} t_n^{r/2} \sqrt{j-j^*}\, 2^{-(j-j^*)(\frac{r}{4}+\frac{1}{2})} \right]^{1/r} + \\
&\quad C \sum_{j=j^*+1}^{j_{\max}} \left[(j-j^*)^{-\frac{r-p}{2}} n^{-\frac{r-p}{2}} M^p 2^{-j(sp+\frac{p}{2}-\frac{r}{2})} \right]^{1/r} \le \\
&\le C \left[2^{j^*(\frac{3r}{4}-1)} n^{-r/2} t_n^{r/2} \right]^{1/r} + C \left[n^{-\frac{r-p}{2}} M^p 2^{-j^*(sp+\frac{p}{2}-\frac{r}{2})} \right]^{1/r}.
\end{aligned}$$

Here we have used that $2sp - r + p > 0$.

Next, for $j \le j^*$ we use (4.15) which gives

$$\sum_{j=j_0}^{j^*} \|D_j\|_r \le C \left[M^{p(1-\rho)} n^{-\rho}\, t_n^{\rho}\, 2^{j[\rho/2 - p(1-\rho)]} \right]^{1/r} \le C \left[M^{p(1-\rho)} n^{-\rho}\, t_n^{\rho}\, 2^{j^*[\rho/2-p(1-\rho)]} \right]^{1/r}$$

since $\rho/2 - p(1-\rho) > 0$ for $\rho = (r-p)/(2-p)$ if $\gamma^{(1)} < \gamma^{(2)}$ and $p < 2$, see (2.7) and (4.16).

Let us take the level j^* again due to (4.22). We get now

$$\begin{aligned}
n^{-r/2} t_n 2^{j^*(\frac{r}{2}-\frac{1}{2})} = n^{-\frac{r-p}{2}} M^p 2^{-j^*(sp+\frac{p}{2}-\frac{r}{2})} &= M^{p(1-\rho)} n^{-\rho}\, t_n^{\rho}\, 2^{j^*[\rho/2-p(1-\rho)]} = \\
&= \left[M^{\gamma^{(2)}} n^{-(1-\gamma^{(2)})/2} t_n^{1/r-\gamma^{(2)}/p} \right]^r
\end{aligned}$$

Since $3r/4 - 1 > r/2 - 1/2$ for $r > 2$, we obtain

$$\sum_{j=j_0}^{j_{\max}} \|D_j\|_r \le C M^{\gamma^{(2)}} n^{-(1-\gamma^{(2)})/2} t_n^{1/r-\gamma^{(2)}/p}.$$

This completes the proof of the theorem.

Appendix

4.4. Proof of Theorem 2.1

The proof follows to Ingster, 1993 and we indicate below only the main points. Given $\sigma = (s,p,q,M)$, let $\varrho(n) = \varrho_{\sigma,r}(n)$ be from (2.8) and let $\varrho'(n)$ be such that $c_\varepsilon = \varrho'(n)/\varrho(n) \to 0$ as $n \to \infty$. To simplify the notation, we will write ϱ_n' instead of $\varrho'(n)$. We have to show that for any tests ϕ_n

$$\liminf_{n\to\infty} \left[\alpha(\phi_n) + \beta_{\sigma,r}(\phi_n, \varrho_n') \right] \ge 1. \tag{4.23}$$

Here, recall, $\beta_{\sigma,r}(\phi_n,\varrho) = \sup_{f\in\mathcal{F}_{\sigma,r}(\varrho)} \mathbf{P}_f(\phi_n = 0)$ and $\mathcal{F}_{\sigma,r}(\varrho) = \{f \in B_{s,p,q}(M) : \|f\|_r \ge \varrho\}$. Au usual, proving such sort of results, one changes the minimax problem by a proper Bayes one. Due to Ingster, 1993, it suffices to construct such measures π_n on the space $B_{s,p,q}$ that

$$\pi_n(\mathcal{F}_{\sigma,r}(\varrho_n')) = 1 - o_n(1) \tag{4.24}$$

and

$$Z_{\pi_n} = d\mathbf{P}_{\pi_n}/d\mathbf{P}_0 \to 1. \tag{4.25}$$

Here $\mathbf{P}_{\pi_n}$ is the Bayes measure corresponding to the prior π_n, see e.g. Lehmann, 1959, and the last convergence is treated in probability under the measure $\mathbf{P}_0$.

