Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Minimax nonparametric hypothesis testingfor small type I errors

Yuri I. Ingster[1], Irina A. Suslina[2]

submitted: 25th August 2003

[1] St. Petersburg State
Electrotechnical University
(LETI)
Department of Mathematics II
Prof. Popov str. 5
197376 St. Petersburg
Russia
E-Mail: ingster@pdmi.ras.ru

[2] St. ptersburg Institute of
Fine Mechanics and Optics
(Technical University)
Department of Mathematics
Sablinskaya str. 14
197101 St. Petersburg
Russia
E-Mail: suslina_i@mail.ru

No. 867
Berlin 2003

2000 *Mathematics Subject Classification.* 62G10, 62G20.

Key words and phrases. Minimax hypothesis testing, nonparametric signal detection, adaptive hypothesis testing, intermediate efficiency.

The work was finished during the stay of Y. Ingster at WIAS.
Research was partially supported by RFFI Grant 02-02-00262, by RFFI-DFG Grnat 02-01-04001 and by Grand NSh-2258.2003.1.

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

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Minimax nonparametric hypothesis testing for small type I errors [*†]

Yuri I. Ingster[‡§¶], Irina A. Suslina [‖]

Abstract

Under the white Gaussian noise model with the noise level $\varepsilon \to 0$, we study minimax nonparametric hypothesis testing problem $H_0 : f = 0$ on unknown function $f \in L_2(0,1)$. We consider alternative sets that are determined a regularity constraint in the Sobolev norm and we suppose that signals are bounded away from the null either in L_2-norm or in L_∞-norm. Analogous problems are considered in the sequence space.

If type I error probability $\alpha \in (0,1)$ is fixed, then these problems were studied in book [13]. In this paper we consider the case $\alpha \to 0$. We obtain either sharp distinguishability conditions or sharp asymptotics of the minimax type II error probability in the problem. We show that if α is "not too small", then there exists natural extension of results [13], whenever if α is "very small", then we obtain classical asymptotics and distinguishability conditions for small α.

Adaptive problems are studied as well.

1 Introduction

1.1 Model

Let us consider minimax nonparametric hypothesis testing problem on a mean of an infinite-dimensional Gaussian random vector

$$X = v + \xi, \quad \xi = (\xi_1, ..., \xi_i, ...), \quad v = (v_1, ..., v_i, ...) \in l^2, \tag{1.1}$$

[*]2000 *Mathematical Subject Classification.* 62G10, 62G20

[†]*Key words and phrases.* Minimax hypothesis testing, nonparametric signal detection, adaptive hypothesis testing, intermediate efficiency.

[‡]St. Petersburg State Electrotechnical University (LETI), Dep. of Mathematics II, Prof. Popov str. 5, 197376, St.Petersburg, Russia

[§]The work was finished during the stay of Yu. Ingster at WIAS.

[¶]Research was partially supported by RFFI Grant 02-01-00262, by RFFI-DFG Grant 02-01-04001 and by Grand NSh-2258.2003.1.

[‖]St.Petersburg Institute of Fine Mechanics and Optics (Technical University), Dep. of Mathematics, Sablinskaya str. 14, 197101, St.Petersburg, Russia

where $\xi_i \sim \mathcal{N}(0,1)$ are i.i.d. and l^2 is the sequence space. Let a set $V \subset l^2$ be given. From an observation X of form (1.1), we test the null-hypothesis $H_0 : v = 0$ against alternative $H_1 : v \in V$.

Sequence model (1.1) is equivalent to the functional white Gaussian noise model

$$dX_\varepsilon(t) = f(t) + \varepsilon dW(t),\ t \in (0,1),\ f \in L_2(0,1), \tag{1.2}$$

where $W(t)$ is the Wiener process, $\varepsilon > 0$ is a noise level. In fact, taking any orthonormal basis $\{\phi_i\}_{i=1}^\infty$ in the Hilbert space $L_2(0,1)$, we pass to the random variables X_i and to the normalized Fourier coefficients v_i,

$$X_i = \varepsilon^{-1}\int_0^1 \phi_i(t)dX_\varepsilon(t), \quad v_i = \varepsilon^{-1}\int_0^1 \phi_i(t)f(t)dt = \varepsilon^{-1}(f,\phi_i). \tag{1.3}$$

Under model (1.2), let a set $F \subset L_2(0,1)$ be given. From an observation $X_\varepsilon(t),\ t \in (0,1)$, we test the null-hypothesis $H_0 : f = 0$ against alternative $H_1 : f \in F$.

For a test ψ [1] we denote $\alpha(\psi) = E_0(\psi),\ \alpha_\varepsilon(\psi) = E_{\varepsilon,0}(\psi)$ type I error probability and $\beta(\psi,v) = E_v(1-\psi),\ \beta(\psi,f) = E_{\varepsilon,f}(1-\psi)$ type II error probability for the alternative $v \in l^2$ or $f \in L_2(0,1)$. Here and later $E_v,\ E_{\varepsilon,f}$ stands for the expectation with respect to the measure $P_v,\ P_{\varepsilon,f}$ that corresponds to observations (1.1), (1.2). For any $\alpha \in (0,1)$ let

$$\Psi_\alpha = \{\psi : \alpha(\psi) \le \alpha\}, \quad \Psi_{\varepsilon,\alpha} = \{\psi : \alpha_\varepsilon(\psi) \le \alpha\}$$

be the sets of all tests of the level α.

Under the sequence Gaussian model (1.1) let $\beta(\psi,V) = \sup_{v\in V}\beta(\psi,v)$ be the maximal type II error probability. Clearly, for any subset $\tilde V \subset V$,

$$\beta(\psi,\tilde V) \le \beta(\psi,V).$$

Set

$$\beta(V,\alpha) = \inf_{\psi\in\Psi_\alpha}\beta(\psi,V).$$

Clearly, for any $\alpha \in (0,1)$ one has

$$0 \le \beta(V,\alpha) \le 1-\alpha; \quad \beta(V,\alpha) \to 1, \quad \text{as} \quad \alpha \to 0;$$

the function $\beta(V,\alpha)$ decreases in α, and, for any subset $\tilde V \subset V$,

$$\beta(\tilde V,\alpha) \le \beta(V,\alpha). \tag{1.4}$$

In particular, if $\tilde V = \{v\},\ v \in V$, then

$$\beta(\{v\},\alpha) = \Phi(T_\alpha - |v|), \quad |v|^2 = \sum_i v_i^2. \tag{1.5}$$

Here and later $\Phi(t)$ stands for distribution function of the standard Gaussian law and T_α is its $(1-\alpha)$-quantile: $\Phi(T_\alpha) = 1-\alpha$.

[1] We call test a measurable function on observation X or X_ε taking values in the interval $[0,1]$.

The inequality (1.4) yields the evident lower bounds

$$\beta(V, \alpha) \geq \sup_{v \in V} \beta(\{v\}, \alpha) = \Phi(T_\alpha - \inf_{v \in V} |v|). \tag{1.6}$$

We use analogous definitions and notations for the functional model (1.2). For this case analogous relations holds true. In particular, for simple alternative $F = \{f\}$ one has $\beta_\varepsilon(\{f\}, \alpha) = \Phi(T_\alpha - \|f\|/\varepsilon)$, where $\|\cdot\|$ is L_2-norm and

$$\beta_\varepsilon(F, \alpha) \geq \sup_{f \in F} \beta_\varepsilon(\{f\}, \alpha) = \Phi(T_\alpha - \varepsilon^{-1} \inf_{f \in F} \|f\|). \tag{1.7}$$

We consider analogous hypothesis testing problems

$$H_0 : v = 0, \quad H_1 : v \in V_R \quad \text{or} \quad H_0 : f = 0, \quad H_1 : f \in F_\varepsilon \subset L_2(0,1)$$

under asymptotic variant of minimax setting assuming $R \to \infty$, $\varepsilon \to 0$.

For wide class of alternatives, these problems are well studied for a fixed level of testing $\alpha \in (0,1)$, see [13] and Section 2 below. However for a lot of practical hypothesis testing problems a statistician wants to have small or very small α. In particular, small or very small α are required in real-time signal detection problems. Under asymptotic approach this corresponds $\alpha = \alpha_R \to 0$ or $\alpha = \alpha_\varepsilon \to 0$ (in what follows limits are assumed either as $R \to \infty$ or as $\varepsilon \to 0$ unless otherwise stated).

Namely, taking a family $\alpha = \alpha_R \in (0,1)$ or $\alpha = \alpha_\varepsilon \in (0,1)$, we are interested the asymptotics of the families $\beta(V_R, \alpha_R)$ or $\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon)$ (up to a vanishing term) and in conditions for minimax distinguishability, i.e., for

$$\beta(V_R, \alpha_R) \to 0, \quad \beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \to 0$$

or for minimax nondistinguishability, i.e., for

$$\beta(V_R, \alpha_R) = 1 - \alpha_R + o(1), \quad \beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) = 1 - \alpha_\varepsilon + o(1).$$

Also if $\beta(V_R, \alpha_R) \to 0$, $\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \to 0$, then we want to construct test procedures ψ_R, ψ_ε providing α_R, α_ε-level distinguishability, i.e., for

$$\psi_R \in \Psi_{\alpha_R}, \ \psi_\varepsilon \in \Psi_{\varepsilon, \alpha_\varepsilon}, \quad \beta(\psi_R, V_R) \to 0, \ \beta(\psi_\varepsilon, F_\varepsilon) \to 0.$$

It follows from well-known asymptotics

$$\Phi(-x) \sim \exp(-x^2/2)/x\sqrt{2\pi}, \quad \text{as} \quad x \to \infty, \tag{1.8}$$

that

$$T_\alpha = \sqrt{2 \log \alpha^{-1}} + o(1), \quad \text{as} \quad \alpha \to 0. \tag{1.9}$$

Using (1.9) we can rewrite (1.6), (1.7) for $\alpha \to 0$:

$$\beta(V_R, \alpha_R) \geq \Phi(\sqrt{2 \log \alpha_R^{-1}} - \inf_{v \in V_R} |v|) + o(1), \tag{1.10}$$

$$\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \geq \Phi(\sqrt{2 \log \alpha_e^{-1}} - \varepsilon^{-1} \inf_{f \in F_\varepsilon} \|f\|) + o(1). \tag{1.11}$$

1.2 Alternatives of interest

It is well known (see [8], [13], Ch. 1) that, under the functional Gaussian model (1.2) in order to obtain minimax distinguishability in nonparametric problem, one needs:

- to remove "small enough" signals;
- to suppose that the set of signals is not too "hudge".

In general, it is impossible to make a "minimax decision" without these assumptions.

To obey the first constraint, i.e. to measure the "size" of the signal, a functional norm is usually used. For instance, L_p-norm, $1 \le p \le \infty$:

$$\|f\|_p = \left(\int_0^1 |f(t)|^p dt\right)^{1/p}, \ 1 \le p < \infty; \quad \|f\|_\infty = \operatorname{ess\,sup}_{0 \le t \le 1} |f(t)|.$$

To obey the second constraint the signal is supposed to belong to some compact set in a Banach space. The typical examples are the classes of smooth functions like Hölder, Sobolev or Besov spaces. To define these classes some semi-norms are usually used. In particular, the Sobolev norm $\|\cdot\|_{\sigma,q}$ is described by two parameters $\sigma,\ q$. Here the parameter $\sigma > 0$ characterizes the level of the smoothness and $q \in [0, \infty]$ characterizes the norm where the smoothness is measured.

In order to specify alternative sets in the functional space, let us consider the norms $\|\cdot\|_{\sigma,q},\ \sigma > 0,\ q \in [1, \infty]$ in a subspace of $L_2(0,1)$. If $\sigma \ge 1$ is an integer, then we assume $f^{(\sigma-1)}$ is absolutely continuous and set

$$\|f\|_{\sigma,q} = \|f\|_q + \|f\|^0_{\sigma,q}, \quad \|f\|^0_{\sigma,q} = \|f^{(\sigma)}\|_q, \tag{1.12}$$

where $f^{(\sigma)}$ is σ-th derivative of the function f and $\|\cdot\|_p$ is L_p-norm. This is the traditional Sobolev norm. For $q = 2$ we can consider the equivalent norm

$$\|f\|^2_{\sigma,2} = \|f\|^2_2 + \|f^{(\sigma)}\|^2_2. \tag{1.13}$$

If $\sigma = l + \tau,\ \tau \in (0,1),\ q \in [1, \infty)$, then we set

$$\|f\|_{\sigma,q} = \|f\|_q + \|f\|^0_{\sigma,q}, \quad \|f\|^0_{\sigma,q} = \sup_{h \in (0,1)} h^{-\tau} \left(\int_0^{1-h} |f^{(l)}(t+h) - f^{(l)}(t)|^q dt\right)^{1/q} \tag{1.14}$$

with evident modification for $q = \infty$. This corresponds to the Nikol'ski norm or Besov norm $\|\cdot\|_{\sigma,q,\infty}$.

Note the following relation (see [13], inequalities (2.81), (2.82)). Let $\eta = \sigma - 1/q > 0$. There exists a constant $c > 0$ such that

$$\|f\|_{\eta,\infty} \le c\|f\|_{\sigma,q}. \tag{1.15}$$

Taking a positive family $r_\varepsilon = o(1)$ and $H > 0$, we consider alternative sets F_ε of the form

$$F_\varepsilon = F(r_\varepsilon, H, \sigma, p, q) = \{f \in L_2(0,1) \,:\, \|f\|_p \geq r_\varepsilon,\ \|f\|_{\sigma,q} \leq H\}.$$

We are interested in the cases $\sigma > 0$ and either $p = q = 2$ or $p = \infty,\ q > 1/\sigma$, i.e., in alternative sets

$$F_\varepsilon = F(r_\varepsilon, H, \sigma) = \{f \in L_2(0,1) : \|f\|_2 \geq r_\varepsilon,\ \|f\|_{\sigma,2} \leq H\}, \tag{1.16}$$

$$F_\varepsilon = F(r_\varepsilon, H, \sigma, q) = \{f \in L_2(0,1) : \|f\|_\infty \geq r_\varepsilon,\ \|f\|_{\sigma,q} \leq H\},\ q > 1/\sigma. \tag{1.17}$$

Note that the results below hold true with change the norms $\|f\|_{\sigma,2},\ \|f\|_{\sigma,q}$ by $\|f\|^0_{\sigma,2},\ \|f\|^0_{\sigma,q}$.

In order to specify alternative sets under the sequence Gaussian model (1.1) take a quantity $p \in (0,\infty]$ and positive family $\rho_R \leq R$. Introduce the norms (quasi-norm for $p < 1$) in the sequence space

$$|v|^p_{\bar{a},p} = \sum_{i=1}^{\infty} |v_i a_i|^p,\ p < \infty; \quad |v|_{\bar{a},\infty} = \sup_{1 \leq i \leq \infty} |v_i a_i|. \tag{1.18}$$

For $a_i = i^\sigma$, we denote this norm by $|\cdot|_{\sigma,p}$. For $\sigma = 0$, we use the notation $|\cdot|_p$.

Taking quantities $p \in (0,\infty],\ q \in (0,\infty],\ s > 0,\ \tau \geq 0,\ R \geq \rho_R > 0$, we consider the alternative sets V_R determined by the inequalities

$$V_R = V(\rho_R, R; \tau, s, p, q) = \{v \in l^2 \,:\, |v|_{\tau,p} \geq \rho_R,\ |v|_{s,q} \leq R\}.$$

This set is l^q-ellipsoid of semi-axes Ri^{-s} with l^p-ellipsoid of semi-axes $\rho_R i^{-\tau}$ removed; if $\tau = 0$, then we remove l^p-ball of radius ρ_R. In this paper we are interested in the cases either $s = \sigma > 0,\ p = q = 2,\ \tau = 0$, or $p = \infty,\ s > \tau \geq 0$, i.e., in alternatives of the form

$$V_R = V(\rho_R, R, \sigma) = \{v \in l^2 \,:\, |v|_2 \geq \rho_R,\ |v|_{\sigma,2} \leq R\}, \quad \sigma > 0, \tag{1.19}$$

$$V_R = V(\rho_R, R; \tau, s, q) = \{v \in l^2 \,:\, |v|_{\tau,\infty} \geq \rho_R,\ |v|_{s,q} \leq R\}, \quad s > \tau \geq 0. \tag{1.20}$$

Since $V_R = \emptyset$ for $R < \rho_R$, we assume $R \geq \rho_R$ later.

If $p = q = 2$, then alternatives (1.16) roughly correspond to alternatives (1.19) with $\rho_R = r_\varepsilon/\varepsilon, R = H/\varepsilon$. Moreover, if $p = 2,\ \sigma > 0$ is an integer and a function $f \in L_2(0,1)$ has 1-periodic σ-smooth extension on R^1, then L_2-norm and the norm (1.13) can be presented in terms of Fourier coefficients $\theta_i = (f, \phi_i)$ under the standard trigonometrical Fourier basis $\{\phi_i\}$ in $L_2(0,1)$. Namely setting $f_\theta(t) = \sum_i \theta_i \phi_i(t)$, we have

$$\|f_\theta\|_2^2 = \sum_{i=1}^{\infty} \theta_i^2, \quad \|f_\theta\|^2_{\sigma,2} = \sum_{i=1}^{\infty} a_i^2 \theta_i^2, \quad a_1 = 1,\ a^2_{2i} = a^2_{2i+1} = (2\pi i)^{2\sigma} + 1,\ i > 1. \tag{1.21}$$

Relation (1.21) determines Sobolev norm in the space $\tilde{W}^\sigma_2$ of 1-periodic σ-smooth functions for all $\sigma > 0$. Using (1.21) we set

$$\tilde{F}_\varepsilon = \tilde{F}_\varepsilon(r, H, \sigma) = \{f_\theta \,:\, |\theta|_2 \geq r,\ |\theta|_{\bar{a},2} \leq H\} \tag{1.22}$$

Relations (1.3) yields the equality

$$\beta(V(\rho_R, R; \bar{a}), \alpha) = \beta(\tilde{F}_\varepsilon, \alpha), \quad \rho_R = r/\varepsilon, \ R = H/\varepsilon,$$

where the set

$$V(\rho_R, R; \bar{a}) = \{v \in l^2 : |v|_2 \geq \rho_R, \ |v|_{\bar{a},2} \leq R\}$$

is analogous to set (1.19) with norm (1.18) for the sequence $\bar{a}$ such that $a_i \sim (\pi i)^\sigma, \ i \to \infty$, in (1.21).

However if $p = \infty$, then the alternatives $F_\varepsilon = F(r_\varepsilon, H, \sigma, q)$ of form (1.17) roughly correspond to the alternatives $V_R = V(\rho_R, R; \tau, s, q)$ of form (1.20) with

$$\tau = 1/2, \quad s = \sigma + 1/2 - 1/q, \quad \rho_R = r_\varepsilon/\varepsilon, \quad R = H/\varepsilon, \tag{1.23}$$

see [13], Sections 2.7, 2.9. This is the main reason why we consider the alternatives (1.20) with $\tau > 0$.

Our aim is to study the asymptotics $\beta(V_R, \alpha_R), \ R \to \infty$ and $\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon), \ \varepsilon \to 0$ in hypothesis testing problems with alternatives (1.19), (1.20), (1.16), (1.17) and to construct asymptotically minimax or consistent families of tests ψ_{R,α_R} or $\psi_{\varepsilon,\alpha_\varepsilon}$ for the case $\alpha_R \to 0$ or $\alpha_\varepsilon \to 0$.

Combining (1.10), (1.11), we obtain the asymptotic lower bounds for $p = 2$

$$\beta(V_R, \alpha_R) \geq \Phi(\sqrt{2 \log \alpha^{-1}} - \rho_R) + o(1), \tag{1.24}$$

$$\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \geq \Phi(\sqrt{2 \log \alpha^{-1}} - r_\varepsilon/\varepsilon) + o(1). \tag{1.25}$$

If $p = \infty$, then (1.10) yields, for alternatives (1.20),

$$\beta(V_R, \alpha_R) \geq \Phi(\sqrt{2 \log \alpha^{-1}} - n^{-\tau}\rho_R) + o(1), \quad n = [(R\rho_R)]^{1/(s-\tau))}], \tag{1.26}$$

where $[n]$ is the integer part of n, because of (see Remark 5.4 in Section 5.6.2 below)

$$\inf_{v \in V_R} |v| = n_R^{-\tau} \rho_R. \tag{1.27}$$

We show later that lower bounds (1.24)–(1.26) are asymptotically sharp for the case when $\alpha_R, \ \alpha_\varepsilon$ decrease fast enough.

Note that distinguishability is possible for large enough $\rho_R = \rho_R(\sigma), \ r_\varepsilon = r_\varepsilon(\sigma)$. The structure of these tests could depend on $\sigma > 0$ which is often unknown to the statistician in practice, and constructed for $\sigma = \sigma_1$ test family could provide pour distinguishability for alternative with differ $\sigma = \sigma_2 \neq \sigma_1$. Therefore we want to construct test families that provide good distinguishability for any σ from wide enough interval $\Sigma = [\sigma_0, \sigma_1]$.

This leads to adaptive setting that first was studied in [21], [22] and corresponds to alternatives of the form

$$V_R(\Sigma) = \bigcup_{\sigma \in \Sigma} V(\rho_R(\sigma), R, \sigma), \tag{1.28}$$

$$F_\varepsilon(\Sigma) = \bigcup_{\sigma \in \Sigma} F(r_\varepsilon(\sigma), H, \sigma), \tag{1.29}$$

where the sets $V(\rho_R(\sigma), R, \sigma), \ F(r_\varepsilon(\sigma), H, \sigma)$ are defined by (1.19), (1.16) with the radii $\rho_R = \rho_R(\sigma), \ r_\varepsilon = r_\varepsilon(\sigma)$ depending on $\sigma \in \Sigma = [\sigma_0, \sigma_1]$.

1.3 Structure of the paper

The paper is structured as follows.

In Section 2 we recall known results on distinguishability conditions for alternatives (1.16), (1.17), (1.29) and on sharp asymptotics for alternatives (1.19), (1.20), (1.28) for fixed $\alpha \in (0,1)$. Mainly these results are contained in [13]. For alternatives (1.19), the sharp asymptotics are presented in terms of solution of specific extreme problem in the sequence space. Also we recall some results from [4], [5], [16] under probability density model. These results show that analogous to (1.25) lower bounds are attained for alternatives analogous to (1.16) with $\alpha = \alpha_\varepsilon$ and $H = H_\varepsilon$ small enough.

In Section 3 we present the main results for $\alpha \to 0$. We show that the quantities

$$\alpha_\varepsilon^* = \exp\left(-\varepsilon^{-2/(2\sigma+1)}\right), \quad \alpha_R^* = \exp\left(-R^{2/(2\sigma+1)}\right)$$

are critical rates for $\alpha_\varepsilon, \alpha_R$ for alternatives (1.16), (1.19) in following sense. If α_R is not too small, namely $\log\alpha_R^{-1} \ll \log(\alpha_R^*)^{-1}$, then we obtain sharp asymptotics for alternatives (1.19) in terms of solution of specific extreme problem in the sequence space that is somewhat different from the noted above. If α_R is small enough, namely $\log\alpha_R^{-1} \gg \log(\alpha_R^*)^{-1}$, then we obtain sharp rates of testing $\rho_R^* = \sqrt{2\log\alpha_R^{-1}}$ that correspond to the lower bounds (1.24). Moreover if α_R is very small, namely, if $\sigma > 1/2, \quad (\log\alpha_R^{-1})^{2\sigma+1}/(\log\log\alpha_R^{-1})^{4\sigma} \gg R^4$, then we show that the lower bounds (1.24) are attained. For alternatives (1.16), we obtain analogous rate asymtptotics. For very small α_ε, we show that the lower bounds (1.25) are sharp.

For adaptive problems, we obtain sharp asymptotics for alternatives (1.28) rates asymptotics for alternatives (1.29).

For alternatives (1.20) and any $\alpha_R \to 0$, we obtain general formula for sharp asymptotics. This yields various corollaries on the rates. For alternatives (1.17) and $\alpha_\varepsilon \to 0$, we obtain analogous rate relations. Roughly, critical rates in α_R, α_ε correspond to $\alpha_R^* = 1/R$, $\alpha_\varepsilon^* = \varepsilon$. Note that the upper bounds are provided by a families of tests that does not depend on parameters determined alternatives (1.20) or (1.17). Therefore we have no any problems on adaptation for these cases.

In Section 4 we formulate some properties of solutions of the extreme problems noted above.

In Section 5 we give the proofs of theorems.

2 Previous results

Let either $\alpha_R = \alpha_\varepsilon \in (0,1)$ be fixed or α_R, α_ε be bounded away from 0 and from 1. Then the problem under consideration was studied intensively, see [13] and references in this book. Note that for $p = q = 2$, $\sigma > 0$ the distinguishability and non-distinguishability conditions were obtained in [6]; in [1] the sharp asymptotics of the quantities $\beta(V(\rho_R, R, \sigma), \alpha)$, $\beta_\varepsilon(\tilde F(r_\varepsilon, H, \sigma), \alpha)$ have been studied. The case $p = \infty$ was studied in [8], [12], [13].

2.1 Distinguishability conditions

Under the functional white Gaussian noise model let us consider the sets $F_\varepsilon = F(r_\varepsilon, H, \sigma)$ determined by (1.16) and let $\varepsilon \to 0$. Introduce the rates

$$r_\varepsilon^* = \varepsilon^{4\sigma/(4\sigma+1)}. \tag{2.1}$$

Then we have distinguishability conditions of the form

$$\beta_\varepsilon(F_\varepsilon, \alpha) \to 1 - \alpha \quad \text{iff} \quad r_\varepsilon / r_\varepsilon^* \to 0; \tag{2.2}$$

$$\beta_\varepsilon(F_\varepsilon, \alpha) \to 0 \quad \text{iff} \quad r_\varepsilon / r_\varepsilon^* \to \infty. \tag{2.3}$$

Moreover, for $r_\varepsilon / r_\varepsilon^* \to \infty$ let us take integer-valued family

$$m = m_\varepsilon \asymp (r^*)^{-1/\sigma} = \varepsilon^{-4/(4\sigma+1)},$$

and consider equispaced partition of the interval $(0, 1]$ into m sub-intervals

$$\delta_{j,m} = (a_{j-1,m}, a_{j,m}]; \quad a_{j,m} = j/m, \quad j = 1, ..., m.$$

Let us take normalized increments of the observing process $X_\varepsilon(t)$ in the sub-intervals

$$X_{j,m} = \varepsilon^{-1} m^{1/2} (X_\varepsilon(a_{j,m}) - X_\varepsilon(a_{j-1,m})) \tag{2.4}$$

and consider χ^2-tests based on the statistics (2.4):

$$\psi_{\varepsilon,\alpha} = \mathbf{1}_{t_{m,\varepsilon} > T_\alpha}, \quad t_{m,\varepsilon} = (2m)^{-1/2} \sum_{1 \le j \le m} (X_{j,m}^2 - 1). \tag{2.5}$$

Then $\alpha_\varepsilon(\psi_{\varepsilon,\alpha}) \le \alpha + o(1)$, and $\beta_\varepsilon(\psi_{\varepsilon,\alpha}, F_\varepsilon) \to 0$, as $r_\varepsilon / r_\varepsilon^* \to \infty$ (see [8], [13], Theorem 3.9 (1)).

