Electronic Journal of Statistics
Vol. 5 (2011) 484–506
ISSN: 1935-7524
DOI: 10.1214/11-EJS615

Estimation and detection of functions from anisotropic Sobolev classes

Yuri Ingster*

Department of Mathematics II
St. Petersburg State Electrotechnical University
5 Professor Popov Street
197376 St.Petersburg
Russia
e-mail: yurii_ingster@mail.ru

and

Natalia Stepanova†

School of Mathematics and Statistics
Carleton University
1125 Colonel By Drive
Ottawa, ON K1S 5B6
Canada
e-mail: nstep@math.carleton.ca

Abstract: We consider the problems of estimating and detecting an unknown function f depending on a multidimensional variable (for instance, an image) observed in the Gaussian white noise. It is assumed that f belongs to anisotropic Sobolev class. The case of a function of infinitely many variables is also considered. An asymptotic study (as the noise level tends to zero) of the estimation and detection problems is done. In connection with the estimation problem, we construct asymptotically minimax estimators and establish sharp asymptotics for the minimax integrated squared risk. In the detection problem, we construct asymptotically minimax tests and provide conditions for distinguishability in the problem.

AMS 2000 subject classifications: Primary 62G10; secondary 62G20.
Keywords and phrases: Nonparametric estimation, nonparametric signal detection, Gaussian white noise, multivariate functions, anisotropic smoothness.

Received December 2010.

1. Introduction

Recently nonparametric estimation and detection of multivariate signals, in a variety of estimation and testing schemes, aroused considerable interest. In this paper we study the problem of estimating and detecting a multivariate function

*Research partly supported by RFBR grant 11-01-00577 and by the grant NSh–4472.2010.1.
†Research supported by an NSERC grant.

$f \in \mathcal{F} \subset L_2([0,1]^d) = L_2^d$, $1 \le d \le \infty$, observed in the Gaussian white noise model

$$X_\varepsilon = f + \varepsilon W, \tag{1.1}$$

where W is a d-dimensional Gaussian white noise, $\varepsilon > 0$ is a small parameter (noise intensity), and $\mathcal{F}$ is a subset of L_2^d that consists of sufficiently smooth functions. In this model, the "observation" is the function $X_\varepsilon \colon L_2^d \to \mathcal{G}$ taking its values in the set $\mathcal{G}$ of normal random variables such that if $\xi = X_\varepsilon(\phi)$, $\eta = X_\varepsilon(\psi)$, where $\phi, \psi \in L_2^d$, then $\mathbf{E}(\xi) = (f, \phi)$, $\mathbf{E}(\eta) = (f, \psi)$, and $\mathbf{Cov}(\xi, \eta) = \varepsilon^2(\phi, \psi)$. For any $f \in L_2^d$, the observation X_ε determines the Gaussian measure $\mathbf{P}_{\varepsilon,f}$ on L_2^d with mean function f and covariance operator $\varepsilon^2 I$, where I is the identity operator (see [4, 17] for references). The corresponding expectation is denoted by $\mathbf{E}_{\varepsilon,f}$. In this paper we study the case of fixed and finite d and the case $d = \infty$.

We assume that f belongs to a Sobolev class $\mathcal{F}$ of functions with *anisotropic* constraints of regularity. One problem of interest is to estimate an unknown signal f using quadratic loss. Another problem of interest is to detect f, that is, to test the hypothesis $H_0 \colon f = 0$ versus a family of nonparametric alternatives of the form $H_{1\varepsilon} \colon f \in \mathcal{F}$, $\|f\|_2 \ge r_\varepsilon$, where $\|\cdot\|_2$ is the L_2-norm and $r_\varepsilon \to 0$ is a positive family.

Let $\{\phi_l\}_{l \in \mathcal{L}}$ be a fixed orthonormal basis in L_2^d, with $\mathcal{L}$ being a countable set. Then model (1.1) can be equivalently represented by the Gaussian sequence space model

$$X_{\varepsilon,l} = \theta_l + \varepsilon \xi_l, \quad \xi_l \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \quad l \in \mathcal{L}, \tag{1.2}$$

where $\theta_l = (f, \phi_l)$ are the Fourier coefficients of f with respect to the basis $\{\phi_l\}_{l \in \mathcal{L}}$ and $X_{\varepsilon,l} = X_\varepsilon(\phi_l)$ are the empirical Fourier coefficients. Then, the problems of interest can be restated, in an obvious way, in terms of the Fourier coefficients.

Anisotropic functional classes were studied in [4, 12, 13, 15] among others in connection with estimating functions of a multidimensional variable. Anisotropic constraints provide for a possible disparity of the inhomogeneous aspect in different directions. We will be interested in Sobolev functional classes described with the aid of Fourier coefficients. In this paper, assuming that f belongs to a Sobolev ball of varying radius, we construct asymptotically minimax estimators and provide optimal rates of convergence and exact asymptotic constants in the estimation problem (see Theorem 3.1 and Remark 3.1), cf. Theorem 1 of [15]. Also, for fixed and finite d we construct a family of asymptotically minimax tests and establish sharp asymptotics of the minimax total error probability in the detection problem (Theorem 3.2).

The estimation and detection problems for infinite-dimensional model (1.1) with $d = \infty$ or $d = d_\varepsilon \to \infty$ were studied in [5, 7–10]. Compared to the d-dimensional case with d being fixed and finite, the problems for infinite-dimensional Gaussian white noise have a much richer analytical content. This motivates the study of the case $d = \infty$. Another argument in favour of studying the infinite-dimensional model is that in reality the true dimensionality may not be known or may vary. Then, the results for infinite d, in conjunction with those for finite d, would form a clearer picture of the state of nature.

In comparison with the case when d is fixed and finite, the problems of estimating and detecting an infinite-dimensional signal are more challenging from a mathematical point of view. So far, we have established logarithmic asymptotics in these two problems (Theorem 3.3). When studying the case of fixed and finite d we use some general results of the minimax theory given by Theorems 2.1 and 2.2. For $d=\infty$ the analysis is completely different and is done by using probabilistic methods for the study of the so-called count function.

The paper is organized as follows. First we introduce anisotropic Sobolev classes with d being fixed and finite and then extend the definition to the infinite-dimensional case (Section 2.1). After that we formulate the problems of interest (Section 2.2), and introduce some general results of the minimax theory used in the subsequent sections (Section 2.3). The main results are collected in Section 3. The last section of the paper, Section 4, is rather diverse and contains the tools of study, auxiliary results, and proof of theorems.

The majority of limits in the paper are taken as $\varepsilon \to 0$. The relation $a_\varepsilon \sim b_\varepsilon$ means $\lim_{\varepsilon\to 0} a_\varepsilon/b_\varepsilon = 1$. The relation $a_\varepsilon \asymp b_\varepsilon$ means that there exist constants $0 < c < C < \infty$ and a number $\varepsilon_0 > 0$ such that $c < a_\varepsilon/b_\varepsilon < C$ for $\varepsilon \in (0, \varepsilon_0)$. Also, if $\lim_{\varepsilon\to 0} a_\varepsilon/b_\varepsilon = \infty$ (or $\lim_{\varepsilon\to 0} a_\varepsilon/b_\varepsilon = 0$) we write $a_\varepsilon \gg b_\varepsilon$ (or $a_\varepsilon \ll b_\varepsilon$).

2. Statement of the problem

2.1. Definition of anisotropic Sobolev balls

Assume that d is fixed and finite and consider functional classes indexed by a smoothness parameter $\sigma = (\sigma_1, \dots, \sigma_d)$, $\sigma_j > 0$, $j = 1, \dots, d$, that are defined by seminorms. Such classes are introduced as follows.

First, assume that σ_j is a positive integer and that f is σ_j-smooth in the jth argument, $j = 1, \dots, d$. For such a function f, define the seminorm $\|f\|_{\sigma,2}$ by

$$\|f\|_{\sigma,2}^2 = \sum_{j=1}^{d} \left\| \frac{\partial^{\sigma_j} f}{\partial x_j^{\sigma_j}} \right\|_2^2, \tag{2.1}$$

where $\partial^{\sigma_j} f/\partial x_j^{\sigma_j}$ is a (generalized) derivative of order σ_j in the jth direction (see, for example, [14, Sec. 4.1]), and denote by $\mathcal{F}_{\sigma,d}$ the anisotropic Sobolev ball, i.e.,

$$\mathcal{F}_{\sigma,d} = \{f \in L_2^d : \|f\|_{\sigma,2} \le 1\}.$$

In a general case of $\sigma = (\sigma_1, \dots, \sigma_d)$ with $\sigma_j > 0$, we assume that all partial derivatives $\frac{\partial^{m_j} f}{\partial x_j^{m_j}}$ of order $0 \le m_j \le [\sigma_j]$, $j = 1, \dots, d$, are 1-periodic, that is, for all $k = 1, \dots, d$,

$$\frac{\partial^{m_j} f}{\partial x_j^{m_j}}(x_1, \dots, x_{k-1}, 0, x_{k+1}, \dots, x_d) = \frac{\partial^{m_j} f}{\partial x_j^{m_j}}(x_1, \dots, x_{k-1}, 1, x_{k+1}, \dots, x_d),$$

and extend (2.1) by means of Fourier expansion. Specifically, let $\{\phi_k(x)\}_{k\in\mathbb{Z}}$ be the standard Fourier basis in $L_2[0,1]$ and let $\{\phi_l(\mathbf{x})\}_{l\in\mathbb{Z}^d}$ be a tensor product basis in L_2^d, i.e., for $\mathbf{x}=(x_1,\dots,x_d)\in[0,1]^d$ and $l=(l_1,\dots,l_d)\in\mathbb{Z}^d$

$$\phi_l(\mathbf{x})=\prod_{k=1}^d \varphi_{l_k}(x_k),$$

where

$$\varphi_0(x)=1,\quad \varphi_l(x)=\sqrt{2}\cos(2\pi lx),\quad \varphi_{-l}(x)=\sqrt{2}\sin(2\pi lx),\quad l>0.$$

