The Annals of Statistics
2012, Vol. 40, No. 3, 1524–1549
DOI: 10.1214/12-AOS1011

MINIMAX SIGNAL DETECTION IN ILL-POSED INVERSE PROBLEMS

BY YURI I. INGSTER[1], THEOFANIS SAPATINAS AND IRINA A. SUSLINA

Saint Petersburg Electrotechnical University, University of Cyprus and Saint Petersburg University of Information Technologies, Mechanics and Optics

Ill-posed inverse problems arise in various scientific fields. We consider the signal detection problem for mildly, severely and extremely ill-posed inverse problems with l^q-ellipsoids (bodies), $q \in (0, 2]$, for Sobolev, analytic and generalized analytic classes of functions under the Gaussian white noise model. We study both rate and sharp asymptotics for the error probabilities in the minimax setup. By construction, the derived tests are, often, nonadaptive. Minimax rate-optimal adaptive tests of rather simple structure are also constructed.

1. Introduction. We consider the detection problem in linear operator equations from noisy data. More precisely, we consider the Gaussian white noise model (GWNM)

$$dY_\varepsilon(t) = Af(t)\,dt + \varepsilon\,dW(t), \qquad t \in D, \tag{1.1}$$

where $A:\mathcal{H} \mapsto L^2(D)$ is a known linear bounded operator, $\mathcal{H} \subset L^2(D)$, $D \subset \mathbb{R}$, W is a standard Wiener process on D, $\varepsilon > 0$ is a small parameter (the noise level) and $f \in L^2(D)$ is the unknown response function (that one needs to detect or estimate).

We consider below the case where A has a kernel with the *singular value decomposition* (SVD) $A(t,x) = \sum_{k\in\mathbb{N}} b_k\psi_k(t)\varphi_k(x)$, in the sense of

$$(Af)(t) = \int_D A(t,x) f(x)\,dx = \sum_{k\in\mathbb{N}} b_k\psi_k(t) \int_D f(x)\varphi_k(x)\,dx, \qquad t \in D,$$

with $b_k > 0$, $k \in \mathbb{N}$, and orthonormal bases $\{\psi_k\}_{k\in\mathbb{N}}$ and $\{\varphi_k\}_{k\in\mathbb{N}}$. (Here, $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers.)

Thus, the GWNM (1.1) generates an equivalent discrete observational model in the sequence space, called the Gaussian sequence model (GSM),

$$y_k = b_k\theta_k + \varepsilon\xi_k, \qquad k \in \mathbb{N}, \tag{1.2}$$

Received May 2011; revised April 2012.
[1]Supported in part by RFBI Grant 11-01-00577 and by Grant NSh-4472.2010.1.
MSC2010 subject classifications. Primary 62G10, 62G20; secondary 62C20.
Key words and phrases. Analytic functions, ill-posed inverse problems, minimax testing, singular value decomposition, Sobolev spaces.

where $y_k = \int_D \psi_k(t)\,dY_\varepsilon(t)$, $b_k > 0$, $\theta_k = \int_D f(t)\varphi_k(t)\,dt$, $\varepsilon > 0$, and $\xi_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, $k \in \mathbb{N}$. The effect of the ill-posedness of the inverse problem is clearly seen in the decay of b_k as $k \to \infty$. As $k \to \infty$, $b_k\theta_k$ usually gets weaker and is then more difficult to detect or estimate $\theta = \{\theta_k\}_{k\in\mathbb{N}}$.

The GSM (1.2) can be rewritten in the (equivalent) form

$$x_k = \theta_k + \varepsilon\sigma_k\xi_k, \qquad k \in \mathbb{N}, \tag{1.3}$$

where $x_k = y_k/b_k$ and $\sigma_k = b_k^{-1} > 0$, $k \in \mathbb{N}$. In this situation, the difficulty of ill-posedness, and hence any asymptotic results, is measured by the rates (type of growth) of σ_k as $k \to \infty$. For polynomial rates, that is, $\sigma_k \asymp k^\beta$, $\beta > 0$, the inverse problem is called *mildly* (*or softly*) *ill-posed*, for exponential rates, that is, $\sigma_k \asymp \exp(\beta k)$, $\beta > 0$, is called *severely ill-posed*, and for the case where $\sigma_{k+1}/\sigma_k \to \infty$ as $k \to \infty$, is called *extremely ill-posed*. Note that an extremely ill-posed inverse problem includes power-exponential rates, that is, $\sigma_k \asymp \exp(\beta k^\gamma)$, $\beta > 0$, $\gamma > 1$.[2]

An important element of the GSMs (1.2) and (1.3) is the prior information about the sequence $\theta = \{\theta_k\}_{k\in\mathbb{N}}$. Successful detection or estimation of the sequence $\theta = \{\theta_k\}_{k\in\mathbb{N}}$ is possible only if its elements θ_k, $k \in \mathbb{N}$, tend to zero sufficiently fast as k tends to infinity, meaning that f in the GWNM (1.1) is sufficiently smooth. A standard smoothness assumption on f is to assume that the sequence $\theta = \{\theta_k\}_{k\in\mathbb{N}}$ belongs to an l^q-ellipsoid (body), $0 < q < \infty$, in l^2, of semi-axes L/a_k, $k \in \mathbb{N}$, that is,

$$\tilde{\Theta} = \tilde{\Theta}_q(a, L) = \left\{\theta \in l^2 : \sum_{k\in\mathbb{N}} |a_k\theta_k|^q \le L^q\right\}, \tag{1.4}$$

where $a = \{a_k\}_{k\in\mathbb{N}}$, $a_k > 0$, $a_k \to \infty$ as $k \to \infty$ and $L > 0$. [Note that the requirement $a_k > 0$, $a_k \to \infty$ as $k \to \infty$, ensures that $\tilde{\Theta}_q(a, L)$ is a compact subset of l^2.] The sequence $a = \{a_k\}_{k\in\mathbb{N}}$ characterizes the "shape" of the ellipsoid while the parameter L characterizes its "size." This means that for large values of k, the elements θ_k, $k \in \mathbb{N}$, will decrease in k and, hence, will be small for large k. In what follows, we consider minimax signal detection in ill-posed problems with l^q-ellipsoids for the range $q \in (0, 2]$.

The functional sets of the form (1.4) that are often used in various ill-posed inverse problems are the Sobolev classes of functions (see [14]) and the classes of analytic functions; see [7]. We also consider a class of generalized analytic

[2]The relation $c_n \asymp d_n$ means that there exist constants $0 < C_1 \le C_2 < \infty$ and n_0 large enough such that $C_1 \le c_n/d_n \le C_2$ for $n \ge n_0$. We say that $c_n(\kappa) \asymp d_n(\kappa)$ uniformly over $\kappa \in \mathcal{K}$, if the similar inequalities hold true for all $\kappa \in \mathcal{K}$ with constants $0 < C_1 \le C_2 < \infty$ and n_0 which do not depend on κ. The relation $c_n \sim d_n$ means that for any $\delta \in (0,1)$ there exists n_0 large enough such that $1 - \delta \le c_n/d_n \le 1 + \delta$ for $n \ge n_0$. The uniform version of the relation $c_n(\kappa) \sim d_n(\kappa)$, $\kappa \in \mathcal{K}$, is defined similarly. Similar notation is used when $0 < \varepsilon \le \varepsilon_0$ for ε_0 small enough.

functions. Then, $\tilde{\Theta}$ in (1.4) takes, respectively, the form

$$\mathcal{W}_q(\alpha, L) = \left\{\theta \in l^2 : \sum_{k\in\mathbb{N}} k^{\alpha q}|\theta_k|^q \leq L^q\right\},$$

$$\mathcal{A}_q(\alpha, L) = \left\{\theta \in l^2 : \sum_{k\in\mathbb{N}} e^{\alpha k q}|\theta_k|^q \leq L^q\right\},$$

$$\mathcal{G}_q(\tau, \alpha, L) = \left\{\theta \in l^2 : \sum_{k\in\mathbb{N}} e^{\alpha k^\tau q}|\theta_k|^q \leq L^q\right\}$$

for some $\alpha > 0$, $\tau \geq 1$ (the case $\tau = 1$ corresponds to the class of analytic functions) and $L > 0$.

Despite the growing number of works for the estimation problem in ill-posed inverse problems under the GWNM (1.1) (see, e.g., [2, 3, 5] and [6]), very little work exists for the corresponding detection problem; see Section 4.3.3 of [11] and [4] (although their results are obtained from models that are neither formulated nor immediately seen as particular ill-posed inverse problems), and [9] for a problem related to the Radon transform (see also, the supplementary material [10], Remark 6.1). Our aim is to present a general framework for the minimax detection study of the aforementioned ill-posed inverse problems. (Nonasymptotic minimax rates of testing for some of the ill-posed inverse problems under consideration were recently studied in [12].)