Let us consider the case $p = q = \infty$, $\sigma > 0$. Introduce the rates

$$r_{\varepsilon,\infty}^* = (\varepsilon^2 \log \varepsilon^{-1})^{\sigma/(2\sigma+1)}. \tag{2.6}$$

Then distinguishability conditions are of the following form (see [8], [13], Theorem 3.9 (2)): there exist constants $0 < C_1 < C_2 < \infty$ such that for all $\alpha \in (0, 1)$ one has

$$\beta_\varepsilon(F_\varepsilon, \alpha) \to 1 - \alpha \quad \text{if} \quad \limsup r_\varepsilon / r_{\varepsilon,\infty}^* < C_0; \tag{2.7}$$

$$\beta_\varepsilon(F_\varepsilon, \alpha) \to 0 \quad \text{if} \quad \liminf r_\varepsilon / r_{\varepsilon,\infty}^* > C_1. \tag{2.8}$$

Moreover, if $\liminf r_\varepsilon / r_{\varepsilon,\infty}^* > C_1$ for large enough C_1, then the relations

$$\alpha_\varepsilon(\psi_\varepsilon) \to 0, \quad \beta_\varepsilon(\psi_\varepsilon, F_\varepsilon) \to 0$$

are provided by the tests based on the thresholding of statistics (2.4):

$$\psi_{\varepsilon,\alpha} = \mathbf{1}_{\mathcal{X}_\varepsilon}, \quad \mathcal{X}_\varepsilon = \{\max_{1 \le j \le m} |X_{j,m}| > \sqrt{2 \log m}\}, \quad m = m_\varepsilon \asymp (r_{\varepsilon,\infty}^*)^{-1/\sigma}. \tag{2.9}$$

For any $\sigma_0,\ \sigma_1,\ 0 < \sigma_0 < \sigma_1 < \infty$, these relations are uniform over $\sigma \in [\sigma_0, \sigma_1]$.

It was shown in [19] that $C_0 = C_1$ under some additional assumptions.

It follows from [13], Theorem 4.7 that the distinguishability conditions (2.7)–(2.8) hold true for $\infty = p > q > 1/\sigma$ with the change the rates (2.6) by

$$r^*_{\varepsilon,\infty} = r^*_{\varepsilon,\infty}(\eta) = (\varepsilon^2 \log \varepsilon^{-1})^{\eta/(2\eta+1)}, \quad \eta = \sigma - 1/q > 0. \tag{2.10}$$

The quantities $m_\varepsilon = m_\varepsilon(\sigma)$ and statistics X_{j,m_ε} depend on σ in (2.9). However there is presented in [13], Section 4.4.4 a test procedure that provides analogous properties and does not depend on σ. This test procedure is based on wavelet transform. Here we give somewhat other test procedure providing analogous properties.

Let us take a family of collections

$$m_l = 2^l,\ J_{\varepsilon,0} \le l \le J_{\varepsilon,1},\quad J_{\varepsilon,0} \asymp (\log \varepsilon^{-1})/\log\log \varepsilon^{-1},\quad J_{\varepsilon,1} \asymp (\log \varepsilon^{-1}) \log\log \varepsilon^{-1} \tag{2.11}$$

and corresponding collections of equispaced partitions of $[0,1]$ into m_l sub-intervals.

Theorem 2.1 *Taking collections (2.11) and corresponding partitions, let us consider family of tests* $\psi^*_\varepsilon = \mathbf{1}_{\mathcal{X}_\varepsilon}$, *where*

$$\mathcal{X}_\varepsilon = \{ \max_{J_{\varepsilon,0} \le l \le J_{\varepsilon,1}} \max_{1 \le j \le m_l} |X_{j,m_l}|/T_l > 1\},\ T_l^2 = 2(cl + \log l),\ c = \log 2,$$

and statistics $X_{j,m}$ *are determined by (2.4). Then* $\alpha_\varepsilon(\psi^*_\varepsilon) \to 0$ *and there exists* $C_1 > 0$ *such that if* $\liminf r_\varepsilon / r^*_{\varepsilon,\infty} > C_1$, *then* $\beta_\varepsilon(\psi^*_\varepsilon, F_\varepsilon) \to 0$, *where* $F_\varepsilon = F(r_\varepsilon, H, \sigma, q)$, $\sigma - 1/q > 0$ *and the rates are defined by (2.10). For any* $\eta_0 > 0,\ \sigma_0,\ \sigma_1,\ 0 < \sigma_0 < \sigma_1 < \infty$, *this relation is uniform over* (σ, q) *such that* $\sigma \in [\sigma_0, \sigma_1],\ \sigma - 1/q > \eta_0$.

Proof of Theorem 2.1 is given in Section 5.9.

So, for any fixed $\alpha \in (0,1)$, the rates $r^*_\varepsilon,\ r^*_{\varepsilon,\infty}$ of form (2.6), (2.10) do not depend on α in the distinguishability conditions (2.7), (2.8). These rates are much slowly than "classical" rates $r^*_\varepsilon = \varepsilon$ corresponding to "known" signal. Therefore, for nonparametric alternatives of form (1.16), (1.17) and for any $\alpha \in (0,1)$, minimax efficiency of testing is much smaller then efficiency of testing for "known" signal.

These results are extended to the probability density model corresponding to i.i.d. sample $X_1, ..., X_N$ with unknown probability density $p(x)$. Let $X_i \in [0,1]$ and $p(x)$ be a probability density on the interval $[0,1]$ with respect to the Lebesgue measure. We test the null-hypothesis $H_0 : p \equiv 1$ against the alternative $H_1 : p = 1 + g,\ g \in G_N$ where the set $G_N = G_N(r_N, H, \sigma)$ consists of the functions g such that

$$\|g\|_2 \ge r_N, \quad \|g\|_{\sigma,2} \le H, \quad (g,1) = 0, \quad \inf_{x \in [0,1]} g(x) \ge -1.$$

This problem is analogous to hypothesis testing problem under the functional white Gaussian noise model with $\varepsilon = N^{-1/2}$ and the results are the same: for any $\alpha \in (0,1)$ the distinguishability conditions, as $N \to \infty$, are analogous to (2.2)–(2.3) with the change r^*_ε of form (2.1) by $r^*_N = N^{-2\sigma/(4\sigma+1)}$ (see [7]).

Extensions of these results for other p, q are given in [12], [13].

2.2 Adaptive rates

Tests (2.5) are determined by integers $m = m_\varepsilon$ and statistics X_{j,m_ε} that depend on σ for $p = q = 2$ and we cannot provide good properties of these tests when the parameter σ is unknown. These lead to adaptive problem which first have been studied in [21], [22] under the functional white Gaussian noise model (1.2) for $p = q = 2$.

Suppose an interval $\Sigma = [\sigma_0, \sigma_1],\ 0 < \sigma_0 < \sigma_1 < \infty$ be given. Taking a family of the functions $r_\varepsilon(\sigma),\ \sigma \in \Sigma$, let us consider alternatives of form (1.29). Let $r^*_\varepsilon(\sigma)$ be the rates determined by (2.1) and set

$$H_\varepsilon(\Sigma) = \inf_{\sigma\in\Sigma} r_\varepsilon(\sigma)/r^*_\varepsilon(\sigma).$$

It follows from (1.4) and results above that the relations $H_\varepsilon(\Sigma) \to \infty$ are necessary in order to obtain distinguishability for alternative (1.29). The problem is: are these conditions sufficient for distinguishability?

The answer is "no" for $p = q = 2$. Namely, introduce *adaptive rates functions*:

$$r^{ad}_\varepsilon(\sigma) = (\varepsilon^4 \log\log \varepsilon^{-1})^{\sigma/(4\sigma+1)}. \tag{2.12}$$

It was shown in [21] that there exist constants $0 < D_1 < D_2 < \infty$ such that

1. *If there exists an interval $\Delta \subset \Sigma$ of positive length such that*

$$\limsup \sup_{\sigma\in\Delta} r_\varepsilon(\sigma)/r^{ad}_\varepsilon(\sigma) < D_1, \tag{2.13}$$

then $\beta_\varepsilon(F_\varepsilon(\Sigma), \alpha) \to 1 - \alpha$.

2. *If*

$$\liminf \inf_{\sigma\in\Sigma} r_\varepsilon(\sigma)/r^{ad}_\varepsilon(\sigma) > D_2, \tag{2.14}$$

then $\beta_\varepsilon(F_\varepsilon(\Sigma), \alpha) \to 0$ for any $\alpha \in (0, 1)$.

Moreover, one can use "Bonferroni method" to construct "adaptive" test procedure. Let us take family of collections (2.11) and corresponding collections of equispaced partitions of $[0, 1]$ into $m_l = 2^l$ sub-intervals. Taking thresholds $C_\varepsilon = 2\sqrt{\log\log \varepsilon^{-1}}$, let us combine test procedures (2.5) for collections m_l, i.e., we set

$$\psi^{ad}_\varepsilon = \mathbf{1}_{\mathcal{X}_\varepsilon},\ \ \mathcal{X}_\varepsilon = \{\max_{J_{\varepsilon,0}\le l\le J_{\varepsilon,1}} t_{m_l,\varepsilon} > C_\varepsilon\}, \tag{2.15}$$

where the statistics $t_{m,\varepsilon}$ are determined by (2.5). Then $\alpha_\varepsilon(\psi^{ad}_\varepsilon) \to 0$ and under the assumption (2.14) with large enough constant D_2, one gets $\beta_\varepsilon(\psi^{ad}_\varepsilon, F_\varepsilon(\Sigma)) \to 0$.

Under the probability density model analogous results were obtained in [10].

So, the lack of knowledge of parameter σ for $p = q = 2$ leads to losses in the rates of testing in $(\log\log \varepsilon^{-1})$-factor. It is the payment for adaptation in the problem (see [21]). It was shown in [11], [13] that one has the same effects for wide class of adaptive problems with asymptotics of Gaussian type.

2.3 Sharp asymptotics

2.3.1 Gaussian asymptotics: $p = q = 2$

This problem has been studied for the alternatives (1.19) under the sequence model (1.1). Let us consider extreme problem

$$u^2(\rho_R, R) = \inf \frac{1}{2}\sum_{i=1}^{\infty} z_i^4 \quad \text{subject to} \tag{2.16}$$

$$\sum_{i=1}^{\infty} z_i^2 \geq \rho_R^2, \quad \sum_{i=1}^{\infty} z_i^2 i^{2\sigma} \leq R^2. \tag{2.17}$$

Let $R \to \infty$. Then, for any $\alpha \in (0,1)$ and any family $\rho_R > 0$, one has

$$\beta(V_R, \alpha) = \Phi(T_\alpha - u(\rho_R, R)) + o(1) \tag{2.18}$$

(see [1], [8]). Moreover, these relations are provided by test families $\psi_{R,\alpha} = \mathbf{1}_{t_R > T_\alpha}$ based on the statistics

$$t_R = \sum_{i=1}^{\infty} w_i(X_i^2 - 1), \quad w_i = z_{i,R}^2 / u(\rho_R, R), \tag{2.19}$$

where $\{z_i\}$ is the extreme sequence and $u^2(\rho, R)$ is the extreme value for the problem (2.16), (2.17); this yields $\sum_i w_{i,R}^2 = 2$.

Some properties of the extreme values $u^2(\rho_R, R)$ and extreme sequence $\{z_i\}$ are given in Section 4.1 (also see [13], Section 4.3.3). In particular for $\rho_R = o(R)$, we have

$$u(\rho_R, R) \sim D_0(\sigma) \rho_R^{2+1/2\sigma} R^{-1/2\sigma}, \tag{2.20}$$

where $D_0(\sigma)$ is a positive continuous function. This yields distinguishability conditions

$$\beta(V_R, \alpha) \to 1 - \alpha \quad \text{iff} \quad \rho_R/\rho_R^* \to 0, \tag{2.21}$$

$$\beta(V_R, \alpha) \to 0 \quad \text{iff} \quad \rho_R/\rho_R^* \to \infty, \tag{2.22}$$

with the rates

$$\rho_R^* = R^{1/(4\sigma+1)}. \tag{2.23}$$

Under the white Gaussian noise model for the sets $F_\varepsilon = \tilde{F}(r_\varepsilon, H, \sigma)$ of form (1.22) and $\varepsilon \to 0$ these yield the results of the form

$$\beta(F_\varepsilon, \alpha) = \Phi(T_\alpha - \tilde{u}) + o(1), \quad \tilde{u} = D_0(\sigma)(r_\varepsilon/H)^{2+1/2\sigma}(\varepsilon/H)^{-2}. \tag{2.24}$$

The results above are extended the probability density model as well (see [2], [8], [9]). The extension of these results for other p, q, τ are given in [12], [13].

2.3.2 Degenerate asymptotics: $p = \infty$

For the alternative $V_R = V(\rho_R, R, \tau, s, q)$ of form (1.20) we have different type of asymptotics (see [8], [12], [13], Theorem 4.5):

$$\beta(V_R, \alpha) = (1-\alpha)\Phi(\sqrt{2\log m_R} - m_R^{-\tau}\rho_R) + o(1), \quad m_R = (R/\rho_R)^{1/(s-\tau)}; \quad (2.25)$$

moreover one can change $\sqrt{2\log m_R}$ by $\sqrt{2s^{-1}\log R}$ in (2.25). This yields distinguishability conditions

$$\beta(V_R, \alpha) \to 1-\alpha \quad \text{if} \quad \limsup \rho_R/\rho_R^* < 1, \quad (2.26)$$
$$\beta(V_R, \alpha) \to 0 \quad \text{if} \quad \liminf \rho_R/\rho_R^* > 1, \quad (2.27)$$

with the rates ρ_R^* defined by the relation

$$\rho_R^* = \Lambda R^{\tau/s}(\log R)^{(s-\tau)/2s}, \quad \Lambda = (2/s)^{(s-\tau)/2s}, \quad (2.28)$$

see [13], Section 4.4.2. Under (1.23) these rates correspond to the rates (2.10).

Moreover, let us consider the randomized tests of the form

$$\psi_{R,\alpha} = \alpha + (1-\alpha)\psi_R,$$

where non-randomized tests $\psi_R = \mathbf{1}_{\mathcal{X}_R}$ are based on thresholding:

$$\mathcal{X}_R = \{X : \sup_i |X_i|/T_{R,i} > 1\}, \quad T_{R,i}^2 = 2\begin{cases} \log N_R, & i \le N_R, \\ \log i + \log\log i, & i > N_R, \end{cases} \quad (2.29)$$

here one can take any $N_R \asymp \log\log R$. Then these tests are asymptotically minimax in the problem, i.e.,

$$\alpha(\psi_{R,\alpha}) \le \alpha + o(1), \quad \beta(\psi_{R,\alpha}, V_R) \le (1-\alpha)\Phi(\sqrt{2\log m_R} - m_R^{-\tau}\rho_R) + o(1).$$

For any $s_0,\ s_1,\ \delta > 0, 0 < \delta < s_0,\ s_0 < s_1 < \infty$, the letter relation is uniform over (s,τ) such that $s \in [s_0, s_1],\ 0 \le \tau \le s - \delta$. Note that these tests do not depend on $s > \tau \ge 0,\ q > 0$.

2.4 Sharp adaptive asymptotics

Sharp adaptive asymptotics were studied in [11], [13] under the sequence model (1.1) for a wide class of alternatives. Taking family of function $\rho_R(\sigma),\ \sigma \in \Sigma$, we consider alternatives $V_R(\Sigma)$ of form (1.28). The results of [11], [13], Section 7.1.3, are of the following form. Let $u^2(\rho, R, \sigma)$ be defined by (2.16), (2.17) and set

$$u_R(\Sigma) = \inf_{\sigma\in\Sigma} u(\rho_R(\sigma), R, \sigma),\ \ H_R(\Sigma) = u_R(\Sigma) - \sqrt{2\log\log R}. \quad (2.30)$$

Then for any $\alpha \in (0,1)$ one has:

1. *Upper bounds:*

$$\beta(V_R(\Sigma), \alpha) \le (1-\alpha)\Phi(-H_R(\Sigma)) + o(1). \quad (2.31)$$

2. *Assume the minimum in (2.30) is "essential", i.e., for any* $\delta > 0$ *there exists nontrivial sub-interval* $\Sigma_0 \subset \Sigma$ *such that*

$$\sup_{\sigma\in\Sigma_0} u(\rho_R(\sigma), R, \sigma) \le u_R(\Sigma) + \delta.$$

Then

$$\beta(V_R(\Sigma), \alpha) \ge \beta(V_R(\Sigma_0), \alpha) \ge (1-\alpha)\Phi(-H_R(\Sigma)) + o(1). \tag{2.32}$$

Moreover, it follows from [13], Section 7.3 that one can use the following construction for tests family $\psi^{ad}_{R,\alpha}$ that provides the upper bounds (2.31). Let $\rho^*_R(\sigma)$ be the quantities such that

$$u_R(\rho^*_R(\sigma), R, \sigma) = u_R(\Sigma) + o(1).$$

Let $\bar z_R(\rho,\sigma) = \{z^4_{i,R}(\rho,\sigma)\}$ be extreme sequence in the problem (2.16), (2.17), i.e.,

$$\frac{1}{2}\sum_i z^4_{i,R}(\rho,\sigma) = u^2_R(\rho,\sigma),$$

and $t_{R,\bar z} = t_{R,\bar z}(X)$ be the statistics determined by (2.19). Let us divide the interval Σ into $M = M_R \asymp (\log R)(\log\log R)^B, B > 1$ sub-intervals $\delta_{R,l} = [\sigma_{R,l-1}, \sigma_{R,l}],\ 1 \le l \le M$ of the length $|\delta_{R,l}| \asymp M^{-1}$ and consider the collections of sequences

$$\bar z_{R,l} = \bar z_R(\rho^*_R(\sigma_{R,l}), \sigma_{R,l-1}),\ l = 1, ..., M_\varepsilon$$

and collections of statistics $t_{R,l} = t_{R,\bar z_{R,l}}$. Set $\psi^{ad}_{R,\alpha} = \alpha + (1-\alpha)\psi^{ad}_R$, where

$$\psi^{ad}_R = \mathbf{1}_{\mathcal{X}_R},\ \mathcal{X}_R = \{X:\ \max_{1\le l\le M_\varepsilon} t_{R,l} > \sqrt{2\log M_\varepsilon}\}.$$

Then

$$\alpha(\psi^{ad}_{R,\alpha}) \le \alpha + o(1), \quad \beta(\psi^{ad}_{R,\alpha}, V_R(\Sigma)) \le (1-\alpha)\Phi(-H_R(\Sigma)) + o(1).$$

2.5 Intermediate efficiency

Let us consider probability density model. Observing i.i.d. sample $X_1, ..., X_N,\ X_i \in [0,1]$ of size $N,\ N \to \infty$ we test the simple hull-hypothesis on the uniformity of a density on the interval $[0,1]$:

$$H_0\ :\ p(x) \equiv 1,\ x \in [0,1],$$

against simple alternative

$$H_1\ :\ p(x) = p_N(x),\ x \in [0,1],$$

that corresponds to a given sequence of densities $p_N(x)$ on the interval $[0,1]$.

Let $p_N(x) = p(x)$ be fixed and a density $p(x)$ be bounded and bounded away from 0, as $N \to \infty$. Let the type I error probability $\alpha = \alpha_N$ be such that the type

II error probability $\beta_N(p, \alpha_N)$ is bounded away from 0 and from 1, as $N \to \infty$. Then it is well known (see [20]) that the logarithmic rate of α_N is determined by the Kullback-Leibler distance between $p_0(x) \equiv 1$ and $p(x)$, namely

$$N^{-1} \log \alpha_N^{-1} \to K(p) = \int_0^1 p(x) \log p(x) dx, \quad \text{as } N \to \infty.$$

This case corresponds to Bahadur's efficiency [20].

From the other hand, let us consider local alternatives of the form

$$H_1 : p(x) = p_N(x) = 1 + \eta N^{-1/2} g(x),$$

where $\eta > 0$, $\|g\|_2 = 1$, $(g, 1) = 0$, $\|g\|_\infty < \infty$. Then, analogously to (1.5), for any $\alpha \in (0, 1)$ one has

$$\beta_N(p_N, \alpha) = \Phi(T_\alpha - \eta) + o(1), \quad \text{as } N \to \infty. \tag{2.33}$$

This case corresponds to Pitman's efficiency [20].

Intermediate efficiency was introduced in [15]. This corresponds to $\alpha_N \to 0$ and alternatives of the form

$$H_1 : p(x) = p_N(x), \quad p_N(x) = 1 + \eta N^{-\zeta} g(x), \quad 0 < \zeta < 1/2. \tag{2.34}$$

It was shown in [4], [5], [16] that the relation analogous to (2.33) holds true for alternatives of form (2.34) with $1/4 < \zeta < 1/2$ and for $\alpha_N \to 0$, $\log \alpha_N^{-1} = o(N)$:

$$\beta_N(p_N, \alpha_N) = \Phi(\sqrt{2 \log \alpha_N^{-1}} - \eta N^{1/2 - \zeta}) + o(1), \quad \text{as } N \to \infty.$$

Moreover, it was shown in these papers that if the function g belongs to the Sobolev space W_2^1, then this relation is provided by Neyman's tests based on the first $n = n_N$ Legendre polynomials (these tests are analogous to χ^2-tests under the Gaussian model) for some sequences $n = n_N$. Various data-driven versions of these tests are studied in these papers. They corresponds to random $n = n(X_1, ..., X_N)$. It is easily seen that the estimations in these papers are uniform over $g \in S_{1,2}^0(H)$ for any Sobolev ball

$$S_{1,2}^0(H) = \{f \in W_2^1, \ \|f\|_{1,2} \le H, \ (f, 1) = 0\}.$$

Set $r_N = \eta N^{-\zeta}$, $H_N = H N^{-\zeta}$, $H > \eta$, $\sigma = 1$,

$$P_N = \{p(x) = 1 + g(x), \ x \in [0, 1], \ \|g\|_2 \ge r_N, \ \|g\|_{\sigma,2} \le H_N, \ (g, 1) = 0\}, \tag{2.35}$$

Then for $1/4 < \zeta < 1/2$, we have the relation

$$\beta_N(P_N, \alpha_N) = \Phi(\sqrt{2 \log \alpha_N^{-1}} - r_N N^{1/2}) + o(1). \tag{2.36}$$

Under the white Gaussian noise model (1.2) with $\varepsilon = N^{-1/2}$ the set (2.35) corresponds to the set $F(r_\varepsilon, H_\varepsilon, 1)$ with

$$r_\varepsilon = \eta \varepsilon^{2\zeta} \to 0, \quad H_\varepsilon = H \varepsilon^{2\zeta} \to 0.$$

Under the sequence model (1.1) this set corresponds to the set $V_R = V(\rho, R, 1)$ with

$$\rho = \eta \varepsilon^{2\zeta - 1} \to \infty, \quad R = H \varepsilon^{2\zeta - 1} \to \infty, \quad 1 < R/\rho \asymp 1.$$

Relation (2.36) corresponds to the equality in inequalities (1.24), (1.25).

3 Main results

We consider the alternatives V_R of form (1.19), (1.20) under the sequence model (1.1). We are interesting in the study of $\rho = \rho_R(\alpha_R)$ such that the family $\beta(V_R, \alpha_R)$ is bounded away from 1 and from 0.

3.1 Sharp and near to sharp asymptotics under the sequence model

3.1.1 The case $p = q = 2$ for not too small α

Let us consider the case $p = q = 2$.

The problem with $\alpha_R \to 0,\ \ R \to \infty$ is of interest for $\rho_R/\rho_R^* \to \infty,\ \ \rho_R \leq R$, where ρ_R^* are defined by (2.23).

First, suppose $\alpha_R \to 0$ but not too fast; namely we assume

$$\alpha_R \to 0, \quad \log \alpha_R^{-1} = O(R^{2/(2\sigma+1)}). \tag{3.1}$$

According to the results of Section 2.1 for $\alpha_R \to 0$, we are interested in the case when distinguishability conditions hold true, however ρ_R increase not too fast:

$$\rho_R R^{-1/(4\sigma+1)} \to \infty, \quad \rho_R = O(R^{1/(2\sigma+1)}). \tag{3.2}$$

Let $\bar{z} = \{z_i\}_{i=1}^{\infty}$ be a nonnegative sequence and

$$P^{\bar{z}} = \prod_i \mathcal{N}(0, z_i^2 + 1)$$

be the Gaussian measure corresponding to independent sequence $X_i \sim \mathcal{N}(0, z_i^2 + 1)$. Let

$$K^2(\bar{z}) = E_{P^{\bar{z}}} \log(dP^{\bar{z}}/dP_0) = \frac{1}{2}\sum_{i=1}^{\infty}(z_i^2 - \log(1 + z_i^2)) \tag{3.3}$$

be the Kullback-Leibler distance between the measures $P^{\bar{z}}$ and P_0. Let us consider extreme problem

$$K^2(\rho, R, \sigma) \quad = \quad K^2(\rho, R) = \inf K^2(\bar{z}) \quad \text{subject to} \tag{3.4}$$

$$\sum_{i=1}^{\infty} z_i^2 \geq \rho^2, \quad \sum_{i=1}^{\infty} z_i^2 i^{2\sigma} \leq R^2 \tag{3.5}$$

(if σ is assumed to be fixed, we omit σ in notations later). Note that, in terms of variables $u_i = z_i^2 \geq 0$, the function $K^2(\bar{z})$ is strictly convex and the set (3.5) is convex. This implies uniqueness of extreme sequence $\bar{z} = \{z_i\},\ z_i \geq 0$.

There are given Section 4.2 some properties of solution of the extreme value $K^2(\rho, R, \sigma)$ and extreme sequence $\bar{z} = \bar{z}_R(\rho_R, \sigma)$ in problem (3.4), (3.5). In particular, we have

$$K(\rho_R, R) \asymp u(\rho_R, R) \asymp \rho_R^{2+1/2\sigma} R^{-1/2\sigma}. \tag{3.6}$$

Remark 3.1 The function $K^2(\rho, R, \sigma)$ is convex in variables (ρ^2, R^2) (see Proposition 2.8 in [13]). Jointed with (3.6), this implies that the function

$$f_R(b_1, b_2) = K^2(b_1\rho_R, b_2 R, \sigma)/K^2(\rho_R, R, \sigma)$$

is uniformly Lipschitzian in $(b_1, b_2) \in D$ over any compact $D \subset R_+^2 = \{b_1 > 0, b_2 > 0\}$ (see [13], Lemma 5.1 in the proof of Proposition 5.6, (4)). In particular, there exists $C = C(D) > 0$ such that, for R large enough,

$$|K^2(b_1\rho_R, b_2 R, \sigma) - K^2(\rho_R, R, \sigma)| \le CK^2(\rho_R, R, \sigma)(|b_1 - 1| + |b_2 - 1|) \quad \forall\ (b_1, b_2) \in D. \tag{3.7}$$

Moreover, the relation (3.7) is uniform over $\sigma \in [\sigma_0, \sigma_1]$, $0 < \sigma_0 < \sigma_1$ such that (3.2) holds for $\sigma = \sigma_1$.