For a function $f(\mathbf{x})=\sum_{l\in\mathbb{Z}^d}\theta_l\phi_l(\mathbf{x})$ with the Fourier coefficients $\theta_l=(f,\phi_l)$ we set

$$\|f\|^2_{\sigma,2}=\sum_{l\in\mathbb{Z}^d} c_l^2\theta_l^2,$$

where for $l=(l_1,\dots,l_d)\in\mathbb{Z}^d$ and $\sigma=(\sigma_1,\dots,\sigma_d)\in\mathbb{R}^d_+$

$$c_l^2=c_l^2(\sigma)=\sum_{j=1}^d(2\pi|l_j|)^{2\sigma_j}. \tag{2.2}$$

When the σ_j's are positive integers this corresponds to (2.1) under the periodic constraints.

We now move on to the case $d=\infty$ and remind the definition of the space $L_2^\infty=L_2([0,1]^\infty)$ of square integrable functions of infinitely many variables (see [8]).

The set $[0,1]^\infty=([0,1]^\infty,\mathcal{B}([0,1]^\infty),\lambda^\infty)$ is viewed as a probability product space with σ-algebra $\mathcal{B}([0,1]^\infty)$ generated by the cylindric sets $\cap_{j\le d}\{x_j\in B_j\}$, $d=1,2,\dots$, where $B_j\subset[0,1]$ are Borel sets, and the product Lebesgue measure λ^∞. Let

$$L_2^\infty=\{f:[0,1]^\infty\to\mathbb{R}:\int_{[0,1]^\infty} f^2(\mathbf{x})\lambda^\infty(d\mathbf{x})<\infty\}.$$

The space L_2^∞ is a Hilbert space with a standard scalar product. Its basis is easy to specify by using the following argument. For each $d\in\mathbb{N}$ the standard projection $P_d:[0,1]^\infty\to[0,1]^d$ generates the standard embedding E_d of the set $\mathcal{F}^d$ of functions defined on $[0,1]^d$ to the set $\mathcal{F}^\infty$ of functions defined on $[0,1]^\infty$. The space L_2^∞ is then the closure of $\cup_{d\in\mathbb{N}}E_dL_2^d$ under the standard embedding

$$E_1L_2^1\subset\cdots\subset E_dL_2^d\subset E_{d+1}L_2^{d+1}\subset\cdots\subset\mathcal{F}^\infty.$$

(Clearly, if $f=\lim_{d\to\infty}f_d$ and $g=\lim_{d\to\infty}g_d$, where $f_d,g_d\in L_2^d$, then

$$(f,g)=\lim_{d\to\infty}(f_d,g_d),$$

so that L_2^∞ is a separable Hilbert space.) Next, define the set $\mathbb{Z}_0^\infty$ that consists of infinite sequences (l_j) with finitely many nonzero terms:

$$\mathbb{Z}_0^\infty=\bigcup_{d=1}^\infty \mathbb{Z}_0^d,\quad \mathbb{Z}_0^d=\{l=(l_1,\dots,l_d,0,\dots,0,\dots)\in\mathbb{Z}^\infty\}.$$

Then, by the fact $L_2^\infty = \overline{\cup_{d\in\mathbb{N}} E_d L_2^d}$ we have $L_2^\infty = \overline{\mathrm{Lin}\left(\{\phi_l(\mathbf{x})\}_{l\in\mathbb{Z}_0^\infty}\right)}$, where for $\mathbf{x} = (x_1, x_2, \ldots) \in [0,1]^\infty$ and $l = (l_1, l_2, \ldots) \in \mathbb{Z}_0^\infty$

$$\phi_l(\mathbf{x}) = \prod_{k=1}^{\infty} \varphi_{l_k}(x_k),$$

and $\mathrm{Lin}\left(\{\phi_l(\mathbf{x})\}_{l\in\mathbb{Z}_0^\infty}\right)$ is the space spanned by the functions $\phi_l(\mathbf{x})$, $l \in \mathbb{Z}_0^\infty$. Thus the orthonormal system $\{\phi_l(\mathbf{x})\}_{l\in\mathbb{Z}_0^\infty}$ form the basis of L_2^∞.

Now we consider smoothness constraints that are applicable to functions of infinitely many variables. First, let $\sigma = (\sigma_1, \sigma_2, \ldots)$ be an infinite sequence of positive integers. Define the semi-norm $\|f\|_{\sigma,2}$ by

$$\|f\|_{\sigma,2}^2 = \sum_{j=1}^{\infty} \left\| \frac{\partial^{\sigma_j} f}{\partial x_j^{\sigma_j}} \right\|_2^2. \tag{2.3}$$

Assume, as before, that f together with all its partial derivatives is 1-periodic, i.e., for all partial derivatives $\frac{\partial^{m_j} f}{\partial x_j^{m_j}}$ of order $0 \le m_j \le \sigma_j$, $j = 1, 2, \ldots$, and for all $k = 1, 2, \ldots$,

$$\frac{\partial^{m_j} f}{\partial x_j^{m_j}}(x_1, \ldots, x_{k-1}, 0, x_{k+1}, \ldots) = \frac{\partial^{m_j} f}{\partial x_j^{m_j}}(x_1, \ldots, x_{k-1}, 1, x_{k+1}, \ldots). \tag{2.4}$$

Let $\theta_l = (f, \phi_l)$ be the Fourier coefficients of f with respect to the Fourier basis $\{\phi_l(\mathbf{x})\}_{l\in\mathbb{Z}_0^\infty}$. Then in terms of the θ_j's,

$$\|f\|_{\sigma,2}^2 = \sum_{l\in\mathbb{Z}_0^\infty} c_l^2 \theta_l^2, \quad c_l^2 = c_l^2(\sigma) = \sum_{j=1}^{\infty} (2\pi|l_j|)^{2\sigma_j}.$$

Thus, in a general case of $\sigma = (\sigma_1, \sigma_2, \ldots)$ with $\sigma_j > 0$, $j = 1, 2, \ldots$, we assume that all partial derivatives $\frac{\partial^{m_j} f}{\partial x_j^{m_j}}$ of order $0 \le m_j \le [\sigma_j]$, $j = 1, 2 \ldots$, are 1-periodic, and set

$$\|f\|_{\sigma,2}^2 = \sum_{l\in\mathbb{Z}_0^\infty} c_l^2 \theta_l^2, \quad c_l^2 = c_l^2(\sigma) = \sum_{j=1}^{\infty} (2\pi|l_j|)^{2\sigma_j}.$$

Under a periodic constraint, the anisotropic Sobolev ball for $d = \infty$ is given by

$$\mathcal{F}_{\sigma,\infty} = \{f \in L_2^\infty : \|f\|_{\sigma,2} \le 1\}.$$

2.2. Estimation and detection over anisotropic Sobolev balls

When dealing with the estimation problem, we follow a familiar pattern. If for an estimator $\hat{f}_\varepsilon$ of f based on the observation X_ε, and a sequence $\delta_\varepsilon \to \infty$,

$$\delta_\varepsilon \times (\text{maximal risk of } \hat{f}_\varepsilon) \le C < \infty \quad \text{for } \varepsilon \text{ sufficiently small},$$

and at the same time for any estimator $\tilde{f}_\varepsilon$ of f based on X_ε,

$$\delta_\varepsilon \times (\text{maximal risk of } \tilde{f}_\varepsilon) \geq c > 0 \quad \text{for } \varepsilon \text{ sufficiently small}$$

then the estimator $\hat{f}_\varepsilon$ is said to be *rate optimal*. The parameter δ_ε controls the best possible rate of convergence. A more delicate problem, called the *sharp optimality* problem, consists of finding the rate optimal estimator whose minimax risk is the smallest possible. That is, if one can find a rate optimal estimator $\hat{f}_\varepsilon$ such that the constants C and c obey the same asymptotics:

$$C = C_\varepsilon = A(1 + o(1)) \quad c = c_\varepsilon = A(1 + o(1)),$$

the estimator $\hat{f}_\varepsilon$ is called *asymptotically minimax*, and A is called an *exact asymptotic constant*.

To be precise, for $1 \leq d \leq \infty$ define the minimax integrated squared risk by

$$R_\varepsilon^2(\mathcal{F}_{\sigma,d}) = \inf_{\tilde{f}_\varepsilon} \sup_{f \in \mathcal{F}_{\sigma,d}} \mathbf{E}_{\varepsilon,f} \|f - \tilde{f}_\varepsilon\|_2^2,$$

where the infimum is taken over all possible estimators $\tilde{f}_\varepsilon$ of f based on the observation X_ε. In this paper, we wish to find the *asymptotically minimax* estimator $\hat{f}_\varepsilon$ of f for which

$$\sup_{f \in \mathcal{F}_{\sigma,d}} \mathbf{E}_{\varepsilon,f} \|f - \hat{f}_\varepsilon\|_2^2 \sim R_\varepsilon^2(\mathcal{F}_{\sigma,d}), \quad \varepsilon \to 0,$$

and establish sharp asymptotics, which includes convergence rates and exact asymptotic constants, for the risk $R_\varepsilon^2(\mathcal{F}_{\sigma,d})$.