The rest of the paper is organized as follows. The general statement of minimax signal detection in ill-posed inverse problems is given in Section 2, while a short description of the main results and a comparison with similar results obtained in the corresponding estimation problems are presented in Section 3. The general methods for the study of minimax signal detection in ill-posed inverse problems with l^2-ellipsoids are given in Section 4.1. In Sections 4.2–4.7, we provide a complete treatment to the minimax signal detection problem for mildly, severely and extremely ill-posed inverse problems with l^q-ellipsoids, $q \in (0, 2]$, for Sobolev, analytic and generalized analytic classes of functions under the GSM (1.2). We study both rate and sharp asymptotics for the error probabilities in the minimax setup. By construction, the derived tests are, often, nonadaptive. In Section 5, for the ill-posed inverse problems under consideration, we also construct minimax rate-optimal adaptive tests of rather simple structure. The proofs along with other relevant material can be found in the supplementary material [10].[3]

2. Signal detection in the GSM: The minimax framework. Consider the GSM (1.2). In order to avoid having a trivial minimax hypothesis testing problem

[3]Some numbering that appears in the text corresponds to numbering in the supplementary material [10], Sections 6–11. Also, some of the references that appear in the reference list are cited in the supplementary material [10].

(i.e., trivial power), one usually needs to remove a neighborhood around the functional parameter under the null hypothesis and to add some additional constraints, that are typically expressed in the form of some regularity conditions, such as constraints on the derivatives, of the unknown functional parameter of interest (see, e.g., [11], Sections 1.3–1.4).

In view of the above observation, the main object of our study is the hypothesis testing problem

$$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \sum_{k\in\mathbb{N}} |a_k\theta_k|^q \le 1, \qquad \sum_{k\in\mathbb{N}} \theta_k^2 \ge r_\varepsilon^2, \tag{2.1}$$

where $\theta = \{\theta_k\}_{k\in\mathbb{N}} \in l^2$, $a = \{a_k\}_{k\in\mathbb{N}}$, $a_k > 0$, $a_k \to \infty$ as $k \to \infty$, $r_\varepsilon > 0$, $r_\varepsilon \to 0$, is a given family, and $q \in (0, 2]$. It means that the set under the alternative corresponds to an l^q-ellipsoid of semi-axes $1/a_k$, $k \in \mathbb{N}$, with an l^2-ball of radius r_ε removed. [For simplicity, in subsequent sections, we focus attention on ellipsoids of the form (1.4) with "size" $L = 1$.]

Consider the sequence $\eta = \{\eta_k\}_{k\in\mathbb{N}}$, $\eta_k = b_k\theta_k = \theta_k/\sigma_k$, $k \in \mathbb{N}$. Recall that, in the ill-posed inverse problems under consideration, $\sigma_k = 1/b_k \to \infty$ or $b_k \to 0$, as $k \to \infty$. Hence, $\eta \in l^2$, and the GSM (1.2) is of the form

$$y_k = \eta_k + \varepsilon\xi_k, \qquad k \in \mathbb{N}. \tag{2.2}$$

Thus, (2.1) can also be written in the following equivalent form:

$$H_0 : \eta = 0 \quad \text{versus} \quad H_1 : \eta \in \Theta_q(r_\varepsilon), \tag{2.3}$$

where the set under the alternative is determined by the constraints

$$\begin{aligned} \Theta_q &= \left\{ \eta \in l^2 : \sum_{k\in\mathbb{N}} |a_k\sigma_k\eta_k|^q \le 1 \right\}, \\ \Theta_q(r_\varepsilon) &= \left\{ \eta \in \Theta_q : \sum_{k\in\mathbb{N}} \sigma_k^2\eta_k^2 \ge r_\varepsilon^2 \right\}; \end{aligned} \tag{2.4}$$

that is, the set under the alternative corresponds to an l^q-ellipsoid of semi-axes $1/(a_k\sigma_k)$, $k \in \mathbb{N}$, with an l^2-ellipsoid of semi-axes r_ε/σ_k, $k \in \mathbb{N}$, removed.

We are therefore interested in the minimax efficiency of the hypothesis testing problem (2.3) and (2.4) for a given family of sets $\Theta_\varepsilon = \Theta_q(r_\varepsilon) \subset l^2$. It is characterized by asymptotics, as $\varepsilon \to 0$, of the minimax error probabilities in the problem at hand. Namely, for a (randomized) test ψ (i.e., a measurable function of the observation $y = \{y_k\}_{k\in\mathbb{N}}$ taking values in [0, 1]), the null hypothesis is rejected with probability $\psi(y)$ and is accepted with probability $1 - \psi(y)$. Let $P_{\varepsilon,\eta}$ be the probability measure for the GSM (2.2), and denote by $E_{\varepsilon,\eta}$ the expectation over this probability measure. Let $\alpha_\varepsilon(\psi) = E_{\varepsilon,0}\psi$ be its type I error probability, and let $\beta_\varepsilon(\Theta_\varepsilon, \psi) = \sup_{\eta\in\Theta_\varepsilon} \beta_\varepsilon(\eta, \psi)$, $\beta_\varepsilon(\eta, \psi) = E_{\varepsilon,\eta}(1 - \psi)$, be its maximal type II error probability. We consider two criteria of asymptotic optimality:

(1) The first one corresponds to the classical Neyman–Pearson criterion. For $\alpha \in (0,1)$, we set $\beta_\varepsilon(\Theta_\varepsilon,\alpha) = \inf_{\psi : \alpha_\varepsilon(\psi)\le\alpha} \beta_\varepsilon(\Theta_\varepsilon,\psi)$. We call a family of tests $\psi_{\varepsilon,\alpha}$ *asymptotically minimax* if $\alpha_\varepsilon(\psi_{\varepsilon,\alpha}) \le \alpha + o(1)$, $\beta_\varepsilon(\Theta_\varepsilon,\psi_{\varepsilon,\alpha}) = \beta_\varepsilon(\Theta_\varepsilon,\alpha) + o(1)$, where $o(1)$ is a family tending to zero (here, and in what follows, unless otherwise stated, all limits are taken as $\varepsilon \to 0$).

(2) The second one corresponds to the total error probabilities. Let $\gamma_\varepsilon(\Theta_\varepsilon,\psi)$ be the sum of the type I and the maximal type II error probabilities, and let $\gamma_\varepsilon(\Theta_\varepsilon)$ be the minimax total error probability, that is, $\gamma_\varepsilon(\Theta_\varepsilon) = \inf_\psi \gamma_\varepsilon(\Theta_\varepsilon,\psi)$, where the infimum is taken over all possible tests. We call a family of tests ψ_ε *asymptotically minimax* if $\gamma_\varepsilon(\Theta_\varepsilon,\psi_\varepsilon) = \gamma_\varepsilon(\Theta_\varepsilon) + o(1)$. It is known that (see, e.g., [11], Chapter 2) that

$$\beta_\varepsilon(\Theta_\varepsilon,\alpha) \in [0, 1-\alpha], \qquad \gamma_\varepsilon(\Theta_\varepsilon) = \inf_{\alpha\in(0,1)} \bigl(\alpha + \beta_\varepsilon(\Theta_\varepsilon,\alpha)\bigr) \in [0,1]. \tag{2.5}$$

We consider the problems of rate and sharp asymptotics for the error probabilities in the minimax setup. The rate optimality problem corresponds to the study of the conditions for which $\gamma_\varepsilon(\Theta_\varepsilon) \to 1$ and $\gamma_\varepsilon(\Theta_\varepsilon) \to 0$ and, under the conditions of the last relation, to the construction of *asymptotically minimax consistent* families of tests ψ_ε, that is, such that $\gamma_\varepsilon(\Theta_\varepsilon,\psi_\varepsilon) \to 0$.