Let us slightly improve the assumptions (3.2), (3.1)

$$\rho_R R^{-1/(4\sigma+1)} \to \infty, \quad \rho_R = o(R^{1/(2\sigma+1)}); \tag{3.8}$$

$$\alpha_R \to 0, \quad \log \alpha_R^{-1} = o(R^{2/(2\sigma+1)}). \tag{3.9}$$

Under (3.8) we have

$$K(\rho_R, R) \sim u(\rho_R, R)/\sqrt{2}. \tag{3.10}$$

Let $\bar{z} = \{z_i\}$ be extreme sequence in the problem (3.4), (3.5). Consider tests of the form

$$\psi_{\alpha_R} = \mathbf{1}_{\{t_R > \log \alpha_R^{-1}\}}, \quad t_R = t_{R,\bar{z}} = \frac{1}{2} \sum_{1 \le i \le n} \left(\frac{z_i^2 X_i^2}{1 + z_i^2} - \log(1 + z_i^2) \right). \tag{3.11}$$

Note that $t_R = \log(dP^{\bar{z}}/dP_0)$ for the extreme sequence $\{z_i\}$.

For tests (3.11) one has

$$\alpha(\psi_{\alpha_R}) \le \alpha_R, \quad \forall\ \alpha_R \in (0, 1). \tag{3.12}$$

Proof of relation (3.12) is given in Section 5.1, Lemma 5.2 (1).

Theorem 3.1 *Under assumptions (3.8) one has:*

(1) Lower bounds.

$$\beta(V_R, \alpha_R) \ge \Phi(\sqrt{2 \log \alpha_R^{-1}} - \sqrt{2} K(\rho_R, R)) + o(1). \tag{3.13}$$

(2) Upper bounds. For the tests (3.11),

$$\beta(\psi_{\alpha_R}, V_R) \le \Phi(\sqrt{2 \log \alpha_R^{-1}} - \sqrt{2} K(\rho_R, R)) + o(1). \tag{3.14}$$

The relations (3.13), (3.14) yield

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2 \log \alpha_R^{-1}} - \sqrt{2} K(\rho_R, R)) + o(1). \tag{3.15}$$

For any $0 < \sigma_0 < \sigma_1 < \infty$, this relation is uniform over $\sigma \in [\sigma_0, \sigma_1]$, such that (3.8) holds with $\sigma = \sigma_1$.

Proof of Theorem 3.1 is given in Sections 5.1, 5.3.

Remark 3.2 It suffices to prove Theorem 3.1 for the case

$$\sqrt{\log \alpha_R^{-1}} = K(\rho_R, R) + O(1). \tag{3.16}$$

In fact, the lower bounds (3.13) are trivial for $K(\rho_R, R) - \sqrt{\log \alpha_R^{-1}} \to \infty$. If $K(\rho_R, R) - \sqrt{\log \alpha_R^{-1}} \to -\infty$, then the lower bounds (3.13) correspond to $\beta(V_R, \alpha_R) \to 1$. If we will prove (3.13) for the case (3.16), then, by passing to $\tilde{\alpha}_R > \alpha_R$, we can get the case (3.16) and $\beta(V_R, \alpha_R) \ge \beta(V_R, \tilde{\alpha}_R) > 1 - \delta$ for any $\delta > 0$. This yields $\beta(V_R, \alpha_R) \to 1$. Analogously, the upper bounds (3.14) are trivial for $K(\rho_R, R) - \sqrt{\log \alpha_R^{-1}} \to -\infty$. If $K(\rho_R, R) - \sqrt{\log \alpha_R^{-1}} \to \infty$, then, by decreasing the families ρ_R to satisfy (3.16), and using the monotonicity of the function $\beta(\psi_{\alpha_R}, V(\rho, R, \sigma))$ in ρ, we easy see that if (3.14) holds true under (3.16), then (3.14) holds true everywhere.

Using Proposition 4.1 (1) and relation (4.15) (see Section 4.2 below) one can see that under (3.16), assumption (3.8) is equivalent to (3.9). For this reason we can use (3.9) in the proof and in applications of Theorem 3.1.

Analogous situation holds for other theorems below and we omit analogous remarks later.

Theorem 3.1 and relation (3.10) yield the following distinguishability conditions.

Corollary 3.1 *Under assumptions (3.8) or (3.9) one has*

$$\beta(V_R, \alpha_R) \to 0 \quad \textit{iff} \quad \sqrt{\log \alpha_R^{-1}} - K(\rho_R, R) \to -\infty,$$
$$\beta(V_R, \alpha_R) \to 1 \quad \textit{iff} \quad \sqrt{\log \alpha_R^{-1}} - K(\rho_R, R) \to \infty.$$

This yields

$$\beta(V_R, \alpha_R) \to 0, \quad \textit{if} \quad \liminf\, u(\rho, R)/\sqrt{2\log \alpha_R^{-1}} > 1,$$
$$\beta(V_R, \alpha_R) \to 1, \quad \textit{if} \quad \limsup\, u(\rho_R, R)/\sqrt{2\log \alpha_R^{-1}} < 1.$$

If the rates of the quantities ρ_R or $\log \alpha_R^{-1}$ are somewhat smaller than in Theorem 3.1, then we can extend the sharp asymptotics from Section 2.3.

Corollary 3.2 *Let* $\alpha_R \to 0$ *and*

$$\rho_R R^{-1/(4\sigma+1)} \to \infty, \quad \rho_R = o(R^{3/(8\sigma+3)}). \tag{3.17}$$

Then

$$\sqrt{2}K(\rho_R, R) = u(\rho_R, R) + o(1), \quad \beta(V_R, \alpha_R) = \Phi(\sqrt{2\log \alpha_R^{-1}} - u(\rho, R)) + o(1). \tag{3.18}$$

Proof of Corollary 3.2 is given in Section 4.2.

Let us consider the case

$$\log \alpha_R^{-1} \asymp R^{2/(2\sigma+1)}, \quad \rho_R \asymp R^{1/(2\sigma+1)}. \tag{3.19}$$

For this case we can extend the distinguishability conditions from Corollary 3.1.

Theorem 3.2 *Assume (3.19) for $\sigma \geq 1/2$ or*

$$\limsup \rho_R R^{-1/(2\sigma+1)} < C(\sigma), \quad C(\sigma) = (2\sigma/(1-2\sigma))^{1/2}(1+1/2\sigma)^{1/2(2\sigma+1)},$$

for $\sigma < 1/2$. Then one has

$$\begin{array}{lll} (1) & \beta(V_R, \alpha_R) \to 1, & \textit{if } \sqrt{\log \alpha_R^{-1}} - K(\rho_R, R) \to \infty, \\ (2) & \beta(\psi_{\alpha_R}, V_R) \to 0, & \textit{if } \sqrt{\log \alpha_R^{-1}} - K(\rho_R, R) \to -\infty. \end{array}$$

Proof of Theorem 3.2 is given in Sections 5.2, 5.4.

3.1.2 The case $p = q = 2$ for small enough α and very small α

Let us consider the case $\rho < R$ and one of two following assumptions holds true

$$R^{-2/(2\sigma+1)} \log \alpha_R^{-1} \to \infty, \tag{3.20}$$

$$\rho_R R^{-1/(2\sigma+1)} \to \infty. \tag{3.21}$$

(analogously to Remark 3.2 the assumptions (3.20) and (3.21) are equivalent for results below). For this case the lower bounds (1.24) are sharp or near to sharp. Namely for an integer-valued family $m = m_R$, let consider χ^2-tests of level α_R

$$\chi^2_{m,\alpha_R} = \mathbf{1}_{\{\chi^2_m > T_{m,\alpha_R}\}}, \quad \chi^2_m = \sum_{i=1}^m X_i^2, \tag{3.22}$$

here and later $T_{m,\alpha}$ is $(1-\alpha)$-quintile of the central chi-square distribution with m degrees of freedom.

Theorem 3.3 *(1) Assume (3.20), i.e., α_R be small enough, and let us take $m \to \infty$ such that*

$$m = o(\log \alpha_R^{-1}), \quad m^{-\sigma} R = o((\log \alpha_R^{-1})^{1/2}) \tag{3.23}$$

(this is possible under (3.20)). If $\liminf \rho_R/\sqrt{2\log \alpha_R^{-1}} > 1$, then $\beta(\chi^2_{m,\alpha_R}, V_R) \to 0$.

(2) Assume α_R be very small, namely,

$$\sigma > 1/2, \quad R^{-4}(\log \alpha_R^{-1})^{2\sigma+1}/(\log\log \alpha_R^{-1})^{4\sigma} \to \infty. \tag{3.24}$$

Let us take $m \to \infty$ such that

$$m \log\log \alpha_R^{-1} = o(\log \alpha_R^{-1})^{1/2}), \quad m^{-2\sigma} R^2 = o((\log \alpha_R^{-1})^{1/2}) \tag{3.25}$$

(this is possible under (3.24)). Then one has

$$\beta(\chi^2_{m,\alpha_R}, V_R) \le \Phi(\sqrt{2\log \alpha_R^{-1}} - \rho_R) + o(1).$$

Combining with (1.24), under (3.24) this yields the sharp asymptotics

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2\log \alpha_R^{-1}} - \rho_R) + o(1). \tag{3.26}$$

Proof of Theorem 3.3 is given in Section 5.5.

Remark 3.3 For any $0 < \sigma_0 < \sigma_1 < \infty$, the relations of Theorem 3.3 are uniform over $\sigma \in [\sigma_0, \sigma_1]$, such that (3.20) or (3.24) hold with $\sigma = \sigma_0$. Therefore taking tests family for $\sigma = \sigma_0$ that does not depend on ρ_R, we obtain distinguishability conditions of Theorem 3.3 (1) or sharp asymptotics (3.26) of Theorem 3.3 (2) uniformly over $\sigma \in [\sigma_0, \sigma_1]$.

3.2 The case $p = \infty$

Let $V_R = V(\rho_R, R, \tau, s, q)$ be alternative of form (1.20), $0 < \rho_R \le R$. Set

$$m_R = (R/\rho_R)^{1/(s-\tau)}, \quad n_R = \left[(R/\rho_R)^{1/(s-\tau)}\right] = m_R + O(1),$$

here and later $[t]$ stands for the integer part of t, i.e., this is the integer k such that $t - 1 < k \le t$.

Theorem 3.4 *Let a family $\alpha_R \to 0$ be given. Then one has*

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2(\log n_R + \log \alpha_R^{-1})} - n_R^{-\tau}\rho_R) + o(1). \tag{3.27}$$

Moreover, let us consider the tests of the form

$$\psi_{R,\alpha_R} = \mathbf{1}_{\mathcal{X}_{R,\alpha_R}}, \quad \mathcal{X}_{R,\alpha_R} = \{X : \sup |X_i|/T_{\alpha_R,i} > 1\}, \tag{3.28}$$

where

$$T^2_{\alpha_R,i} = 2(\log \alpha_R^{-1} + \log i + \log\log(i+1)).$$

Then for R large enough, one has

$$\alpha(\psi_{R,\alpha_R}) \le \alpha_R, \quad \beta(\psi_{R,\alpha_R}, V_R) \le \Phi(\sqrt{2(\log n_R + \log \alpha_R^{-1})} - n_R^{-\tau}\rho_R) + o(1). \tag{3.29}$$

For any $0 < s_0 < s_1$, $\delta > 0$, the relations above are uniform over (s, τ) such that $s \in [s_0, s_1]$ and $0 \le \tau \le s - \delta$.

Proof of Theorem 3.4 is given in Section 5.6.

Note that tests (3.28) do not depend on $s > 0$, $q > 0$. Also one can verify that if either $\tau = 0$ or $\log(\alpha_R/m_R)^{-1} \asymp m_R^{-2\tau}\rho_R^2 = o(R^{2/(2s+1)})$, then $m_R^{-\tau}\rho_R =$

$n_R^{-\tau}\rho_R + o(1)$, and we can change n_R by m_R in (3.27), (3.29). For $\rho_R = o(R)$ we have $m_R \to \infty$, $n_R \sim m_R$, and Theorem 3.4 yields distinguishability conditions

$$\beta(V_R, \alpha_R) \to 1 \quad \text{if} \quad \limsup \rho_R/\rho^*_{R,\alpha_R} > 1, \tag{3.30}$$

$$\beta(V_R, \alpha_R) \to 0 \quad \text{if} \quad \liminf \rho_R/\rho^*_{R,\alpha_R} < 1, \tag{3.31}$$

where

$$\rho^*_{R,\alpha_R} = m_R^{\tau}\sqrt{2(\log m_R + \log \alpha_R^{-1})}. \tag{3.32}$$

If $\rho_R = o(R)$ and $\log(\alpha_R/m_R)^{-1} \asymp m_R^{-2\tau}\rho_R^2$, then

$$m_R \sim R^{1/s}\left(2(\log m_R + \log \alpha_R^{-1})\right)^{-1/2s}, \quad \log m_R \sim \frac{\log R}{s} - \frac{\log\log \alpha_R^{-1}}{2s}. \tag{3.33}$$

Using (3.33) and considerations analogous to Remark 3.2 we easy obtain the following corollaries.

Corollary 3.3

(1) Let $\alpha_R \to 0$, $\quad \log \alpha_R^{-1} = o(\sqrt{\log R})$. *Then we get asymptotics analogous to (2.25) that do not depend on* α_R*:*

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2\log m_R} - m_R^{-\tau}\rho_R) + o(1) = \Phi(\sqrt{2s^{-1}\log R} - m_R^{-\tau}\rho_R) + o(1).$$

Let $\log \alpha_R^{-1} = d\sqrt{\log R}$. *Then*

$$\beta(V_R, \alpha_R) = \Phi\left(\sqrt{2s^{-1}\log R} + d\sqrt{s/2} - m_R^{-\tau}\rho_R\right) + o(1)$$

Let $\alpha_R \to 0$, $\log \alpha_R^{-1} = o(\log R)$. *Then we have distinguishability conditions (3.30), (3.31) with the rates* $\rho^*_{R,\alpha_R} = \rho^*_R$ *defined by (2.28) that do not depend on* α_R.

(2) Let $(\log R)^2 = o(\log \alpha_R^{-1})$. *Then*

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2\log \alpha_R^{-1}} - n_R^{-\tau}\rho_R) + o(1). \tag{3.34}$$

Let $(\log R)^2 = o(\log \alpha_R^{-1})$, $\log \alpha_R^{-1} = o(R^{2/(2s+1)})$. *Then*

$$\beta(V_R, \alpha_R) = \Phi(\sqrt{2\log \alpha_R^{-1}} - m_R^{-\tau}\rho_R) + o(1).$$

Let $\log \alpha_R^{-1} \sim d(\log R)^2$. *Then*

$$\beta(V_R, \alpha_R) = \Phi\left(\sqrt{2\log \alpha_R^{-1}} + \frac{1}{s\sqrt{2d}} - m_R^{-\tau}\rho_R\right) + o(1).$$

(3) Let $\log R = o(\log \alpha_R^{-1})$, $\rho_R = o(R)$. *Then we have distinguishability conditions (3.30), (3.31) with the rates* $\rho^*_{R,\alpha_R} = m_R^{\tau}\sqrt{2\log \alpha_R^{-1}}$.

(4) Let $\log \alpha_R^{-1} = d\log R$. *Then*

$$\beta(V_R, \alpha_R) = \Phi\left(\sqrt{2(d+s^{-1})\log R} - m_R^{-\tau}\rho_R\right) + o(1).$$

Let $\log \alpha_R^{-1} \sim d \log R$. *Then we have distinguishability conditions (3.30), (3.31) with the rates*

$$\rho^*_{R,\alpha_R} = \rho^*_R = \Lambda R^{\tau/s} (\log R)^{(s-\tau)/2s}, \quad \Lambda = \left(2(d+s^{-1})\right)^{(s-\tau)/2s},$$

which do not depend on α_R.

Relation (3.34) corresponds to the equality in inequality (1.26).

3.3 Distinguishability conditions for the functional model

Let us consider alternatives (1.16), (1.17) under model (1.2). We assume $\log \alpha_\varepsilon^{-1} = o(\varepsilon^{-2})$ below.

3.3.1 The case $p = q = 2$

Set

$$\delta_\varepsilon = \delta_\varepsilon(\sigma) = \varepsilon^{2/(2\sigma+1)} \log \alpha_\varepsilon^{-1}. \tag{3.35}$$

Introduce critical rates which depend on α_ε:

$$r^*_{\varepsilon,\alpha_\varepsilon} = r^*_{\varepsilon,\alpha_\varepsilon}(\sigma) = (\varepsilon^4 \log \alpha_\varepsilon^{-1})^{\sigma/(4\sigma+1)}. \tag{3.36}$$

Observe that if $\delta_\varepsilon = O(1)$, then $r^*_{\varepsilon,\alpha_\varepsilon} \to 0$, if $\delta_\varepsilon = o(1)$, then

$$r^*_{\varepsilon,\alpha_\varepsilon} = o(\varepsilon^{2\sigma/(2\sigma+1)}), \tag{3.37}$$

and if $\delta_\varepsilon \asymp 1$, then we can take

$$r^*_{\varepsilon,\alpha_\varepsilon} = \varepsilon(\log \alpha_\varepsilon^{-1})^{1/2}. \tag{3.38}$$

Let us take an integer-valued family $m = m_\varepsilon \to \infty$ and consider equispaced partition of the interval $(0,1]$ into m_ε sub-intervals

$$\delta_{j,m} = (a_{j-1,m}, a_{j,m}]; \ a_{j,m} = j/m, \ j = 1, ..., m.$$

Let be $X_{j,m}$ are normalized increments of the observing process in the sub-intervals

$$X_{j,m} = \varepsilon^{-1} m^{1/2} (X_\varepsilon(a_{j,\varepsilon}) - X_\varepsilon(a_{j-1,\varepsilon})), \ j = 1, ..., m. \tag{3.39}$$

The random variables $X_{j,m}$ are independent standard Gaussian under H_0. Let us consider $\chi^2_{m_\varepsilon,\alpha_\varepsilon}$ based on statistics (3.39):

$$\chi^2_{m,\alpha} = \mathbf{1}_{\{\chi^2_m > T_{m,\alpha}\}}, \quad \chi^2_m = \sum_{j=1}^m X^2_{j,m}. \tag{3.40}$$

Recall that $T_{m,\alpha}$ stands for $(1-\alpha)$-quantile of the chi-square distribution with m degree of freedom.

Theorem 3.5 *Let $\delta_\varepsilon = O(1)$ (this corresponds to not too small α_ε), and $r^*_{\varepsilon,\alpha_\varepsilon}$ be defined either by (3.36) or by (3.38) for $\delta_\varepsilon \asymp 1$.*

(1) Lower bounds. There exists a constant $C_0 \in (0,\infty)$ such that if $\limsup r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon} < C_0$, then $\beta_\varepsilon(F_\varepsilon,\alpha_\varepsilon) \to 1$.

(2) Upper bounds. There exists a constant $C_1 \in (0,\infty)$ such that if $\liminf r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon} > C_1$, then $\beta_\varepsilon(F_\varepsilon,\alpha_\varepsilon) \to 0$. Moreover, let us take $m = m_\varepsilon \sim (r^*_\varepsilon)^{-1/\sigma}$ and the family of chi-square tests $\chi^2_{m_\varepsilon,\alpha_\varepsilon}$ of form (3.40). This yields $\alpha_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}) = \alpha_\varepsilon$. Then one has $\beta_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}, F_\varepsilon) \to 0$, as $\liminf r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon} > C_1$.

For any $0 < \sigma_0 < \sigma_1 < \infty$ such that $\delta_\varepsilon(\sigma_1) = O(1)$, these relations are uniform over $\sigma \in [\sigma_0,\sigma_1]$.

Proof of Theorem 3.5 is given in Sections 5.7, 5.8.

Note that analogous statements with the rates (3.36) were established in [17], [3] for the case $\log\alpha_\varepsilon^{-1} = O(\log\varepsilon^{-1})$.

Let us consider the case $\delta_\varepsilon \to \infty$ (this is analogous to the first relation (3.20)). Let $\phi_k(t),\ k \ge 0$ be L_2-normalized Legendre polynomials of degree k that provide an orthonormal basis in $L_2(0,1)$. Taking an integer-valued family $m = m_\varepsilon \to \infty$ and equispaced partition of the interval $(0,1]$ into m_ε sub-intervals $\delta_{j,m}$, we set

$$\phi_{jk,m}(t) = m^{1/2}\phi_k(mt-j+1)\mathbf{1}_{\delta_{j,m}}.$$

This is an orthonormal basis in $L_2(\delta_{j,m})$. Take an integer $l \ge 0$. Consider orthonormal system in $L_2(0,1)$ of the form $\{\phi_{jk,m},\ 0 \le k \le l,\ 1 \le j \le m\}$ and statistics

$$X_{jk,m} = \varepsilon^{-1}\int_{\delta_{j,m}} \phi_{jk,m}(t)dX_\varepsilon(t),\ 0 \le k \le l,\ 1 \le j \le m. \tag{3.41}$$

Note that random variables $X_{jk,m}$ are independent standard Gaussian under H_0. Let us consider chi-square tests based on statistics (3.41):

$$\chi^2_{m(l+1),\alpha} = \mathbf{1}_{\{\chi^2_{m,l} > T_{m(l+1),\alpha}\}}, \quad \chi^2_{m,l} = \sum_{j=1}^{m}\sum_{k=0}^{l} X^2_{jk,m}, \tag{3.42}$$

where, as above, $T_{m(l+1),\alpha}$ is $(1-\alpha)$-quantile of chi-square distribution with $m(l+1)$ degrees of freedom. Note that if $l = 0$, then we obtain tests (3.40).

The following theorem is analogous to Theorem 3.3.

Theorem 3.6

(1) Let $\delta_\varepsilon \to \infty$ and $r^*_{\varepsilon,\alpha_\varepsilon}$ be defined by (3.38). Then one can take $C_1 = \sqrt{2}$ in Theorem 3.5, (2) (recall that under (3.38) one can take $C_0 = \sqrt{2}$ in Theorem 3.5, (1) by (1.25)).

Moreover, let $\sigma = l + \tau,\ \tau \in (0,1],\ l \ge 0$ be an integer. Consider the chi-square tests $\chi^2_{m_\varepsilon(l+1),\alpha_\varepsilon}$ of form (3.42) with $m = m_\varepsilon \to \infty$ such that

$$m = o(\log\alpha_\varepsilon^{-1}), \quad m^{-\sigma}\varepsilon^{-1} = o((\log\alpha_\varepsilon^{-1})^{1/2}) \tag{3.43}$$

(this is possible for $\delta_\varepsilon \to \infty$; compare with (3.23)). Then $\beta_\varepsilon(\chi^2_{m_\varepsilon(l+1),\alpha_\varepsilon}, F_\varepsilon) \to 0$, as $\liminf r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon} > \sqrt{2}$.

(2) Suppose

$$\sigma > 1/2, \quad \varepsilon^4(\log\alpha_\varepsilon^{-1})^{2\sigma+1}/(\log\log\alpha_\varepsilon^{-1})^{4\sigma} \to \infty, \tag{3.44}$$

Let us take $m = m_\varepsilon \to \infty$ such that

$$m\log\log\alpha_\varepsilon^{-1} = o(\log\alpha_\varepsilon^{-1})^{1/2}), \quad m^{-2\sigma}\varepsilon^{-2} = o((\log\alpha_\varepsilon^{-1})^{1/2})$$

(this is possible under (3.44), compare with (3.25)). Then

$$\beta_\varepsilon(\chi^2_{m_\varepsilon(l+1),\alpha_\varepsilon}, F_\varepsilon) \le \Phi(\sqrt{2\log\alpha_\varepsilon^{-1}} - r_\varepsilon/\varepsilon) + o(1).$$

Combining with (1.25) under (3.44), this yields the sharp asymptotics

$$\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) = \Phi(\sqrt{2\log\alpha_\varepsilon^{-1}} - r_\varepsilon/\varepsilon) + o(1). \tag{3.45}$$

Proof of Theorem 3.6 is given in Section 5.10.

Remark 3.4 Analogously to Remark 3.3 for any $0 < \sigma_0 < \sigma_1 < \infty$ such that (3.43) or (3.44) hold with $\sigma = \sigma_0$, the relations of Theorem 3.6 are uniform over $\sigma \in [\sigma_0, \sigma_1]$. Therefore taking tests family for $\sigma = \sigma_0$ that does not depend on r_ε, we obtain distinguishability conditions of Theorem 3.6 (1) or sharp asymptotics (3.45) of Theorem 3.6 (2) uniformly over $\sigma \in [\sigma_0, \sigma_1]$.

So, if $\alpha_\varepsilon \to 0$, then rates (3.36), (3.38) depend essentially on α_ε. In the case $\delta_\varepsilon = o(1)$ one can consider the factor $\log\alpha_\varepsilon^{-1}$ in rates (3.36) as the payment for small type I error with respect to rates (2.1).

3.3.2 The case $p = \infty$

Let us consider the case $p = \infty$, $\eta = \sigma - 1/q > 0$. Introduce the rates

$$r^*_{\varepsilon,\alpha_\varepsilon,\infty} = (\varepsilon^2\log(\varepsilon\alpha_\varepsilon)^{-1})^{\eta/(2\eta+1)} \asymp (\varepsilon^2\log(\hat\alpha_\varepsilon)^{-1})^{\eta/(2\eta+1)}, \quad \hat\alpha_\varepsilon = \min(\varepsilon, \alpha_\varepsilon). \tag{3.46}$$

Under (1.23) rates (3.46) correspond to rates (3.32), i.e., $\varepsilon^{-1}r^*_{\varepsilon,\alpha_\varepsilon,\infty} \asymp \rho^*_{R,\alpha_R}$.

Theorem 3.7 *There exist constants $0 < C_0 < C_1 < \infty$ such that*

$$\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \to 1 \quad \text{if} \quad \limsup r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon,\infty} < C_0, \tag{3.47}$$

$$\beta_\varepsilon(F_\varepsilon, \alpha_\varepsilon) \to 0 \quad \text{if} \quad \liminf r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon,\infty} > C_1. \tag{3.48}$$

Moreover let us consider family of tests

$$\psi_{\varepsilon,\alpha_\varepsilon} = \mathbf{1}_{\mathcal{X}_{\varepsilon,\alpha_\varepsilon}}, \quad \mathcal{X}_{\varepsilon,\alpha_\varepsilon} = \{\sup_{l\ge 1}\max_{1\le j\le m_l}|X_{j,m_l}|/T_l > 1\}, \quad m_l = 2^l, \tag{3.49}$$

where $T_l = (2(cl + \log\alpha_\varepsilon^{-1} + \log l))^{1/2}$, $c = \log 2$ and the statistics $X_{j,m}$ are determined by (3.39). Then $\alpha_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}) \le \alpha_\varepsilon$ for small enough $\varepsilon > 0$, and if C_1 is large enough in the right-hand side of (3.48), then one has $\beta_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}, F_\varepsilon) \to 0$

For any $\eta_0 > 0$, $0 < \sigma_0 < \sigma_1 < \infty$, these relations are uniform over $\sigma \in [\sigma_0, \sigma_1]$, q such that $\eta = \sigma - 1/q \ge \eta_0$.

Proof of Theorem 3.7 is given in Section 5.11.

3.4 Adaptive setting

3.4.1 Sequence space

Let us consider the sequence model (1.1). Let the parameter σ be unknown and an interval $\Sigma = [\sigma_0, \sigma_1]$ be given. Taking a family of the functions $\rho_R(\sigma)$, $\sigma \in \Sigma$, we consider alternatives of form (1.28). Let a family $\alpha_R \to 0$, $\log \alpha_R^{-1} \leq R^2$, be given. Set

$$\tau_R = \log\log \alpha_R^{-1} / \log R \leq 2$$

and assume that there exists a limit

$$\tau = \lim_{R\to\infty} \tau_R; \quad 0 \leq \tau \leq 2.$$

We set $\sigma(\tau) = \infty$ for $\tau = 0$ and $\sigma(\tau) = 1/\tau - 1/2$ for $\tau > 0$. Also we set $\sigma_1(\tau) = 2/\tau - 1/2$.