For the proof of these statements, we pass to the sequence space model and identify the function f with the set of the corresponding wavelet coefficients $\theta = \{\theta_I, I \in \mathcal{I}\}$. Under the Bayes approach, these coefficients are random and the measure π_n describes their joint distribution.

Below we consider priors with a special levelwise structure. More precisely, given $j \in \mathcal{J}^n$ and two numbers $\mu \in [0,1]$ and $v > 0$, introduce a prior $\pi_{j,\mu,v}$ such that $\theta_I = 0$ for all $I \notin \mathcal{I}_j$ and $\{\theta_I, I \in \mathcal{I}_j\}$ are iid with the three-points distribution of the form

$$\pi_{j,\mu,v}(\theta_I = 0) = 1 - \mu, \qquad \pi_{j,\mu,v}(\theta_I = \pm v) = \mu/2, \qquad I \in \mathcal{I}_j. \tag{4.26}$$

Further we are constructing a sequence (j_n, μ_n, v_n) in such a way that the corresponding priors $\pi_n = \pi_{j_n,\mu_n,v_n}$ obey (4.24) and (4.25). Following to Ingster, 1993, we use the following three families of priors π_{j_n,μ_n,v_n}:

$\pi_n^{(1)}$: $\mu_n^{(1)} = 1$, $\quad v_n^{(1)} = n^{-1/2}2^{-j_n/4}c_n^{1/2}$;
$\pi_n^{(2)}$: $\mu_n^{(2)} = 2^{-j_n/2}$, $v_n^{(2)} = n^{-1/2}c_n^{1/2}$;
$\pi_n^{(3)}$: $\mu_n^{(3)} = 2^{-j_n}$, $\quad v_n^{(3)} = (n/\log n)^{-1/2}c_n^{1/2}$.

Here $c_n = \varrho_n'/\varrho_{\sigma,r}(n) = o_n(1)$.

Lemma 4.6 (Ingster, 1993). *Let $j_n \to \infty$ and $c_n \to 0$. Then (4.25) holds true for each from three indicated above families of priors $\pi_n = \pi_n^{(i)}$, $i = 1,2,3$.*

Next, it is easy to observe using ISO2 and ISO3 that the condition (4.24) for the priors $\pi_n^{(i)}$, $i = 1,2,3$ can be represented in the form

$$\pi_n(\|D_{j_n}\theta\|_p \le CM2^{-j_n(s+1/2-1/p)}) = 1 - o_n(1),$$
$$\pi_n(\|D_{j_n}\theta\|_r \ge C2^{-j_n(1/2-1/r)}\varrho_n') = 1 - o_n(1),$$

where $D_j\theta = \{\theta_I, I \in \mathcal{I}_j\}$ means the subvector of θ corresponding to jth level. Since θ_I are iid within the level j, one gets

$$E_{\pi_n}\|D_{j_n}\theta\|_r^r = N_{j_n}\mu_n v_n^r = 2^{j_n}\mu_n v_n^r, \qquad j > j_0,$$

and similarly for $\|D_{j_n}\theta\|_p$. Using these equalities and again the iid structure of $\{\theta_I, I \in \mathcal{I}_j\}$, one can reduce the above statements to the following conditions, see Ingster, 1993,

$$M^{-p} 2^{j_n(sp+p/2)} \mu_n v_n^p c_n^{p/2} \leq C, \tag{4.27}$$

$$2^{j_n r/2} \mu_n v_n^r c_n^{r/2} \geq C(\varrho_n')^r \tag{4.28}$$

It remains to show that these conditions are fulfilled under a proper choice of j_n. Define $j_n^{(i)}$ for $i = 1, 2, 3$ by the relations

$$\begin{aligned} 2^{j_n^{(1)}} &= (M^2 n)^{\frac{2}{4s+1}}, \\ 2^{j_n^{(2)}} &= (M^2 n)^{\frac{1}{2s+1-1/p}}, \\ 2^{j_n^{(3)}} &= (M^2 n / \log n)^{\frac{1}{2s+1-2/p}}. \end{aligned}$$