Now, we turn to the detection problem. For a meaningful minimax testing problem, the alternative hypothesis must have some neighborhood of the null hypothesis removed. Therefore, for $r_\varepsilon > 0$ and $1 \leq d \leq \infty$, we put

$$\mathcal{F}_{\sigma,d}(r_\varepsilon) = \{f \in \mathcal{F}_{\sigma,d} : \|f\|_2 \geq r_\varepsilon\},$$

and consider testing the hypotheses

$$H_0 \colon f = 0 \quad \text{vs.} \quad H_{1\varepsilon} \colon f \in \mathcal{F}_{\sigma,d}(r_\varepsilon).$$

When dealing with the detection problem, we judge the quality of testing by using the minimax criterion based on the total error probability. For a test ψ_ε based on the observation X_ε, define the error probabilities

$$\begin{aligned}
\alpha_\varepsilon(\psi_\varepsilon) &= \mathbf{E}_{\varepsilon,0}\psi_\varepsilon, \\
\beta_\varepsilon(\psi_\varepsilon, f) &= \mathbf{E}_{\varepsilon,f}(1 - \psi_\varepsilon), \\
\gamma_\varepsilon(\psi_\varepsilon, f) &= \alpha_\varepsilon(\psi_\varepsilon) + \beta_\varepsilon(\psi_\varepsilon, f).
\end{aligned}$$

The maximum probability of type II error is then given by

$$\beta_\varepsilon(\psi_\varepsilon, \mathcal{F}_{\sigma,d}(r_\varepsilon)) = \sup_{f \in \mathcal{F}_{\sigma,d}(r_\varepsilon)} \beta_\varepsilon(\psi_\varepsilon, f).$$

The quantity

$$\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = \inf_{\psi_\varepsilon} \gamma_\varepsilon(\psi_\varepsilon, \mathcal{F}_{\sigma,d}(r_\varepsilon)),$$

where

$$\gamma_\varepsilon(\psi_\varepsilon, \mathcal{F}_{\sigma,d}(r_\varepsilon)) = \alpha_\varepsilon(\psi_\varepsilon) + \beta_\varepsilon(\psi_\varepsilon, \mathcal{F}_{\sigma,d}(r_\varepsilon)),$$

and the infimum is taken over all tests ψ_ε based on X_ε, is called the *minimax total error probability*. A family of tests ψ_ε^* is called *asymptotically minimax* if

$$\gamma_\varepsilon(\psi_\varepsilon^*, \mathcal{F}_{\sigma,d}(r_\varepsilon)) = \gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) + o(1), \quad \varepsilon \to 0.$$

We are interested in finding asymptotics of $\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon))$ and determining the structure of asymptotically minimax tests. In the context of signal detection problem, this is called the sharp optimality problem.

It is always true that $0 \le \gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \le 1$. If the parameter r_ε in the alternative hypothesis is too close to zero then $\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 1$ as $\varepsilon \to 0$, and one cannot distinguish between the null hypothesis and the alternative. Therefore the knowledge of the smallest r_ε for which $\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 0$ as $\varepsilon \to 0$ is important. If there exists a family $r_\varepsilon^* = r_\varepsilon^*(\mathcal{F}_{\sigma,d}) \to 0$ such that

$$\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 1 \quad \text{if} \quad r_\varepsilon / r_\varepsilon^* \to 0 \quad \text{and} \quad \gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 0 \quad \text{if} \quad r_\varepsilon / r_\varepsilon^* \to \infty,$$

then the family r_ε^* is called the *separation rate*. Thus, another problem of interest to us is to find asymptotics for the separation rate r_ε^*.

From a technical point of view, it is more convenient to deal with ellipsoids in sequence spaces rather than Sobolev balls in functional spaces. In the sequence space of Fourier coefficients, the ball $\mathcal{F}_{\sigma,d}$ with fixed and finite d corresponds to the ellipsoid

$$\Theta_{\sigma,d} = \{\theta = (\theta_l)_{l \in \mathbb{Z}^d} : \sum_{l \in \mathbb{Z}^d} c_l^2 \theta_l^2 \le 1\}, \tag{2.5}$$

and for $d = \infty$ to the ellipsoid

$$\Theta_{\sigma,\infty} = \{\theta = (\theta_l)_{l \in \mathbb{Z}_0^\infty} : \sum_{l \in \mathbb{Z}_0^\infty} c_l^2 \theta_l^2 \le 1\}.$$

The estimation problem then transforms to constructing asymptotically minimax estimator $\hat{\theta}_\varepsilon$ of θ using the data $X_{\varepsilon,l}$ in model (1.2), and establishing exact asymptotics for the minimax squared risk associated to the ellipsoid $\Theta_{\sigma,d}$, $1 \le d \le \infty$:

$$R_\varepsilon^2(\Theta_{\sigma,d}) = \inf_{\tilde{\theta}_\varepsilon} \sup_{\theta \in \Theta_{\sigma,d}} \|\theta - \tilde{\theta}_\varepsilon\|^2 \sim \sup_{\theta \in \Theta_{\sigma,d}} \|\theta - \hat{\theta}_\varepsilon\|^2.$$

It is well known that $R_\varepsilon(\Theta_{\sigma,d}) \asymp \varepsilon^{2/(2+\sigma^{-1})}$ (see [4, Sec. 16.3] and [15, Sec. 3]). Moreover, the *sharp* asymptotic relation for $R_\varepsilon(\Theta_{\sigma,d})$ exists (see Theorem 1 of [15]). In Sections 3 and 4, we shall state and prove a similar result for the ball $\mathcal{F}_{\sigma,d}(M_\varepsilon) = \{f \in L_2^d : \|f\|_{\sigma,2} \le M_\varepsilon\}$ of a varying radius $M_\varepsilon > 0$ such that

$\varepsilon/M_\varepsilon \to 0$, as $\varepsilon \to 0$ (see Remark 3.1), and provide the asymptotically minimax estimator of f. We do that to illustrate our approach, which is somewhat different (and shorter) compared to the one in [15], and goes in parallel with deriving sharp asymptotics in the detection problem.

In the detection problem the set $\mathcal{F}_{\sigma,d}(r_\varepsilon)$, $1 \le d < \infty$, that specifies the alternative hypothesis corresponds to the ellipsoid with a small ball removed:

$$\Theta_{\sigma,d}(r_\varepsilon) = \{\theta = (\theta_l)_{l\in\mathbb{Z}^d} : \sum_{l\in\mathbb{Z}^d} \theta_l^2 c_l^2(\sigma) \le 1 \text{ and } \sum_{l\in\mathbb{Z}^d} \theta_l^2 \ge r_\varepsilon^2\}. \tag{2.6}$$

If $M = M_\varepsilon$ is a positive constant such that $r_\varepsilon/M \to 0$ as $\varepsilon \to 0$, then the results obtained for $\Theta_{\sigma,d}(r_\varepsilon)$ are immediately extended to the set $\Theta_{\sigma,d}(r_\varepsilon, M)$ defined similarly to (2.6) with $\sum_{l\in\mathbb{Z}^d} \theta_l^2 c_l^2(\sigma) \le M^2$ in place of $\sum_{l\in\mathbb{Z}^d} \theta_l^2 c_l^2(\sigma) \le 1$ (see Remark 3.1).

When $d = \infty$ the set

$$\mathcal{F}_{\sigma,\infty}(r_\varepsilon) = \{f \in \mathcal{F}_{\sigma,\infty} : \|f\|_2 \ge r_\varepsilon\},$$

where $\mathcal{F}_{\sigma,\infty} = \{f \in L_2^\infty : \|f\|_{\sigma,2}^2 \le 1\}$, takes the form

$$\Theta_{\sigma,\infty}(r_\varepsilon) = \{\theta = (\theta_l)_{l\in\mathbb{Z}_0^\infty} : \sum_{l\in\mathbb{Z}_0^\infty} \theta_l^2 c_l^2(\sigma) \le 1 \text{ and } \sum_{l\in\mathbb{Z}_0^\infty} \theta_l^2 \ge r_\varepsilon^2\}.$$

In both cases, the hypotheses to be tested become $H_0 \colon \theta = 0$ versus $H_{1\varepsilon} \colon \theta \in \Theta_{\sigma,d}(r_\varepsilon)$.

2.3. Some general results

When estimating and detecting an infinite-dimensional vector from an ellipsoid in a sequence space, the sharp asymptotics of the minimax squared risk and minimax error probabilities are obtained by solving alike extremal problems. Solutions to these problems, in implicit form, nowadays constitute standard results of the minimax theory. The first of these results connected to the estimation problem is largely due to Pinsker [16] (see also [1, Ch. 7], [7, Sec. 2.2], and [18, Sec. 3.1]); the second one, dealing with detection of a signal and obtained for the first time in a slightly different setup in Theorem 1 of [2], is a combination of statements from [6] (see [7, Sec. 2] for details).

In what follows, we use notation $(x)_+ = \max(x, 0)$, $x \in \mathbb{R}$.