Let $\sigma_0 > \sigma(\tau)$. Then relation (3.20) is fulfilled uniformly over Σ. In view of Remark 3.3 taking test family from Theorem 3.3 (1) for $\sigma = \sigma_0$, we obtain distinguishability conditions with $\rho_R = \inf_{\sigma\in\Sigma} \rho_R(\sigma)$:

$$\text{if} \quad \liminf \frac{\rho_R}{\sqrt{2\log \alpha_R^{-1}}} < 1, \quad \text{then} \quad \beta(V_R(\Sigma), \alpha_R) \to 1,$$

$$\text{if} \quad \limsup \frac{\rho_R}{\sqrt{2\log \alpha_R^{-1}}} > 1, \quad \text{then} \quad \beta(V_R(\Sigma), \alpha_R) \to 0.$$

Moreover if $\sigma_0 > \sigma_1(\tau)$, then the relation (3.24) fulfilled uniformly over Σ. Analogously, we have sharp asymptotics (3.26) that provided by tests family from Theorem 3.3 (2) for $\sigma = \sigma_0$.

Let $\sigma_1 < \sigma(\tau)$. Then the relation (3.9) is fulfilled uniformly over Σ. It follows from results of Section 3.1.1 that tests procedures depend essentially on σ. Therefore we have adaptive problem in this case.

Theorem 3.8 *Let $\sigma_1 < \sigma(\tau)$. Let $K^2(\rho_R, R, \sigma)$ be defined by (3.4), (3.5). Set*

$$K_R(\Sigma) = \inf_{\sigma\in\Sigma} K(\rho_R(\sigma), R, \sigma), \;\; H_R(\Sigma, \alpha_R) = \sqrt{2}\left(K_R(\Sigma) - \sqrt{\log\log R + \log \alpha_R^{-1}}\right). \tag{3.50}$$

(1) Upper bounds

$$\beta(V_R(\Sigma), \alpha_R) \leq \Phi(-H_R(\Sigma, \alpha_R)) + o(1). \tag{3.51}$$

(2) Lower bounds. For the case $\log \alpha_R^{-1} = O((\log\log R)^2)$, assume that the infimum in (3.50) is "essential", i.e., for any $\delta > 0$ there exists nontrivial subinterval $\Sigma_0 \subset \Sigma$ such that

$$\sup_{\sigma\in\Sigma_0} K(\rho_R(\sigma), R, \sigma) \leq K_R(\Sigma) + \delta.$$

Then

$$\beta(V_R(\Sigma), \alpha_R) \geq \beta(V_R(\Sigma_0), \alpha_R) \geq \Phi(-H_R(\Sigma, \alpha_R)) + o(1). \tag{3.52}$$

Let us describe the structure of test procedure ψ_{R,α_R} that provides the upper bounds (3.51). Let $\rho^*_R(\sigma)$ be the quantities such that

$$K_R(\rho^*_R(\sigma), R, \sigma) = K_R(\Sigma) + o(1).$$

Let $\bar{z}_R(\rho,\sigma)$ be extreme sequence in the problem (3.4), (3.5). Let $t_{R,\bar{z}} = t_{R,\bar{z}}(X)$ be the statistics determined by (3.11). Let us divide the interval Σ into $M = M_R$ sub-intervals,

$$\delta_{R,l} = [\sigma_{R,l-1}, \sigma_{R,l}], \quad 1 \le l \le M, \quad M \asymp (\log R)(K_R(\Sigma))^B, \quad B > 1,$$

of the length $|\delta_{R,l}| \asymp M^{-1}$ and consider collections of sequences

$$\bar{z}_{R,l} = \bar{z}_R(\rho^*_R(\sigma_{R,l}), \sigma_{R,l-1}), \quad l = 1, ..., M$$

and collections of statistics $t_{R,l} = t_{R,\bar{z}_{R,l}}$. Set

$$\psi^{ad}_{R,\alpha_R} = \mathbf{1}_{\mathcal{X}_{R,\alpha_R}}, \ \mathcal{X}_{R,\alpha_R} = \left\{X : \max_{1\le l\le M} t_{R,l} > (\log \alpha_R^{-1} + \log M)\right\}. \tag{3.53}$$

Then one has

$$\alpha(\psi^{ad}_{R,\alpha_R}) \le \alpha_R, \quad \beta(\psi^{ad}_{R,\alpha_R}, V_R(\Sigma)) \le \Phi(-H_R(\Sigma, \alpha_R)) + o(1). \tag{3.54}$$

Proofs of Theorem 3.8 and relations (3.54) are given in Section 5.12.

Taking into account Corollary 3.2, we get

Corollary 3.4

(1) Let $(\log \alpha_R^{-1})^2 = o(\log\log R)$*. Then we can change in Theorem 3.8 the quantity* $H_R(\Sigma, \alpha_R)$ *by the quantity* $H_R(\Sigma)$ *defined by (2.30) that does not depend on* α_R*.*

Let $\log \alpha_R^{-1} = o(\log\log R)$*. Then we have distinguishability conditions that does not depend on* α_R*:*

$$\textit{if} \ \ \liminf \frac{u_R(\Sigma)}{\sqrt{2\log\log R}} < 1, \quad \textit{then} \ \ \beta(V_R(\Sigma), \alpha_R) \to 1,$$
$$\textit{if} \ \ \limsup \frac{u_R(\Sigma)}{\sqrt{2\log\log R}} > 1, \quad \textit{then} \ \ \beta(V_R(\Sigma), \alpha_R) \to 0,$$

where the quantity $u_R(\Sigma)$ *is defined by (2.30).*

(2) Let $\log \alpha_R^{-1} \gg (\log\log R)^2$*. Then in Theorem 3.8, we can change the quantity* $H_R(\Sigma, \alpha_R)$ *by the quantity* $\tilde{H}_R(\Sigma, \alpha_R^{-1}) = \sqrt{2}(K_R(\Sigma) - \sqrt{\log \alpha_R^{-1}})$*.*

Let $\log \alpha_R^{-1} \gg \log\log R$*. Then we have distinguishability conditions:*

$$\textit{if} \ \ \liminf \frac{K_R(\Sigma)}{\sqrt{\log \alpha_R^{-1}}} < 1, \quad \textit{then} \ \ \beta(V_R(\Sigma), \alpha_R) \to 1,$$
$$\textit{if} \ \ \limsup \frac{K_R(\Sigma)}{\sqrt{\log \alpha_R^{-1}}} > 1, \quad \textit{then} \ \ \beta(V_R(\Sigma), \alpha_R) \to 0.$$

3.4.2 Functional space

Let us consider functional model (1.2). For given interval $\Sigma = [\sigma_0, \sigma_1]$ and a family of functions $r_\varepsilon(\sigma)$, $\sigma \in \Sigma$, we consider alternatives of form (1.29). For a family $\alpha_\varepsilon \to 0$ analogously to above, we set

$$\tau_\varepsilon = \log\log\alpha_\varepsilon^{-1}/\log\varepsilon^{-1}, \quad \tau_\varepsilon \le 2 + o(1).$$

Assume there exists a limit

$$\tau = \lim_{\varepsilon\to 0}\tau_\varepsilon \in [0,2],$$

and set $\sigma(\tau) = \infty$ for $\tau = 0$ and $\sigma(\tau) = 1/\tau - 1/2$ for $\tau > 0$.

Taking into account Remark 3.4, we consider the case $\sigma_1 < \sigma(\tau)$. Let the quantities $\delta_\varepsilon = \delta_\varepsilon(\sigma)$ be defined by (3.35). The assumption $\sigma_1 < \sigma(\tau)$ yields $\delta_\varepsilon(\sigma) \to 0$ uniformly over $\sigma \in \Sigma$.

Let us define adaptive rate function:

$$r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma) = (\varepsilon^4(\log\log\varepsilon^{-1} + \log\alpha_\varepsilon^{-1}))^{\sigma/(4\sigma+1)}. \tag{3.55}$$

Note that (compare with (2.12) and (3.36))

$$r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma) \asymp \begin{cases} (\varepsilon^4(\log\log\varepsilon^{-1}))^{\sigma/(4\sigma+1)}, & \text{as } \alpha_\varepsilon^{-1} \le \log\varepsilon^{-1}, \\ (\varepsilon^4(\log\alpha_\varepsilon^{-1}))^{\sigma/(4\sigma+1)}, & \text{as } \alpha_\varepsilon^{-1} > \log\varepsilon^{-1}. \end{cases} \tag{3.56}$$

Theorem 3.9 *Let $\sigma_1 < \sigma(\tau)$. There exists constants $0 < C_0 < C_1 < \infty$ such that:*

(1) Lower bounds. If there exists an interval $\Delta \subset \Sigma$ of positive length such that

$$\limsup\sup_{\sigma\in\Delta} r_\varepsilon(\sigma)/r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma) < C_0,$$

then $\beta_\varepsilon(F_\varepsilon(\Sigma), \alpha_\varepsilon) \to 1$;

(2) Upper bounds. If

$$\liminf\inf_{\sigma\in\Sigma} r_\varepsilon(\sigma)/r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma) > C_1, \tag{3.57}$$

then $\beta_\varepsilon(F_\varepsilon(\Sigma), \alpha_\varepsilon) \to 0$.

The following test procedure provides upper bounds (3.57). Let us take families $J_{\varepsilon,0} < J_{\varepsilon,1}$ such that

$$2^{-\sigma_1(J_{\varepsilon,0}+1)} = r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma_1), \quad 2^{-\sigma_0(J_{\varepsilon,1}-1)} = r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma_0).$$

For all integer l, $J_{\varepsilon,0} < l < J_{\varepsilon,1}$, let us take the collection of statistics $\chi^2_{m_l}$, $m_l = 2^l$ of form (3.40). Set

$$\psi^{ad}_{\varepsilon,\hat\alpha_\varepsilon} = \mathbf{1}_{\mathcal{X}_\varepsilon}, \quad \mathcal{X}_\varepsilon = \{X : \max_{J_{\varepsilon,0}<l<J_{\varepsilon,1}} \chi^2_{m_l}/T_{m_l,\hat\alpha_\varepsilon} > 1\} \tag{3.58}$$

with $\hat\alpha_\varepsilon = \alpha_\varepsilon/M$, $M = J_{\varepsilon,1} - J_{\varepsilon,0}$; recall that $T_{m,\alpha}$ stands for $(1-\alpha)$-quantile of the chi-square distribution with m degree of freedom.

Then for large enough C_1 in (3.57), one has

$$\alpha_\varepsilon(\psi^{ad}_{\varepsilon,\alpha_\varepsilon}) \le \alpha_\varepsilon, \quad \beta_\varepsilon(\psi^{ad}_{\varepsilon,\alpha_\varepsilon}, F_\varepsilon(\Sigma)) \to 0. \tag{3.59}$$

Proof of Theorem 3.9. By (3.56), for the case $\alpha_\varepsilon^{-1} \le \log \varepsilon^{-1}$ the lower bounds of Theorem follow directly from the statement (1) in Section 2.2 (see (2.13)). For the case $\alpha_\varepsilon^{-1} > \log \varepsilon^{-1}$ the lower bounds of Theorem follow from the lower bounds of Theorems 3.5. The proof of the upper bounds (3.59) for family of tests (3.58) is given in Section 5.13.

So, for the case $\alpha_\varepsilon^{-1} \le \log \varepsilon^{-1}$ one has no additional payments for small type I errors in the rates of testing; more precisely, this payment is included into the payment for adaptation. From the other hand, for the case $\alpha_\varepsilon^{-1} > \log \varepsilon^{-1}$ one has no additional payments for adaptation: this payment is included into the payment for small enough type I errors.

4 Some properties of extreme problems

4.1 Extreme problem (2.16), (2.17)

The results of this section are contained in [13], Section 4.3.

Using the Lagrange multipliers rule for a convex extreme problem (2.16), (2.17) in terms of variables $u_i = z_i^2 \ge 0$ and returning to variables $z_i \ge 0$,, one can easy describe nonnegative extreme sequence $\{z_{i,R}\}$ and the extreme value $u_R^2 = u^2(\rho_R, R)$. We have

$$u_R^2 = \frac{1}{2} n z^4 S_{0,n}; \quad z_{i,R} = z(1 - x_i^{2\sigma})_+^{1/2}, \quad x_i = i/n,$$

here $t_+ = t$ for $t \ge 0$ and $t_+ = 0$ for $t < 0$. The quantities $z = z_R > 0,\ n = n_R > 0$ are determined by the equations

$$n z^2 S_{1,n} = \rho_R^2, \quad n^{1+2\sigma} z^2 S_{2,n} = R^2.$$

The quantities $S_{l,n},\ l = 0, 1, 2$ are determined by the relations

$$\begin{aligned} S_{0,n} &= n^{-1} \sum_{1 \le i \le n} (1 - x_i^{2\sigma})^2 = S_0(\sigma) + O(n^{-1}), \\ S_{1,n} &= n^{-1} \sum_{1 \le i \le n} (1 - x_i^{2\sigma}) = S_1(\sigma) + O(n^{-1}), \\ S_{2,n} &= n^{-1} \sum_{1 \le i \le n} x_i^{2\sigma}(1 - x_i^{2\sigma}) = S_2(\sigma) + O(n^{-1}), \end{aligned}$$

where

$$S_0(\sigma) = \int_0^1 (1 - x^{2\sigma})^2 dx, \quad S_1(\sigma) = \int_0^1 (1 - x^{2\sigma}) dx, \quad S_2(\sigma) = \int_0^1 x^{2\sigma}(1 - x^{2\sigma}) dx. \tag{4.1}$$

Assume $\rho_R = o(R)$. Then $n_R \to \infty$ and we have the relations

$$n \sim D_1(\sigma)(R/\rho_R)^{1/\sigma}, \quad z \sim D_2(\sigma)\rho_R^{1+1/2\sigma} R^{-1/2\sigma}, \quad u(\rho_R, r) \sim D_0(\sigma)\rho_R^{2+1/2\sigma} R^{-1/2\sigma}, \tag{4.2}$$

where the functions $D_l(\sigma) > 0,\ l = 0, 1, 2$ are continuous Lipshician and bounded away from 0 over $\sigma \in [\sigma_0, \sigma_1]$ for any $0 < \sigma_0 < \sigma_1$. Therefore for any $\alpha \in (0, 1)$, we get

$$\beta(V_R, \alpha) = \Phi(T_\alpha - \tilde{u}) + o(1), \quad \tilde{u} = D_0(\sigma)\rho_R^{2+1/2\sigma} R^{-1/2\sigma}. \tag{4.3}$$

For any $0 < \sigma_0 < \sigma_1 < \infty$, the relations above are uniform over $\sigma \in [\sigma_0, \sigma_1]$.

4.2 Extreme problem (3.4), (3.5)

We give the outline of the study of the extreme problem (3.4), (3.5) for $\rho_R < R$ assuming $R \to \infty$. Using the Lagrange multipliers rule for a convex extreme problem (3.4), (3.5) in terms of variables $u_i = z_i^2 \geq 0$ and returning to variables $z_i \geq 0$, one can write the equations for extreme sequence $\{z_i\}$ in the problem

$$\frac{z_i^2}{1 + z_i^2} = \lambda - \mu i^{2\sigma} + C_i, \quad \lambda \geq 0, \quad \mu \geq 0, \quad C_i \geq 0, \quad C_i z_i^2 = 0; \tag{4.4}$$

if either $\lambda = 0$ or $\mu = 0$, then we have the strict inequality in the first or in the second inequalities (3.5), if $z_i > 0$, then $C_i = 0$. It is easily seen that we can take $\lambda > 0,\ \mu > 0$ and $C_i = 0$, when $\lambda - \mu i^{2\sigma} \geq 0$. Setting

$$\lambda = z^2,\ \mu = z^2 n^{-2\sigma}, \tag{4.5}$$

we can rewrite the equations (4.4) in terms of variables

$$z^2 = z_R^2 \in (0, 1 + \kappa_n), \quad \kappa_n = (n^{2\sigma} - 1)^{-1}, \quad n = n_R > 1.$$

We get

$$z_i = 0, \quad i \geq n, \quad z_i^2 = \frac{z^2(1 - x_i^{2\sigma})}{1 - z^2(1 - x_i^{2\sigma})}, \quad x_i = i/n, \quad 1 \leq i < n; \tag{4.6}$$

the variables $z,\ n$ are determined by the relations

$$\sum_{i=1}^{\infty} z_i^2 = n z^2 S_1(n, z) = \rho_R^2, \tag{4.7}$$

$$\sum_{i=1}^{\infty} z_i^2 i^{2\sigma} = n^{1+2\sigma} z^2 S_2(n, z) = R^2, \tag{4.8}$$

and

$$K^2(\rho_R, R) = \frac{1}{2} n z^4 S_0(n, z).$$

Here, as $n \to \infty$,

$$S_1(n, z) = n^{-1} \sum_{1 \leq i \leq n} \frac{1 - x_i^{2\sigma}}{1 - z^2(1 - x_i^{2\sigma})} = \tilde{S}_1(z) + O(n^{-1}); \tag{4.9}$$

$$S_2(n, z) = n^{-1} \sum_{1 \leq i \leq n} \frac{x_i^{2\sigma}(1 - x_i^{2\sigma})}{1 - z^2(1 - x_i^{2\sigma})} = \tilde{S}_2(z) + O(n^{-1}); \tag{4.10}$$

$$\begin{aligned} S_0(n, z) &= n^{-1} z^{-4} \sum_{1 \leq i \leq n} \left(\frac{z^2(1 - x_i^{2\sigma})}{1 - z^2(1 - x_i^{2\sigma})} + \log(1 - z^2(1 - x_i^{2\sigma})) \right) \\ &= \tilde{S}_0(z) + O(n^{-1}). \end{aligned} \tag{4.11}$$

In (4.9)–(4.11) we set

$$\tilde{S}_1(z) = \int_0^1 \frac{1 - x^{2\sigma}}{1 - z^2(1 - x^{2\sigma})} dx; \quad \tilde{S}_2(z) = \int_0^1 \frac{x^{2\sigma}(1 - x^{2\sigma})}{1 - z^2(1 - x^{2\sigma})} dx, \quad (4.12)$$

$$\tilde{S}_0(z) = z^{-4} \int_0^1 \left(\frac{z^2(1 - x^{2\sigma})}{1 - z^2(1 - x^{2\sigma})} + \log(1 - z^2(1 - x^{2\sigma})) \right) dx. \quad (4.13)$$

The integrals in (4.12)–(4.13) converge for any $z \in (0, 1 - b),\ b > 0$, and as $z \to 0$,

$$\tilde{S}_l(z) = S_l(\sigma) + O(z^2),\ l = 1, 2; \quad \tilde{S}_0(z) = \frac{1}{2} S_0(\sigma) + O(z^2),$$

where the quantities $S_l(\sigma),\ l = 0, 1, 2$ are defined by (4.1).

Note that for any $R > 0,\ \rho_R \in (0, R)$ there exist unique $z = z_R,\ n = n_R$ determined by relations (4.7), (4.8).

Let $n \to \infty$ and z be bounded away from 1. These yield z_i^2 are bounded and

$$\rho_R^2 \asymp nz^2, \quad R^2 \asymp n^{2\sigma+1} z^2; \quad n \asymp (R/\rho_R)^{1/\sigma} \to \infty, \quad z^{2\sigma} \asymp \rho_R^{2\sigma+1} R^{-1}, \quad (4.14)$$

$$K^2(\rho_R, R) \asymp nz^4 \asymp \rho_R^{4+1/\sigma} R^{-1/\sigma}. \quad (4.15)$$

Observe that if n is bounded and $1 + \kappa_n - z^2$ is bounded away from 0, then using (4.6), (4.8), (4.10) we get $R = O(1)$, which is impossible for $R \to \infty$. Also one easily seen that

$$R^2 \sim \rho_R^2 \sim 2K^2(\rho_R, R), \quad \text{if } \ 1 + \kappa_n - z^2 \to 0 \ \text{ and } \ n = O(1). \quad (4.16)$$

Let us study the asymptotics for the case $n \to \infty,\ z \to 1,\ z^2 < 1 + \kappa_n$. Setting

$$z^2 = 1 - \delta, \quad \tau = n^{2\sigma}\delta/(1 - \delta)), \quad -1 < \tau = o(n^{2\sigma}), \quad z^2 = \frac{n^{2\sigma}}{\tau + n^{2\sigma}}, \quad (4.17)$$

we can rewrite (4.6) in the form

$$z_i^2 = \frac{\tau + n^{2\sigma}}{\tau + i^{2\sigma}} - 1, \quad 1 \le i \le n. \quad (4.18)$$

Assume $\tau \to \infty,\ \tau = o(n^{2\sigma})$. Set $m = \tau^{1/2\sigma} \sim n\delta^{1/2\sigma}$. Rewriting (4.7) we have

$$\rho_R^2 = \sum_{1 \le 1 \le n} z_i^2 = (\tau + n^{2\sigma})\tau^{1/2\sigma - 1} S_{n,m} - n + O(1), \quad S_{n,m} = \frac{1}{m} \sum_{1 \le i \le n} \frac{1}{1 + (i/m)^{2\sigma}}. \quad (4.19)$$

We can replace the normalized sums $S_{n,m}$ in (4.19) by the integrals

$$S_{n,m} = I_{n,m} + O(m^{-1}), \quad I_{n,m} = \int_0^{n/m} \frac{dx}{1 + x^{2\sigma}} \sim \begin{cases} \delta^{1-1/2\sigma}/(1 - 2\sigma), & \sigma < 1/2 \\ \log \delta^{-1}, & \sigma = 1/2 \\ c(\sigma), & \sigma > 1/2 \end{cases},$$

where $c(\sigma) = \int_0^\infty (1 + t^{2\sigma})^{-1} dt,\ \sigma > 1/2$. This yields

$$\rho_R^2 \sim \begin{cases} 2n\sigma/(1 - 2\sigma), & \sigma < 1/2 \\ n \log(n/\tau), & \sigma = 1/2 \\ n^{2\sigma}\tau^{1/2\sigma - 1} c(\sigma), & \sigma > 1/2 \end{cases}. \quad (4.20)$$

Next, assume $-1 < \tau < B$, for some $B > 0$. Set

$$\rho_R^2 = z_1^2 + \tilde\rho_R^2, \quad \tilde\rho_R^2 = \sum_{2\le i\le n} z_i^2, \quad z_1^2 = \frac{n^{2\sigma}-1}{\tau+1}. \tag{4.21}$$

Using (4.18) we have

$$\tilde\rho_R^2 = (\tau+n^{2\sigma})S_{n,\tau} - n + O(1), \;\; S_{n,\tau} = \sum_{2\le i\le n} \frac{1}{\tau+i^{2\sigma}} \sim \begin{cases} n^{1-2\sigma}/(1-2\sigma), & \sigma<1/2 \\ \log n, & \sigma = 1/2 \end{cases}.$$

If $\sigma > 1/2$, then $S_{n,\tau} \asymp n$. Therefore if $-1 < \tau \le B$, then

$$\rho_R^2 \sim z_1^2 + \begin{cases} 2n\sigma/(1-2\sigma), & \sigma<1/2 \\ n\log n, & \sigma=1/2 \end{cases}; \quad \rho_R^2 \asymp z_1^2 + n^{2\sigma}, \quad \sigma > 1/2. \tag{4.22}$$

It follows from (4.20), (4.22) that $\rho_R^2/n \to \infty$, as $z \to 1,\ n \to \infty$ for $\sigma \ge 1/2$.
Also using (4.18) and evaluations above one can verify that, for any $b \in (0,1)$,

$$\min_{1\le i\le bn} z_i^2 \ge (b^{-2\sigma}-1)(1+o(1)), \qquad \sum_{1\le i\le bn} z_i^2 \asymp \rho_R^2, \quad \text{as} \quad z\to 1,\ n\to\infty.$$

Since $\forall\ c>0\ \exists\ d>0$ such that $x - \log(1+x) > dx$ for $x > c$, this yields

$$K^2(\rho_R, R) = \frac{1}{2}\sum_i (z_i^2 - \log(1+z_i^2)) \asymp \rho_R^2, \quad \text{as} \quad z\to 1,\ n\to\infty. \tag{4.23}$$

On the other hand, using (4.17), (4.18) and rewriting (4.8), we have

$$\begin{aligned} R^2 &= \sum_{1\le i\le n} i^{2\sigma} z_i^2 = \sum_{1\le i\le n} \left(n^{2\sigma} - \tau z_i^2 - i^{2\sigma}\right) \\ &= \frac{2\sigma}{2\sigma+1} n^{2\sigma+1} + O(n^{2\sigma}) - \tau\rho_R^2 \sim \frac{2\sigma}{2\sigma+1} n^{2\sigma+1} - \tau z_1^2, \end{aligned} \tag{4.24}$$

since $\sum_{1\le i\le n} i^{2\sigma} = n^{2\sigma+1}/(2\sigma+1) + O(n^{2\sigma})$ and $\tau\tilde\rho_R^2 = o(n^{2\sigma+1})$ by (4.20), (4.22).

It is easily seen from (4.20), (4.21), (4.24) that if τ is bounded away from -1, then the item z_1^2 is not essential for the rates of ρ_R^2, R^2. This yields

$$\rho_R R^{-1/(2\sigma+1)} \to \infty \ \text{ for } \sigma \ge 1/2, \quad \liminf \rho_R R^{-1/(2\sigma+1)} \ge C(\sigma) \ \text{ for } \sigma < 1/2, \tag{4.25}$$

where

$$C(\sigma) = (2\sigma/(1-2\sigma))^{1/2}(1+1/2\sigma)^{1/2(2\sigma+1)}. \tag{4.26}$$

Let $\tau \to -1$. Then the item z_1^2 is essential for the rates of ρ_R^2 and it may be essential for the rates of R^2. Relations (4.25), (4.26) hold true, for $z \to 1,\ n \to \infty$.

Relations (4.14)–(4.16), (4.23), (4.25) yield the following statements.

Proposition 4.1

(1) Let either $K^2(\rho_R, R) = o(R^{2/(2\sigma+1)})$ or $\rho_R = o(R^{1/(2\sigma+1)})$. Then $z \to 0$, $n \to \infty$ and relations (4.14)–(4.15) hold true.

(2) Let $\rho_R \asymp R^{1/(2\sigma+1)}$ and either $\sigma \ge 1/2$ or $\sigma < 1/2$, $\limsup \rho_R R^{-1/(2\sigma+1)} < C(\sigma)$. Then $n \to \infty$, the quantities z are bounded away from 0 and from 1 and relations (4.14)–(4.15) hold true as well.

Assume $\rho_R = o(R^{1/(2\sigma+1)})$. By Proposition 4.1 (1), this yields $z \to 0$ and

$$K^2(\rho_R, R) = u^2(\rho_R, R)(1/2 + O(z^2)), \quad K^2(\rho_R, R) = \frac{1}{4} \sum_{1 \le i \le n} z_i^4 (1 + O(z^2)). \quad (4.27)$$

Proof of Corollary 3.2. Relation (4.15) and the second relation (3.17) yield $z^2 K(\rho_R, R) = o(1)$. Therefore relations (3.18) follow from (4.27) and (3.15). □

5 Proofs of Theorems

5.1 Proof of Theorem 3.1 (1)

It suffices to consider the case

$$K(\rho_R, R) = \sqrt{\log \alpha_R^{-1}} + O(1). \quad (5.1)$$

and assumptions (3.9) and (3.8) are equivalent under (5.1) (see Remark 3.2).