Straightforward calculation shows that with these choice, one gets for each of priors $\pi_n^{(i)} = (j_n^{(i)}, \mu_n^{(i)}, v_n^{(i)})$, $i = 1, 2, 3$,

$$\begin{aligned} M^{-p} 2^{j_n^{(i)}(sp+p/2)} \mu_n |v_n^{(i)}|^p &= 1, \\ 2^{j_n^{(i)} r/2} \mu_n^{(i)} |v_n^{(i)}|^r &= |\varrho^{(i)}(n)|^r. \end{aligned}$$

this implies (4.27) and (4.28) since $c_n \to 0$ as $n \to \infty$ and the theorem follows.

4.5. Proof of Lemmas 4.1–4.4

First we recall that all the statement of these lemmas relate to the case of null hypothesis that is we observe a pure noise and all the considered statistics are sums of iid random variables.

The statements of Lemma 4.1 are obvious since the normalized empirical wavelet coefficients Y_I coincides with Gaussian errors ξ_I. Statements (i) and (ii) of Lemmas 4.2 through 4.4 are also straightforward. To prove assertion (iii) of these lemmas we use the following general technical assertion.

Lemma 4.7. *Let random variables $\eta_{i,n}$, $i = 1, \dots, n$, be independent identically distributed and bounded,*

$$|\eta_{i,n}| \leq C_1 (\log n)^k \tag{4.29}$$

with some positive k and C_1. If also

$$d_n^2 \geq C_2 n^{-1/4} \tag{4.30}$$

where $d_n^2 = \mathbf{D}\eta_{1,n}$, then for any constant $C_3 > 0$, one has for n large enough and any $a < C_3 \log n$

$$\mathbf{P}\left(\sum_{i=1}^{n} (\eta_{i,n} - \mathbf{E}\eta_{i,n}) > a d_n \sqrt{n}\right) \leq 2e^{-a^2/2}.$$

Proof. Denote

$$\begin{aligned} b_n &= \mathbf{E}\eta_{1,n}, \\ \mu &= \frac{a}{d_n\sqrt{n}}. \end{aligned}$$

We apply the Chebyshev exponential inequality

$$\begin{aligned} \mathbf{P}\left(\sum_{i=1}^{n}\eta_{i,n} > nb_n + ad_n\sqrt{n}\right) &\le \exp\{-\mu(nb_n + ad_n\sqrt{n})\}\,\mathbf{E}\exp\left\{\sum_{i=1}^{n}\mu\eta_{i,n}\right\} = \\ &= \exp\{-\mu n b_n - a^2 + nI(\mu)\} \end{aligned}$$

where

$$I(\mu) = \log \mathbf{E}e^{\mu\eta_{i,n}}.$$

By Taylor' expansion

$$I(\mu) = I(0) + \mu I'(0) + \frac{\mu^2}{2}I''(0) + \frac{\mu^3}{6}I'''(\theta\mu)$$

with some $|\theta| \le 1$. Easy to check using (4.29) that

$$\begin{aligned} I(0) &= 0, \\ I'(0) &= b_n, \\ I''(0) &= d_n^2, \\ I'''(\theta\mu) &\le 6(C_1\log n)^{3k} \end{aligned}$$

and therefore by (4.30)

$$n\frac{\mu^3}{6}|I'''(\theta\mu)| \le \frac{a^3(C_1\log n)^{3k}}{d_n^3\sqrt{n}} \le \frac{C_1^3(C_3\log n)^{3k+3}}{C_2^{3/2}n^{1/8}} \to 0, \qquad n\to\infty.$$

Since $a < C_3\log n$, this yields for n large enough

$$\mathbf{P}\left(\sum(\eta_{i,n} - b_n) > ad_n\sqrt{n}\right) \le e^{-a^2/2+o_n(1)} < 2e^{-a^2/2}.$$