For the set of the form (2.4), we use the notation $\gamma_\varepsilon(\Theta_q(r_\varepsilon)) = \gamma_\varepsilon(r_\varepsilon)$ and $\beta_\varepsilon(\Theta_q(r_\varepsilon),\alpha) = \beta_\varepsilon(r_\varepsilon,\alpha)$, and we are interested in the minimal decreasing rates for the sequence r_ε such that $\gamma_\varepsilon(r_\varepsilon) \to 0$. Namely, we say that the positive sequence $r_\varepsilon^* \to 0$ is a *separation rate* if

$$\gamma_\varepsilon(r_\varepsilon) \to 1, \qquad \beta_\varepsilon(r_\varepsilon,\alpha) \to 1-\alpha \qquad \text{for any } \alpha \in (0,1) \qquad \text{as } r_\varepsilon/r_\varepsilon^* \to 0, \tag{2.6}$$

$$\gamma_\varepsilon(r_\varepsilon) \to 0, \qquad \beta_\varepsilon(r_\varepsilon,\alpha) \to 0 \qquad \text{for any } \alpha \in (0,1) \qquad \text{as } r_\varepsilon/r_\varepsilon^* \to \infty. \tag{2.7}$$

In other words, it means that, for small ε, one can detect all sequences $\eta \in \Theta_q(r_\varepsilon)$ if the ratio $r_\varepsilon/r_\varepsilon^*$ is large, whereas if this ratio is small, then it is impossible to distinguish between the null and the alternative hypothesis, with small minimax total error probability. Hence, the rate optimality problem corresponds to finding the separation rates r_ε^* and to constructing asymptotically minimax consistent families of tests.

On the other hand, the sharp optimality problem corresponds to the study of the asymptotics of the quantities $\beta_\varepsilon(\Theta_\varepsilon,\alpha)$, $\gamma_\varepsilon(\Theta_\varepsilon)$ (up to vanishing terms) and to the construction of asymptotically minimax families of tests $\psi_{\varepsilon,\alpha}$ and ψ_ε, respectively. We shall see (see Section 4.1) that often the sharp asymptotics are of Gaussian type, that is,

$$\beta_\varepsilon(r_\varepsilon,\alpha) = \Phi\bigl(H^{(\alpha)} - u_\varepsilon\bigr) + o(1), \qquad \gamma_\varepsilon(r_\varepsilon) = 2\Phi(-u_\varepsilon/2) + o(1), \tag{2.8}$$

where Φ is the standard Gaussian distribution function, and $H^{(\alpha)}$ is its $(1-\alpha)$-quantile, that is, $\Phi(H^{(\alpha)}) = 1-\alpha$. The quantity $u_\varepsilon = u_\varepsilon(r_\varepsilon)$ is the value of the specific extreme problem (3.1) on the sequence space l^2, and the extreme sequence of this problem determines the structure of the asymptotically minimax families of tests $\psi_{\varepsilon,\alpha}$ and ψ_ε. Moreover, we shall see that if $u_\varepsilon(r_\varepsilon) \to \infty$, then $\gamma_\varepsilon(r_\varepsilon) \to 0$, $\beta_\varepsilon(r_\varepsilon, \alpha) \to 0$, and if $u_\varepsilon(r_\varepsilon) \to 0$, then $\gamma_\varepsilon(r_\varepsilon) \to 1$, $\beta_\varepsilon(r_\varepsilon, \alpha) \to 1-\alpha$ for any $\alpha \in (0,1)$; that is, the family $u_\varepsilon(r_\varepsilon)$ characterizes *distinguishability* in the testing problem. The separation rates r_ε^* are usually determined by the relation $u_\varepsilon(r_\varepsilon^*) \asymp 1$ (see, e.g., [8, 11]). Hence, sharp and rate optimality problems correspond to the study of the extreme problem (3.1) and of the asymptotics of the family $u_\varepsilon(r_\varepsilon)$.

3. Minimax signal detection in ill-posed inverse problems: A short description of some of the main results. Sharp and rate optimality results for the specific ill-posed inverse problems under consideration are discussed in detail in Section 4. We give below a short description of the corresponding results for mildly and severely ill-posed inverse problems with l^q-ellipsoids, $q \in (0,2]$, for Sobolev and analytic classes of functions.

We consider the hypothesis testing problem (2.3) and (2.4) in the GSM (2.2). For the "standard" case $q = 2$, consider the extreme problem

$$u_\varepsilon^2 = u_\varepsilon^2(r_\varepsilon) = \frac{1}{2\varepsilon^4} \inf_{\eta \in \Theta(r_\varepsilon)} \sum_{k \in \mathbb{N}} \eta_k^4. \tag{3.1}$$

Suppose that $\Theta(r_\varepsilon) \neq \varnothing$ and $u_\varepsilon > 0$, and let there exist an extreme sequence $\{\tilde\eta_k\}_{k\in\mathbb{N}}$ in the extreme problem (3.1). Denote[4]

$$w_k = \frac{\tilde\eta_k^2}{\sqrt{2\sum_{k\in\mathbb{N}} \tilde\eta_k^4}}, \qquad k \in \mathbb{N}, \qquad w_0 = \sup_{k\in\mathbb{N}} w_k, \tag{3.2}$$

and consider the families of test statistics and tests

$$t_\varepsilon = \sum_{k\in\mathbb{N}} w_k\big((y_k/\varepsilon)^2 - 1\big), \qquad \psi_{\varepsilon,H} = \mathbb{1}_{\{t_\varepsilon > H\}}, \tag{3.3}$$

where $\mathbb{1}_A$ denotes indicator function of a set A.

The key tool for the study of the above mentioned hypothesis testing problem is (the general) Theorem 4.1. It shows that the family $u_\varepsilon = u_\varepsilon(r_\varepsilon)$ determines distinguishability in the problem; if $w_0 = o(1)$, then it also determines the sharp asymptotics (2.8). The rate optimal tests correspond to a weighted χ^2-statistic t_ε of the form (3.3). The main reason is that, for $q = 2$, the problem (3.1) is quadratically convex or can be reduced to a convex problem.[5] The key idea is that the χ^2-distance between P_0 the probability measure of the observations under the null

[4]The values of $\tilde\eta_k$, w_k, $k \in \mathbb{N}$, and w_0 depend on ε, that is, $\tilde\eta_k = \tilde\eta_{k,\varepsilon}$, $w_k = w_{k,\varepsilon}$, $k \in \mathbb{N}$, and $w_0 = w_{0,\varepsilon}$.

[5]Rate optimal tests of simpler structure are also given in the supplementary material [10], Sections 7–9.

hypothesis and the mixture P_{π_η} over the product symmetric two-points prior π_η for a point $\eta \in \Theta(r_\varepsilon)$ is characterized by the quantity $\sum_{k\in\mathbb{N}} \eta_k^4$, which leads to the extreme problem (3.1). Moreover, if $w_0 = o(1)$, then the Bayesian log-likelihood ratio is asymptotically Gaussian under P_0, that is,

$$\log(dP_{\pi_{\tilde\eta}}/dP_0) = -u_\varepsilon^2/2 + u_\varepsilon t_\varepsilon + \delta_\varepsilon,$$

where u_ε is given by (3.1), t_ε is given by (3.3) and is asymptotically standard Gaussian and $\delta_\varepsilon \to 0$ in P_0-probability. This yields the lower bounds of the Gaussian form. On the other hand, the choice of the optimal coefficients of the test statistic t_ε given by (3.3) leads to a maximin problem that is reduced to the extreme problem (3.1) by convexity arguments. This yields the corresponding upper bounds. On the other hand, if $w_0 \neq o(1)$, then the test statistic t_ε given by (3.3) is not asymptotically Gaussian. This is due to that fact that, in this case, a_k and b_k, $k \in \mathbb{N}$, converge fast enough as k increases, implying that only a small number of observations is important. [Often, but not always, using embedding properties, these results can be extended to the case $q \in (0, 2)$; see Remark 4.11.]

The asymptotics of the quality of testing $u_\varepsilon(r_\varepsilon)$ as $r_\varepsilon \to 0$ is presented in Table 1, where $c_1 = c_1(\alpha, \beta) > 0, c_2 = c_2(\alpha, \beta) > 0$ are some constants. We have the sharp asymptotics of the form (2.8) for mildly ill-posed inverse problems with Sobolev and analytic classes of functions, while the derived asymptotically minimax tests are based on weighted χ^2-statistics with weights $w_k \geq 0, \sum_k w_k^2 = 1/2$ and are determined by the extreme problem (3.1). For severely ill-posed inverse problems, however, we do not have sharp asymptotics of minimax error probabilities but, instead, we get distinguishability conditions [i.e., $\gamma_\varepsilon(r_\varepsilon) \to 0$ if and only if $u_\varepsilon(r_\varepsilon) \to \infty$ and $\gamma_\varepsilon(r_\varepsilon) \to 1$ iff $u_\varepsilon(r_\varepsilon) \to 0$]. The main reason is that the weights are not "uniformly small," that is, there exist a few coefficients w_k that are bounded away from zero, and, hence, in this case, we do not have asymptotic Gaussianity.