Let us consider Bayesian hypothesis testing problem on a probability measure P corresponding to random vector X:

$$H_0 \,:\, P = P_0, \quad H_1 \,:\, P = P_{\pi_R}, \quad (5.2)$$

where $P_{\pi_R} = \int P_v \pi_R(dv)$ is the mixture over the prior π_R. Denote $\beta(P_{\pi_R}, \alpha_R)$ the minimum of the type II error probability in the problem (5.2) for given type I error probability α_R. It suffices to verify that

$$\beta(P_{\pi_R}, \alpha_R) = \Phi(\sqrt{2 \log \alpha_R^{-1}} - \sqrt{2} K(\rho_R, R)) + o(1); \quad (5.3)$$
$$\pi_R(V_R) \to 1, \quad V_R = V(\rho_R, R, \sigma). \quad (5.4)$$

In fact, let $\hat{\pi}_R$ be conditional measure with respect to the condition $v \in V_R$, i.e. $\hat{\pi}_R(A) = \pi_R(A \cap V_R)/\pi_R(V_R)$ and $\hat{P}_R = P_{\hat{\pi}_R}$. Since $\hat{\pi}_R$ is supported on V_R, for any $\alpha \in (0,1)$ we have $\beta(V_R, \alpha_R) \ge \beta(\hat{P}_R, \alpha_R)$. On the other hand, it follows from [13], Proposition 2.1, and inequalities (2.32), (2.49) that, for any $\alpha \in (0,1)$,

$$|\beta(P_{\pi_R}, \alpha_R) - \beta(\hat{P}_R, \alpha_R)| \le \frac{1}{2} |P_{\pi_R} - \hat{P}_R|_1 \le \frac{1}{2} |\pi_R - \hat{\pi}_R|_1 \le 1 - \pi_R(V_R),$$

where $|\cdot|_1$ is the total variation distance.

Let us consider extreme problem (3.4), (3.5) with slightly changed quantities

$$\tilde{\rho}_R^2 = \rho_R^2(1+\delta), \quad \tilde{R}^2 = R^2(1-\delta). \quad (5.5)$$

We take $\delta = \delta_R$ such that

$$z^{-4} n^{-1} \gg \delta^2 \gg n^{-1}, \quad (5.6)$$

where $n = n_R$, $z = z_R$ be the quantities determined by (4.7), (4.8). The first relation in (5.6) and (3.7) yield $K(\tilde{\rho}_R, \tilde{R}) = K(\rho_R, R) + o(1)$, and we can change $K(\rho_R, R)$ by $K(\tilde{\rho}_R, \tilde{R})$ in relation (5.3).

Let $\{z_i\}$ be the extreme sequence in the changed problem (3.4), (3.5) and

$$\pi_R = \mathcal{N}(0, \{z_i^2\}) = \prod_i \mathcal{N}(0, z_i^2) \tag{5.7}$$

be the Gaussian measure on $(l^2, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra of subsets in l^2. This corresponds to independent Gaussian coordinates $v_i \sim \mathcal{N}(0, z_i^2)$ of random mean vector $v \in l^2$. Note that

$$P_{\pi_R} = \int P_v \pi_R(dv) = \prod_i \mathcal{N}(0, z_i^2 + 1) \tag{5.8}$$

is the Gaussian measure that corresponds to independent $X_i \sim \mathcal{N}(0, z_i^2 + 1)$. The log-likelihood ratio $t_R = \log dP_{\pi_R}/dP_0$ is of form (3.11). According to Neyman–Pearson's Lemma the quantity $\beta(\alpha, P_{\pi_R})$ is of the form

$$\beta(\alpha, P_{\pi_R}) = P_{\pi_R}(t_R < T_{R,\alpha}), \tag{5.9}$$

where $T_{R,\alpha}$ is $(1-\alpha)$-quintile of the statistic t_R in P_0-probability, i.e.,

$$P_0(t_R \geq T_{R,\alpha}) = \alpha.$$

Denote

$$E_{\pi,R} = E_{P_{\pi_R}} t_R = \frac{1}{2}\sum_i (z_i^2 - \log(1 + z_i^2)),$$

$$\sigma_{\pi,R}^2 = \mathrm{Var}_{P_{\pi_R}} t_R = \frac{1}{2}\sum_i z_i^4, \quad \lambda_R = (t_R - E_{\pi,R})/\sigma_{\pi,R}.$$

Clearly,

$$E_{\pi,R} = K^2(\tilde{\rho}_R, \tilde{R}), \quad \sigma_{\pi,R}^2 \sim 2E_{\pi,R} \tag{5.10}$$

(the latter relation follows from (4.27)).

Let z_i be determined by (4.6) with $n = n_R$, $z = z_R$ determined by (4.7), (4.8) for the changed extreme problem (they are of the same rates as for the original problem and we use the same notation).

Lemma 5.1 *Let $n \to \infty$ and $\sup_i z_i^2 = O(1)$ (by (4.14), these hold under (3.8)). Then $\lambda_R \to \xi \sim \mathcal{N}(0,1)$, in P_{π_R}-probability.*

Proof. In P_{π_R}-probability, the random variables t_R are distributed as

$$K^2(\tilde{\rho}_R, \tilde{R}) + \frac{1}{2}\sum_i Y_i, \quad Y_i = z_i^2(\xi_i^2 - 1), \quad \xi_i \sim \mathcal{N}(0,1) \text{ are i.i.d.}$$

Note that $\mu_i^4 = EY_i^4 = 60z_i^8$. It follows from (4.6) that

$$\sum_i z_i^4 \asymp nz^4, \quad \sum_i z_i^8 \asymp nz^8,$$

and the Lyapunov ratio Λ_R is of the rate

$$\Lambda_R = \frac{\sum_i \mu_i^4}{(\mathrm{Var}_{P_{\pi_R}} t_R)^2} \asymp n^{-1} \to 0.$$

This yields the statement of the lemma. $\square$

The next lemma is formulated in more general form that we need for the proof of Theorem 3.1 (1), because we'll use it for the proof of Theorem 3.4 below.

Lemma 5.2 *Let $t_R = \log dP_{\pi_R}/dP_0$ be a log-likelihood ratio (not necessarily for the priors of form (5.7)), and $T_{R,\alpha}$ be $(1-\alpha)$-quantile of t_R under P_0. Then*

(1) For any T_R one has

$$\log(P_0(t_R > T_R)) \le -T_R.$$

This yields $T_{R,\alpha} \le \log \alpha^{-1}$.

(2) Let $\lambda_R = (t_R - E_{\pi,R})/\sigma_{\pi,R}$, where $E_{\pi,R}$, $\sigma_{\pi,R} \to \infty$ be a quantities (not necessarily defined by above) such that $\lambda_R \to \zeta$ in P_{π_R}-probability. Let $T_R = E_{\pi,R} + a_R\sigma_{\pi,R}$, $a_R \to a \in R$ and assume, for the random variable ζ and the quantity a,

$$\log(Ee^{h(a-\zeta)}\mathbf{1}_{\{\zeta > a\}}) = o(h), \quad as \quad h \to \infty. \tag{5.11}$$

Then

$$\log(P_0(t_R > T_R)) = -T_R + o(\sigma_{\pi,R}). \tag{5.12}$$

If $\log \alpha_R^{-1} = E_{\pi,R} + a_R\sigma_{\pi,R}$, $a_R \to a$ (by (5.10), this corresponds to assumption (5.1) for priors (5.7)), then this yields

$$T_{R,\alpha_R} = \log \alpha_R^{-1} + o(\sigma_{\pi,R}). \tag{5.13}$$

Note that Lemma 5.2 (1) yields relation (3.12).

Proof. First, since the measure P_{π_R} is absolute continuous with respect to P_0, we have $E_{P_0}e^{t_R} = E_{P_0}(dP_{\pi_R}/dP_0) = 1$. Using the Markov inequality we get

$$P_0(t_R > T_R) = P_0(e^{t_R} > e^{T_R}) \le e^{-T_R}E_{P_0}e^{t_R} = e^{-T_R}.$$

This implies Lemma 5.2 (1).

Next, we can write $t_R = T_R + \sigma_{\pi,R}\eta_R$, where $\eta_R = \lambda_R - a_R \to \zeta - a$ in P_{π_R}-probability. Moreover,

$$\begin{aligned} P_0(t_R > T_R) &= E_{P_{\pi_R}} e^{-t_R}\mathbf{1}_{\{t_R > T_R\}} = e^{-T_R}E_{P_{\pi_R}} e^{-\sigma_{\pi,R}\eta_R}\mathbf{1}_{\{\eta_R > 0\}} \\ &= e^{-T_R}E_{P_{\pi_R}} Z_R^{\sigma_{\pi,R}}, \quad Z_R = e^{-\eta_R}\mathbf{1}_{\{\eta_R > 0\}}. \end{aligned} \tag{5.14}$$

Note that the random variables Z_R are bounded and $Z_R \to Z = e^{-\zeta+a}\mathbf{1}_{\{\zeta > a\}}$ in P_{π_R}-probability and $EZ > 0$. Therefore $E_{P_{\pi_R}} Z_R^h/EZ^h \to 1$ for any $h > 0$. This yields there exists a family $h_R \to \infty$ such that $E_{P_{\pi_R}} Z_R^{h_R}/EZ^{h_R} \to 1$. Assuming $h_R = o(\sigma_{\pi,R})$ and using the inequality $EY^b \ge (EY)^b$ for $Y \ge 0$, $b > 1$ and (5.11), we have

$$\log(E_{P_{\pi_R}} Z_R^{\sigma_{\pi,R}}) \ge \frac{\sigma_{\pi,R}}{h_R}\log(E_{P_{\pi_R}} Z_R^{h_R}) = \frac{\sigma_{\pi,R}}{h_R}(\log(EZ^{h_R}) + o(1)) = o(\sigma_{\pi,R}). \tag{5.15}$$

Lemma 5.2 (1) and relations (5.14)–(5.15) imply Lemma 5.2 (2). $\square$

Remark 5.1 Note that if λ_R is asymptotically standard Gaussian, i.e., $\zeta = \xi \sim \mathcal{N}(0,1)$, then relation (5.11) holds true. In fact, direct calculation and (1.8) give

$$Ee^{h(a-\xi)}\mathbf{1}_{\{\xi>a\}} = e^{ha+h^2/2}\Phi(-a-h) \sim e^{-a^2/2}/(a+h)\sqrt{2\pi}. \qquad (5.16)$$

In view of Lemma 5.1, we can apply Lemma 5.2 to the problem under consideration.

Moreover, relation (5.11) holds true for the case $a \geq H$ and $\zeta = \xi_H$ is lower H-truncated standard Gaussian variable $\xi \sim \mathcal{N}(0,1)$:

$$\xi_H = H + (\xi - H)\mathbf{1}_{\{\xi>H\}} = \begin{cases} \xi, & \xi > H, \\ H, & \xi \leq H, \end{cases} \quad P(\xi_H < t) = \begin{cases} \Phi(t), & t \geq H, \\ 0, & t < H. \end{cases} \qquad (5.17)$$

In fact, by $a \geq H$ we have $\xi_H = \xi$ for $\xi_H > a$ and we repeat calculations (5.16).

Remark 5.2 Let us take sequences $m \to \infty,\ z = z_m > 0$ and consider chi-square statistics χ_m^2 of form (3.22) and Gaussian measure $\pi = \mathcal{N}(0,\{z\})$ corresponding to $z_i = z,\ i = 1,...,m;\ z_i = 0$ for $i > m$. Then the statistics t_R are of the form

$$t_R = \frac{z^2}{2(1+z^2)}\chi_m^2 - \frac{m}{2}\log(1+z^2);$$

the quantities $E_{\pi,R},\ \sigma_{\pi,R}^2$ are of the form

$$E_{\pi,R} = \frac{m}{2}(z^2 - \log(1+z^2)), \quad \sigma_{\pi,R}^2 = \frac{m}{2}z^4.$$

Analogously to the proof of Lemma 5.1 we see that the random variable λ_R is asymptotically $\mathcal{N}(0,1)$-Gaussian in P_{π_R}-probability. Setting $T_R = E_{\pi,R}$ and applying Lemma 5.2, we obtain the large deviation inequality for chi-square statistics χ_m^2

$$\log P_0(\chi_m^2 > m(1+z^2)) \leq -\frac{m}{2}(z^2 - \log(1+z^2)); \qquad (5.18)$$

$$\log P_0(\chi_m^2 > m(1+z^2)) = -\frac{m}{2}(z^2 - \log(1+z^2)) + o(z^2\sqrt{m}). \qquad (5.19)$$

Let us prove relation (5.3). Using (5.9), (5.10) (5.13) and Lemma 5.1 we have

$$\beta(P_{\pi_R}, \alpha_R) = P_{\pi_R}(t_R < T_{R,\alpha}) = P_{\pi_R}\left(\lambda_R < \frac{T_{R,\alpha} - E_{\pi,R}}{\sigma_{\pi,R}}\right) =$$
$$\Phi\left(\frac{\log\alpha_R^{-1} - K^2(\tilde{\rho}_R, \tilde{R}) + o(K(\tilde{\rho}_R, \tilde{R}))}{\sqrt{2}K(\tilde{\rho}_R, \tilde{R})}\right) + o(1).$$

We can write

$$\log\alpha_R^{-1} - K^2(\tilde{\rho}_R, \tilde{R}) = \left(\sqrt{\log\alpha_R^{-1}} - K(\tilde{\rho}_R, \tilde{R})\right)\left(\sqrt{\log\alpha_R^{-1}} + K(\tilde{\rho}_R, \tilde{R})\right).$$

Under the assumption (5.1) the second factor is $2K(\tilde{\rho}_R, \tilde{R}) + O(1)$. This yields relation (5.3).

In order to verify relation (5.4), note that

$$1 - \pi_R(V_R) \le \pi_R(S_1 < \rho_R^2) + \pi_R(S_2 > R^2); \quad S_1 = \sum_i v_i^2, \quad S_2 = \sum_i i^{2\sigma} v_i^2.$$

By construction of π_R, we have

$$\begin{aligned}
E_{\pi_R}(S_1) &= \sum_i z_i^2 = (1+\delta)\rho_R^2, \quad \rho_R^2 \asymp n z^2, \\
E_{\pi_R}(S_2) &= \sum_i i^{2\sigma} z_i^2 = (1-\delta)R^2, \quad R^2 \asymp n^{1+2\sigma} z^2, \\
\mathrm{Var}_{\pi_R}(S_1) &= 2\sum_i z_i^4 \asymp n z^4, \quad \mathrm{Var}_{\pi_R}(S_2) = 2\sum_i i^{4\sigma} z_i^4 \asymp n^{1+4\sigma} z^4.
\end{aligned}$$

Therefore using the Chebyshev inequality we get

$$\begin{aligned}
\pi_R(S_1 < \rho_R^2) &\le \pi_R(|S_1 - E_{\pi_R}(S_1)| > \delta\rho_R^2) \le \mathrm{Var}_{\pi_R}(S_1)/\delta^2\rho_R^4 \asymp (n\delta^2)^{-1} \to 0, \\
\pi_R(S_2 > R^2) &\le \pi_R(|S_2 - E_{\pi_R}(S_2)| > \delta R^2) \le \mathrm{Var}_{\pi_R}(S_2)/\delta^2 R^4 \asymp (n\delta^2)^{-1} \to 0.
\end{aligned}$$

These yield relation (5.4). □

5.2 Proof of Theorem 3.2 (1)

It suffices to show that for any $\varepsilon > 0$ there exist $C > 0$, $R_0 > 0$ such that if $R > R_0$ and $\sqrt{\log \alpha_R^{-1}} - K(\rho_R, R) > C$, then $\beta(V_R, \alpha_R) > 1 - \varepsilon$.

The proof follows to the scheme of Section 5.1 and we note the differences only. Note that $n = n_R \to \infty$ and $z = z_R$ are bounded away from 0 and from 1 in the case (see Proposition 4.1 (2)). First, we take $\delta = Bn^{-1/2}$ in (5.5) with B such that $\pi_R(V_R) > 1 - \varepsilon/2$ for large enough R (these correspond to evaluations in the end of Section 5.1). Under this choice we get $K(\tilde\rho_R, \tilde R) \le K(\rho_R, R) + B_1$ for some $B_1 = B_1(B)$. Note that $E_{\pi,R} = K^2(\tilde\rho_R, \tilde R) \asymp \sigma_{\pi,R}^2$ in the case under consideration. Other evaluations are analogous to above. □

5.3 Proof of Theorem 3.1 (2)

Analogously to Section 5.1 it suffices to assume (5.1) (see Remark 3.2). We study the distributions of statistics t_R determined by (3.11) under alternatives $v \in V_R$. Set

$$E_R(v) = E_v t_R, \quad \sigma_R^2(v) = \mathrm{Var}_v t_R, \quad \lambda_{v,R} = \frac{t_R - E_R(v)}{\sigma_R(v)}.$$

Since

$$E_v X_i^2 = 1 + v_i^2, \quad \mathrm{Var}_v X_i^2 = 2 + 4v_i^2,$$

we have

$$E_R(v) = \frac{1}{2}\sum_i \left(\frac{z_i^2}{1+z_i^2}(1+v_i^2) - \log(1+z_i^2) \right), \tag{5.20}$$

$$\Delta E_R(v) = E_v t_R - K^2(\rho_R, R) = \frac{1}{2}\sum_i \frac{z_i^2}{1+z_i^2}(v_i^2 - z_i^2), \tag{5.21}$$

$$\sigma_R^2(v) = \frac{1}{2}\sum_i \frac{z_i^4}{(1+z_i^2)^2}(1+2v_i^2). \tag{5.22}$$

Lemma 5.3 *For the set $V_R = V(\rho_R, R, \sigma)$, one has*

$$\inf_{v \in V_R} \Delta E_R(v) \geq 0.$$

Proof. Taking into account (4.4), (4.5) one has, for all i,

$$\frac{z_i^2}{1+z_i^2} = \lambda - \mu i^{2\sigma} + C_i, \quad \lambda > 0, \quad \mu > 0, \quad C_i \geq 0, \quad C_i z_i^2 = 0.$$

Therefore

$$2\Delta E_R(v) = \sum_i \frac{z_i^2}{1+z_i^2}(v_i^2 - z_i^2) = \lambda \sum_i (v_i^2 - z_i^2) - \mu \sum_i i^{2\sigma}(v_i^2 - z_i^2) + \sum_i C_i(v_i^2 - z_i^2).$$

Recalling definition (1.19) of the set $V_R = V(\rho_R, R, \sigma)$ and relations (4.7), (4.8), we have

$$\sum_i (v_i^2 - z_i^2) \geq 0, \quad \sum_i i^{2\sigma}(v_i^2 - z_i^2) \leq 0, \quad \sum_i C_i(v_i^2 - z_i^2) = \sum_i C_i v_i^2 \geq 0.$$

This yields the statement of Lemma. □

Recall that under assumption (3.8), relation (4.6) holds and $z \to 0$. These yield $g_R = \sup_i z_i^2 \to 0$, and we have

$$\sum_i \frac{z_i^4}{(1+z_i^2)^2} \leq 2\sigma_R^2(v) \leq \sum_i \frac{z_i^4}{(1+z_i^2)^2} + 2g_R \sum_i \frac{z_i^2 v_i^2}{1+z_i^2} =$$
$$\sum_i \frac{z_i^4}{(1+z_i^2)^2} + 2g_R \sum_i \frac{z_i^2}{1+z_i^2}(v_i^2 - z_i^2) + 2g_R \sum_i \frac{z_i^4}{1+z_i^2}.$$

In view of (4.27),

$$\sum_i \frac{z_i^4}{(1+z_i^2)^2} \sim \sum_i \frac{z_i^4}{1+z_i^2} \sim 4K^2(\rho_R, R).$$

This yields that, uniformly over $v \in l^2$,

$$(2+o(1))K^2(\rho_R, R) \leq \sigma_R^2(v) \leq 2K^2(\rho_R, R)(1+o(1)) + o(\Delta E_R(v)). \tag{5.23}$$

Assume $\Delta E_R(v)/K(\rho_R, R) \to \infty$. It follows from the Chebyshev inequality and (5.23) that

$$\begin{aligned}\beta(\psi_{\alpha_R}, v) &= P_v(t_R < \log \alpha^{-1}) = P_v(E_R(v) - t_R \geq E_R(v) - \log \alpha_R^{-1}) \\ &\leq \frac{\sigma_R^2(v)}{(E_R(v) - \log \alpha^{-1})^2} \to 0,\end{aligned}$$

because of $E_R(v) - \log\alpha^{-1} = \Delta E_R(v) + O(K(\rho_R, R))$.

Assume

$$\Delta E_R(v) = O(K(\rho_R, R)). \tag{5.24}$$

By (5.23) this yields

$$\sigma_R^2(v) \sim 2K^2(\rho_R, R) \tag{5.25}$$

In P_v-probability, the random variable $\lambda_{v,R}$ is distributed as

$$\lambda_R \sim \sum_i (a_{R,i}U_i + b_{R,i}V_i), \quad U_i = (\xi_i^2 - 1), \quad V_i = \xi_i, \quad \xi_i \sim \mathcal{N}(0,1) \text{ i.i.d}, \tag{5.26}$$

$$\begin{aligned} a_{R,i} &= \frac{z_i^2}{2\sigma_R(v)(1+z_i^2)} \geq 0, \quad b_{R,i} = \frac{z_i^2 v_i}{\sigma_R(v)(1+z_i^2)}, \\ A_R &= \sum_i a_{R,i}^2, \quad B_R = \sum_i b_{R,i}^2, \quad 2A_R + B_R = 1, \end{aligned} \tag{5.27}$$

and by (4.15),

$$a_R^2 = \max_i a_{R,i}^2 = O(z^4/K^2(\rho_R, R)) = O(n^{-1}) \to 0.$$

Lemma 5.4 *Let random variables variables λ_R be of form (5.26), (5.27) and $a_R^2 = \sup_i a_{R,i}^2 = o(1)$. Then λ_R are asymptotically standard Gaussian.*

Proof. Let $A_R = o(1)$. Clearly,

$$E\left(\sum_i a_{R,i}U_i\right)^2 = 2A_R \to 0, \quad B_R = 1 + o(1), \quad \sum_i b_{R,i}V_i \sim \mathcal{N}(0, B_R).$$

This yields Lemma (5.4) for the case $A_R = o(1)$. Let $A_R \asymp 1$. Set

$$I_R = \{i \,:\, b_{R,i}^2 > a_R\}, \ |I_R| = \#I_R,$$

and consider the representation $\lambda_R = \tilde{\lambda}_R^{(1)} + \tilde{\lambda}_R^{(2)} + \delta_R$, where

$$\tilde{\lambda}_R^{(1)} = \sum_{i \notin I_R} (a_{R,i}U_i + b_{R,i}V_i) = \sum_{i \notin I_R} W_i, \quad \tilde{\lambda}_R^{(2)} = \sum_{i \in I_R} b_{R,i}V_i, \quad \delta_R = \sum_{i \in I_R} a_{R,i}U_i.$$

Note that

$$\mathrm{Var}(\delta_R) = 2\sum_{i \in I_R} a_{R,i}^2 \leq 2a_R^2|I_R|, \quad 1 \geq \sum_{i \in I_R} b_{R,i}^2 > a_R|I_R|.$$

This yields $\mathrm{Var}(\delta_R) = o(1)$, $\delta_R \to 0$. Clearly, $\tilde{\lambda}_R^{(1)}$ and $\tilde{\lambda}_R^{(2)}$ are independent and $\tilde{\lambda}_R^{(2)} \sim \mathcal{N}(0, \tilde{B}_R)$, where $\tilde{B}_R = \sum_{i \in I_R} b_{R,i}^2$. Observe that, for some $B > 0$,

$$\begin{aligned} &\sum_{i \notin I_R} E_v W_i^4 \leq B \sum_{i \notin I_R} (a_{R,i}^4 + b_{R,i}^4) \leq B\left(a_R^2 \sum_i a_{R,i}^2 + a_R \sum_i b_{R,i}^2\right) = o(1), \\ &\mathrm{Var}(\tilde{\lambda}_R^{(1)}) = \sum_{i \notin I_R} (2a_{R,i}^2 + b_{R,i}^2) \asymp 1, \end{aligned}$$

because of

$$A_R = \sum_i a^2_{R,i} \asymp 1, \quad \sum_{i\in I_R} a^2_{R,i} = o(1).$$

This yields the Lyapunov condition and asymptotic normality of $\tilde\lambda^{(1)}_{v,R}$. $\square$

In view of Lemma 5.4 and (5.25), under (5.24) we have, uniformly over $v \in l^2$,

$$\beta(\psi_{\alpha_R}, v) = P_v(t_R < \log\alpha^{-1}) = \Phi((\log\alpha^{-1} - E_R(v))/\sigma_R(v)) + o(1) =$$
$$\Phi((\log\alpha^{-1} - K^2(\rho_R, R) - \Delta E_R(v))/\sqrt{2}K(\rho_R, R)) + o(1). \tag{5.28}$$

Using Lemma 5.3 and (5.28), we have

$$\beta(\psi_{\alpha_R}, V_R) \le \Phi((\log\alpha^{-1} - K^2(\rho_R, R))/\sqrt{2}K(\rho_R, R)) + o(1).$$

At last, note that under assumption (5.1)

$$\log\alpha^{-1} - K^2(\rho_R, R) = (\sqrt{\log\alpha^{-1}} - K(\rho_R, R))(\sqrt{\log\alpha^{-1}} + K(\rho_R, R)) \sim$$
$$2(\sqrt{\log\alpha^{-1}} - K(\rho_R, R))K(\rho_R, R).$$

This yields the statement of Theorem. $\square$

Remark 5.3 Let us describe deeper sense of Lemma 5.3 that corresponds to extreme properties of the Gaussian priors (5.7) in the mixture (5.8).

The statement of Lemma 5.3 corresponds to the inequality

$$\inf_{v\in V_R} \sum_i E_{v_i} \log(dP_{\pi_i}/dP_0) \ge \sum_i E_{P_{\pi_i}} \log(dP_{\pi_i}/dP_0).$$

Clearly this follows from the inequality

$$\inf_{\bar\mu\in\Pi_R} \sum_i E_{P_{\mu_i}} \log(dP_{\pi_i}/dP_0) \ge \sum_i E_{P_{\pi_i}} \log(P_{\pi_i}/dP_0), \tag{5.29}$$

where Π_R is a set that consists of sequences of priors $\bar\mu = \{\mu_i(du)\}$ such that

$$\sum_i E_{\mu_i} u^2 \ge \rho_R^2, \quad \sum_i i^{2\sigma} E_{\mu_i} u^2 \le R^2. \tag{5.30}$$

Let $\mathcal{P}$ be the set of probability measures on the real line that are absolutely continuous with respect to the standard Gaussian measure P_0. Let us consider the functions

$$\phi(P,Q) = E_P \log(dQ/dP_0), \quad P, Q \in \mathcal{P}, \quad \varphi(P) = \phi(P,P), \quad P \in \mathcal{P}$$

(possibly, $\phi(P) = \infty$ or $\phi(P,Q) = \infty$). It is easily seen that $\phi(P,Q)$ is convex (linear) in P and concave in Q. Also since the function $\log(x)$ is concave and using Jensen's inequality, we have

$$\phi(P,Q) - \phi(P,P) = E_P \log(dQ/dP) \le \log E_P(dQ/dP) = 0.$$

Therefore

$$\varphi(P) = \sup_{Q \in \mathcal{P}} \phi(P, Q),$$

and $\varphi(P)$ is convex in P.