Theorem 2.1. *Let $E_\varepsilon^2(\sigma, d)$ be the value of the extremal problem on the set of real-valued bilateral sequences $\{v_l : l \in \mathbb{Z}^d\}$:*

$$E_\varepsilon^2(\sigma, d) = \varepsilon^2 \sup \sum_{l\in\mathbb{Z}^d} \frac{v_l^2}{v_l^2 + 1} \quad \textit{subject to} \quad \sum_{l\in\mathbb{Z}^d} c_l^2 v_l^2 \le \varepsilon^{-2}, \tag{2.7}$$

where $c_l^2 = c_l^2(\sigma)$ are given by (2.2). Then

$$R_\varepsilon(\mathcal{F}_{\sigma,d}) \le E_\varepsilon(\sigma, d). \tag{2.8}$$

The extremal sequence $\{\hat{v}_l^2\}_{l\in\mathbb{Z}^d}$ *in* (2.7) *is of the form*

$$\hat{v}_l^2 = (T/c_l - 1)_+$$

(we formally set $\hat{v}_l^2 = \infty$ *if* $c_l = 0$*), where the quantity* $T = T_\varepsilon > 0$ *satisfies*

$$\sum_{c_l<T} c_l^2 \hat{v}_l^2 = T^2 \sum_{c_l<T} \left(c_l/T - (c_l/T)^2\right) = \varepsilon^{-2}, \tag{2.9}$$

and the value of the problem is

$$E_\varepsilon^2(\sigma,d) = \varepsilon^2 \sum_{c_l<T} (1 - c_l/T). \tag{2.10}$$

Suppose that $T_\varepsilon \to \infty$*. Then* $R_\varepsilon(\mathcal{F}_{\sigma,d}) \geq E_\varepsilon(\sigma,d)(1+o(1))$*. Jointly with* (2.8) *this yields*

$$R_\varepsilon(\mathcal{F}_{\sigma,d}) \sim E_\varepsilon(\sigma,d). \tag{2.11}$$

The asymptotically minimax estimator is a weighted projection-type estimator

$$\hat{f}_\varepsilon(\mathbf{x}) = \sum_{c_l<T} c_{\varepsilon,l} X_\varepsilon(\phi_l)\phi_l(\mathbf{x}), \quad c_{\varepsilon,l} = 1 - c_l/T, \quad \mathbf{x} \in [0,1]^d.$$

Theorem 2.2. *Let* $u_\varepsilon^2(\mathcal{F}_{\sigma,d}(r_\varepsilon))$ *be the value of the extremal problem*

$$u_\varepsilon^2(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = \inf \frac{1}{2} \sum_{l\in\mathbb{Z}^d} v_l^4 \quad \text{subject to} \quad \sum_{l\in\mathbb{Z}^d} v_l^2 \geq (r_\varepsilon/\varepsilon)^2, \quad \sum_{l\in\mathbb{Z}^d} c_l^2 v_l^2 \leq \varepsilon^{-2}, \tag{2.12}$$

where $c_l^2 = c_l^2(\sigma)$ *are given by* (2.2)*. Then*

$$\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 1, \quad \text{as} \quad u_\varepsilon^2(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 0.$$

Moreover, if $r_\varepsilon \to 0$*, then*

$$\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = 2\Phi(-u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon))/2) + o(1). \tag{2.13}$$

The extremal sequence is of the form

$$\hat{v}_l^2 = u_0^2 \left(1 - (T/c_l)^2\right)_+,$$

where the quantities $u_0 = u_{0,\varepsilon} > 0$ *and* $T = T_\varepsilon > 0$ *are determined by the equations*

$$\sum_{l\in\mathbb{Z}^d} \hat{v}_l^2 = u_0^2 \sum_{c_l<T} \left(1 - (c_l/T)^2\right) = (r_\varepsilon/\varepsilon)^2, \tag{2.14}$$

$$\sum_{l\in\mathbb{Z}^d} c_l^2 \hat{v}_l^2 = u_0^2 \sum_{c_l<T} c_l^2 \left(1 - (c_l/T)^2\right) = (1/\varepsilon)^2, \tag{2.15}$$

$$T \geq r_\varepsilon^{-1} \to \infty, \tag{2.16}$$

and the value of the problem is

$$u_\varepsilon^2(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = \frac{1}{2}\sum_{l\in\mathbb{Z}^d} \hat{v}_l^4 = \frac{1}{2}u_0^4 \sum_{c_l<T} \left(1-(c_l/T)^2\right)^2. \tag{2.17}$$

The asymptotically minimax test is $\psi_\varepsilon = \mathbb{I}(t_\varepsilon > u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon))/2)$ *and is based on the* χ^2*-type test statistic*

$$t_\varepsilon = w_\varepsilon^{-1}\sum_{l\in\mathbb{Z}^d} w_{\varepsilon,l}(\tilde{X}_{\varepsilon,l}^2 - 1), \quad w_{\varepsilon,l} = \left(1-(c_l/T)^2\right)_+, \quad w_\varepsilon^2 = \frac{1}{2}\sum_{l\in\mathbb{Z}^d} w_{\varepsilon,l}^2,$$

where $\tilde{X}_{\varepsilon,l} = \varepsilon^{-1}X_\varepsilon(\phi_l)$.

The definition of separation rate $r_\varepsilon^* = r_\varepsilon^*(\mathcal{F}_{\sigma,d})$ implies that r_ε^* is determined by the relation $u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon^*)) \asymp 1$. Thus, the family $u_\varepsilon = u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon))$ characterizes the distinguishability in the problem.

3. Main results

3.1. Sharp asymptotics for fixed d

Based on Theorems 2.1 and 2.2 we now establish two results that solve the sharp optimality problems in connection with estimating and detecting a multivariate signal f. We keep the notation of Theorems 2.1 and 2.2, and put

$$\sigma^{-1} = \sigma^{-1}(d) = \sum_{j=1}^d \sigma_j^{-1}.$$

Theorem 3.1. *Assume that the dimension* $d < \infty$ *and the smoothness parameter* $\sigma = (\sigma_1, \dots, \sigma_d) \in \mathbb{R}_+^d$*, are fixed. Then as* $\varepsilon \to 0$

$$E_\varepsilon(\sigma, d) \sim c(\sigma,d)\varepsilon^{2/(2+\sigma^{-1})}, \quad T^2 \sim E_\varepsilon^{-2}(\sigma,d)(\sigma^{-1}+2)/\sigma^{-1},$$

where the exact asymptotic constant $c(\sigma, d)$ *is given by the formula*

$$c^2(\sigma,d) = \left(\left(1+2/\sigma^{-1}\right)^{\sigma^{-1}/2} \frac{2\prod_{j=1}^d \Gamma(1+1/(2\sigma_j))}{\pi^d\sigma^{-1}(\sigma^{-1}+1)\Gamma(\sigma^{-1}/2)}\right)^{2/(2+\sigma^{-1})}. \tag{3.1}$$

Theorem 3.2. *Assume that the dimension* $d < \infty$ *and the smoothness parameter* $\sigma = (\sigma_1, \dots, \sigma_d) \in \mathbb{R}_+^d$*, are fixed. Then as* $\varepsilon \to 0$

$$u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \sim C(\sigma,d) r_\varepsilon^{2+\sigma^{-1}/2}\varepsilon^{-2}, \quad T^2 \sim r_\varepsilon^{-2}(\sigma^{-1}+4)/\sigma^{-1},$$

where the exact asymptotic constant $C(\sigma, d)$ *is given by the formula*

$$C^2(\sigma,d) = \frac{\pi^d(\sigma^{-1})^{\sigma^{-1}/2}(\sigma^{-1}+2)\Gamma(1+\sigma^{-1}/2)}{(\sigma^{-1}+4)^{1+\sigma^{-1}/2}\prod_{j=1}^d \Gamma(1+1/(2\sigma_j))}. \tag{3.2}$$

Remark 3.1. By using rescaling arguments, it is straightforward to extend the results of Theorems 3.1 and 3.2 to the case of anisotropic Sobolev ball

$$\mathcal{F}_{\sigma,d}(M) = \{f \in L_2^d : \|f\|_{\sigma,2} \le M\}$$

of a radius $M = M_\varepsilon > 0$ such that $\varepsilon/M \to 0$ as $\varepsilon \to 0$ (plus the assumption that $r_\varepsilon/M \to 0$ as $\varepsilon \to 0$ in the detection problem). For such a ball, Theorems 3.1 and 3.2 remain valid with the constant $c_M(\sigma, d)$ in place of $c(\sigma, d)$ and the constant $C_M(\sigma, d)$ in place of $C(\sigma, d)$, where

$$c_M(\sigma, d) = M^{\sigma^{-1}/(2+\sigma^{-1})} c(\sigma, d), \quad C_M(\sigma, d) = M^{-\sigma^{-1}/2} C(\sigma, d).$$