The lemma is proved. □

Now we show how e.g. statement (iii) of Lemma 4.4 can be proved with the help of this result. Let $A_n = 2\sqrt{\log n}$. Given j, set for $I \in \mathcal{I}_j$ and $\lambda \in \Lambda_j$

$$\eta_I = |\xi_I|^r\mathbf{1}(\lambda < |\xi_I| < A_n).$$

Denote $b_n = \mathbf{E}\eta_I$, $d_n^2 = \mathbf{D}\eta_I$. Straightforward calculation show that

$$\begin{aligned} &|b_n - b_r(\lambda)| \le \mathbf{E}|\xi|^r\mathbf{1}(|\xi| > A_n) \le Cn^{-2}, \\ &|d_n^2 - d_r^2(\lambda)| \le Cn^{-2}. \end{aligned}$$

Since obviously $\sum_{I\in\mathcal{I}_j}\eta_I \le S_{j,r}(\lambda)$, then in view of Lemma 4.4,iv, the desirable assertion follows directly from Lemma 4.7.

The statement (iii) of Lemma 4.3 is a particular case of this result with $\lambda = 0$. The same is true for Lemma 4.2,iii, where, moreover, $r = 2$.

4.6. Proof of Propositions 4.1 – 4.4

We start with Proposition 4.1. Let $\sqrt{n}|\theta_I| > 4\sqrt{\log N_j}$ for some $j \in \mathcal{J}^n$ and $I \in \mathcal{I}_j$. Then

$$\mathbf{P}_f(\phi_{n,\infty} = 0) \leq \mathbf{P}_f(|\sqrt{n}\theta_I + \xi_I| < 2\sqrt{\log N_j}) \leq \mathbf{P}_f(|\xi_I| > 2\sqrt{\log N_j}) = o_n(1)$$

and the assertion follows.

4.6.1. Proof of Proposition 4.2

Let $\bar{S}_j \geq 4d_2t_n\sqrt{N_j}$ for some $j \in \mathcal{J}^n$. Obviously

$$\mathbf{P}_f(\phi_{n,2} = 0) \leq \mathbf{P}_f(S_j < N_j + 2d_2t_n\sqrt{N_j})$$

and using the Cauchy-Schwarz inequality

$$\begin{aligned}\mathbf{P}_f(\phi_{n,2} = 0) &= \mathbf{P}_f(-(S_j - \mathbf{E}_f S_j) > \mathbf{E}_f S_j - N_j - 2d_2t_n\sqrt{N_j}) \leq \\ &\leq \frac{\mathbf{D}_j S_j}{(\mathbf{E}_f S_j - N_j - 2d_2t_n\sqrt{N_j})^2}.\end{aligned}$$

Since ξ_I are independent for different I, one gets easily

$$\begin{aligned}\mathbf{E}_f S_j &= \sum_{I\in\mathcal{I}_j} E_f|\sqrt{n}\theta_I + \xi_I|^2 = \sum_{I\in\mathcal{I}_j}(n\theta_I^2 + 1) = \bar{S}_j + N_j, \\ \mathbf{D}_f S_j &= \sum_{I\in\mathcal{I}_j} D_f|\sqrt{n}\theta_I + \xi_I|^2 = \sum_{I\in\mathcal{I}_j}(4n\theta_I^2 + 2) = 4\bar{S}_j + 2N_j.\end{aligned}$$

Now the above estimate can be rewritten in the form

$$\mathbf{P}_f(\phi_{n,2} = 0) \leq \frac{4\bar{S}_j + 2N_j}{(\bar{S}_j - 2d_2t_n\sqrt{N_j})^2}.$$

Since the function $g(x) = \frac{a+bx}{(x-c)^2}$ is for any positive a, b, c monotonously decreasing for $x > c$, one gets by $\bar{S}_j \geq 4d_2t_n\sqrt{N_j}$

$$\mathbf{P}_f(\phi_{n,2} = 0) \leq \frac{4 \cdot 4d_2t_n\sqrt{N_j} + 2N_j}{(2d_2t_n\sqrt{N_j})^2} \leq \frac{1 + o_n(1)}{4t_n^2} = o_n(1).$$

Here we have used that $t_n \to \infty$ as $n \to \infty$.