Let $P = P_\pi, \ Q = P_\mu$ be a mixture over priors π, μ on $(R^1, \mathcal{B})$. Then we set

$$\phi(\pi, \mu) = \phi(P_\pi, P_\mu), \quad \varphi(\pi) = \varphi(P_\pi).$$

It follows from above that $\phi(\pi, \mu)$ is convex (linear) in π and concave in μ,

$$\sup_{\mu} \phi(\pi, \mu) = \varphi(\pi), \tag{5.31}$$

and $\varphi(\pi)$ is convex in π as well.

Let us rewrite the inequality (5.29) in the form

$$\inf_{\bar{\mu} \in \Pi_R} \sum_i \phi(\mu_i, \pi_i) \geq \sum_i \varphi(\pi_i), \tag{5.32}$$

where $\pi_i = N(0, z_i^2)$ are Gaussian measures that correspond to the extreme sequence in (3.4), (3.5) (clearly $\bar{\pi} \in \Pi_R$).

In order to verify (5.32), let us try to maximize the left-hand side of (5.32) over $\bar{\pi}$, i.e., consider maximin extreme problem

$$H(\rho_R, R) = \sup_{\bar{\pi}} \inf_{\bar{\mu} \in \Pi_R} \sum_i \phi(\mu_i, \pi_i). \tag{5.33}$$

Using convex properties of this problem and applying minimax theorem we can replace the supremum and infimum in (5.33). Using (5.31) we have

$$\begin{aligned} H(\rho_R, R) &= \inf_{\bar{\mu} \in \Pi_R} \sup_{\bar{\pi}} \sum_i \phi(\mu_i, \pi_i) = \inf_{\bar{\mu} \in \Pi_R} \sum_i \sup_{\pi_i} \phi(\mu_i, \pi_i) \\ &= \inf_{\bar{\mu} \in \Pi_R} \sum_i \varphi(\mu_i) = \inf_{\bar{z} \in V_R} \sum_i h(z_i), \end{aligned} \tag{5.34}$$

where

$$h(z) = \inf_{\mu} \{\varphi(\mu) \, : \, E_\mu u^2 = z^2\} \tag{5.35}$$

Let us show that the extreme measure μ^* in (5.35) is the Gaussian measure

$$\mu^* = N(0, z^2). \tag{5.36}$$

In view of relations above, this yields the inequality (5.32). Since

$$\varphi(\mu^*) = h(z) = (z^2 - \log(1 + z^2))/2,$$

the extreme problem (5.34) is the same as the extreme problem (3.4), (3.5); also $H(\rho_R, R) = K(\rho_R, R)$ in (5.33) and this equals to the right-hand side of (5.32).

In order to verify (5.36) note that the function $\varphi(\mu)$ is of the form

$$\varphi(\mu) = E_{P_0}(g(x, \mu) \log g(x, \mu)), \quad g(x, \mu) = \int \exp(-u^2/2 + xu) \mu(du).$$

The function $\varphi(\mu)$ is strictly convex and the constraint set in (5.35) is convex as well. It follows from the method of subdifferentials (this corresponds to a formal derivative $\partial\varphi/\partial\mu$, compare with [13], Section A.6) and the Kuhn–Tucker Theorem that it suffices to verify the following relation: there exist constants A, B such that

$$\frac{\partial\varphi}{\partial\mu}(\mu^*, u) = E_{P_0}\left(\exp(-u^2/2 + xu)\log g(x,\mu^*)\right) + 1 = Au^2 + B \tag{5.37}$$

(the constant B corresponds to the constraints $E_\mu 1 = 1$). However for $\mu^* = N(0, z^2)$, we have

$$g(x,\mu^*) = \frac{1}{\sqrt{1+z^2}}\exp\left(\frac{z^2x^2}{2(1+z^2)}\right),$$
$$E_{P_0}\left(\exp(-u^2/2 + xu)\log g(x,\mu^*)\right) = -\frac{1}{2}\log(1+z^2) + \frac{z^2}{2(1+z^2)}(1+u^2).$$

This yields (5.37).

5.4 Proof of Theorem 3.2 (2)

It suffices to show that for any $\varepsilon > 0$ there exist $C > 0,\ R_0 > 0$ such that if $R > R_0$ and $\sqrt{\log\alpha_R^{-1}} - K(\rho_R, R) < -C$, then $\beta(\psi_{\alpha_R}, V_R) < \varepsilon$. The proof follows to the scheme of Section 5.3. Recall that $n \to \infty$ and z are bounded away from 0 and from 1 in the case (see Proposition 4.1 (2)). For this reason the relation (5.23) is changed by

$$b_1K^2(\rho_R, R) \le \sigma_R^2(v) \le b_2K^2(\rho_R, R) + b_3\Delta E_R(v),$$

for large enough R and some positive constants $b_l,\ l = 1, 2, 3$. Lemma 5.3 holds true as well and other evaluations are analogous. □

5.5 Proof of Theorem 3.3

Define the quantities $T_R = m(1 + z_R^2),\ z_R$ by the relation

$$L(T_R, m) = m(z_R^2 - \log(1 + z_R^2))/2 = \log\alpha_R^{-1}, \tag{5.38}$$

i.e.,

$$T_R = m + 2\log\alpha_R^{-1} + m\log(1 + z_R^2), \quad T_R \ge T_{m,\alpha}; \tag{5.39}$$

the latter inequality follows (5.18). It suffices to consider the case $\log\alpha_R^{-1} \asymp \rho_R^2$. Since $m = o(\log\alpha_R^{-1})$, we have

$$\begin{gathered} z_R^2 - \log(1 + z_R^2) = 2m^{-1}\log\alpha_R^{-1} \to \infty,\ z_R^2 \sim 2m^{-1}\log\alpha_R^{-1},\ z_R \to \infty, \\ \log(1 + z_R^2) \asymp \log z_R,\ m\log(1 + z_R^2)/\log\alpha_R^{-1} \asymp z_R^{-2}\log z_R \to 0. \end{gathered} \tag{5.40}$$

It suffices to show that under assumption of Theorem 3.3 (1),

$$\sup_{v \in V_R} P_v(\chi_m^2 < T_R) \to 0, \tag{5.41}$$

and under assumption of Theorem 3.3 (2),

$$\sup_{v\in V_R} P_v(\chi_m^2 < T_R) \le \Phi(\sqrt{2\log\alpha_R^{-1}} - \rho_R) + o(1). \tag{5.42}$$

It is easily seen that the function $f(h) = P_{hv}(\chi_m^2 < T)$ decreases in $h > 0$ and it suffices to consider the case when

$$\rho^2(v) = \sum_{i=1}^{\infty} v_i^2 = \rho_R^2. \tag{5.43}$$

Set

$$\rho_m^2(v) = \sum_{i=1}^{m} v_i^2, \quad \lambda_{m,v} = \frac{\chi_m^2 - m - \rho_m^2(v)}{\sigma_m(v)}, \quad \sigma_m^2(v) = 2m + 4\rho_m^2(v).$$

Applying Lemma 5.4 we see that the random variables $\lambda_{m,v}$ are asymptotically standard Gaussian in P_v-probability, uniformly over $v \in l^2$. Therefore

$$\begin{aligned} P_v(\chi_m^2 < T_R) &= P_v\left(\lambda_{m,v} < \frac{T_R - m - \rho_m^2(v)}{\sigma_m(v)}\right) = \Phi\left(\frac{T_R - m - \rho_m^2(v)}{\sigma_m(v)}\right) + o(1) \\ &= \Phi\left(\frac{2\log\alpha_R^{-1} - \rho_m^2(v) + m\log(1+z_R^2)}{\sigma_m(v)}\right) + o(1). \end{aligned} \tag{5.44}$$

Let us estimate $\rho_m^2(v)$ for $v \in V_R$. By (5.43) we have

$$\rho_m^2(v) = \rho_R^2 - \sum_{i=m+1}^{\infty} v_i^2, \quad \sum_{i=m+1}^{\infty} v_i^2 \le m^{-2\sigma} \sum_{i=m+1}^{\infty} i^{2\sigma} v_i^2 \le m^{-2\sigma} R^2.$$

This yield

$$\rho_R^2 \ge \rho_m^2(v) \ge \rho_R^2 - m^{-2\sigma} R^2. \tag{5.45}$$

Under assumptions of Theorem 3.3 using (5.40) we have

$$m^{-2\sigma} R^2 = o(\log\alpha_R^{-1}), \ \ m\log(1+z_R^2) = o(\log\alpha_R^{-1}), \ \ \sigma_m(v) \sim 2\rho_m(v) \sim 2\rho_R. \tag{5.46}$$

For any $\delta > 0$, if $\liminf 2(\log\alpha_R^{-1})/\rho_R^2 < 1 - \delta$, then (5.44), (5.46) yield (5.41).

Under assumptions of Theorem 3.3 (2) it suffices to assume $\sqrt{2\log\alpha_R^{-1}} - \rho_R = O(1)$. In this case using relations (3.25), (5.40), (5.45), (5.46) we have

$$\begin{aligned} 2\log\alpha_R^{-1} - \rho_m^2(v) &= (\sqrt{2\log\alpha_R^{-1}} - \rho_R)(\sqrt{2\log\alpha_R^{-1}} + \rho_R) + o(\rho_R) \\ &= (2\sqrt{2\log\alpha_R^{-1}} + O(1))(\sqrt{2\log\alpha_R^{-1}} - \rho_R); \\ m\log(1+z^2) &\asymp m\log((\log\alpha_R^{-1})/m) = o(\sqrt{\log\alpha_R^{-1}}). \end{aligned}$$

since $\log\alpha_R^{-1}/m \to \infty$. In view of the first relation (3.25), (5.44) and the third relation (5.46), these yield relations (5.42). $\square$

5.6 Proof of Theorem 3.4

5.6.1 Lower bounds

Set $\tilde{\rho}_R = n_R^{-\tau}\rho_R$. It suffices to consider the case

$$\sqrt{2(\log n_R + \log \alpha_R^{-1})} - \tilde{\rho}_R = O(1). \tag{5.47}$$

This yields

$$\log \alpha_R^{-1} = \tilde{\rho}_R^2/2 - \log n_R + O(\tilde{\rho}_R) \to \infty, \quad \tilde{\rho}_R \geq \sqrt{2\log n_R} + O(1) \to \infty. \tag{5.48}$$

Since $\beta(V_R, \alpha)$ decreases in $\alpha > 0$, it follows from (2.25) that, for any $\alpha_R \to 0$,

$$\beta(V_R, \alpha_R) \geq \Phi(\sqrt{2\log m_R} - m_R^{-\tau}\rho_R) + o(1).$$

If $\log \alpha_R^{-1} = o(\sqrt{\log m_R})$, then

$$m_R \to \infty, \quad \log n_R = \log m_R + o(1), \quad \tilde{\rho}_R = m_R^{-\tau}\rho_R + o(1). \tag{5.49}$$

This yields the required lower bounds.

Let $\{e_j,\ j = 1, 2, ...\}$ be the standard basis in l^2, i.e.,

$$e_j = \{e_{ij}\}, \quad e_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad 1 \leq i < \infty.$$

Observe that

$$v^* = \tilde{\rho}_R e_{n_R} \in V_R, \quad |v^*| = \tilde{\rho}_R. \tag{5.50}$$

It follows from (1.10), (5.50) that

$$\beta(V_R, \alpha_R) \geq \Phi(\sqrt{2\log \alpha_R^{-1}} - \tilde{\rho}_R) + o(1).$$

If $(\log m_R)^2 = o(\log \alpha_R^{-1})$, then this yields the required lower bounds.

Therefore it suffices to assume, for any $B > b > 0$ and R large enough,

$$b\sqrt{\log m_R} < \log \alpha_R^{-1} < B(\log m_R)^2, \tag{5.51}$$

which yields (5.49). Denote $n_1 = [n_R(1 - 1/\log(n_R\alpha_R^{-1}))]$, where $[t]$ is an integer part of t. Let us take the collections

$$V_{n_R} = \{v_j = \rho_R j^{-\tau} e_j,\ n_1 + 1 \leq j \leq n_R\}, \ \ \tilde{V}_{n_R} = \{\tilde{v}_j = \rho_R n_1^{-\tau} e_j,\ n_1 + 1 \leq j \leq n_R\}.$$

Clearly, $V_{n_R} \subset V_R$ for any $q > 0$ and, since $\tilde{v}_{ij} \geq v_{ij}\ \forall i$, it is easily seen that $\beta(V_{n_R}, \alpha_R) \geq \beta(\tilde{V}_{n_R}, \alpha_R)$ (compare with Proposition 2 and Lemma 3.1 in [14]).

Therefore it suffices to verify that

$$\beta(\tilde{V}_{n_R}, \alpha_R) \geq \Phi(\sqrt{2(\log m_R + \log \alpha_R^{-1})} - \tilde{\rho}_R) + o(1).$$

Set

$$k = n_R - n_1 = \frac{n_R}{\log(n_R \alpha_R^{-1})} + o(1), \quad u_R = \rho_R n_1^{-\tau}; \quad n_R > k > \frac{c n_R}{(\log n_R)^2}, \qquad (5.52)$$

for some $c > 0$ under (5.51). Take the priors

$$\pi_R = k^{-1} \sum_{i=1}^{k} \delta_{\tilde{v}_{i+n_1}},$$

where δ_v is Dirac mass at the point $v \in l^2$. This yields $\pi_R(\tilde{V}_{n_R}) = 1$. Under (5.47), (5.51), (5.52) we have

$$u_R = \tilde{\rho}_R + o(1), \quad \sqrt{\log k + \log \alpha_R^{-1}} = \sqrt{\log m_R + \log \alpha_R^{-1}} + o(1), \qquad (5.53)$$

and it suffices to verify that, for the Bayesian hypothesis testing problem (5.2),

$$\beta(\alpha_R, P_{\pi_R}) \geq \Phi(\sqrt{2(\log k + \log \alpha_R^{-1})} - u_R) + o(1); \quad P_{\pi_R} = \frac{1}{k} \sum_{i=1}^{k} P_{\tilde{v}_i + n_1} \qquad (5.54)$$

(compare with Section 5.1). The likelihood ratio is of the form

$$L_{\pi_R} = \frac{dP_{\pi_R}}{dP_0}(X) = \frac{1}{k} \sum_{i=1}^{k} \exp(-u_R^2/2 + u_R X_{i+n_1}). \qquad (5.55)$$

Lemma 5.5 *Let $\xi \sim \mathcal{N}(0,1)$, $\sqrt{2 \log k + \log \alpha_R^{-1}} - u_R = O(1)$, $k \to \infty$. Then*

$$L_{\pi_R} = k^{-1} \exp(u_R^2/2 + u_R \xi_R) + \eta_R + o(1), \quad \xi_R \to \xi \sim \mathcal{N}(0,1), \qquad (5.56)$$

in P_{π_R}-probability, where $\eta_R = \Phi(\sqrt{2 \log k} - u_R)$.

Proof. It is easily seen that for any $t \in R$ one has

$$P_{\pi_R}(L_{\pi_R} < t) = k^{-1} \sum_{i=1}^{k} P_{\tilde{v}_{i+n_1}}(L_{\pi_R} < t) = P_{w^*}(L_{\pi_R} < t), \quad w^* = \tilde{v}_{n_1+1}$$

and it suffices to verify that (5.56) holds in P_{w^*}-probability. On the other hand, in P_{w^*}-probability, the random variables L_{π_R} are distributed as

$$\frac{1}{k} \exp(u_R^2/2 + u_R \xi_1) + \frac{k-1}{k} \eta_{R,k-1}, \quad \eta_{R,k-1} = \frac{1}{k-1} \sum_{i=2}^{k} \exp(-u_R^2/2 + u_R \xi_i),$$

where ξ_i are i.i.d. standard Gaussian. It follows from [13], Proposition 4.10 and Corollary 4.5 that, in probability,

$$\eta_{R,k-1} = \Phi(\sqrt{2 \log(k-1)} - u_R) + o(1)$$

(to apply Corollary 4.5 one can take $u_\varepsilon = u_R$, $w_{\varepsilon,i} = u_\varepsilon - D_\varepsilon = \sqrt{2 \log(k-1)}$, $p_{\varepsilon,i} = (k-1)^{-1}$ for $2 \leq i \leq k$ and $w_{\varepsilon,i} = \infty$ in other cases). This yields the lemma. □

Set

$$H_R = u_R - \sqrt{2 \log k}, \quad E_{\pi,R} = -\log k + u_R^2/2, \quad \lambda_R = (\log L_{\pi_R} - E_{\pi,R})/u_R. \qquad (5.57)$$

Under (5.47) one has $H \geq O(1)$, and $H_R = O(1)$ for $\log \alpha_R^{-1} \asymp (\log m_R)^{1/2}$.

Lemma 5.6 *Assume (5.56) with* $\eta_R \geq \Phi(\sqrt{2\log k} - u_R)$, $\eta_R = O(1)$ *in* P_{π_R}-*probability. Then one has under (5.47), in* P_{π_R}-*probability,*

(1) if $H_R \to \infty$, *then* $\lambda_R \to \xi \sim \mathcal{N}(0,1)$;

(2) if $H_R \to H \in R^1$, *then* $\lambda_R \to \xi_{-H}$, *where* ξ_{-H} *is determined by (5.17).*

Proof. Let $H_R \to \infty$. Then

$$E_{\pi,R} = H_R\sqrt{2\log k} + H_R^2/2 \gg u_R, \quad E_{\pi,R} + tu_R \to \infty \quad \forall\, t \in R.$$

Using (5.56), for any $t \in R$ we have

$$\begin{aligned}
&P_{\pi_R}(\lambda_R < t) = P_{\pi_R}\left(L_{\pi_R} < \exp(E_{\pi,R} + tu_R)\right) = \\
&P\left(\exp(E_{\pi,R} + u_R\xi_R) < \exp(E_{\pi,R} + tu_R) - \eta_R + o(1)\right) = \\
&P(e^{u_R\xi_R} < e^{tu_R}(1+\kappa_R)) = P(\xi_R < t + \zeta_R) = \Phi(t) + o(1),
\end{aligned} \tag{5.58}$$

where, in P_{π_R}-probability,

$$\kappa_R = \frac{-\eta_R + o(1))}{\exp(E_{\pi,R} + tu_R)} = o(1), \quad \zeta_R = \frac{\log(1+\kappa_R)}{u_R} = o(1). \tag{5.59}$$

Let $H_R \to H \in R^1$. Then

$$E_{\pi,R} + tu_R = u_R(H_R + t) + O(1) \to \begin{cases} \infty, & t > -H, \\ -\infty, & t < -H. \end{cases}$$

For any $t > -H$, using (5.56) analogously to above we have (5.58), (5.59). If $t < -H$, then $P_{\pi_R}(\lambda_R < t) = P_{\pi_R}\left(L_{\pi_R} < \exp(E_{\pi,R} + tu_R)\right) \to 0$. $\square$

Let us return to the proof of (5.54). By (5.48) it suffices to consider the case

$$\log \alpha_R^{-1} = E_{\pi,R} + a_R u_R, \quad a_R \to a \in R^1,$$

and either $H_R \to \infty$ or $H_R \to H \in R^1$. Observe that if $H_R \to H$, then we have

$$\log \alpha_R^{-1} = (H_R + a_R)u_R - H_R^2/2 = (H+a)u_R + o(u_R) \to \infty.$$

This yields $a + H \geq 0$, and if $a + H = 0$, then we go to the case $\log \alpha_R^{-1} = o(\sqrt{\log m_R})$ that was considered above. Therefore it suffices to assume $a + H > 0$, if $H_R = O(1)$. Taking into account Lemmas 5.5, 5.6 and Remark 5.1 we can apply Lemma 5.2 with $\sigma_{\pi,R} = u_R$ and $E_{\pi,R}$ defined by (5.57). This yields the relation $T_{R,\alpha_R} = \log \alpha_R^{-1} + o(u_R)$ for $(1-\alpha_R)$-quantile of the log-likelihood ratio $t_R = \log L_{\pi_R}$. Furthermore, using Lemma 5.6 we have

$$\begin{aligned}
&\beta(\alpha_R, P_{\pi_R}) = P_{\pi_R}(t_R \leq T_{R,\alpha_R}) = P_{\pi_R}\left(E_{\pi,R} + \lambda_R\sigma_{\pi,R} \leq \log \alpha_R^{-1} + o(u_R)\right) = \\
&P_{\pi_R}\left(\lambda_R \leq \frac{\log \alpha_R^{-1} - E_{\pi,R}}{u_R} + o(1)\right) = P_{\pi_R}(\lambda_R \leq a_R + o(1)) \to \Phi(a).
\end{aligned}$$

On the other hand, under (5.47), (5.53) we have

$$\begin{aligned}
a_R \;&=\; \frac{\log \alpha_R^{-1} + \log k - u_R^2/2}{u_R} = \left(\sqrt{2(\log \alpha_R^{-1} + \log k)} - u_R\right) \\
&\times \left(\frac{\sqrt{2(\log \alpha_R^{-1} + \log k)} + u_R}{2u_R}\right) \sim \sqrt{2(\log \alpha_R^{-1} + \log k)} - u_R.
\end{aligned}$$

This yields (5.54). $\square$

5.6.2 Upper bounds

It suffices to consider the case $q = \infty$ that corresponds to the "widest" alternative. For tests (3.28) and α_R small enough, using (1.8) we get the first relation (3.29):

$$\alpha(\psi_{R,\alpha_R}) = P_0(\mathcal{X}_{R,\alpha_R}) \le \sum_{i=1}^{\infty} P_0(|X_i| > T_{\alpha_R,i}) \le 2\sum_{i=1}^{\infty} \Phi(-T_{\alpha_R,i})$$
$$\sim \frac{2}{\sqrt{2\pi}} \sum_{i=1}^{\infty} \frac{\exp(-T^2_{\alpha_R,i}/2)}{T_{\alpha_R,i}} \le \frac{\alpha_R}{\sqrt{\pi}} \sum_{i=1}^{\infty} \frac{1}{i\log(i+1)\sqrt{\log i + \log \alpha_R^{-1}}} < \alpha_R.$$

Next, by construction of the tests (3.28) and since $T_i = T_{\alpha_R,i}$ increases in i, one has, for any $v = \{v_i\} \in l^2$ and any $m > 0$,

$$\beta(\psi_{R,\alpha_R}, v) \le \inf_i P_v(|X_i| \le T_i) \le \Phi(\min_{i\le m}(T_i - |v_i|)) \le \Phi(T_{[m]} - \max_{i\le m}|v_i|). \quad (5.60)$$

Lemma 5.7 *Set $m = m_R = (R/\rho_R)^{1/(s-\tau)}$, $n_R = [m_R]$, $n_{R,1} = n_R + 1 = m_R(1+\delta)$ for some $\delta > 0$. Then one has the inequality*

$$\inf_{v\in V_R} \max_{i\le m_R} |v_i| \ge \rho_R n_R^{-\tau}. \quad (5.61)$$

Proof of the lemma (compare with Lemma 4.2 in [13]). Fix $v = (v_1, ..., v_n, ...) \in V_R$ and let $i \ge n_{R,1}$. Since

$$\sup_i i^\tau |v_i| \ge \rho_R, \quad \sup_i i^s|v_i| \le R, \quad s > \tau \ge 0,$$

we have

$$\frac{i^\tau|v_i|}{\rho_R} = \frac{i^s|v_i|}{\rho_R} i^{\tau-s} \le \frac{R}{\rho_R} i^{\tau-s} = \left(\frac{m}{i}\right)^{s-\tau} \le \left(\frac{1}{1+\delta}\right)^{s-\tau}.$$

Therefore the supremum $\sup_i i^\tau|v_i|$ is attained in some $i_0 \le n_R$, and we have

$$\max_{i\le m_R} |v_i| \ge |v_{i_0}| = (|v_{i_0}|i_0^\tau) i_0^{-\tau} \ge \rho_R n_R^{-\tau}. \quad \square$$

Since $T_{\alpha_R,n} = \sqrt{2(\log\alpha_R^{-1} + \log n)} + o(1)$, using (5.60) and (5.61) we have the second relation (3.29):

$$\beta(\psi_{R,\alpha_R}, V_R) \le \Phi(T_{\alpha_R,n_R} - n_R^{-\tau}\rho_R) = \Phi(\sqrt{2(\log n_R + \log\alpha_R^{-1})} - n_R^{-\tau}\rho_R) + o(1). \quad \square$$

Remark 5.4 It follows from (5.50) and Lemma 5.7 that

$$n_R^{-\tau}\rho_R \ge \inf_{v\in V_R} |v| \ge \inf_{v\in V_R} \sup_i |v_i| \ge \inf_{v\in V_R} \max_{i\le m_R} |v_i| \ge \rho_R n_R^{-\tau}.$$

By (1.10), this yields relation (1.26).

5.7 Proof of Theorem 3.5 (1)

For rates (3.38), taking $C_0 = \sqrt{2}$ we get Theorem 3.5 (1) directly from inequality (1.25). Therefore we need to consider the case $\delta_\varepsilon = o(1)$. It suffices to assume $r_\varepsilon \asymp r^*_{\varepsilon,\alpha_\varepsilon}$.