Indeed, setting $\tilde{c}_l = c_l/M$ transforms the ellipsoid $\Theta_{\sigma,d}(M) = \{\theta = (\theta_l)_{l\in\mathbb{Z}^d} : \sum_{l\in\mathbb{Z}^d} c_l^2\theta_l^2 \le M^2\}$ into the ellipsoid $\tilde{\Theta}_{\sigma,d} = \{\theta = (\theta_l)_{l\in\mathbb{Z}^d} : \sum_{l\in\mathbb{Z}^d} \tilde{c}_l^2\theta_l^2 \le 1\}$. If now $E_\varepsilon^2(\sigma, d, c)$ stands for the value of extremal problem (2.7) and $\tilde{\varepsilon} = \varepsilon/M \to 0$, then by Theorem 3.1, as $\varepsilon \to 0$,

$$E_\varepsilon(\sigma, d, \tilde{c}) = M E_{\tilde{\varepsilon}}(\sigma, d, c) \sim c_M(\sigma, d)\varepsilon^{2/(2+\sigma^{-1})}.$$

Similarly, if $u_{\varepsilon,c}^2(\mathcal{F}_{\sigma,d}(r_\varepsilon))$ stands for the value of extremal problem (2.12) and $\tilde{r}_\varepsilon = r_\varepsilon/M \to 0$, then by Theorem 3.2, as $\varepsilon \to 0$

$$u_{\varepsilon,\tilde{c}}(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = u_{\tilde{\varepsilon},c}(\mathcal{F}_{\sigma,d}(\tilde{r}_\varepsilon)) \sim C_M(\sigma, d) r_\varepsilon^{2+\sigma^{-1}/2}\varepsilon^{-2}.$$

Remark 3.2. Theorem 3.1 together with Remark 3.1 extends Theorem 1 of [15] to the case of Sobolev ball $\mathcal{F}_{\sigma,d}(M_\varepsilon)$ of varying radius $M_\varepsilon \gg \varepsilon$ (with the assumption that $M \gg r_\varepsilon$ in the detection problem). In addition, Theorems 3.1 and 3.2 extend Theorems 3 and 4 of [7]. Indeed, when $\sigma_1 = \ldots = \sigma_d = \sigma > 0$, our results coincide with those of Theorems 3 and 4 of [7] for the norm $\|f\|_{\sigma,2}^2 = \sum_{l\in\mathbb{Z}^d} c_l^2\theta_l^2$, where $c_l^2 = \sum_{j=1}^d (2\pi|l_j|)^{2\sigma}$. In this special (isotropic) case, the constants $c(\sigma, d)$ and $C(\sigma, d)$ are monotone in d; decreasing and increasing, respectively. In a general (anisotropic) case, when $d = d_\varepsilon \to \infty$ or $d = \infty$, under the assumption $\sum_{j=1}^\infty \sigma_j^{-1} < \infty$, formulas (3.1) and (3.2) yield

$$c(\sigma, d) \asymp \pi^{-d/(2+\sigma^{-1})}, \quad C(\sigma, d) \asymp \pi^{d/2}.$$

Generally speaking, these asymptoics are not usable because Theorems 3.1 and 3.2 are proved for fixed d. For example, in the isotropic case, the limiting behaviour of $R_\varepsilon(\mathcal{F}_{\sigma,d})$ and $r_\varepsilon^*(\mathcal{F}_{\sigma,d})$ for $d = d_\varepsilon \gg \log(\varepsilon^{-1})$ is completely different compared to the case $d = d_\varepsilon = o(\log(\varepsilon^{-1}))$ (see [7, Sec. 2.4]).

Remark 3.3. Theorems 3.1 and 3.2 imply as $\varepsilon \to 0$

$$R_\varepsilon(\mathcal{F}_{\sigma,d}) \asymp \varepsilon^{2/(2+\sigma^{-1})}, \quad r_\varepsilon^*(\mathcal{F}_{\sigma,d}) \asymp \varepsilon^{4/(4+\sigma^{-1})}. \tag{3.3}$$

Indeed, the separation rate $r_\varepsilon^* = r_\varepsilon^*(\mathcal{F}_{\sigma,d})$ must satisfy $u_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon^*)) \asymp 1$, which, in view of Theorem 3.2, leads to the required asymptotics for $r_\varepsilon^*(\mathcal{F}_{\sigma,d})$. By (2.13) this asymptotics yields $\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon)) \to 0$ for $r_\varepsilon \gg r_\varepsilon^* \asymp \varepsilon^{4/(4+\sigma^{-1})}$. Next, using (2.11) and Theorem 3.1, we have $R_\varepsilon(\mathcal{F}_{\sigma,d}) \sim c(\sigma, d)\varepsilon^{2/(2+\sigma^{-1})}$, and the rate asymptotics for $R_\varepsilon(\mathcal{F}_{\sigma,d})$ follows immediately.

3.2. Logarithmic asymptotics for $d = \infty$

The study of infinite-dimensional estimation and detection problems is done under the assumption

$$\sigma^{-1} = \sum_{j=1}^{\infty} \sigma_j^{-1} < \infty. \tag{3.4}$$

We have the following theorem.

Theorem 3.3. *Assume that (3.4) is satisfied. Then as* $\varepsilon \to 0$

$$\log R_\varepsilon(\mathcal{F}_{\sigma,\infty}) \sim \frac{2\log\varepsilon}{2+\sigma^{-1}}, \quad \log r_\varepsilon^*(\mathcal{F}_{\sigma,\infty}) \sim \frac{4\log\varepsilon}{4+\sigma^{-1}}.$$

Remark 3.4. Observe that the log-asymptotics for $d = \infty$ in Theorem 3.3 are similar to those for $1 \le d < \infty$ (see (3.3)).

4. Proofs of Theorems

4.1. Proofs of Theorems 3.1 and 3.2

The proofs of Theorems 3.1 and 3.2 follow the pattern of [7, Sec. 3]. Let us start with Theorem 3.2.

Proof of Theorem 3.2. We need to study equations (2.14), (2.15), and (2.17) under condition (2.16). Denote

$$I_1 = I_1(\sigma, d) = \sum_{c_l<T} \left(1 - (c_l/T)^2\right), \tag{4.1}$$

$$I_2 = I_2(\sigma, d) = \sum_{c_l<T} (c_l/T)^2 \left(1 - (c_l/T)^2\right), \tag{4.2}$$

$$I_0 = I_0(\sigma, d) = \sum_{c_l<T} \left(1 - (c_l/T)^2\right)^2 = I_1 - I_2. \tag{4.3}$$

It follows from (2.14), (2.15), and (2.17) that

$$T^2 = r_\varepsilon^{-2} I_1/I_2, \tag{4.4}$$

$$u_\varepsilon^2(\mathcal{F}_{\sigma,d}(r_\varepsilon)) = \frac{1}{2}(r_\varepsilon/\varepsilon)^4 I_0/I_1^2. \tag{4.5}$$

Let us study the asymptotic behaviour of I_k, $k = 0, 1, 2$, when $T \to \infty$. For this put

$$x_{l_j} = \frac{2\pi l_j}{T^{1/\sigma_j}} = \frac{l_j}{m_j}, \quad j = 1, \ldots, d, \tag{4.6}$$

where $m_j = \frac{T^{1/\sigma_j}}{2\pi}$. Then, recalling (2.2), as $T \to \infty$,

$$\begin{aligned} I_1 &= \sum_{c_l<T} \left(1 - \sum_{j=1}^d (2\pi|l_j|)^{2\sigma_j}/T^2\right) \\ &= m_1 \dots m_d \sum_{\sum |x_{l_j}|^{2\sigma_j}<1} \left(1 - \sum_{j=1}^d |x_{l_j}|^{2\sigma_j}\right) m_1^{-1} \dots m_d^{-1} \\ &\sim \frac{T^{\sigma^{-1}}}{(2\pi)^d} \int_{D_{\sigma,d}(1)} \left(1 - \sum_{j=1}^d |x_j|^{2\sigma_j}\right) dx_1 \dots dx_d, \end{aligned} \tag{4.7}$$

where

$$D_{\sigma,d}(1) = \{\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d : \sum_{j=1}^d |x_j|^{2\sigma_j} < 1\}.$$

Similarly, as $T \to \infty$,

$$I_2 \sim \frac{T^{\sigma^{-1}}}{(2\pi)^d} \int_{D_{\sigma,d}(1)} \sum_{j=1}^d |x_j|^{2\sigma_j} \left(1 - \sum_{j=1}^d |x_j|^{2\sigma_j}\right) dx_1 \dots dx_d, \tag{4.8}$$

$$I_0 \sim \frac{T^{\sigma^{-1}}}{(2\pi)^d} \int_{D_{\sigma,d}(1)} \left(\sum_{j=1}^d |x_j|^{2\sigma_j} \left(1 - \sum_{j=1}^d |x_j|^{2\sigma_j}\right)\right)^2 dx_1 \dots dx_d. \tag{4.9}$$

Next, making the change of variables in the integrals in (4.7)–(4.9):

$$y_j = x_j^{2\sigma_j}, \quad j = 1, \dots, d,$$

and denoting by Σ_d the d-dimensional simplex, i.e,

$$\Sigma_d = \{\mathbf{y} = (y_1, \dots, y_d) \in \mathbb{R}^d : y_j \geq 0, \sum_{j=1}^d y_j \leq 1\},$$

we get

$$I_1 \sim \frac{T^{\sigma^{-1}}}{(2\pi)^d \sigma_1 \dots \sigma_d} \int_{\Sigma_d} \left(1 - \sum_{j=1}^d y_j\right) y_1^{1/(2\sigma_1)-1} \dots y_d^{1/(2\sigma_d)-1} dy_1 \dots dy_d, \tag{4.10}$$

$$I_2 \sim \frac{T^{\sigma^{-1}}}{(2\pi)^d \sigma_1 \dots \sigma_d} \int_{\Sigma_d} \sum_{j=1}^d y_i \left(1 - \sum_{j=1}^d y_j\right) y_1^{1/(2\sigma_1)-1} \dots y_d^{1/(2\sigma_d)-1} dy_1 \dots dy_d, \tag{4.11}$$

$$I_0 \sim \frac{T^{\sigma^{-1}}}{(2\pi)^d \sigma_1 \dots \sigma_d} \int_{\Sigma_d} \left(1 - \sum_{j=1}^d y_j\right)^2 y_1^{1/(2\sigma_1)-1} \dots y_d^{1/(2\sigma_d)-1} dy_1 \dots dy_d. \tag{4.12}$$