Furthermore, the separation rates r_ε^* as $\varepsilon \to 0$ are presented in Table 2. (Similar nonasymptotic minimax rates are recently given in [12].) Note that, despite the fact that we have no sharp asymptotics for severely ill-posed inverse problems, we get the sharp separation rates for severely ill-posed inverse problems with Sobolev f, when $a_k \sim k^\alpha$ and $b_k \asymp \exp(-\beta k)$, $k \in \mathbb{N}$ [i.e., $\gamma_\varepsilon(r_\varepsilon) \to 0$ as $\lim r_\varepsilon/r_\varepsilon^* > 1$ and $\gamma_\varepsilon(r_\varepsilon) \to 1$ as $\lim r_\varepsilon/r_\varepsilon^* < 1$].

TABLE 1
The asymptotics $u_\varepsilon(r_\varepsilon)$ as $r_\varepsilon \to 0$

Detection problem	**Sobolev classes ($a_k = k^\alpha$)**	**Analytic classes ($a_k = \exp\{\alpha k\}$)**
Mildly ill-posed ($\sigma_k = k^\beta$)	$c_1\varepsilon^{-2} r_\varepsilon^{(4\alpha+4\beta+1)/2\alpha}$	$c_2\varepsilon^{-2} r_\varepsilon^2 (\log r_\varepsilon^{-1})^{-2\beta-1/2}$
Severely ill-posed ($\sigma_k = \exp\{\beta k\}$)	$\varepsilon^{-2} r_\varepsilon^2 e^{-2\beta r_\varepsilon^{-1/\alpha}}$	$\varepsilon^{-2} r_\varepsilon^{2(\alpha+\beta)/\alpha}$

TABLE 2
The separation rates r_ε^* as $\varepsilon \to 0$

Detection problem	Sobolev classes ($a_k = k^\alpha$)	Analytic classes ($a_k = \exp\{\alpha k\}$)
Mildly ill-posed ($\sigma_k = k^\beta$)	$\varepsilon^{4\alpha/(4\alpha+4\beta+1)}$	$\varepsilon(\log \varepsilon^{-1})^{\beta+1/4}$
Severely ill-posed ($\sigma_k = \exp\{\beta k\}$)	$((\log \varepsilon^{-1})/\beta)^{-\alpha}$	$\varepsilon^{\alpha/(\alpha+\beta)}$

We do not present the results for extremely ill-posed inverse problems with generalized analytic f in this short description because the asymptotics are more complicated in this case. Roughly speaking, these asymptotics are determined by a piecewise linear function $u_\varepsilon^{\text{lin}} = u_\varepsilon^{\text{lin}}(r_\varepsilon)$ and, principally, seem to be of a new type; see Section 4.6 and remarks therein for details. (Moreover, in subsequent sections, we consider this case only for $q = 2$.)

Consider now the "sparse" case $q \in (0, 2)$. Then, the results noted above still hold true for severely ill-posed inverse problems with Sobolev f or analytic f. For mildly ill-posed inverse problems with analytic f, we also get the same separation rates r_ε^*. However, the situation for mildly ill-posed inverse problems with Sobolev f is more delicate. More precisely, let $\alpha > 0$, $\beta > 0$ and set $\lambda = (\alpha + \beta)/2 - \beta/q$. If $\lambda > 0$, then the sharp asymptotics are of the Gaussian type (2.8) with

$$u_\varepsilon = c_3 \varepsilon^{-(2\alpha+1/q-1/2)/(\alpha+\beta(1-2/q))} r_\varepsilon^{(2(\alpha+\beta)+1/q)/(\alpha+\beta(1-2/q))}$$

for some constant $c_3 = c_3(\alpha, \beta, q) > 0$, while the separation rates r_ε^* are of the form

$$r_\varepsilon^* = \varepsilon^{(2\alpha+1/q-1/2))/(2(\alpha+\beta)+1/q)}.$$

The corresponding rate optimal tests are of more complicated structure and are based on a different extreme problem; see Section 4.7.

On the other hand, if $\lambda \le 0$, then the sharp asymptotics are of the following degenerate type:

$$\beta_\varepsilon(\alpha) = (1-\alpha)\Phi(-D_\varepsilon) + o(1), \qquad \gamma_\varepsilon = \Phi(-D_\varepsilon) + o(1),$$

where $D_\varepsilon = n_\varepsilon^{-\beta} r_\varepsilon/\varepsilon - \sqrt{2\log(n_\varepsilon)}$, $n_\varepsilon = r_\varepsilon^{-1/\alpha}$, while the separation rates r_ε^* are of the form

$$r_\varepsilon^* = \Lambda \varepsilon^{\alpha/(\alpha+\beta)} \big(\log(\varepsilon^{-1})\big)^{\alpha/2(\alpha+\beta)}, \qquad \Lambda = \big(2/(\alpha+\beta)\big)^{\alpha/2(\alpha+\beta)}.$$

The corresponding rate optimal tests are based on a simple thresholding rule.

It seems natural to compare the separation rates r_ε^* in the detection problem with the minimax accuracy R_ε in the corresponding estimation problem using loss functions that correspond to the norm which is used for bounding the alternative away from zero. We compare below the above mentioned minimax rates of testing with the corresponding minimax rates of estimation. The minimax estimation

TABLE 3
The estimation rates $R_\varepsilon = R_\varepsilon(\mathcal{F})$ as $\varepsilon \to 0$

Estimation problem	**Sobolev classes ($a_k = k^\alpha$)**	**Analytic classes ($a_k = \exp\{\alpha k\}$)**
Mildly ill-posed ($\sigma_k = k^\beta$)	$\varepsilon^{2\alpha/(2\alpha+2\beta+1)}$	$\varepsilon(\log\varepsilon^{-1})^{\beta+1/2}$
Severely ill-posed ($\sigma_k = \exp\{\beta k\}$)	$(\log\varepsilon^{-1})^{-\alpha}$	$\varepsilon^{\alpha/(\alpha+\beta)}$

problem for the GWNM (1.1) [or, equivalently, for the GSM (1.2)] was studied very intensively in statistical ill-posed inverse problems; see, for example, [1] (and references therein), [2, 3, 5] and [6]. The main object of the study is the minimax quadratic risk, defined by $R_\varepsilon^2(\mathcal{F}) = \inf_{\hat f}\sup_{f\in\mathcal{F}} E_{\varepsilon,f}\|\hat f - f\|^2$, where the infimum is taken over all possible estimators $\hat f$ of f, based on observations from the GWNM (1.1).

For the main types of the ill-posed inverse problems and classes of functions under consideration, with $q = 2$, the estimation rates $R_\varepsilon = R_\varepsilon(\mathcal{F})$ as $\varepsilon \to 0$ are presented in Table 3; see, for example, [1]. For mildly ill-posed inverse problems with Sobolev f, $q \in (0,2)$, one has, for $\lambda > 0$,

$$R_\varepsilon \asymp \varepsilon^{(\alpha-1/2+1/q)/(\alpha+\beta+1/q)},$$

while, for $\lambda < 0$, the estimation rates R_ε coincide with the separation rates r_ε^* in the corresponding detection problem; see, for example, [11], Section 2.8, and references therein. Observe that the minimax rates of testing are faster than the corresponding minimax rates of estimation (as it is common in nonparametric inference, see, e.g., [11], Sections 2.10 and 3.5.1), except for the cases of mildly ill-posed inverse problems with Sobolev f, $q \in (0,2)$ and $\lambda \le 0$, and the cases of severely ill-posed inverse problems with Sobolev f or analytic f, $q = 2$. [To the best of our knowledge, we are not aware of any minimax estimation results with the case of severely ill-posed inverse problems for Sobolev f or analytic f, $q \in (0,2)$.]

Returning to the signal detection problem, note that, except for the mildly ill-posed inverse problems with Sobolev f, $q \in (0,2)$ and $\lambda \le 0$, or analytic f, the aforementioned separation rates r_ε^* still hold true for the *known* parameters (α,β,q) associated with the classes of functions and the ill-posed inverse problems under consideration. For these cases, the rate-optimal tests also depend on the parameters (α,β,q). In practice, the parameters α and q associated with the considered functional classes are typically unknown and, very often, the statistician is not confident about the value of the parameter β associated with the sequence b_k, $k \in \mathbb{N}$. For unknown parameters $(\alpha,\beta,q) \in \Sigma \subset \mathbb{R}_+^2 \times (0,2] := (0,\infty)\times(0,\infty)\times(0,2]$, we have the so-called *adaptive* problems: in order to distinguish between the null hypothesis and the "combined" alternative, which

TABLE 4
The adaptive separation rates $r_\varepsilon^{\mathrm{ad}}$ as $\varepsilon \to 0$

Detection problem	**Sobolev classes** ($\boldsymbol{a_k = k^\alpha}$)	**Analytic classes** ($\boldsymbol{a_k = \exp\{\alpha k\}}$)
Mildly ill-posed ($\sigma_k = k^\beta$)	$(\tilde{\varepsilon}_1)^{4\alpha/(4\alpha+4\beta+1)}$	$\varepsilon(\log \varepsilon^{-1})^{\beta+1/4}$
Severely ill-posed ($\sigma_k = \exp\{\beta k\}$)	$((\log \varepsilon^{-1})/\beta)^{-\alpha}$	$(\tilde{\varepsilon}_2)^{\alpha/(\alpha+\beta)}$

corresponds to a wide enough compact set $\Sigma \subset \mathbb{R}_+^2 \times (0, 2]$, it does not suffice to just require $u_\varepsilon = u_\varepsilon(r_\varepsilon(\alpha, \beta, q), \alpha, \beta, q) \to \infty$ for all $(\alpha, \beta, q) \in \Sigma$; instead, one needs that it should tend to ∞ faster than some family $u_\varepsilon^{\mathrm{ad}} \to \infty$, which is a "payment" for adaptation; see [13].