Let us take integer-valued family $m = m_\varepsilon$, an integer $d > \sigma$ and d-differentiable function $\phi(t),\ t \in R^1$ supported on $[0,1]$, $\|\phi\|_2 = 1$. Set

$$\phi_{\varepsilon,i}(t) = m^{1/2}\phi(mt - i + 1),\ i = 1,...,m; \quad f_\varepsilon(t,\theta) = \sum_{i=1}^{m} \theta_i \phi_{\varepsilon,i}(t);,\ \theta \in R^m.$$

Clearly, the functions $\phi_{\varepsilon,i}(t)$ have disjoint supports, the functions $f_\varepsilon(t,\theta)$ are supported on $[0,1]$. One can verify that

$$\|f_\varepsilon(\cdot,\theta)\|_2 = |\theta|; \quad \|f_\varepsilon(\cdot,\theta)\|_{\sigma,2} \le cm^\sigma|\theta| \tag{5.62}$$

where $c = c(\phi, d)$ is a positive constant and $|\cdot|$ is the Euclidean norm in R^m (see [13], inequality (2.80) and Lemma 3.8, for the inequality in (5.62)). Let us take

$$m = m_\varepsilon \sim (2cr_\varepsilon/H)^{-1/\sigma} \to \infty, \tag{5.63}$$

by (3.37), where c is the constant from (5.62). Under the assumption $\delta_\varepsilon = o(1)$ this yields

$$\log \alpha_\varepsilon^{-1} = o(m_\varepsilon). \tag{5.64}$$

Introduce the set

$$F_{\varepsilon,m} = \{f_\varepsilon(\cdot,\theta)\ :\ \theta \in R^m,\ |\theta| = r_\varepsilon\},$$

corresponding to the sphere of radius r_ε in R^m. It follows from (5.62) that $F_{\varepsilon,m} \subset F_\varepsilon$. This yields $\beta_\varepsilon(F_{\varepsilon,m},\alpha_\varepsilon) \le \beta_\varepsilon(F_\varepsilon,\alpha_\varepsilon)$. On the other hand, passing to random variables $X_i = \varepsilon^{-1}\int_0^1 \phi_{\varepsilon,i}(t)dX_\varepsilon(t)$ and to parameters $v_i = \varepsilon^{-1}(f,\phi_{\varepsilon,i}),\ i = 1,...,m$ we see that

$$\beta_\varepsilon(F_{\varepsilon,m},\alpha_\varepsilon) = \beta(S^{m-1}(r_\varepsilon/\varepsilon),\alpha_\varepsilon);$$

the last quantity corresponds to testing of the hypothesis $v = 0$ against alternative $v \in S^{m-1}(r_\varepsilon/\varepsilon)$ under n-dimensional Gaussian model $X = v + \xi,\ v \in R^m$; here $S^{m-1}(\rho)$ is the sphere of radius ρ in R^m. It is well known (see [13], Example 2.2 in Section 2.3) that

$$\beta(S^{m-1}(\rho),\alpha) = G(T_{m,\alpha},\rho^2),$$

where $G(t,\rho^2) = P_v(\chi^2_m < t),\ v \in R^m,\ |v| = \rho$ is the distribution function of non-central chi-square distribution with m degrees of freedom and parameter of non-centrality ρ^2, and $T_{m,\alpha}$ is $(1-\alpha)$-quantile of the central chi-square distribution with m degrees of freedom, i.e., $P_0(\chi^2_m > T_{m,\alpha}) = \alpha$. Therefore it suffices to verify that there exists $C_0 > 0$ such that if

$$\limsup r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon} < C_0, \tag{5.65}$$

then $\beta(S^{m-1}(\rho),\alpha_\varepsilon) = P_v(\chi^2_m < T_{m,\alpha_\varepsilon}) \to 1$ uniformly over $v \in R^m,\ |v| = r_\varepsilon/\varepsilon$.

For all $v \in R^m$, $|v| = \rho$ one has

$$E_v(\chi^2_m) = m + \rho^2, \quad \mathrm{Var}_v(\chi^2_m) = 2m + 4\rho^2.$$

Using the Chebyshev inequality we see that (5.65) follows from the relation

$$(T_{m_\varepsilon,\alpha_\varepsilon} - m_\varepsilon - \rho^2_\varepsilon)/\sqrt{2m_\varepsilon + 4\rho^2_\varepsilon} \to \infty, \quad \rho_\varepsilon = r_\varepsilon/\varepsilon. \tag{5.66}$$

Since $\delta_\varepsilon = o(1)$, under (5.63) we have $\rho^2_\varepsilon = o(m_\varepsilon)$. Recalling (5.19), let us take $z = z_\varepsilon$ such that

$$\log \alpha_\varepsilon^{-1} = m_\varepsilon(z^2 - \log(1+z^2))/2 + o(z^2 m_\varepsilon^{1/2}).$$

Using (5.64) we see that

$$z = o(1),\ T_{m_\varepsilon,\alpha_\varepsilon} = m_\varepsilon(1+z^2),\ \log \alpha_\varepsilon^{-1} \sim m_\varepsilon z^4/4 \tag{5.67}$$

and the left-hand side of (5.66) is of the rate $(m_\varepsilon z^2 - \rho^2_\varepsilon)/\sqrt{2m_\varepsilon}$. Therefore it suffices to verify that $\limsup \rho^2_\varepsilon/m_\varepsilon z^2 < 1$. By (5.67) this is equivalent $\limsup r^4_\varepsilon/\varepsilon^4 m_\varepsilon \log \alpha_\varepsilon^{-1} < 4$. By (5.63) the last relation follows from (5.65) with $C_0 = (4(H/2c)^{1/\sigma})^{\sigma/(4\sigma+1)}$. □

5.8 Proof of Theorem 3.5 (2)

The proof follows to the scheme of Section 5.5. We consider chi-square tests of form (3.40). Take $z = z_\varepsilon$, $T_\varepsilon = m_\varepsilon(1+z^2)$ such that $L(T_\varepsilon, m_\varepsilon) = \log \alpha_\varepsilon^{-1}$; the function $L(T,m)$ is defined by (5.38). Analogously to (5.39) using (5.18), (5.19) we have

$$T_\varepsilon = m_\varepsilon + 2\log \alpha_\varepsilon^{-1} + m_\varepsilon \log(1+z^2_\varepsilon), \quad T_\varepsilon \geq T_{m_\varepsilon,\alpha_\varepsilon}.$$

The letter relation yields $P_{\varepsilon,0}(\chi^2_{m_\varepsilon} \geq T_\varepsilon) \leq \alpha_\varepsilon$. It suffices to verify that, uniformly over $f \in F_\varepsilon$,

$$P_{\varepsilon,f}(\chi^2_{m_\varepsilon} < T_\varepsilon) \to 0, \quad \text{as } \varepsilon \to 0. \tag{5.68}$$

It suffices to consider the case

$$\|f\|_2 = r_\varepsilon \asymp r^*_{\varepsilon,\alpha}. \tag{5.69}$$

Let us consider the orthonormal projection to the subspace that consists of the step functions

$$Pr_m f = \sum_{j=1}^m f_{j,m} \mathbf{1}_{\delta_{\mathbf{j},\mathbf{m}}}, \quad f_{j,m} = m\int_{\delta_{j,m}} f(t)dt.$$

Set

$$\rho_\varepsilon(f) = \|Pr_{m_\varepsilon} f\|_2/\varepsilon, \quad \sigma^2_\varepsilon(f) = 2m + 4\rho^2_\varepsilon(f), \quad \lambda_{\varepsilon,f} = (\chi^2_{m_\varepsilon} - m_\varepsilon - \rho^2_\varepsilon(f))/\sigma_\varepsilon(f).$$

Applying Lemma 5.4 we see that the random variables $\lambda_{\varepsilon,f}$ are asymptotically standard Gaussian in $P_{\varepsilon,f}$-probability, uniformly over $f\in L_2(0,1)$. Therefore

$$P_{\varepsilon,f}(\chi^2_{m_\varepsilon}<T_\varepsilon)=\Phi\left(\frac{2\log\alpha_\varepsilon^{-1}-\rho_\varepsilon^2(f)+m_\varepsilon\log(1+z_\varepsilon^2)}{\sigma_\varepsilon(f)}\right)+o(1). \tag{5.70}$$

Note the following statement: for any $\sigma_0>0$ there exist constants $B_1>0,\ B_2>0$ such that for any $p\in[1,\infty],\ \sigma\in(0,\sigma_0),\ f\in L_p(0,1),\ \|f\|_{\sigma,p}<\infty$ and any integer $m\geq 1$ one has

$$\|Pr_mf\|_p\geq B_1\|f\|_p-B_2m^{-\sigma}\|f\|_{\sigma,p}. \tag{5.71}$$

(this corresponds to Proposition 2.16 in [13]). For any $B>0$ we can take $C_1>0$ such that

$$B_1C_1-HB_2>B. \tag{5.72}$$

Set $\rho_\varepsilon^*=r^*_{\varepsilon,\alpha_\varepsilon}/\varepsilon$. Recall that we take the quantities $m_\varepsilon\sim(r^*_{\varepsilon,\alpha_\varepsilon})^{-1/\sigma}$. Using (5.71) for $p=2$ and by definitions (1.16) of the set F_ε for small enough $\varepsilon>0$ and uniformly over $f\in F_\varepsilon$,

$$\|Pr_{m_\varepsilon}f\|_2\geq B_1r_\varepsilon-HB_2r_\varepsilon^*(1+o(1))>Br_\varepsilon^*,\quad \rho_\varepsilon(f)>B\rho_\varepsilon^*. \tag{5.73}$$

First, suppose $\delta_\varepsilon=o(1)$, (3.36) and take $B^2>2$ in (5.72). Note that $(\rho_\varepsilon^*)^4\sim m_\varepsilon\log\alpha_\varepsilon^{-1}$, and under (5.69),

$$\begin{aligned}
&z_\varepsilon^2-\log(1+z_\varepsilon^2)=2m_\varepsilon^{-1}\log\alpha_\varepsilon^{-1}\asymp\delta_\varepsilon^{(4\sigma+2)/(4\sigma+1)}\to 0,\quad z_\varepsilon\to 0,\\
&(\rho_\varepsilon^*)^{-2}\log\alpha_\varepsilon^{-1}\sim(\rho_\varepsilon^*)^2/m_\varepsilon\sim\delta_\varepsilon^{(2\sigma+1)/(4\sigma+1)}\to 0,\quad \sigma_\varepsilon(f)\sim\sqrt{2m_\varepsilon},\\
&m_\varepsilon\log(1+z_\varepsilon^2)/\sigma_\varepsilon(f)\sim z_\varepsilon^2\sqrt{m_\varepsilon/2}\sim\sqrt{m_\varepsilon(z_\varepsilon^2-\log(1+z_\varepsilon^2))}=\sqrt{2\log\alpha_e^{-1}};\\
&(\rho_\varepsilon^*)^2/\sigma_\varepsilon(f)\sim(\rho_\varepsilon^*)^2/\sqrt{2m_\varepsilon}\sim\sqrt{(\log\alpha_e^{-1})/2};\ 2\log\alpha_e^{-1}/\sigma_\varepsilon(f)=o(\sqrt{\log\alpha_\varepsilon^{-1}}).
\end{aligned}$$

Therefore the argument in the braces (5.70) is $\sqrt{(\log\alpha_\varepsilon^{-1})/2}(2-B^2+o(1))\to-\infty$. This yields (5.68).

Next, suppose $\delta_\varepsilon\to\delta>0$ and let the rates be defined by (3.38). We have

$$(\rho_\varepsilon^*)^2=\log\alpha_\varepsilon^{-1},\quad z_\varepsilon^2-\log(1+z_\varepsilon^2)=2m_\varepsilon^{-1}\log\alpha_\varepsilon^{-1}\sim 2\delta_\varepsilon^{(4\sigma+2)/(4\sigma+1)},\quad \sigma(f)\asymp\rho_\varepsilon^*.$$

This yields $\liminf z>z_0(\delta,\sigma)>0$ and there exists a constant $M=M(\delta,\sigma)>0$ such that, for small enough $\varepsilon>0$,

$$m_\varepsilon\log(1+z_\varepsilon^2)<Mm_\varepsilon(z_\varepsilon^2-\log(1+z_\varepsilon^2))=2M\log(\alpha_\varepsilon^{-1}).$$

We get

$$2\log\alpha_\varepsilon^{-1}-\rho_\varepsilon^2(f)+m_\varepsilon\log(1+z_\varepsilon^2)\leq(\rho_\varepsilon^*)^2(2-B^2+2M).$$

Therefore taking $B^2>2+2M$ we see that the argument in the braces (5.70) tends to $-\infty$. This yields (5.68). $\square$

5.9 Proof of Theorem 2.1

First, using relation (1.8) we have

$$\alpha_\varepsilon(\psi_\varepsilon^*) \le \sum_{l=J_{\varepsilon,0}}^{J_{\varepsilon,1}} \sum_{j=1}^{m_l} P_{\varepsilon,0}(|X_{j,m_l}| \ge T_l) = 2 \sum_{l=J_{\varepsilon,0}}^{J_{\varepsilon,1}} m_l \Phi(-T_l) \asymp \sum_{l=J_{\varepsilon,0}}^{J_{\varepsilon,1}} l^{-3/2} \to 0.$$

Next, note that if

$$\phi,\ f \in L_2(0,1), \quad \|\phi\| = 1, \quad X = \varepsilon^{-1} \int_0^1 \phi(t) dX_\varepsilon(t), \quad v = \varepsilon^{-1}(f, \phi), \quad T > 0,$$

then we have

$$P_{\varepsilon,f}(|X| < T) = \Phi(T - |v|) - \Phi(-T - |v|) \le \Phi(T - |v|).$$

Setting

$$\phi_{j,m_l} = m_l^{1/2} \delta_{j,m_l}, \quad v_{jl}(f) = \varepsilon^{-1} m_l^{1/2} \int_{\delta_{j,m_l}} f(t) dt,$$

observe that

$$\max_{1 \le j \le m_l} |v_{jl}(f)| = \varepsilon^{-1} m_l^{-1/2} \|Pr_{m_l} f\|_\infty,$$

and for any $f \in L_2(0,1)$ we have

$$\beta_\varepsilon(\psi_\varepsilon^*, f) \le \min_{J_{\varepsilon,0} \le l \le J_{\varepsilon,1}} \min_{1 \le j \le m_l} P_{\varepsilon,f}(|X_{j,m_l}| \le T_l) \le$$
$$\min_{J_{\varepsilon,0} \le l \le J_{\varepsilon,1}} \min_{1 \le j \le m_l} \Phi(T_l - |v_{jl}(f)|) = \min_{J_{\varepsilon,0} \le l \le J_{\varepsilon,1}} \Phi(T_l - \varepsilon^{-1} m_l^{-1/2} \|Pr_{m_l} f\|_\infty). \quad (5.74)$$

Inequality (1.15) yields the embedding

$$F_\varepsilon(r_\varepsilon, H, \sigma, q) \subset F_\varepsilon(r_\varepsilon, H_1, \eta, \infty), \quad H_1 = cH, \quad \eta = \sigma - 1/q > 0. \quad (5.75)$$

By (5.75) it suffices to consider the case $q = \infty$ with the change (σ, H) by (η, H_1).

Let $f \in F_\varepsilon = F_\varepsilon(r_\varepsilon, H_1, \eta, \infty)$. Let us take $l = l_\varepsilon(\eta)$ such that $m_l = 2^l \sim (r_{\varepsilon,\infty}^*)^{-1/\eta}$, where $r_{\varepsilon,\infty}^*$ are determined by (2.10). Clearly,

$$J_{\varepsilon,0} \le l \sim h_1(\eta) \log \varepsilon^{-1} \le J_{\varepsilon,1}, \quad T_l \sim h_2(\eta) \sqrt{\log \varepsilon^{-1}}.$$

where

$$h_1(\eta) = 2/(2\eta + 1) \log 2, \quad h_2(\eta) = 2(2\eta + 1)^{-1/2}. \quad (5.76)$$

Using inequality (5.71) for $p = \infty$ and taking C_1 such that

$$B_1 C_1 - B_2 H_1 = B, \quad B > C h_2(\eta), \quad C > 1, \quad (5.77)$$

we see that, for ε small enough uniformly over $f \in F_\varepsilon$,

$$\varepsilon^{-1} m_l^{-1/2} \|Pr_{m_l} f\|_\infty \ge \varepsilon^{-1} (r_{\varepsilon,\infty}^*)^{1+1/2\eta} (B + o(1)) = (B + o(1)) \sqrt{\log \varepsilon^{-1}} > C T_l.$$

Thus the argument of the function Φ in (5.74) is no larger then $(1 - C) T_l \to -\infty$. This yields $\beta_\varepsilon(\psi_\varepsilon^*, F_\varepsilon) \to 0$. $\square$

5.10 Proof of Theorem 3.6

Set $\tilde{m}_\varepsilon = (l+1)m_\varepsilon$. Analogously to Section 5.8, we take $z = z_\varepsilon$, $T_\varepsilon = \tilde{m}_\varepsilon(1+z^2)$ such that $L(T_\varepsilon, \tilde{m}_\varepsilon) = \log \alpha_\varepsilon^{-1}$. We have

$$T_\varepsilon = \tilde{m}_\varepsilon + 2\log\alpha_\varepsilon^{-1} + \tilde{m}_\varepsilon \log(1+z_\varepsilon^2), \quad T_\varepsilon \geq T_{\tilde{m}_\varepsilon, \alpha_\varepsilon}.$$

The letter relation yields $P_{\varepsilon,0}(\chi^2_{m_\varepsilon,l} \geq T_\varepsilon) \leq \alpha_\varepsilon$. Let us consider the orthonormal projection of the space $L_2(0,1)$ to the subspace that consists of the piecewise polynomial of degree $\leq l$ (no necessary continuous) functions that correspond to the partition of $[0,1)$ to sub-intervals $\delta_{j,m} = [(j-1)/m, j/m]$, $j = 1, ..., m$

$$Pr_{m,l} f = \sum_{j=1}^{m}\sum_{k=0}^{l} f_{jk,m}\phi_{jk,m}, \quad f_{jk,m} = (f, \phi_{jk,m}).$$

Set

$$\rho_\varepsilon(f) = \|Pr_{m_\varepsilon,l} f\|_2/\varepsilon, \quad \sigma_\varepsilon^2(f) = 2\tilde{m}_\varepsilon + 4\rho_\varepsilon^2(f), \quad \lambda_{\varepsilon,f} = (\chi^2_{m_\varepsilon,l} - \tilde{m}_\varepsilon - \rho_\varepsilon^2(f))/\sigma_\varepsilon(f).$$

Applying Lemma 5.4 we see that the random variables $\lambda_{\varepsilon,f}$ are asymptotically standard Gaussian in $P_{\varepsilon,f}$-probability, uniformly over $f \in L_2(0,1)$. Analogously to (5.70) we have

$$P_{\varepsilon,f}(\chi^2_{\tilde{m}_\varepsilon,l} < T_\varepsilon) = \Phi\left(\frac{2\log\alpha_\varepsilon^{-1} - \rho_\varepsilon^2(f) + \tilde{m}_\varepsilon \log(1+z_\varepsilon^2)}{\sigma_\varepsilon(f)}\right) + o(1). \tag{5.78}$$

Lemma 5.8 *There exists $B = B(\sigma) > 0$ such that for any $f \in L_2(0,1)$, $\|f\|_{\sigma,2} < \infty$, and any integers $m > 0$, $l \geq 0$, one has*

$$\|Pr_{m,l} f\|_2^2 \geq \|f\|_2^2 - B^2 m^{-2\sigma}\|f\|_{\sigma,2}^2. \tag{5.79}$$

Proof of the lemma. Observe the following approximation property.

Proposition 5.1 *Let $\sigma = l + \tau$, $\tau \in (0,1]$. There exists $B = B(\sigma) > 0$ such that for any $f \in L_2(0,1)$, $\|f\|_{\sigma,2} < \infty$, one can find a piecewise polynomial $p_{m,l}$ of degree $\leq l$ satisfying*

$$\|f - p_{m,l}\|_2 \leq B m^{-\sigma}\|f\|_{\sigma,2}^0. \tag{5.80}$$

Proof of the proposition. Let $\tau \in (0,1)$ and $\|f\|_{\sigma,q}^0$ be defined by (1.14). Let $l \geq 1$. For each $j = 1, ..., m$ let us take polynomials in $t \in \delta_{m,j}$ of degree $\leq l$:

$$p_{j,l}(t,x) = \sum_{s=0}^{l} f^{(s)}(x)\frac{(t-x)^s}{s!}, \quad p_j(t) = m\int_{\delta_{m,j}} p_{j,l}(t,x)dx.$$

Applying the integral Taylor formula,

$$f(t) - p_{j,l}(t,x) = C\int_x^t (f^{(l)}(u) - f^{(l)}(x))(t-u)^{l-1}du, \quad C = ((l-1)!)^{-1},$$

we have, for $x, t \in \delta_{m,j}$,

$$\begin{aligned} f(t) - p_j(t) &= Cm \int_{\delta_{m,j}} \int_x^t (f^{(l)}(u) - f^{(l)}(x))(t-u)^{l-1} du\, dx, \\ |f(t) - p_j(t)| &\leq \frac{C}{m^{l-2}} \int_{\delta_{m,j}} \int_{\delta_{m,j}} \frac{|f^{(l)}(u) - f^{(l)}(x)|}{|u-x|^\tau} |u-x|^\tau du\, dx. \end{aligned}$$

Applying the Cauchy inequality for $t \in \delta_{m,j}$, we get

$$(f(t) - p_j(t))^2 \leq \frac{C^2}{m^{2(l-2)}} \int_{\delta_{m,j}} \int_{\delta_{m,j}} \frac{(f^{(l)}(u) - f^{(l)}(x))^2}{|u-x|^{2\tau}} du\, dx$$

$$\times \int_{\delta_{m,j}} \int_{\delta_{m,j}} |u-x|^{2\tau} du\, dx \leq \frac{C^2}{m^{2\sigma-2}} \int_{\delta_{m,j}} \int_{\delta_{m,j}} \frac{(f^{(l)}(u) - f^{(l)}(x))^2}{|u-x|^{2\tau}} du\, dx.$$

Set $p_{m,l}(t) = p_j(t)$ for $t \in \delta_{m,j}$. We have, for $u = x + h,\ |h| \leq m^{-1}$,

$$\begin{aligned} &\|f - p_{m,l}\|_2^2 = \sum_{j=1}^m \int_{\delta_{m,j}} (f(t) - p_j(t))^2 dt \\ &\leq \frac{C^2}{m^{2\sigma-1}} \sum_{j=1}^m \int_{\delta_{m,j}} \int_{\delta_{m,j}} \frac{(f^{(l)}(u) - f^{(l)}(x))^2}{|u-x|^{2\tau}} du\, dx \\ &\leq \frac{2C^2}{m^{2\sigma-1}} \int_0^{m^{-1}} \left(\int_0^{1-h} \frac{(f^{(l)}(x+h) - f^{(l)}(x))^2}{h^{2\tau}} dx \right) dh \leq \frac{2C^2}{m^{2\sigma}} (\|f\|_{\sigma,2}^0)^2. \end{aligned}$$

If $l = 0,\ \tau \in (0,1)$, then we set $p_j(t) = m \int_{\delta_{m,j}} f(x)dx$ and repeat the estimations.

Let $\sigma = l + 1 > 0$ be an integer and $\|f\|_{\sigma,2}^0$ be defined by (1.12). Then we set $p_j(t) = p_{j,l}(t, x_j),\ x_j = (j-1)/m$. Applying the Taylor formula

$$f(t) - p_{j,l}(t, x_j) = C \int_{x_j}^t f^{(\sigma)}(u)(t-u)^l du, \quad C = 1/l!,$$

and the Cauchy inequality once again we have, for $t \in \delta_{m,j}$,

$$(f(t) - p_j(t))^2 = C^2 \left(\int_{x_j}^t f^{(\sigma)}(u)(t-u)^{\sigma-1} du \right)^2 \leq C^2 m^{-2\sigma+1} \int_{\delta_{m,j}} (f^{(\sigma)}(u))^2 du,$$

$$\|f - p_{m,l}\|_2^2 = \sum_{j=1}^m \int_{\delta_{m,j}} (f(t) - p_j(t))^2 dt \leq C^2 m^{-2\sigma} (\|f\|_{\sigma,2}^0)^2. \qquad \square$$

Clearly, the orthonormal projection provides better approximation property in $L_2(0,1)$. This yields (5.80) with the change $p_{m,l}$ by the orthonormal projection of f to the space of piecewise polynomial functions:

$$\|f - Pr_{m,l} f\|_2 \leq B m^{-\sigma} \|f\|_{\sigma,2}. \tag{5.81}$$

Then we use the equality

$$\|f\|_2^2 = \|Pr_{m,l} f\|_2^2 + \|f - Pr_{m,l} f\|_2^2.$$

Jointed with (5.81), this yields (5.79). $\square$

Next considerations repeat the proof of Theorem 3.3 with the change ρ_R by $r_\varepsilon/\varepsilon$, $m =$ by $\tilde{m}_\varepsilon$, $\rho_m(v)$ by $\rho_\varepsilon(f)$ and R by $H\varepsilon^{-1}$. We consider the case $\|f\|_2 = r_\varepsilon$ and apply inequality (5.79) instead of (5.45). $\square$

5.11 Proof of Theorem 3.7

5.11.1 Lower bounds

Analogously to Section 5.7 let us take integer-valued family $n = n_\varepsilon$, an integer $s > \sigma$ and s-differentiable function $\phi(t),\ t \in R^1$ supported on $[0,1]$,

$$\|\phi\|_2 = 1, \quad \|\phi\|_\infty = d, \quad \|\phi\|_{q,\sigma} = a,$$

for some $d > 0,\ a = a(\sigma) > 0$. Setting $f_{\varepsilon,i}(t) = r_\varepsilon \phi(nt - i + 1)/d, \quad i = 1, ..., n$, we see that the functions $f_{\varepsilon,i}$ have disjoin supports on $[0,1]$ and

$$\|f_{\varepsilon,i}\|_2 = r_\varepsilon n^{-1/2}/d, \quad \|f_{\varepsilon,i}\|_\infty = r_\varepsilon, \quad \|f_{\varepsilon,i}\|_{q,\sigma} \le a r_\varepsilon n^{\sigma - 1/q}/d = a r_\varepsilon n^\eta/d.$$

Take $n \sim (a r_\varepsilon/dH)^{-1/\eta}$ such that $a r_\varepsilon n^\eta/d \le H$. We have $f_{\varepsilon,i} \in F_\varepsilon \ \ \forall\ i = 1, ..., n$. Let us take a prior π_ε on $L_2(0,1)$ and consider corresponding mixture

$$\pi_\varepsilon = n^{-1}\sum_{i=1}^n \delta_{f_{\varepsilon,i}}, \quad P_{\pi_\varepsilon} = \int P_{\varepsilon,f}\pi_\varepsilon(df) = n^{-1}\sum_{i=1}^n P_{\varepsilon,f_{\varepsilon,i}}.$$

Since $\pi_\varepsilon(F_\varepsilon) = 1$, it suffices to obtain the lower bounds for the quantities $\beta_\varepsilon(\alpha_\varepsilon, P_{\pi_\varepsilon})$ in the Bayesian hypothesis testing problem on a measure P, which generates observations X_ε of form (1.2),

$$H_0\ :\ P = P_{\varepsilon,0} \quad \text{against} \quad H_1\ :\ P = P_{\pi_\varepsilon}. \tag{5.82}$$

Hypothesis testing problem (5.82) is equivalent to

$$H_0\ :\ P = P_0 \text{ against } H_1\ :\ P = P_{\pi_{\varepsilon,n}} = n^{-1}\sum_{i=1}^n P_{\rho_\varepsilon e_i}, \tag{5.83}$$

where $\{e_i\}_{i=1}^n$ is the standard basis in R^n,

$$\pi_{\varepsilon,n} = n^{-1}\sum_{i=1}^n \delta_{\rho_\varepsilon e_i}, \quad \rho_\varepsilon = r_\varepsilon n^{-1/2}/d\varepsilon.$$

Let us verify the inequality

$$\beta_\varepsilon(\alpha_\varepsilon, P_{\pi_\varepsilon}) \ge \Phi(\sqrt{2\log n + \log \alpha_\varepsilon^{-1}} - \rho_\varepsilon) + o(1). \tag{5.84}$$

To prove (5.84) it suffices assume $\sqrt{2\log n + \log \alpha_\varepsilon^{-1}} - \rho_\varepsilon = O(1)$. However hypothesis testing problem analogous to (5.83) has been studied in Section 5.6.1 and inequality (5.84) corresponds to (5.54) with the change u_R, k, α_R by $\rho_\varepsilon, n, \alpha_\varepsilon$ that was established in Section 5.6.1 under the same constraints.