The integrals on the right-hand sides of (4.10)–(4.12) can be calculated using the Liouville formula (see, for example, [3, Ch. XVIII])

$$\int_{\Sigma_d} \phi(x_1+\cdots+x_d)x_1^{p_1-1}\dots x_d^{p_d-1}dx_1\dots dx_d$$
$$= \frac{\Gamma(p_1)\dots\Gamma(p_d)}{\Gamma(p_1+\cdots+p_d)}\int_0^1 \phi(u)u^{p_1+\cdots p_d-1}du,$$

where $p_i>0$, $i=1,\dots d$, and the integral on the right-hand side is absolutely convergent. Applying the Liouville formula, we get as $T\to\infty$

$$\begin{aligned} I_1 &\sim \frac{T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d\sigma_1\dots\sigma_d\Gamma(\sigma^{-1}/2)}\int_0^1(1-u)u^{\sigma^{-1}/2-1}du \\ &= \frac{T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d\sigma_1\dots\sigma_d\Gamma(2+\sigma^{-1}/2)}; \\ I_2 &\sim \frac{T^{\sigma^{-1}}\sigma^{-1}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d\sigma_1\dots\sigma_d\Gamma(\sigma^{-1}/2)}\int_0^1(1-u)u^{\sigma^{-1}/2}du \\ &= \frac{T^{\sigma^{-1}}\sigma^{-1}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d\sigma_1\dots\sigma_d(\sigma^{-1}+4)\Gamma(2+\sigma^{-1}/2)}, \end{aligned}$$

and

$$I_0 = I_1 - I_2 \sim \frac{4T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d(\sigma^{-1}+4)\sigma_1\dots\sigma_d\Gamma(2+\sigma^{-1}/2)}.$$

From this

$$I_1/I_2 \sim (\sigma^{-1}+4)/\sigma^{-1},$$

and by (4.4)

$$T \sim r_\varepsilon^{-1}\left((\sigma^{-1}+4)/\sigma^{-1}\right)^{1/2}. \tag{4.13}$$

Next, using relation (4.13) and the identity $\Gamma(x+1)=x\Gamma(x)$, we have

$$\begin{aligned} I_0/I_1^2 &\sim \frac{2(2\pi)^d\sigma_1\dots\sigma_d\Gamma(2+\sigma^{-1}/2)}{T^{\sigma^{-1}}(\sigma^{-1}+4)\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))} \\ &\sim \frac{\pi^d r_\varepsilon^{\sigma^{-1}}(\sigma^{-1})^{\sigma^{-1}/2}(\sigma^{-1}+2)\Gamma(1+\sigma^{-1}/2)}{(\sigma^{-1}+4)^{1+\sigma^{-1}/2}\Gamma(1+1/(2\sigma_1))\dots\Gamma(1+1/(2\sigma_d))}. \end{aligned}$$

Whence, by (4.5) and in view of Theorem 2.2, we arrive at the statement of Theorem 3.2. □

Proof of Theorem 3.1. Similarly to the proof of Theorem 3.2, we need to study equations (2.9) and (2.10) as $T\to\infty$. Denote

$$J_1 = J_1(\sigma,d) = \sum_{c_l<T}(1-c_l/T), \tag{4.14}$$

$$J_2 = J_2(\sigma,d) = \sum_{c_l<T}(c_l/T)(1-c_l/T). \tag{4.15}$$

Let x_{l_j} be defined in (4.6), and let $D_{\sigma,d}$ and Σ_d be as before. Applying the Liouville formula, we have as $T \to \infty$

$$
\begin{aligned}
J_1 &= \sum_{c_l<T} \left(1 - \left(\sum_{j=1}^{d} (2\pi|l_j|)^{2\sigma_j}/T^2\right)^{1/2}\right) \\
&= m_1 \dots m_d \sum_{\sum |x_{l_j}|^{2\sigma_j}<1} \left(1 - \left(\sum_{j=1}^{d} |x_{l_j}|^{2\sigma_j}\right)^{1/2}\right) m_1^{-1} \dots m_d^{-1} \\
&\sim \frac{T^{\sigma^{-1}}}{(2\pi)^d} \int_{D_{\sigma,d}(1)} \left(1 - \left(\sum_{j=1}^{d} |x_j|^{2\sigma_j}\right)^{1/2}\right) dx_1 \dots dx_d, \\
&= \frac{T^{\sigma^{-1}}}{(2\pi)^d \sigma_1 \dots \sigma_d} \int_{\Sigma_d} \left(1 - \left(\sum_{j=1}^{d} y_j\right)^{1/2}\right) y_1^{1/(2\sigma_1)-1} \dots y_d^{1/(2\sigma_d)-1} dy_1 \dots dy_d \\
&= \frac{T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d \sigma_1 \dots \sigma_d \Gamma(\sigma^{-1}/2)} \left(\int_0^1 u^{\sigma^{-1}/2-1} du - \int_0^1 u^{(\sigma^{-1}-1)/2} du\right) \\
&= \frac{2T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d \sigma^{-1}(\sigma^{-1}+1)\sigma_1 \dots \sigma_d \Gamma(\sigma^{-1}/2)}.
\end{aligned}
$$

Similar calculations for the sum J_2 yield as $T \to \infty$

$$
\begin{aligned}
J_2 &\sim \frac{T^{\sigma^{-1}}}{(2\pi)^d} \int_{D_{\sigma,d}(1)} \left(\left(\sum_{j=1}^{d} |x_j|^{2\sigma_j}\right)^{1/2} - \sum_{j=1}^{d} |x_j|^{2\sigma_j}\right) dx_1 \dots dx_d, \\
&= \frac{T^{\sigma^{-1}}}{(2\pi)^d \sigma_1 \dots \sigma_d} \int_{\Sigma_d} \left(\left(\sum_{j=1}^{d} y_j\right)^{1/2} - \sum_{j=1}^{d} y_j\right) y_1^{1/(2\sigma_1)-1} \dots y_d^{1/(2\sigma_d)-1} dy_1 \dots dy_d \\
&= \frac{T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d \sigma_1 \dots \sigma_d} \left(\int_0^1 u^{(\sigma^{-1}-1)/2} du - \int_0^1 u^{\sigma^{-1}/2} du\right) \\
&= \frac{2T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}{(2\pi)^d (\sigma^{-1}+1)(\sigma^{-1}+2)\sigma_1 \dots \sigma_d \Gamma(\sigma^{-1}/2)},
\end{aligned}
$$

and hence

$$J_1/J_2 \sim (\sigma^{-1}+2)/\sigma^{-1}.$$

Next, in view of (2.9) and (2.10),

$$E_\varepsilon^2(\sigma,d) = \varepsilon^2 J_1, \quad T^2 = E_\varepsilon^{-2}(\sigma,d) J_1/J_2 \sim E_\varepsilon^{-2}(\sigma,d)(\sigma^{-1}+2)/\sigma^{-1},$$

where $T = T_\varepsilon$ satisfies

$$T^2 = \varepsilon^{-2}/J_2 \sim \frac{(2\pi)^d(\sigma^{-1}+1)(\sigma^{-1}+2)\sigma_1 \dots \sigma_d \Gamma(\sigma^{-1}/2)}{2\varepsilon^2 T^{\sigma^{-1}}\Gamma(1/(2\sigma_1))\dots\Gamma(1/(2\sigma_d))}.$$

From this, using again the identity $\Gamma(x+1) = x\Gamma(x)$,

$$T^2 \quad \sim \quad \left(\frac{\pi^d(\sigma^{-1}+1)(\sigma^{-1}+2)\Gamma(\sigma^{-1}/2)}{2\varepsilon^2\Gamma(1+1/(2\sigma_1))\dots\Gamma(1+1/(2\sigma_d))} \right)^{2/(2+\sigma^{-1})},$$

and the statement of Theorem 3.1 follows. □

4.2. Proof of Theorem 3.3

The proof of Theorem 3.3 utilizes the so-called count function and largely consists of studying its properties.

4.2.1. Count function

An important role in the analysis of the infinite-dimensional case is played by the *count function* $N(t)$ which is defined for any $t > 0$ as follows:

$$N(t) = \text{card}\{\mathcal{N}(t)\}, \quad \mathcal{N}(t) = \{l \in \mathbb{Z}_0^\infty : c_l \leq t\}.$$

The count function can be thought of as the distribution function of the coefficients c_l. It satisfies $N(t) \to \infty$ as $t \to \infty$, and determines rate asymptotics of integrated squared risk in the estimation problem and of separation rate in the detection problem. More precisely, for the estimation problem (see, for example, [10, Sec. 2])

$$R_\varepsilon(\mathcal{F}_{\sigma,\infty}) \asymp T^{-1}, \quad \text{where} \quad \varepsilon^2 T^2 N(T) \asymp 1, \tag{4.16}$$

and for the detection problem

$$r_\varepsilon^*(\mathcal{F}_{\sigma,\infty}) \asymp T^{-1}, \quad \text{where} \quad \varepsilon^4 T^4 N(T) \asymp 1. \tag{4.17}$$

In addition, under certain regularity constraints on $N(t)$, this function controls sharp asymptotics of the minimax integrated squared risk $R_\varepsilon^2(\mathcal{F}_{\sigma,\infty})$ and the minimax total error probability $\gamma_\varepsilon(\mathcal{F}_{\sigma,d}(r_\varepsilon))$ (see, for example, [10, Sec. 2] for details).