4.6.2. Proof of Proposition 4.4

Let $r \geq 2$ and let for some $j \in \mathcal{J}^n$ and $\lambda \in \Lambda_j$,

$$\bar{S}_{j,r}(\lambda) \geq 4t_n d_r(\lambda)\sqrt{N_j}. \tag{4.31}$$

Recall also that $|\sqrt{n}\theta_I| \le 4\sqrt{\log N_j}$ for all $I \in \mathcal{I}_j$ due to the conditions of the proposition. One gets similarly to the above

$$\begin{aligned}
&\mathbf{P}_f(\phi_{n,r,t} = 0) \le \mathbf{P}_f(S_{j,r}(\lambda) < b_r(\lambda)N_j + 3d_r(\lambda)t_n\sqrt{N_j}) = \\
&= \mathbf{P}_f(-(S_{j,r}(\lambda) - \mathbf{E}_f S_{j,r}(\lambda)) > \mathbf{E}_f S_{j,r}(\lambda) - b_r(\lambda)N_j - 3d_r(\lambda)t_n\sqrt{N_j}) \le \\
&\le \frac{\mathbf{D}_f S_{j,r}(\lambda)}{(\mathbf{E}_f S_{j,r}(\lambda) - b_r(\lambda)N_j - 3d_r(\lambda)t_n\sqrt{N_j})^2}.
\end{aligned}$$

To estimate $\mathbf{E}_f S_{j,r}(\lambda)$ and $\mathbf{D}_f S_{j,r}(\lambda)$, we use the following technical statements.

Lemma 4.8. *Let $\lambda \ge \sqrt{r}$ and let ξ be standard normal. Then there are some positive constants C_1, C_2 such that*

(*i*) *for any x*

$$\mathbf{D}|x+\xi|^r \mathbf{1}(|x+\xi| > \lambda) \le C_1 \left[\lambda^r \mathbf{E}|x+\xi|^r \mathbf{1}(|x+\xi| > \lambda) + |x|^{2r}\mathbf{1}(|x| > \lambda)\right];$$

(*ii*) *for any x*

$$\mathbf{E}|x+\xi|^r \mathbf{1}(|x+\xi| > \lambda) \ge \mathbf{E}|\xi|^r \mathbf{1}(|\xi| > \lambda);$$

(*iii*) *if $|x| \ge \lambda$, then*

$$\mathbf{E}|x+\xi|^r \mathbf{1}(|x+\xi| > \lambda) \ge |x|^r + \mathbf{E}|\xi|^r \mathbf{1}(|\xi| > \lambda);$$

(*iv*) $$\mathbf{D}|\xi|^r \mathbf{1}(|\xi| > \lambda) \ge C_2 \lambda^r \mathbf{E}|\xi|^r \mathbf{1}(|\xi| > \lambda).$$

Proof of statements (i), (iii), (iv) is straightforward and left to the reader. Statement (ii) follows from Andersen's lemma, see Ibragimov and Khasminskii, 1981[Lemma 2.12.1].

Using Lemma 4.8, i, and the condition $|\sqrt{n}\theta_I| \le 4\sqrt{\log N_j}$ we have

$$\mathbf{D}_f S_{j,r}(\lambda) \le C_1 \lambda^r \mathbf{E}_f S_{j,r}(\lambda) + C_1 (4\log N_j)^{r/2} \bar{S}_{j,r}(\lambda).$$

Next, by Lemma 4.8, ii, iii, and (4.31)

$$\mathbf{E}_f S_{j,r}(\lambda) \ge b_r(\lambda)N_j + \bar{S}_{j,r}(\lambda) \ge b_r(\lambda)N_j + 4d_r(\lambda)t_n\sqrt{N_j}.$$