Using (5.84) we see that suffices to verify that

$$\limsup \rho_\varepsilon/\sqrt{2(\log n + \log \alpha_\varepsilon^{-1})} < 1. \tag{5.85}$$

To prove (5.85) we can assume $r_\varepsilon \asymp r^*_{\varepsilon,\alpha_\varepsilon\infty}$. In this case one can see that

$$\log \alpha_\varepsilon^{-1} + \log n \asymp \log \alpha_\varepsilon^{-1} + \log \varepsilon^{-1}. \tag{5.86}$$

If $\limsup r_\varepsilon/r^*_{\varepsilon,\alpha_\varepsilon\infty} < C_0$ for C_0 small enough, then (5.85) follows from (5.86). $\square$

5.11.2 Upper bounds

For test family (3.49) using relation (1.8) we have, for a_ε small enough,

$$\alpha_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}) \le \sum_{l=1}^{\infty}\sum_{j=1}^{m_l} P_{\varepsilon,0}(|X_{j,m_l}| \ge T_l) = 2\sum_{l=1}^{\infty} m_l \Phi(-T_l) \sim$$
$$\frac{2}{\sqrt{2\pi}}\sum_{l=1}^{\infty} 2^l \exp(-T_l^2/2)/T_l \le \frac{\alpha_\varepsilon}{\sqrt{\pi}}\sum_{l=1}^{\infty}(cl + \log \alpha_\varepsilon^{-1})^{-1/2} l^{-1} < \alpha_\varepsilon.$$

Next considerations are analogous to Section 5.9. In view of (5.75) it suffices to consider the case $q = \infty$ with the change σ by $\eta = \sigma - 1/q$ and H by $H_1 = cH$. Let $f \in F_\varepsilon = F_\varepsilon(r_\varepsilon, H_1, \eta, \infty)$. We have relation (5.74) and take $l = l_\varepsilon(\eta)$ such that $2^l \sim (r^*_{\varepsilon,\alpha_\varepsilon,\infty})^{-1/\eta}$, where $r^*_{\varepsilon,\alpha_\varepsilon,\infty}$ are defined by (3.46). We get

$$T_l = h(\eta)\sqrt{\log \varepsilon^{-1} + (\eta + 1/2)\log \alpha_\varepsilon^{-1})} + o(1), \quad h(\eta) = 2(2\eta+1)^{-1/2}.$$

Using inequality (5.71) for $p = \infty$ and taking C_1 such that

$$B = B_1 C_1 - B_2 H_1 > \max(h(\eta), \sqrt{2}),$$

we see that, uniformly over $f \in F_\varepsilon$,

$$\varepsilon^{-1} m_l^{-1/2} \|Pr_{m_l} f\|_\infty \ge \varepsilon^{-1}(r^*_{\varepsilon,\alpha_\varepsilon,\infty})^{1+1/2\eta}(B + o(1)) \sim B\sqrt{\log(\alpha_\varepsilon \varepsilon)^{-1}}.$$

Therefore the argument of the function Φ tends to $-\infty$ in (5.74). This yields $\beta_\varepsilon(\psi_{\varepsilon,\alpha_\varepsilon}, F_\varepsilon) \to 0$. $\square$

5.12 Proof of Theorem 3.8

5.12.1 Lower bounds

Our considerations follow to [13], Section 7.2. It suffices to consider $\Sigma = \Sigma_0$ and

$$\sup_{\sigma\in\Sigma} K^2(\rho_R(\sigma), R, \sigma) = K_R^2(\Sigma) + o(1);\ K_R(\Sigma) = \sqrt{\log\log R + \log \alpha_R^{-1}} + O(1). \tag{5.87}$$

Moreover it suffices to assume

$$\log \alpha_R^{-1} = O((\log\log R)^2), \quad \tau = 0, \quad \sigma(\tau) = \infty, \tag{5.88}$$

since the required lower bounds follows from Theorem 3.2, (1) for the case $\log \alpha_R^{-1}/(\log\log R)^2) \to \infty$. Under (5.87), (5.88) we have

$$K_R(\Sigma) > \sqrt{\log\log R} + O(1) \to \infty, \quad K_R(\Sigma) = O(\log\log R), \tag{5.89}$$

relations (3.9), (4.14)–(4.27) hold uniformly over $\sigma \in \Sigma$. Moreover relation (3.17) is valid uniformly over $\sigma_R = \sigma \in \Sigma$ and using Corollary 3.2, (3.18) we can change the quantities $\sqrt{2}K(\rho_R(\sigma), R)$ by the quantities $u(\rho_R(\sigma), R$ defined by (2.16), (2.17).

Note also that it suffices assume $\log \alpha_R^{-1} > b\sqrt{\log\log R}$, $b > 0$. In fact, if $\log \alpha_R^{-1} = o(\sqrt{\log\log R})$, then the lower bounds easy follow from (2.32), because of $H_R(\Sigma) = H_R(\alpha_R, \Sigma) + o(1)$ in this case.

Take an integer-valued family

$$M = M_R \to \infty, \quad M_R \asymp (\log R)/(\log\log R)^B, \quad B > 1. \tag{5.90}$$

This yields

$$\sqrt{\log M + \log \alpha_R^{-1}} = \sqrt{\log\log R + \log \alpha_R^{-1}} + o(1).$$

Take collections

$$\sigma_{R,l} = \sigma_0 + l h_R, \quad 1 \le l \le M, \quad h_R = (\sigma_1 - \sigma_0)/M \to 0.$$

For each l, let us consider extreme problem (3.4), (3.5) with $\sigma = \sigma_{R,l}$, $\rho_R = \rho_R(\sigma_{R,l})$ and with the change (5.5), (5.6), where the quantities $z = z_{R,l}$, $n = n_{R,l}$ are determined by relations analogous to (4.7)–(4.13). We omit the index R below to simplify the notation. Let $\bar{z}_l = \{z_{l,i}\}$ be the extreme sequence in the problem. Recall that by (5.87), (5.89), one has, uniformly over (l,i),

$$z_{l,i} = 0, \ i \ge n_l, \quad z_{l,i}^2 \sim z_l^2(1 - (i/n_l)^{2\sigma_l}), \ 1 \le i < n_l, \tag{5.91}$$

$$u_l^2 = \frac{1}{2}\sum_i z_{l,i}^4 = 2K_R^2(\Sigma) + o(1), \tag{5.92}$$

$$n_l z_l^2 \asymp \rho_l^2, \quad n_l^{1+2\sigma_l} z_l^2 \asymp R^2, \quad n_l z_l^4 \asymp K_R^2(\Sigma). \tag{5.93}$$

Relations (5.91)–(5.93) yield (see (4.14), (4.15))

$$z_l^{2\sigma} \asymp \rho_l^{2\sigma_l+1} R^{-1} = o(R^{-b}), \quad n_l \asymp (R/\rho_l)^{1/\sigma_l} \asymp (R/K_R^{1/2}(\Sigma))^{\phi(\sigma_l)} \gg R^b \tag{5.94}$$

for some $b > 0$, where $\phi(\sigma) = (\sigma + 1/4)^{-1}$. For $j \ne l$, $j, l = 1, ..., M$, we set

$$n^+ = \max(n_j, n_l), \quad n^- = \min(n_j, n_l).$$

It follows from (5.90), (5.94) that there exists $c > 0$ such that for any $H > 0$,

$$\log(n^+/n^-) \ge c(\log\log R)^B \gg 2H \log\log R, \tag{5.95}$$

$$\sum_i z_{j,i}^2 z_{l,i}^2 \le z_j^2 z_l^2 n^- \asymp K_R^2(\Sigma)(n^-/n^+)^{1/2} = o((\log R)^{-H}). \tag{5.96}$$

Let $\pi_l = \mathcal{N}(0, \{z_l^2\})$ be the Gaussian measure on l^2 of form (5.7) that corresponds to independent $v_i \sim \mathcal{N}(0, z_{l,i}^2)$. Take measures P_{π_l} and mixtures $\pi = \pi_R,\ P_\pi$ of the form

$$\pi = \frac{1}{M}\sum_{l=1}^{M}\pi_l, \quad P_\pi = \frac{1}{M}\sum_{l=1}^{M}P_{\pi_l}, \quad P_{\pi_l} = \int P_v \pi_l(dv).$$

It follows from consideration in Section 5.1 that

$$\pi_l(V(\rho(\sigma_l), R, \sigma_{R,l}) = 1 + o(1) \quad \forall l,\ 1 \le l \le M,$$

see (5.4). This yields

$$\pi(V(\Sigma)) = \frac{1}{M}\sum_{l=1}^{M}\pi_l(V(\Sigma)) \ge \frac{1}{M}\sum_{l=1}^{M}\pi_l(V(\rho(\sigma_l), R, \sigma_l) = 1 + o(1).$$

Therefore it suffices to prove the lower bounds for the Bayesian hypothesis testing problem on a probability measure P that generates random observations $X = \{X_i\}$:

$$H_0 : P = P_0, \quad H_1 : P = P_\pi.$$

Recall that, by (5.8), P_{π_l} is the Gaussian measure corresponding to independent $X_i \sim \mathcal{N}(0, z_{l,i}^2 + 1)$ and the likelihood ratio $L = L_\pi = dP_\pi/dP_0$ is of the form

$$L = \frac{1}{M}\sum_{l=1}^{M}L_l, \quad L_l = dP_{\pi_l}/dP_0 = \exp(t_l), \quad t_l = \frac{1}{2}\sum_i\left(\frac{z_{l,i}^2 X_i^2}{1+z_{l,i}^2} - \log(1+z_{l,i}^2)\right).$$

Next considerations follow to the scheme of Section 5.6.1. The following lemma is analogous to Lemma 5.5.

Lemma 5.9 *Assume (5.87)–(5.89), (5.90). Let $\xi \sim \mathcal{N}(0,1)$, $u = \sqrt{2}K_R(\Sigma)$. One has, in P_π-probability,*

$$L_\pi = M^{-1}e^{u^2/2+u\xi_R} + \eta_R,\ \xi_R \to \xi,\ \eta_R \ge \Phi(\sqrt{2\log M} - u) + o(1),\ \eta_R = O(1). \quad (5.97)$$

Proof. For any $t \in R$ one has

$$P_\pi(L_\pi < t) = M^{-1}\sum_{l=1}^{M}P_{\pi_l}(L_\pi < t)$$

and it suffices to verify that (5.97) holds in P_{π_l}-probability for all $l,\ 1 \le l \le M$.

Fix l (next consideration are uniform over $l,\ 1 \le l \le M$). In P_{π_l}-probability, the random variables X_i^2 are distributed as $\xi_i^2(z_{l,i}^2 + 1)$, where $\xi_i \sim \mathcal{N}(0,1)$ are i.i.d. Therefore the random variables t_l are P_{π_l}-distributed as

$$t_l \sim \sum_i w_i + E_l, \quad w_i = \frac{1}{2}z_{l,i}^2(\xi_i^2 - 1), \quad E_l = K^2(\bar{z}_l),$$

where $K^2(\bar{z}) = K_R^2(\Sigma) + o(1)$ are defined by (3.3). By (5.92),

$$E(t_l) = E_l, \quad \mathrm{Var}(t_l) = \frac{1}{2} \sum_{1 \le i \le n_l} z_{l,i}^4 = 2E_l + o(1),$$

here and later expectations and variances correspond to the measure P_{π_l}. The Lyapunov ratio is of the form

$$\Lambda = \frac{\sum_i E(w_i^4)}{\left(\sum_i E(w_i^2)\right)^2} = O(n_l^{-1}) = o(R^{-b}), \ b > 0.$$

Therefore using (5.89), (5.91)–(5.93) we easily get, in P_{π_l}-probability,

$$\xi_R = (\log L_l - u^2/2)/u \to \xi \sim \mathcal{N}(0,1).$$

Set

$$\eta_R = \frac{1}{M} \sum_{j \ne l} L_j = \frac{1}{M} \sum_{j \ne l} \exp(t_j).$$

It remain to verify that, in P_{π_l}-probability,

$$\eta_R \ge \Phi(\sqrt{2 \log M} - u) + o(1), \quad \eta_R = O(1). \tag{5.98}$$

Simple calculation and (5.96) give, for $j \ne l$,

$$E(L_j) = \prod_i (1 - z_{j,i}^2 z_{l,i}^2)^{-1/2} \to 1, \quad \text{as } \sum_i z_{j,i}^2 z_{l,i}^2 \to 0.$$

This yields $E(\eta_R) = 1 + o(1)$, and by $\eta_R \ge 0$, we get the second relation (5.98). To verify the first relation (5.98) note that random variables $t_j, \ j \ne l$ are P_{π_l}-distributed as

$$t_j \sim \sum_i w_{ij} + E_{lj} + \Delta E_{lj}, \quad w_{ij} = \frac{1}{2} z_{j,i}^2 (\xi_i^2 - 1) \frac{z_{l,i}^2 + 1}{z_{j,i}^2 + 1}; \quad E(w_{ij}) = 0,$$

where

$$E_{lj} = \frac{1}{2} \sum_i \left(\frac{z_{j,i}^2}{z_{j,i}^2 + 1} - \log(z_{j,i}^2 + 1) \right), \quad \Delta E_{lj} = \frac{1}{2} \sum_i \frac{z_{j,i}^2 z_{l,i}^2}{z_{j,i}^2 + 1}, \quad \xi_i \sim \mathcal{N}(0,1) \ \text{i.i.d.}$$

By (5.92), (5.96), it is easily seen that

$$\Delta E_{lj} \le \sum_i z_{j,i}^2 z_{l,i}^2 \to 0, \quad E_{lj} = -\frac{1}{4} \sum_i z_{j,i}^4 + o(1) = -u_j^2/2 + o(1), \tag{5.99}$$

$$\mathrm{Var}(t_j) = \frac{1}{2} \sum_i z_{j,i}^4 + o(1) = u_j^2 + o(1) \asymp n_j z_j^4, \quad \mu_{ij} = E w_{ij}^4 \asymp z_{ji}^8, \tag{5.100}$$

$$r_{jk} = \mathrm{Cov}(t_j, t_k) = O\left(\sum_i z_j^2 z_k^2 \right) = o(1), \quad j \ne k, \ j \ne l, \ k \ne l. \tag{5.101}$$

The first relation (5.98) is contained in the proof of Lemma 7.2, relation (7.24) in [13] with the change P_0 by P_{π_l} and ε by R. Note that the proof of relation (7.24) in [13] does not use the structure of the statistics $L_l = dP_{\pi_l}/dP_0$ but its P_0-distributions only. Let us verify the assumptions of the lemma. Our case corresponds to

$$\omega_{\varepsilon,j} = \sqrt{2\log M},\ u_{\varepsilon,j} = u_j,\ D_\varepsilon = u - \sqrt{2\log M},\ l_{\varepsilon,j} = t_j,\ \rho_{\varepsilon,jk} = r_{jk}/u_j u_k.$$

These yield relations (7.19) and (7.20) with $p_{\varepsilon,l} = M^{-1}$ in [13]. Relation (7.21) in [13] follows from (5.101). Relations (7.22) follow from (5.99)–(5.100) and from two-dimensional version of the Bahr-Essen inequality (compare with the proof of Lemma 7.3 in [13]). The evaluation of the Lyapunov ratio is analogous to one in the proof of Lemma 5.1 above. □

Next considerations repeat ones from Section 5.6.1 with the change k by M and u_R by $u = \sqrt{2}K_R(\Sigma)$. This yields the required lower bounds. □

5.12.2 Upper bounds

We study test family ψ^{ad}_{R,α_R} determined by (3.53). Set

$$\alpha^*_R = \alpha_R/M, \quad T_R = \log((\alpha^*_R)^{-1}).$$

The first relation (3.54) easily follows from Lemma 5.2 (1):

$$\alpha(\psi^{ad}_{R,\alpha_R}) \le \sum_{l=1}^{M} P_0\left(t_{R,l} > T_R\right) \le M\exp\left(-T_R\right) = \alpha_R.$$

To verify the second relation (3.54) it suffices to assume

$$K(\rho_R(\sigma), R, \sigma) = K(\Sigma) \quad \forall\, \sigma \in \Sigma; \quad H_R(\Sigma, \alpha_R) = O(1). \tag{5.102}$$

This yields $\rho^*_R(\sigma) = \rho_R(\sigma)$. Note that, under (5.102), one has

$$V(\rho, R, \sigma') \subset V(\rho, R, \sigma''),\ K(\rho, R, \sigma') \ge K(\rho, R, \sigma''),\ \rho_R(\sigma') \le \rho_R(\sigma'') \text{ for } \sigma' > \sigma''. \tag{5.103}$$

Take $v = v_R \in V_R(\Sigma)$ and let $v \in V_R(\delta_{R,l})$, i.e.,

$$v_R \in V(\rho_R(\sigma_R), R, \sigma_R), \quad \sigma_R \in \delta_{R,l},\ l \in \{1, ..., M_R\}.$$

It follows from (5.103) that

$$V(\rho_R(\sigma_R), R, \sigma_R) \subset V(\rho_R(\sigma_{R,l}), R, \sigma_{R,l-1}).$$

It suffices to verify that

$$\sup_{v \in V(\rho_R(\sigma_{R,l}), R, \sigma_{R,l-1})} P_v(t_{R,l} \le T_R) \le \Phi(\sqrt{2\log((\alpha^*_R)^{-1})} - \sqrt{2}K_R(\Sigma)) + o(1). \tag{5.104}$$

By the construction of statistics $t_{R,l}$, the left-hand side of (5.104) is the left-hand side of (3.14) for the set $V_R = V(\rho_R(\sigma_{R,l}), R, \sigma_{R,l-1})$ and with the change α_R by α_R^*. Under (5.102) it is easily seen that assumptions (3.8), (3.9) are fulfilled for $\sigma_{R,l} \le \sigma_1 < \sigma(\tau)$ (see Remark 3.2). Applying Theorem 3.1 (2) we get the upper bounds of the type (5.104) with the change $K_R(\Sigma)$ by $K_R = K(\rho_R(\sigma_{R,l}), R, \sigma_{R,l-1})$. By (5.102), it remains to verify that

$$K_R(\Sigma) = K(\rho_R(\sigma_{R,l}), R, \sigma_{R,l}) \le K(\rho_R(\sigma_{R,l}), R, \sigma_{R,l-1}) + o(1). \tag{5.105}$$

Lemma 5.10 *Assume (3.2) with $\sigma = \sigma' \in \Sigma$ and let $n = n_R(\sigma')$ be the quantity determined by (4.7), (4.8). Let $\sigma'' = \sigma' + \delta_R$, $\delta_R > 0$, $\delta_R \log n = o(1)$. Then there exists $C > 0$ such that, for R large enough uniformly over $\sigma' \in \Sigma$,*

$$K(\rho_R, R, \sigma') \ge K(\rho_R, Rn^{\delta_R}, \sigma'') \ge K(\rho_R, R, \sigma'')(1 - C(n^{\delta_R} - 1)). \tag{5.106}$$

Proof. Let $\bar{z} = \bar{z}_R(\rho_R, \sigma')$ be the extreme sequence in the problem (3.4), (3.5) with $\sigma = \sigma'$, i.e., $K(\bar{z}) = K(\rho_R, R, \sigma')$. Since $z_i = 0$ for $i > n$, the sequence $\bar{z}$ satisfies the constraint (3.5) with the change σ' by σ'' and R by Rn^δ. Therefore $K(\bar{z}) \ge K(\rho_R, Rn^\delta, \sigma'')$. This yields the first inequality (5.106). The second inequality (5.106) follows from Remark 3.1, (3.7). $\square$

We apply Lemma 5.10 to the case $\sigma' = \sigma_{R,l-1}$, $\sigma'' = \sigma_{R,l}$. By (4.14) we have $\log n = O(\log R)$. By the choice of $M \asymp (\log R)(K_R(\Sigma))^B$, $B > 1$, we have

$$\delta_R \asymp 1/M, \quad n_R^\delta - 1 \sim \delta_R \log n \asymp \log n/M = O((K_R(\Sigma))^{-B}).$$

This yields inequality (5.105). $\square$

5.13 Proof of upper bounds of Theorem 3.9

Since $P_0(\chi^2_m > T_{m,\alpha}) = \alpha$, we have the first relation (3.59):

$$\alpha(\psi_{\varepsilon,\hat{\alpha}_\varepsilon}) \le \sum_{J_{\varepsilon,0} < l < J_{\varepsilon,0}} P_0(\chi^2_{m_l} > T_{m_l,\hat{\alpha}_\varepsilon}) < M\hat{\alpha}_\varepsilon = \alpha_\varepsilon.$$

Let us verify the second relation (3.59). It suffices to assume $r_\varepsilon(\sigma) = C_1 r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma)$, where C_1 is large enough. For an integer l,

$$l_0 \le l \le l_1 \quad l_0 \in (J_{\varepsilon,0}, J_{\varepsilon,0} + 1], \quad l_1 \in [J_{\varepsilon,1} - 1, J_{\varepsilon,1}),$$

let $\sigma^{(l)}$ be determined by relation

$$2^{-\sigma^{(l)}l} = r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma^{(l)}); \quad \sigma^{(l_0)} \ge \sigma_1, \quad \sigma^{(l_1)} \le \sigma_0.$$

Let $\sigma \in [\sigma^{(l)}, \sigma^{(l-1)}]$ (next consideration are uniform over $l_0 \le l \le l_1$ and over $\sigma \in [\sigma^{(l)}, \sigma^{(l-1)}]$). It suffices to verify that

$$\sup_{f \in F(r_\varepsilon(\sigma), H, \sigma)} P_{\varepsilon,f}(\chi^2_{m_l} \le T_{m_l,\hat{\alpha}_\varepsilon}) \to 0. \tag{5.107}$$

Since $\log\log\alpha_\varepsilon^{-1} \le (2+o(1))\log\varepsilon^{-1}$, we have $M = J_{\varepsilon,1} - J_{\varepsilon,0} \asymp J_{\varepsilon,1} \asymp \log\varepsilon^{-1}$. Note that

$$b(l) = l\sigma^{(l)} - (l-1)\sigma^{(l-1)} = -1/4, \quad r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma^{(l-1)})/r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma^{(l)}) = 2^{b(l)} = 2^{-1/4}.$$

This yields $r_\varepsilon(\sigma^{(l-1)}) \asymp r_\varepsilon(\sigma^{(l)})$ and there exist $D>0$ such that $F(r_\varepsilon(\sigma),H,\sigma) \subset F(Dr_\varepsilon(\sigma^{(l)}),H,\sigma^{(l)})$. Relation (5.107) follows from

$$\sup_{f\in F(Dr_\varepsilon(\sigma^{(l)}),H,\sigma^{(l)})} P_{\varepsilon,f}(\chi^2_{m_l} \le T_{m_l,\hat\alpha_\varepsilon}) \to 0. \tag{5.108}$$

However for C_1 large enough, relation (5.108) follows directly from Theorem 3.5 (2) that is applied to the case $\sigma = \sigma^{(l)}$, $\alpha_\varepsilon = \hat\alpha_\varepsilon$. In fact, since $\sigma^{(l)} \le \sigma_1 + o(1)$ and $\sigma_1 < \sigma(\tau)$, we have $\delta_\varepsilon = o(1)$:

$$\log\hat\alpha_\varepsilon^{-1} = \log\alpha_\varepsilon^{-1} + \log M = o(\varepsilon^{-2/(2\sigma^{(l)}+1)}).$$

Taking into account (3.36), (3.55), we have

$$\begin{aligned} r^*_{\varepsilon,\hat\alpha_\varepsilon}(\sigma^{(l)}) &= \left(\varepsilon^4(\log\alpha_\varepsilon^{-1} + \log M)\right)^{\sigma^{(l)}/(4\sigma^{(l)}+1)} \\ &\sim \left(\varepsilon^4(\log\alpha_\varepsilon^{-1} + \log\log\varepsilon^{-1})\right)^{\sigma^{(l)}/(4\sigma^{(l)}+1)} = r^{ad}_{\varepsilon,\alpha_\varepsilon}(\sigma^{(l)}). \qquad \square \end{aligned}$$

References

[1] Ermakov, M.S. Minimax detection of a signal in a Gaussian white noise. Theory Probab. Appl., v. 35 (1990), 667–679.

[2] Ermakov, M.S. Minimax nonparametric testing of a hypotheses on a distribution density. Theory Probab. Appl., v. 39 (1994), 3, 396–416.

[3] Hoffmann M, Lepski O.V. Random rates in anisotropic regression (with discussion). Ann. Statist., 30 (2002), v. 2, 325–396.

[4] Inglot, T. and Ledwina, T. Asymptotic optimality of data-driven Neyman's tests for uniformity. Ann. Statist., v. 24 (1996), 5, 1996–2019.

[5] Inglot, T., Kallenberg, W.C.M. and Ledwina, T. Vanishing shortcoming and asymptotic relative efficiency. Ann. Statist., v. 28 (2000), 1, 215–238.

[6] Ingster, Yu.I. Minimax nonparametric detection of signals in the white Gaussian noise. Problems Inform. Transmission, v. 18 (1982), 130–140.

[7] Ingster, Yu.I. Asymptotic minimax testing of nonparametric hypotheses on the distribution density of an independent sample. Zapiski Nauchn. Seminar. LOMI, v. 136 (1984), 74–96. (In Russian.)

[8] Ingster, Yu. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I, II, III. Math. Methods of Statistics, v. 2 (1993), 2, 85–114, 3, 171–189, 4, 249–268.

[9] Ingster, Yu.I. Minimax hypotheses testing on a probability density for ellipsoids in l_p. Theory Probab. Appl., v. 39 (1994), 3, 417–430.

[10] Ingster, Yu.I. Adaptive chi-square tests. Zapiski Nauchn. Seminar. POMI, Probability and Statistics. 2., v. 244 (1997), 150–166.

[11] Ingster, Yu.I. Adaptation in minimax non-parametric hypothesis testing. Preprint WIAS No. 419 (1998), Berlin.

[12] Ingster, Yu.I. and Suslina, I.A. Minimax nonparametric hypothesis testing for ellipsoids and Besov bodies. ESAIM: Probability and Statistics, v. 4 (2000), 53–135.

[13] Ingster, Yu.I. and Suslina, I.A. Nonparametric Goodness-of-Fit Testing under Gaussian Model. Lectures Notes in Statistics, v. 169, Springer-Verlag, New York, 2002

[14] Ingster, Yu.I. and Suslina, I.A. On a detection of a signal of known shape in multichannel system. Zapiski Nauchn. Sem. POMI, v. 294 (2002), 88-112.

[15] Kallenberg, W.C.M. Intermediate efficiency, theory and examples. Ann. Statist., v. 11 (1983), 170–182.

[16] Kallenberg, W.C.M. The penalty in data driven Neyman's tests. Math. Methods of Statist., v. 11 (2002), 1, 323–340

[17] Lepski O.V. How to improve accuracy of estimation. Math. Methods of Statist., v. 8 (1999), 4, 1–46.

[18] Lepski, O.V. and Spokoiny, V.G. Minimax nonparametric hypothesis testing: the case of an inhomogeneous alternative. Bernoulli, v. 5 (1999), 333–358.

[19] Lepski, O.V. and Tsybakov, A.B. Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. Probability Theory and Related Fields, v. 117 (2000), 1, 17–48.

[20] Nikitin Ya.Yu. Asymptotic Efficiency of Nonparametric Tests. Cambridge Univ. Press, 1995.

[21] Spokoiny, V.G. Adaptive hypothesis testing using wavelets. Ann. Stat., v. 24 (1996), 6, 2477–2498.

[22] Spokoiny, V.G. Adaptive and spatially adaptive testing of nonparametric hypothesis. Mathematical Methods of Statistics, v. 7 (1998), 3, 245–273.