Adaptive rate optimality results for the specific ill-posed inverse problems under consideration are discussed in detail in Section 5. Below, we give a short description of these results for mildly and severely ill-posed inverse problems with Sobolev f or analytic f, $q \in (0, 2]$.

For mildly ill-posed inverse problems with Sobolev f, $\lambda \le 0$, or analytic f, one has $u_\varepsilon^{\mathrm{ad}} \asymp 1$, while for Sobolev f, with $\lambda > 0$, one has $u_\varepsilon^{\mathrm{ad}} = \sqrt{\log\log \varepsilon^{-1}}$. On the other hand, for severely ill-posed inverse problems with Sobolev f or analytic f (as well as for extremely ill-posed inverse problems for generalized analytic f, $q = 2$), one has $u_\varepsilon^{\mathrm{ad}} = \log\log \varepsilon^{-1}$. These yield the *adaptive separation rates* $r_\varepsilon^{\mathrm{ad}}$ as $\varepsilon \to 0$ presented in Table 4 (here $q = 2$ for mildly ill-posed problems with Sobolev f), where $\tilde{\varepsilon}_1 = \varepsilon\sqrt[4]{\log\log \varepsilon^{-1}}$ and $\tilde{\varepsilon}_2 = \varepsilon\sqrt{\log\log \varepsilon^{-1}}$. On the other hand, for the "sparse" case $q \in (0, 2)$ and $\lambda > 0$, for mildly ill-posed inverse problems with Sobolev f, the adaptive separation rates $r_\varepsilon^{\mathrm{ad}}$ are of the form

$$r_\varepsilon^{\mathrm{ad}} = \tilde{\varepsilon}_2^{(2\alpha+1/q-1/2))/(2(\alpha+\beta)+1/q)}.$$

As we shall see in Section 5, the rate-optimal adaptive tests are of rather simple structure for all problems under consideration [except for the mildly ill-posed problems with Sobolev f, $q \in (0, 2)$]: they are based on combinations of tests based on a grid of centered and normalized statistics of χ^2-type and on simple thresholding. For mildly ill-posed problems with Sobolev f, $q \in (0, 2)$, the rate-optimal adaptive tests are, however, more complicated; see [11], Chapter 7.

Finally, we mention that, in most cases, the arguments for $q = 2$ do not work for $q \in (0, 2)$ [e.g., for mildly ill-posed inverse problems with Sobolev f, $q \in (0, 2)$]; instead, we have different extreme problems in the space of sequences of probability measures. The key ideas for the proofs of the main results for $0 < q < 2$, as well as possible extensions to $0 < q \le \infty$ and to a wider range of sets under the alternative (i.e., to replace the constraint in the ℓ^2-norm by a constraint in the ℓ^p-norms, $0 < p \le \infty$), are discussed in some detail in the supplementary material [10], Section 10.

4. Minimax signal detection in ill-posed inverse problems: Rate and sharp asymptotics. In this section, we consider the GSM (2.2) and the hypothesis testing problem (2.3) and (2.4).

4.1. *A general result for l^q-ellipsoids: The "standard" case $q = 2$.*

THEOREM 4.1. *Let $q = 2$, let u_ε be determined by the extreme problem* (3.1), *let the coefficients w_k, $k \in \mathbb{N}$, and w_0 be as in* (3.2) *and consider the family of tests $\psi_{\varepsilon,H}$ given by* (3.3).

(1) (a) *If $u_\varepsilon \to 0$, then $\beta_\varepsilon(r_\varepsilon, \alpha) \to 1-\alpha$ for any $\alpha \in (0, 1)$ and $\gamma_\varepsilon(r_\varepsilon) \to 1$; that is, minimax testing is impossible. If $u_\varepsilon = O(1)$, then* $\liminf \beta_\varepsilon(r_\varepsilon, \alpha) > 0$ *for any $\alpha \in (0, 1)$ and* $\liminf \gamma_\varepsilon(r_\varepsilon) > 0$; *that is, minimax consistent testing is impossible.*
(b) *If $u_\varepsilon \asymp 1$ and $w_0 = o(1)$, then the family of tests $\psi_{\varepsilon,H}$ of the form* (3.3) *with $H = H^{(\alpha)}$ and $H = u_\varepsilon/2$ are asymptotically minimax, that is,*

$$\alpha_\varepsilon(\psi_{\varepsilon,H^{(\alpha)}}) \le \alpha + o(1), \qquad \beta_\varepsilon\bigl(\Theta(r_\varepsilon), \psi_{\varepsilon,H^{(\alpha)}}\bigr) = \beta_\varepsilon(r_\varepsilon, \alpha) + o(1),$$

$$\gamma_\varepsilon\bigl(\Theta(r_\varepsilon), \psi_{\varepsilon,u_\varepsilon/2}\bigr) = \gamma_\varepsilon(r_\varepsilon) + o(1),$$

and the sharp asymptotics (2.8) *hold true.*
(2) *If $u_\varepsilon \to \infty$, then the family of tests $\psi_{\varepsilon,H}$ of the form* (3.3) *with $H \sim cu_\varepsilon$ are asymptotically minimax consistent for any $c \in (0, 1)$, that is, $\gamma_\varepsilon(\Theta(r_\varepsilon), \psi_{\varepsilon,T_\varepsilon}) \to 0$.*

The proof is given in the supplementary material [10], Section 11.1.
Theorem 4.1 shows that the asymptotics of the quality of testing is determined by the asymptotics of values u_ε of the the extreme problem (3.1). This latter problem is studied by using Lagrange multipliers. Then the extreme sequence in the above mentioned extreme problem is of the form

$$\tilde{\eta}_k^2 = z_0^2\sigma_k^2\bigl(1 - Aa_k^2\bigr)_+, \qquad k \in \mathbb{N}, \tag{4.1}$$

where $(a)_+ = \max\{0, a\}$, and the quantities $z_0 = z_{0,\varepsilon}$ and $A = A_\varepsilon$ are determined by the equations

$$\sum_{k\in\mathbb{N}} \sigma_k^2\tilde{\eta}_k^2 = r_\varepsilon^2, \qquad \sum_{k\in\mathbb{N}} a_k^2\sigma_k^2\tilde{\eta}_k^2 = 1. \tag{4.2}$$

Note that the quantity A determines the *efficient dimension m* in the specific ill-posed inverse problems considered below: if a_k is an increasing sequence (it is assumed further), the efficient dimension is a quantity $m = m_\varepsilon \in [1, \infty)$ such that $Aa_{[m]}^2 \le 1 < Aa_{[m]+1}^2$.

REMARK 4.1. Since, in the ill-posed problems under consideration, $\sigma_k \to \infty$ as $k \to \infty$, it is immediate that $\sum_{k\in\mathbb{N}} \sigma_k^4 = \infty$. Under this condition, and the fact that $a_k \to \infty$ as $k \to \infty$, one can see that, for r_ε small enough, the equations in (4.2) have a unique solution; see Proposition 11.2 in the supplementary material [10], Section 11.9.

REMARK 4.2. Let $u_\varepsilon = u_\varepsilon(r_\varepsilon)$ be the value of the extreme problem (3.1) with $a = \{a_k\}_{k\in\mathbb{N}}$ and $\sigma = \{\sigma_k\}_{k\in\mathbb{N}}$ associated with $\Theta_\varepsilon = \Theta(r_\varepsilon)$ given by (2.4), and let $\tilde{u}_\varepsilon = \tilde{u}_\varepsilon(r_\varepsilon)$ be the corresponding value of the extreme problem similar to (3.1) with $\tilde{a} = Ca = \{Ca_k\}_{k\in\mathbb{N}}$ and $\tilde{\sigma} = D\sigma = \{D\sigma_k\}_{k\in\mathbb{N}}$ in (2.4), for some positive constants C and D. Then it is easily seen that the relation $\tilde{u}_\varepsilon(r_\varepsilon) = (CD)^{-2}u_\varepsilon(Cr_\varepsilon)$ holds true.