4.2.2. Probability measures

Due to (4.16) and (4.17) finding rate asymptotics of $R_\varepsilon(\mathcal{F}_{\sigma,\infty})$ and $r_\varepsilon^*(\mathcal{F}_{\sigma,\infty}(r_\varepsilon))$ requires studying the properties of $N(t)$ as $t \to \infty$. A general method consists of defining a family of prior distributions $\mathbf{P}_h$ on the set of indices $\mathbb{Z}_0^\infty$ and investigating the behaviour of the function $N(t) = \text{card}\{l \in \mathbb{Z}_0^\infty : c_l \leq t\}$ using probabilistic and analytical tools.

First, let us define a family of probability measures $\mathbf{P}_h$, depending on a positive parameter h, of the form

$$\mathbf{P}_h(l) = \prod_{j=1}^{\infty} \mathbf{P}_{h,j}(l_j), \quad l \in \mathbb{Z}^\infty.$$

To this end, define the random variables

$$Y_j(k) = (2\pi|k|)^{2\sigma_j}, \quad j = 1, 2, \dots, \quad k \in \mathbb{Z},$$

and view the coefficients

$$c_l^2 = \sum_{j=1}^{\infty} (2\pi|l_j|)^{2\sigma_j} = \sum_{j=1}^{\infty} Y_j(l_j) \stackrel{\text{def}}{=} S(l), \quad l \in \mathbb{Z}_0^\infty,$$

as realizations of the random variables $S(l)$, $l \in \mathbb{Z}_0^\infty$. Then

$$N(t) = \text{card}\{l \in \mathbb{Z}_0^\infty : S(l) \le t^2\}.$$

Next, for $h > 0$ define the probability measures $\mathbf{P}_{h,j}$ on $\mathbb{Z}$ by

$$\mathbf{P}_{h,j}(k) = \exp(-hY_j(k) - Z_j(h)), \quad Z_j(h) = \log\left(\sum_{k \in \mathbb{Z}} \exp(-hY_j(k))\right),$$

and put

$$\mathbf{P}_h(l) = \prod_{j=1}^{\infty} \mathbf{P}_{h,j}(l_j) = \exp(-hS(l) - Z(h)),$$

where

$$Z(h) = \sum_{j=1}^{\infty} Z_j(h) = \sum_{j=1}^{\infty} \log\left(1 + G_j(h)\right), \quad G_j(h) = 2\sum_{k=1}^{\infty} \exp(-h(2\pi k)^{2\sigma_j}). \tag{4.18}$$

Using the arguments similar to those in [9, p. 16], it is readily shown that for all $h > 0$

$$\mathbf{P}_h(\mathbb{Z}_0^\infty) = \lim_{d \to \infty} \mathbf{P}_h(\mathbb{Z}_0^d) = 1,$$

so that $\mathbf{P}_h(l)$ is a probability measure on $\mathbb{Z}^\infty$ with support on $\mathbb{Z}_0^\infty$.

Setting

$$H = t^2$$

leads to the representation, for any $h > 0,\ t > 0$,

$$\begin{aligned} N(t) &= \text{card}\{l \in \mathbb{Z}_0^\infty : c_l \le t\} = \text{card}\{l \in \mathbb{Z}_0^\infty : S(l) \le H\} \\ &= e^{Z(h)+hH} \sum_{l \in \mathbb{Z}_0^\infty : S(l) \le H} e^{h(S(l)-H)} \mathbf{P}_h(l) = e^{Z(h)+hH} I_h, \end{aligned}$$

where $I_h = \mathbf{E}_h\left(e^{h(S-H)} \mathbb{I}_{\{S \le H\}}\right) \le 1$. Hence

$$\log N(t) \le Z(h) + ht^2, \quad \forall\, h > 0,\ t > 0. \tag{4.19}$$

4.2.3. Auxiliary results

This section contains some auxiliary results that will be used in the proof of Theorem 3.3.

Lemma 4.1. *Let the sequence* $(\sigma_j)_{j\geq 1}$ *be non-decreasing and let* (3.4) *hold true. Then*

$$\sigma_j/j \to \infty, \quad j \to \infty.$$

Proof. The lemma is easily proved by contradiction. Suppose there exist a subsequence $(\sigma_{j_k})_{k\geq 1}$ and a constant $B > 0$ such that

$$\sigma_{j_k} \leq B j_k, \quad k = 1, 2, \ldots.$$

Then for $j_k/2 \leq j \leq j_k$ we have $\sigma_j \leq 2Bj$, and for k sufficiently large

$$\sum_{j_k/2\leq j\leq j_k} \sigma_j^{-1} \geq \frac{1}{2B} \sum_{j_k/2\leq j\leq j_k} \frac{1}{j} \sim \frac{1}{2B} \int_{j_k/2}^{j_k} \frac{dx}{x} = \frac{\log 2}{2B} > 0. \tag{4.20}$$

On the other hand, in view of (3.4),

$$\sum_{j_k/2\leq j\leq j_k} \sigma_j^{-1} < \sum_{j\geq j_k/2} \sigma_j^{-1} \to 0, \quad k \to \infty,$$

which contradicts (4.20). Hence, the lemma follows. □

In the sequel, without loss of generality the sequence $(\sigma_j)_{j\geq 1}$ is assumed non-decreasing.

By (4.19) the upper bound for $\log N(t)$ is controlled by the term $Z(h) + ht^2$. The following lemma establishes the asymptotic behavior of $Z(h)$.

Lemma 4.2. *As* $h \to 0$

$$Z(h) \sim \frac{\sigma^{-1}}{2} \log(h^{-1}).$$

Proof. The key point is to split $Z(h)$ appropriately into the main term, which gives the required asymptotics, and the remainder. We have (see (4.18))

$$Z(h) = \sum_{1\leq j\leq J} \log(1 + G_j(h)) + \sum_{j>J} \log(1 + G_j(h)) \stackrel{\text{def}}{=} S_1 + S_2, \tag{4.21}$$

where parameter $J = J(h) \to \infty$ as $h \to 0$ is chosen to have

$$G_j(h) = \begin{cases} C_j h^{-1/(2\sigma_j)} + O(1) & , \quad \text{if } j \leq J, \\ B_j e^{-h(2\pi)^{2\sigma_j}} & , \quad \text{if } j > J, \end{cases}$$

with some constants $C_j \asymp 1$ and $B_j \asymp 1$. Such a choice of J is possible because for "small" σ_j's, setting $m_j = h^{-1/(2\sigma_j)}$,

$$\begin{aligned} G_j(h) &= 2m_j \sum_{k=1}^{\infty} \exp\left(-(2\pi(k/m_j))^{2\sigma_j}\right) m_j^{-1} \\ &= 2m_j \int_0^{\infty} e^{-(2\pi x)^{2\sigma_j}}\, dx + O(1) = \frac{\Gamma(1/(2\sigma_j))}{2\sigma_j \pi h^{1/(2\sigma_j)}} + O(1), \end{aligned}$$

and for "large" σ_j's, the function $G_j(h)$ is approximated by the first term $2e^{-h(2\pi)^{2\sigma_j}}$.

To be precise, let the parameter $J = J(h) \to \infty$ as $h \to 0$ be such that

$$\delta_1 \stackrel{\text{def}}{=} h\sigma_J (2\pi)^{2\sigma_J} \asymp 1. \tag{4.22}$$

This yields

$$\delta \stackrel{\text{def}}{=} \delta_1/\sigma_J = h(2\pi)^{2\sigma_J} = o(1),$$

and using Lemma 4.1

$$J \ll \sigma_J \asymp \log(h^{-1}). \tag{4.23}$$

Then the first sum in (4.21) equals

$$\begin{aligned} S_1 &= \sum_{1\le j\le J} \log(1 + G_j(h)) = \sum_{1\le j\le J} \log\left(C_j h^{-1/(2\sigma_j)} + O(1)\right) \\ &= \sum_{1\le j\le J} \log\left(h^{-1/(2\sigma_j)}\right) + \sum_{1\le j\le J} \log\left(C_j + O(1)h^{1/(2\sigma_j)}\right) \\ &= \frac{\log(h^{-1})}{2} \sum_{1\le j\le J} \sigma_j^{-1} + O(J) = \frac{\sigma^{-1}\log(h^{-1})}{2}(1 + o(1)). \end{aligned} \tag{4.24}$$

It remains to show that $S_2 = o(\log(h^{-1}))$. By the inequality $e^y - e^x \ge e^x(y - x)$, which is the same as $e^z - 1 \ge z$, we have

$$(2\pi)^{2\sigma_j} - (2\pi)^{2\sigma_J} \ge (2\pi)^{2\sigma_J} c(\sigma_j - \sigma_J), \quad c = 2\log(2\pi).$$