Hence as above in the proof of Proposition 4.3

$$\mathbf{P}_f(\phi_{n,r,t} = 0) \le \frac{C_1\lambda^r b_r(\lambda)N_j + 4C_1[\lambda^r + (4\log N_j)^{r/2}]d_r(\lambda)t_n\sqrt{N_j}}{d_r^2(\lambda)t_n^2 N_j}.$$

To complete the proof we note that by Lemma 4.4,iv, $d_r(\lambda) \ge N_j^{-1/4}$ and by definition $\lambda < \sqrt{0.5\log N_j}$ for $\lambda \in \Lambda_j$. Using also Lemma 4.8,iv, we get

$$\mathbf{P}_f(\phi_{n,r,t} = 0) \le \frac{C_1\lambda^r b_r(\lambda)}{d_r^2(\lambda)t_n^2} + \frac{5C_1(4\log N_j)^{r/2}}{d_r(\lambda)t_n\sqrt{N_j}} = o_n(1).$$

Proposition 4.3 can be proved in the same line with inessential modifications. We omit the details.

References

BURNASHEV, M. (1979). On the minimax detection of an inaccurately known signal in a white gaussian noise background. *Theory Probab. Appl.*, 24:107 – 119.

COHEN, A., DAUBECHIES, I., JAWERTH, B., AND VIAL, P. (1993a). Multiresolution analysis, wavelets, and fast algorithms on an interval. *Comptes Rendus Acad. Sci. Paris (A)*, 316:417–421.

COHEN, A., DAUBECHIES, I., AND VIAL, P. (1993b). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis (A)*, 1:54–81.

DAUBECHIES, I. (1992). *Ten Lectures on Wavelets.* SIAM, Philadelphia.

DONOHO, D. AND JOHNSTONE, I. (1995). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455.

DONOHO, D., JOHNSTONE, I., KERKYACHARIAN, G., AND PICARD, D. (1995). Wavelet shrinkage: asymptopia? *J. Royal Statist. Soc. (Ser.B)*, 57:301–369.

ERMAKOV, M. (1990). Minimax detection of a signal in a white gaussian noise. *Theory Probab. Appl.*, 35:667 – 679.

IBRAGIMOV, I. AND KHASMINSKII, R. (1977). One problem of statistical estimation in gaussian white noise. *Soviet Math. Dokl.*, 236(4):1351 – 1354.

IBRAGIMOV, I. AND KHASMINSKII, R. (1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

INGSTER, Y. (1982). Minimax nonparametric detection of signals in white gaussian noise. *Problems Inform. Transmission*, 18:130 – 140.

INGSTER, Y. (1986). Minimax testing of nonparametric hypothesis on a distribution density in l_p-metrics. *Theory Probab. Appl.*, 32:333 – 337.

INGSTER, Y. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. *Math. Methods of Statist.*, 2,3,4:85 – 114,171 – 189,249 – 268.

JUDITSKY, A. (1995). Adaptive wavelet estimators. *Math. Methods of Statist.*, ?:? in press.

LEHMANN, E. (1959). *Testing Statistical Hypothesis.* Wiley.

LEPSKI, O., MAMMEN, E., AND SPOKOINY, V. (1994). Ideal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selection. *Ann. Statist.* to appear.

LEPSKI, O. AND SPOKOINY, V. (1995). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. Technical Report 44, Humboldt Univ., Berlin. (submitted in *Bernoulli*).

MANN, H. AND WALD, A. (1942). On the choice of the number of intervals in the application of the chi-square test. *Ann. Math. Stat.*, 13:306–317.

MEYER, Y. (1990). *Ondlettes.* Herrmann, Paris.

NEMIROVSKI, A. (1985). On nonparametric estimation of smooth regression function.

Sov. J. Comput. Syst. Sci., 23(6):1–11.

SPOKOINY, V. (1995). Adaptive hypothesis testing using wavelets. Technical Report 176, Weierstrass Institute, Berlin. (submitted in *Ann. Statist.*).

TRIEBEL, H. (1992). *Theory of function spaces. II.* Birkhauser, Basel.