4.2. *Application to mildly ill-posed inverse problems with the Sobolev class of functions.* Consider first the "standard" case $q = 2$.

THEOREM 4.2. *Let $q = 2$, $a_k = k^\alpha$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.*

(a) *The sharp asymptotics* (2.8) *hold with the value u_ε of the extreme problem* (3.1) *determined by* (11.10) *in* [10].

(b) *The asymptotically minimax family of tests $\psi_{\varepsilon,H}$ are determined by the family of test statistics t_ε given by* (3.3) *with w_k, $k \in \mathbb{N}$, and w_0 as in* (3.2), *and with $\{\tilde{\eta}_k\}_{k\in\mathbb{N}}$ given by* (11.9) *with m determined by* (11.10) *in* [10].

(c) *The separation rates are of the form*

$$r_\varepsilon^* = \varepsilon^{4\alpha/(4\alpha+4\beta+1)}. \tag{4.3}$$

The proof is given in the supplementary material [10], Section 11.3.

REMARK 4.3. It follows from the evaluations of the functions J_0, J_1 and J_2 used to express (4.2) (see (11.2) in [10]) that their asymptotics are determined by the tails of the sequences $a_k = k^\alpha$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$. For this reason, in view of Remark 4.2, we get the sharp asymptotics (11.11) in [10] for the sequences $a_k \sim k^\alpha$ and $\sigma \sim k^\beta$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$, and similar rate asymptotics for the sequences $a_k \asymp k^\alpha$, $\sigma \asymp k^\beta$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$. In both cases, the separation rates are still of the form (4.3).

Unlike the case $q = 2$, the "sparse" case $q \in (0, 2)$ is not directly linked to Theorem 4.1; it will be considered separately in Section 4.7.

4.3. *Application to severely ill-posed inverse problems with the class of analytic functions.*

THEOREM 4.3. *Let $q \in (0, 2]$, $a_k = \exp(\alpha k)$ and $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$.*

(a) *The asymptotically minimax consistent family of tests $\psi_{\varepsilon,H}$ are determined by the family of test statistics t_ε given by* (3.3) *with w_k, $k \in \mathbb{N}$, as in* (3.2), *and with $\{\tilde{\eta}_k\}_{k\in\mathbb{N}}$ given by* (11.17) *with m determined by* (11.18) *in* [10].

(b) *The separation rates are of the form*

$$r_\varepsilon^* = \varepsilon^{\alpha/(\alpha+\beta)}. \tag{4.4}$$

The proof is given in the supplementary material [10], Section 11.5.

REMARK 4.4. We do not consider sharp asymptotics in this case, since the assumption $w_0 = o(1)$ does not hold for $\beta > 0$ in the case $q = 2$.

REMARK 4.5. Similar to Remark 4.3, the asymptotics (11.19) in [10] hold true for the sequences $a_k \asymp \exp(\alpha k)$ and $\sigma_k \asymp \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$. In this case, the separation rates are still of the form (4.4). Remark 4.4 still applies to these cases, too.

4.4. *Application to severely ill-posed inverse problems with the Sobolev class of functions.*

THEOREM 4.4. *Let* $q \in (0, 2], a_k = k^\alpha$ *and* $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.

(a) *The asymptotically minimax consistent family of tests* $\psi_{\varepsilon,H}$ *are determined by the family of test statistics* t_ε *given by* (3.3) *with* w_k, $k \in \mathbb{N}$, *as in* (3.2), *and with* $\{\tilde{\eta}_k\}_{k\in\mathbb{N}}$ *given by* (11.20) *with* m *determined by* (11.21) *in* [10].

(b) *The separation rates are of the form*

$$r_\varepsilon^* = \big((\log(\varepsilon^{-1}))/\beta\big)^{-\alpha}. \tag{4.5}$$

The proof is given in the supplementary material [10], Section 11.6.

REMARK 4.6. A stronger result is possible in this case. In view of (11.21) in [10], the relation (4.5) determines *sharp separation rates* r_ε^* in the following sense:

(a) if $\liminf(r_\varepsilon/r_\varepsilon^*) > 1$, then $u_\varepsilon \to \infty$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 0$;

(b) if $\limsup(r_\varepsilon/r_\varepsilon^*) < 1$, then $u_\varepsilon \to 0$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 1$, and the minimax testing is impossible.

Moreover the relation $(r_\varepsilon^*)^{-1/\alpha} = ((\log(\varepsilon^{-1}) - \alpha \log\log(\varepsilon^{-1}))/\beta) + O(1)$, determines the sharper separation rates r_ε^* in the following sense:

(c) if $\liminf(r_\varepsilon^{-1/\alpha} - (r_\varepsilon^*)^{-1/\alpha}) = -\infty$, then $u_\varepsilon \to \infty$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 0$;

(d) if $\limsup(r_\varepsilon^{-1/\alpha} - (r_\varepsilon^*)^{-1/\alpha}) = +\infty$, then $u_\varepsilon \to 0$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 1$, and the testing is asymptotically impossible.

REMARK 4.7. We do not consider sharp asymptotics in this case, since the assumption $w_0 = o(1)$ does not hold for $\beta > 0$ in the case $q = 2$.

REMARK 4.8. Similar to Remark 4.3, the asymptotics (11.22) in [10] and the sharp separation rates (4.5) mentioned in Remark 4.6 hold true for the sequences $\sigma_k \asymp \exp(\beta k)$, $k \in \mathbb{N}$, $\beta > 0$. The dependence on the sequence $\{a_k\}_{k\in\mathbb{N}}$ is, however, more delicate. One can actually show that the sharp separation rates (4.5) mentioned in Remark 4.6 are still of the same form for $a_k \sim k^\alpha$, $k \in \mathbb{N}$, $\alpha > 0$. Remark 4.7 still applies to these cases too.

4.5. *Application to mildly ill-posed inverse problems with the class of analytic functions*. Here, we consider the "standard" case $q = 2$. [The "sparse" case $q \in (0, 2)$ will be discussed in Remark 4.11.]

THEOREM 4.5. *Let $q = 2$, $a_k = \exp(\alpha k)$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.*

(a) *The sharp asymptotics* (2.8) *hold with the the value u_ε of the extreme problem* (3.1) *determined by* (11.29) *in* [10].

(b) *The asymptotically minimax family of tests are determined by the test statistics t_ε given by* (3.3) *with w_k, $k \in \mathbb{N}$ and w_0 as in* (3.2), *and with $\{\tilde{\eta}_k\}_{k\in\mathbb{N}}$ given by* (11.24) *with m determined by* (11.30) *in* [10].

(c) *The separation rates are of the form*

$$r_\varepsilon^* = \varepsilon\bigl(\log \varepsilon^{-1}\bigr)^{\beta+1/4}. \tag{4.6}$$

The proof is given in the supplementary material [10], Section 11.7.

REMARK 4.9. It is also easy to see that, uniformly over $(\alpha, \beta) \in \Sigma$, for any compact set $\Sigma \subset \mathbb{R}_+^2$, the efficient dimension $m = m_\varepsilon(\alpha, \beta)$ satisfies

$$\begin{aligned} m_\varepsilon(\alpha, \beta) &\sim \frac{2\log(\varepsilon^{-1}) - \log(u_\varepsilon)}{2\alpha} \\ &\asymp \log\bigl(\varepsilon^{-1}\bigr) \qquad \text{as } \log(u_\varepsilon) = o\bigl(\log\bigl(\varepsilon^{-1}\bigr)\bigr). \end{aligned} \tag{4.7}$$

REMARK 4.10. As in Remark 4.3, the asymptotics (11.30) in [10] hold true for the sequences $a_k \sim \exp(\alpha k)$ and $\sigma_k \sim k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$. Similar rate asymptotics hold true for the sequences $a_k \asymp \exp(\alpha k)$ and $\sigma_k \asymp k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$. In both cases, the separation rates are still of the form (4.6).