Therefore using $\log(1 + x) \le x$ and noting that

$$G_j(h) \sim 2\exp\left(-h(2\pi)^{2\sigma_j}\right), \quad h \to 0,$$

we get

$$\begin{aligned} S_2 &= \sum_{j>J} \log(1 + G_j(h)) \le 4 \sum_{j>J} \exp\left(-h(2\pi)^{2\sigma_j}\right) \\ &= 4\exp(-h(2\pi)^{2\sigma_J}) \sum_{j>J} \exp(-h\left((2\pi)^{2\sigma_j} - (2\pi)^{2\sigma_J}\right) \end{aligned}$$

$$\le 4\sum_{j>J}\exp(-h(2\pi)^{2\sigma_J}c(\sigma_j-\sigma_J)) = 4\exp(c\delta_1)\sum_{j>J}\exp(-c\delta\sigma_j)$$
$$\le 4\exp(c\delta_1)\sum_{j>J}\exp(-c\delta Aj) \le \frac{4\exp(c\delta_1)\exp(-c\delta AJ)}{(1-\exp(-cA\delta))}$$
$$\le \frac{4\exp(c\delta_1)}{(1-\exp(-cA\delta))},$$

where, thanks to Lemma 4.1, we use the inequality $\sigma_j > Aj$, $A \to \infty$. If $A\delta$ is bounded away from zero then $S_2 = O(1)$ and we are done. If $A\delta = o(1)$ then, recalling that $\delta_1 \asymp 1$ and $A \to \infty$, we may continue

$$S_2 \le \frac{4\exp(c\delta_1)}{(1-\exp(-cA\delta))} \sim \frac{4\exp(c\delta_1)}{cA\delta} = \frac{4\exp(c\delta_1)}{c\delta_1}\frac{\sigma_J}{A}$$
$$\asymp \frac{\sigma_J}{A} = o(\sigma_J) = o(\log(h^{-1})),$$

where the last equality is due to (4.23). Thus

$$Z(h) = \frac{\sigma^{-1}\log(h^{-1})}{2}(1+o(1)),$$

and the lemma is proved. □

4.2.4. Proof of Theorem 3.3

The proofs of both relations are similar. Let us prove the second one. This is done in two steps, the lower bound part and the upper bound part. The lower bound part is reduced to the finite-dimensional case. To simplify notation, we write r_ε^* for $r_\varepsilon^*(\mathcal{F}_{\sigma,\infty})$.

By Remark 3.3,

$$\frac{\log r_\varepsilon^*}{\log\varepsilon} \ge \frac{\log r_\varepsilon^*(\mathcal{F}_{\sigma,d})}{\log\varepsilon} \sim \frac{4}{4+\sigma^{-1}(d)}, \quad \sigma^{-1}(d) = \sum_{j=1}^{d}\sigma_j^{-1}.$$

Then, in view of (3.4), passage to the limit as $d \to \infty$ gives the lower bound:

$$\frac{\log r_\varepsilon^*}{\log\varepsilon} \ge \frac{4}{4+\sigma^{-1}}(1+o(1)). \tag{4.25}$$

The upper bound part consists of showing that for sufficiently small $\varepsilon > 0$

$$\frac{\log r_\varepsilon^*}{\log\varepsilon} \le \frac{4}{4+\sigma^{-1}}(1+o(1)). \tag{4.26}$$

By (4.17)

$$\log r_\varepsilon^* \sim -\log T, \quad \text{where} \quad 4\log\varepsilon + \log N(T) \sim -4\log T,$$

and hence

$$\frac{\log r_\varepsilon^*}{\log \varepsilon} \sim 1 + \frac{\log N(T)}{4 \log \varepsilon}.$$

Therefore for the validity of (4.26) it suffices to show that for sufficiently small ε

$$\log N(T) \le \frac{4\sigma^{-1} \log(\varepsilon^{-1})}{4 + \sigma^{-1}}(1 + o(1)), \tag{4.27}$$

where $T = T_\varepsilon$ satisfies $\varepsilon^4 T^4 N(T) \asymp 1$.

From (4.19) for any $h > 0$

$$\log N(T) \le Z(h) + hT^2, \quad T^2 \asymp \frac{1}{\varepsilon^2 N^{1/2}(T)}, \tag{4.28}$$

where by Lemma 4.2,

$$Z(h) \sim \frac{\sigma^{-1}}{2} \log(h^{-1}), \quad h \to 0.$$

Up to a vanishing term, the right-hand side of inequality in (4.28) is equal to $(\sigma^{-1}/2) \log(h^{-1}) + hT^2$. The minimum of the latter (as a function of h) is attained at the point $h = \sigma^{-1}/(2T^2)$. In other words, the minimum occurs when $h \asymp \varepsilon^2 N^{1/2}(T)$, see (4.28). In this case $hT^2 \asymp 1$, and

$$\log(h^{-1}) \sim 2 \log(\varepsilon^{-1}) - \frac{1}{2} \log N(T).$$

Therefore, using (4.28)

$$\begin{aligned} \log N(T) &\le \frac{\sigma^{-1}}{2} \log(h^{-1})(1 + o(1)) + O(1) \\ &= \left(\sigma^{-1} \log(\varepsilon^{-1}) - \frac{\sigma^{-1}}{4} \log N(T)\right)(1 + o(1)) + O(1). \end{aligned}$$

From this inequality (4.27), and hence the upper bound (4.26), follows. Combining (4.25) and (4.26), we get the required asymptotic expression for r_ε^*.

For the estimation problem the proof is completely analogous, cf. (4.16) and (4.17). The proof of Theorem 3.3 is completed. □

Acknowledgements

We would like to thank Associate Editor and anonymous referee for their comments and suggestions.

References

[1] Efromovich, S. (1999), *Nonparametric Curve Estimation: Methods, Theory and Applications*, Springer, New York. MR1705298

[2] Ermakov, M. S. (1991), Minimax detection of a signal in Gaussian white noise, *Theory of Probability and Its Applications* **35** (4), 667–679. MR1090496

[3] Fichtengoltz, G. M. (1977), *Differential- und Integralrechnung*, Vol. III, VEB Deutscher Verlag der Wissenschaften, Berlin.

[4] Ibragimov, I. A. and Khasminskii, R. Z. (1997), Some estimation problems in infinite dimensional Gaussian white noise, in *Festschrift for Lucien Le Cam. Research Papers in Probability and Statistics*, Springer-Verlag, New York, pp. 275–296. MR1462950

[5] Ingster, Yu. I. and Stepanova, N. (2009), Estimation and detection of functions from weighted tensor product spaces, *Mathematical Methods of Statistics* **18**, 310–340. MR2608165

[6] Ingster, Yu. I. and Suslina I. A. (2003), *Nonparametric Goodness-of-Fit Testing under Gaussian Models*. Lecture Notes in Statistics, Vol. **169**, Springer-Verlag, New York.

[7] Ingster, Yu. I. and Suslina I. A. (2005), On estimation and detection of smooth function of many variables, *Mathematical Methods of Statistics* **14**, 299–331. MR2195328

[8] Ingster, Yu. I. and Susulina I. A. (2006), On estimation and detection of functions of infinitely many variables, *Journal of Mathematical Sciences* **139** (3), 6548–6561. MR2214536

[9] Ingster, Yu. I. and Suslina I. A. (2007), Estimation and detection of high-variable functions from Sloan–Woźniakowski space. *Mathematical Methods of Statistics* **16**, 318–353. MR2378279

[10] Ingster, Yu. I. and Suslina I. A. (2007), On estimation and detection of a function from tensor product spaces (in Russian). *Zapiski Nauchnyh Seminarov POMI* **351**, 180–218. (English translation in *Journal of Mathematical Sciences* (2008) **152**, 897–920.) MR2742908

[11] Ingster, Yu. I. and Suslina I. A. (2009), Adaptive detection of a high-variable function (in Russian). *Zapiski Nauchnyh Seminarov POMI*, **368**, 156–170. (English translation in *Journal of Mathematical Sciences* (2010) **167**, 522–530.) MR2749190

[12] Kerkyacharian, G., Lepski, O., and Picard, D. (2001), Nonliner estimation in anisotropic multi-index denoising, *Probability Theory and Related Fields* **121**, 137–170. MR1863916

[13] Kerkyacharian, G., Lepski, O., and Picard, D. (2008), Nonliner estimation in anisotropic multi-index denoising. Sparse case, *Theory of Probability and Its Applications* **52** (1), 58–77. MR2354574

[14] Nikolskii, S. M. (1969), *Approximation of Functions of Several Variables and Imbedding Theorems* (in Russian), 1st ed., Nauka, Moscow. (Elglish translation: Springer-Verlag, Berlin, 1975.) MR0374877

[15] Nussbaum, M. (1983), Optimal filtering of a function of many variables in Gaussian white noise, *Problems of Information Transmission* **19** (2), 105–111. MR0744616

[16] Pinsker, M. S. (1980), Optimal filtration of square-integrable signals in Gaussian noise, *Problems of Information Transmission* **16**, 120–133. MR0624591

[17] Skorohod, A. V. (1974), *Integration in Hilbert Spaces*, Springer-Verlag, Berlin–New York. MR0466482

[18] Tsybakov, A. B. (2009), *Introduction to Nonparametric Estimation*, Springer Science+Business Media, New York. MR2724359