Recent publications of the
Weierstraß–Institut für Angewandte Analysis und Stochastik

Preprints 1995

205. Michael H. Neumann: On the effect of estimating the error density in nonparametric deconvolution.

206. Wolfgang Dahmen, Angela Kunoth, Reinhold Schneider: Operator equations, multiscale concepts and complexity.

207. Annegret Glitzky, Konrad Gröger, Rolf Hünlich: Free energy and dissipation rate for reaction diffusion processes of electrically charged species.

208. Jörg Schmeling: A dimension formula for endomorphisms – The Belykh family.

209. Alfred Liemant: Leitfähigkeit eindimensionaler periodischer elektrischer Netze.

210. Günter Albinus: A thermodynamically motivated formulation of the energy model of semiconductor devices.

211. Dmitry Ioffe: Extremality of the disordered state for the Ising model on general trees.

212. Stefan Seelecke: Equilibrium thermodynamics of pseudoelasticity and quasiplasticity.

Preprints 1996

213. Björn Sandstede: Stability of N–fronts bifurcating from a twisted heteroclinic loop and an application to the FitzHugh–Nagumo equation.

214. Jürgen Sprekels, Songmu Zheng, Peicheng Zhu: Asymptotic behavior of the solutions to a Landau–Ginzburg system with viscosity for martensitic phase transitions in shape memory alloys.

215. Yuri I. Ingster: On some problems of hypothesis testing leading to infinitely divisible distributions.

216. Grigori N. Milstein: Evaluation of moment Lyapunov exponents for second order linear autonomous SDE.

217. Hans Günter Bothe: Shift spaces and attractors in non invertible horse shoes.

218. Gianfranco Chiocchia, Siegfried Prößdorf, Daniela Tordella: The lifting line equation for a curved wing in oscillatory motion.

219. Pavel Krejčí, Jürgen Sprekels: On a system of nonlinear PDE's with temperature–dependent hysteresis in one–dimensional thermoplasticity.

220. Boris N. Khoromskij, Siegfried Prößdorf: Fast computations with the harmonic Poincaré–Steklov operators on nested refined meshes.

221. Anton Bovier, Véronique Gayrard: Distribution of overlap profiles in the one–dimensional Kac–Hopfield model.

222. Jürgen Sprekels, Dan Tiba: A duality–type method for the design of beams.

223. Wolfgang Dahmen, Bernd Kleemann, Siegfried Prößdorf, Reinhold Schneider: Multiscale methods for the solution of the Helmholtz and Laplace equation.

224. Herbert Gajewski, Annegret Glitzky, Jens Griepentrog, Rolf Hünlich, Hans–Christoph Kaiser, Joachim Rehberg, Holger Stephan, Wilfried Röpke, Hans Wenzel: Modellierung und Simulation von Bauelementen der Nano– und Optoelektronik.

225. Andreas Rathsfeld: A wavelet algorithm for the boundary element solution of a geodetic boundary value problem.

226. Sergej Rjasanow, Wolfgang Wagner: Numerical study of a stochastic weighted particle method for a model kinetic equation.

227. Alexander A. Gushchin: On an information–type inequality for the Hellinger process.

228. Dietmar Horn: Entwicklung einer Schnittstelle für einen DAE–Solver in der chemischen Verfahrenstechnik.

229. Oleg V. Lepski, Vladimir G. Spokoiny: Optimal pointwise adaptive methods in nonparametric estimation.

230. Bernd Kleemann, Andreas Rathsfeld, Reinhold Schneider: Multiscale methods for boundary integral equations and their application to boundary value problems in scattering theory and geodesy.

231. Jürgen Borchardt, Ludger Bruell, Friedrich Grund, Dietmar Horn, Frank Hubbuch, Tino Michael, Horst Sandmann, Robert Zeller: Numerische Lösung großer strukturierter DAE–Systeme der chemischen Prozeßsimulation.

232. Herbert Gajewski, Klaus Zacharias: Global behaviour of a reaction–diffusion system modelling chemotaxis.

233. Fréderic Guyard, Reiner Lauterbach: Forced symmetry breaking perturbations for periodic solutions.