REMARK 4.11. The rate asymptotics of Theorem 4.5 hold true uniformly in $q \in [\delta, 2]$, for any $\delta \in (0, 2)$. Indeed, in view of the embedding

$$\Theta_q \subset \Theta_2, \qquad \Theta_q(r_\varepsilon) \subset \Theta_2(r_\varepsilon) \tag{4.8}$$

for the sets defined by (2.4), it suffices to establish the lower bounds. Let $\Theta_q^\alpha(L) = \{\eta \in l^2 : \sum_{k\in\mathbb{N}} |\exp(\alpha k) k^\beta \eta_k|^q \le L^q\}$. [Note that the set $\Theta_q^\alpha(1)$ just corresponds to the set under the alternative considered above.] For any $\alpha_1 = \alpha + \delta, \delta > 0$, we have the embedding

$$\Theta_2^{\alpha_1}(cL) \subset \Theta_q^{\alpha}(L), \qquad c = c(q,\delta) = \bigl(\exp\bigl(2q\delta/(2-q)\bigr) - 1\bigr)^{(2-q)/2q} \tag{4.9}$$

and $c(q,\delta) \to \exp(\delta)$ as $q \nearrow 2$. Using Hölder's inequality, the above embedding follows easily on noting that

$$\begin{aligned}\sum_{k\in\mathbb{N}} |e^{\alpha k} k^{\beta} \eta_k|^q &\le \biggl(\sum_{k\in\mathbb{N}} (e^{\alpha k} k^{\beta} \eta_k)^2\biggr)^{2/q} \biggl(\sum_{k\in\mathbb{N}} e^{-2kq\delta/(2-q)}\biggr)^{1-q/2} \\ &= \biggl(c^{-2} \sum_{k\in\mathbb{N}} (e^{\alpha k} k^{\beta} \eta_k)^2\biggr)^{2/q}.\end{aligned}$$

Since the separation rates from Theorem 4.5 do not depend on α and c, in view of Remark 4.10, the rate asymptotics of Theorem 4.5 hold true uniformly in $q \in [\delta, 2]$, for any $\delta \in (0, 2)$.

4.6. *Application to extremely ill-posed inverse problems with the class of generalized analytic functions*. We consider the case $q = 2$ only. Assume that $\{a_k\}_{k\in\mathbb{N}}$ and $\{\sigma_k\}_{k\in\mathbb{N}}$ are increasing sequences such that

$$\lim_{k\to\infty} \sigma_{k+1}/\sigma_k \to \infty, \qquad \liminf_{k\to\infty} a_{k+1}/a_k = c, \qquad c \in (1, \infty]. \tag{4.10}$$

In order to describe the asymptotics of the value $u_\varepsilon = u_\varepsilon(r)$ of the extreme problem (3.1), we introduce the following functions.

Let $m \in \mathbb{N}$, $m \ge 2$, $\Delta_m^* = [1/a_m, 1/a_{m-1}]$, and for $r < 1/a_1$ take $m = m(r) \ge 2$ such that $r \in \Delta_m^*$. Consider now the piecewise quadratic (in r^2) function defined by

$$\bigl(u_\varepsilon^*(r)\bigr)^2 = \frac{1}{2\varepsilon^4 (a_m^2 - a_{m-1}^2)^2} \biggl(\frac{(a_m^2 r^2 - 1)^2}{\sigma_{m-1}^4} + \frac{(1 - a_{m-1}^2 r^2)^2}{\sigma_m^4}\biggr), \tag{4.11}$$

and the piecewise linear (in r^2) function defined by

$$u_\varepsilon^{\mathrm{lin}}(r) = \frac{1}{\varepsilon^2 (a_m^2 - a_{m-1}^2)} \biggl(\frac{a_m^2 r^2 - 1}{\sigma_{m-1}^2} + \frac{1 - a_{m-1}^2 r^2}{\sigma_m^2}\biggr). \tag{4.12}$$

THEOREM 4.6. *Let $u_\varepsilon = u_\varepsilon(r)$ be the value of the extreme problem* (3.1). *Let $(u_\varepsilon^*(r))^2$ be the piecewise quadratic (in r^2) function defined by* (4.11), *and let $u_\varepsilon^{\mathrm{lin}}(r)$ be the piecewise linear (in r^2) function defined by* (4.12), *where $\{a_k\}_{k\in\mathbb{N}}$ and $\{\sigma_k\}_{k\in\mathbb{N}}$ be increasing sequences satisfying* (4.10).

(a) (*Sharp asymptotics of u_ε.*) *The families $u_\varepsilon(r)$ and $u_\varepsilon^*(r)$ are related by*

$$u_\varepsilon(r_\varepsilon) \sim u_\varepsilon^*(r_\varepsilon) \qquad \text{as } r_\varepsilon \to 0. \tag{4.13}$$

(b) (*Rate asymptotics of* u_ε.) *The families* $u_\varepsilon(r)$ *and* $u_\varepsilon^{\mathrm{lin}}(r)$ *are related by*

$$u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\bigl(1/2+o(1)\bigr)\le u_\varepsilon(r_\varepsilon)\le u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\bigl(1/\sqrt{2}+o(1)\bigr)\qquad \text{as } r_\varepsilon\to 0. \tag{4.14}$$

(c) (*Distinguishability conditions*.) *Consider the GSM* (2.2) *and the hypothesis testing problem* (2.3) *and* (2.4). *Then*

$$\gamma_\varepsilon(r_\varepsilon)\to 0 \quad \text{if and only if} \quad u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\to\infty;$$
$$\gamma_\varepsilon(r_\varepsilon)\to 1 \quad \text{if and only if} \quad u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\to 0.$$

The proof is given in the supplementary material [10], Section 11.9.

REMARK 4.12. It is easy to see that relation (4.14) is true uniformly over all sequences $\{a_k\}_{k\in\mathbb{N}}$ and $\{\sigma_k\}_{k\in\mathbb{N}}$ such that $\sigma_{k+1}/\sigma_k\ge B_k$, $B_k\to\infty$, and $a_{k+1}/a_k>c$, as $k\ge k_0$, $k_0\ge 1$.

REMARK 4.13. We do not consider sharp asymptotics in this case, since the assumption $w_0=o(1)$ does not hold under assumption (4.10).

REMARK 4.14. The relation $u_\varepsilon^{\mathrm{lin}}(r_\varepsilon^*)\asymp 1$ determines the separation rates r_ε^* that are rather sharp in the follows sense. Let $r_\varepsilon^*=a_m^{-1}$ for some $m\in\mathbb{N}, m\to\infty$, and let $r^2=(1+b)(r_\varepsilon^*)^2\in(a_m^{-2},a_{m-1}^{-2}), b>0$. Then, one has $u_\varepsilon^{\mathrm{lin}}(r)=u_\varepsilon^{\mathrm{lin}}(r_\varepsilon^*)(1+k_m b)$, where, as $m\to\infty$,

$$\begin{aligned} k_m &= \frac{\sigma_m^2}{\sigma_{m-1}^2}\frac{1-(\sigma_{m-1}a_{m-1}/\sigma_m a_m)^2}{1-(a_{m-1}/a_m)^2} \\ &\sim \frac{\sigma_m^2}{\sigma_{m-1}^2(1-(a_{m-1}/a_m)^2)}\asymp\frac{\sigma_m^2}{\sigma_{m-1}^2}\to\infty. \end{aligned}$$

Therefore, in order to obtain $u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\to\infty$, it suffices to take $r_\varepsilon=r_\varepsilon^*(1+\delta)$ for any $\delta>0$. On the other hand, let $r^2=(1-b)(r_\varepsilon^*)^2\in(a_{m+1}^{-2},a_m^{-2}), b\in(0,1)$. Then, similarly, one has $u_\varepsilon^{\mathrm{lin}}(r)=u_\varepsilon^{\mathrm{lin}}(r_\varepsilon^*)(1-l_m b)$, where, as $m\to\infty$,

$$l_m=\frac{1-(\sigma_m a_m/\sigma_{m+1}a_{m+1})^2}{1-(a_m/a_{m+1})^2}\sim\frac{1}{1-(a_m/a_{m+1})^2}\asymp 1.$$

If $a_{m+1}/a_m\to\infty$ as $m\to\infty$, then, in order to obtain $u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\to 0$, one needs to take r_ε such that $r_\varepsilon/r_\varepsilon^*\to 0$, and if $a_{m+1}/a_m\to c, 1<c<\infty$, then, for $u_\varepsilon^{\mathrm{lin}}(r_\varepsilon)\to 0$, one needs to take $r_\varepsilon<r_\varepsilon^*/c$. Thus, the conditions for distinguishability and nondistinguishability could be nonsymmetric in these problems.

REMARK 4.15. Let us consider the example

$$a_k=\exp\bigl(\alpha k^\tau\bigr),\qquad \alpha>0,\tau\ge 1,\qquad \sigma_k=\exp\bigl(\beta k^\gamma\bigr),\qquad \beta>0,\gamma>1.$$

For the moment, let us forget that $m \in \mathbb{N}$ and define $m = m(r) \in \mathbb{R}_+$ by the equality $r = a_m^{-1}$, that is, $m(r) = (\log(r^{-1})/\alpha)^{1/\tau}$. Set also

$$\hat{u}_\varepsilon(r) = (\varepsilon a_{m(r)}\sigma_{m(r)})^{-2} = (r/\varepsilon)^2 \exp(-2\beta(\log(r^{-1})/\alpha)^{\gamma/\tau}).$$

Observe that $\hat{u}_\varepsilon(r) = u_\varepsilon^{\text{lin}}(r)$ as $r = a_m^{-1}$, $m \in \mathbb{N}$. On the other hand, one can check that the function $\hat{u}_\varepsilon(r)$ is a convex function in r^2 for $r > 0$ small enough. Therefore, $\hat{u}_\varepsilon(r) < u_\varepsilon^{\text{lin}}(r)$ as $r \neq a_m^{-1}$ for any $m \in \mathbb{N}$, and the condition $\hat{u}_\varepsilon(r_\varepsilon) \to \infty$ implies $u_\varepsilon^{\text{lin}}(r_\varepsilon) \to \infty$. However, it is possible that $u_\varepsilon^{\text{lin}}(r_\varepsilon) \to \infty$ when $\hat{u}_\varepsilon(r_\varepsilon) = O(1)$, in general. For instance, let $\tau = \gamma$. Then

$$\hat{u}_\varepsilon(r) = \varepsilon^{-2} r^{2+2\beta/\alpha}.$$

If $r_\varepsilon = a_m^{-1}$ and $u_\varepsilon^{\text{lin}}(r_\varepsilon) \asymp 1$, then it was noted in Remark 4.14 that $u_\varepsilon^{\text{lin}}(r_\varepsilon(1+\delta)) \to \infty$ for any $\delta > 0$, but $\hat{u}_\varepsilon(r_\varepsilon(1+\delta)) \asymp 1$. The same holds for $\gamma < \tau$.

The relation $\hat{u}_\varepsilon(\tilde{r}_\varepsilon) \asymp 1$ determines the family $\tilde{r}_\varepsilon$. Note that if $r_\varepsilon/\tilde{r}_\varepsilon \to \infty$, then $\hat{u}_\varepsilon(r_\varepsilon) \to \infty$, and since $u_\varepsilon^{\text{lin}}(r) \geq \hat{u}_\varepsilon(r)$, this yields $u_\varepsilon^{\text{lin}}(r_\varepsilon) \to \infty$ and $\gamma_\varepsilon(r_\varepsilon) \to 0$ by Theorem 4.6. However, this family is not a family of separation rates, at least if $\gamma \leq \tau$, because the condition $r_\varepsilon/\tilde{r}_\varepsilon \to 0$ does not guaranty that $\gamma_\varepsilon(r_\varepsilon) \to 1$.

More precisely, there exists a sequence $\varepsilon_m \to 0$ and $\hat{r}_m = o(\tilde{r}_{\varepsilon_m})$ such that $\gamma_{\varepsilon_m}(\hat{r}_m) \to 0$. In fact, observe that if $\gamma \leq \tau$, then the function $\hat{u}_\varepsilon(r)$ satisfies [uniformly over $\varepsilon > 0$ since ε^2 is a factor in $u_\varepsilon(r)$]

$$\hat{u}_\varepsilon(Br) \asymp \hat{u}_\varepsilon(r) \quad \text{iff} \quad B \asymp 1, \qquad r \to 0.$$

Take a sequence $m \to \infty$ and put $r_m = a_m^{-1}$, $(r_m^{(1)})^2 = r_m^2(1+\delta_m)$, $\hat{r}_m^2 = r_m^2(1+\delta)$, where $\delta_m \to 0$, $\delta_m\sigma_m^2/\sigma_{m-1}^2 \to \infty$, $\delta > 0$. Observe that similarly to evaluations in Remark 4.14, one has, uniformly over $\varepsilon > 0$,

$$\hat{u}_\varepsilon(\hat{r}_m) \asymp \hat{u}_\varepsilon(r_m^{(1)}) \asymp \hat{u}_\varepsilon(r_m) = u_\varepsilon^{\text{lin}}(r_m) \ll u_\varepsilon^{\text{lin}}(r_m^{(1)}) \ll u_\varepsilon^{\text{lin}}(\hat{r}_m), \qquad m \to \infty.$$

Take now $\tilde{r}_m$ and ε_m such that $\hat{u}_{\varepsilon_m}(\tilde{r}_m) \asymp u_{\varepsilon_m}^{\text{lin}}(r_m^{(1)}) \asymp 1$. This implies $\hat{u}_{\varepsilon_m}(\tilde{r}_m) \gg \hat{u}_{\varepsilon_m}(r_m)$ and $\tilde{r}_m \gg r_m \asymp \hat{r}_m$. By construction, we see that the sequence $\tilde{r}_m$ satisfies $\hat{u}_{\varepsilon_m}(\tilde{r}_m) \asymp 1$ and $\hat{r}_m = o(\tilde{r}_m)$, but $u_{\varepsilon_m}^{\text{lin}}(\hat{r}_m) \to \infty$, which yields $\gamma_{\varepsilon_m}(\hat{r}_m) \to 0$.

4.7. *Mildly ill-posed inverse problems with l^q-ellipsoids for Sobolev classes of functions*: *The "sparse" case $q \in (0,2)$.* Unlike the "standard" case $q = 2$, the sharp and rate optimality results for the "sparse" case $q \in (0,2)$ are of different nature and are not directly linked with Theorem 4.1, but can be obtained from a hitherto unknown link with results obtained in another context and presented in Sections 4.4.2–4.4.3 of [11]. For completeness and an immediate access to these results, we formulate and present them below.

Consider the extreme problem

$$u_\varepsilon^2 = 2 \inf \sum_{i \in \mathbb{N}} h_i^2 \sinh^2(z_i^2/2), \tag{4.15}$$

where the infimum is taken over sequences (h_i, z_i), $h_i \in [0, 1]$, $z_i \geq 0$, $i \in \mathbb{N}$, such that

$$\sum_{i\in\mathbb{N}} i^{2\beta} h_i z_i^2 \geq (\tilde{r}_\varepsilon^2/\varepsilon^2), \qquad \sum_{i\in\mathbb{N}} i^{q(\alpha+\beta)} h_i z_i^q \leq (1/\varepsilon^q), \tag{4.16}$$

where $\tilde{r}_\varepsilon = r_\varepsilon(1-\delta_\varepsilon)$, $\delta_\varepsilon > 0$, $\delta_\varepsilon \to 0$, $\delta_\varepsilon \log(\varepsilon^{-1}) \to \infty$.

Set $\lambda = (\alpha+\beta)/2 - \beta/q$. If $\lambda > 0$, then there exist extreme sequences $h_{i,\varepsilon} \in (0, 1]$, $z_{i,\varepsilon} > 0$, in the problem (4.15) and (4.16), and we have the asymptotics of the form

$$u_\varepsilon^2 \sim c_0 n h_0^2, \tag{4.17}$$

where the quantities $n = n_\varepsilon$ and $h_0 = h_{0,\varepsilon}$ are determined by the relations

$$c_1 n^{\beta+1/2} h_0^{1/2} \sim r_\varepsilon/\varepsilon c_2 n^{\alpha+\beta+1/q} h_0^{1/q} \sim 1/\varepsilon \tag{4.18}$$

for some constants $c_l = c_l(\alpha, \beta, q) > 0$, $l = 0, 1, 2$, which, in turn, imply

$$u_\varepsilon \sim c_3 \varepsilon^{-(2\alpha+1/q-1/2)/(\alpha+\beta(1-2/q))} r_\varepsilon^{(2(\alpha+\beta)+1/q)/(\alpha+\beta(1-2/q))} \tag{4.19}$$

for some constant $c_3 = c_3(\alpha, \beta, q) > 0$. (The quantity $n = n_\varepsilon \to \infty$ plays the role of the "efficient dimension" in the problem.)

Set $Q_{\varepsilon,i} = \sqrt{2(\log i + \log\log i + 2\log\log(\varepsilon^{-1}))}$ and consider the events

$$\mathcal{Y}_\varepsilon = \Big\{ y = \{y_i\}_{i\in\mathbb{N}} : \Big(\sup_{i\in\mathbb{N}} |y_i| / (\varepsilon Q_{\varepsilon,i})\Big) > 1 \Big\} \tag{4.20}$$

and the following families of test statistics:

$$l_\varepsilon(y) = u_\varepsilon^{-1} \sum_{i\in\mathbb{N}} h_{\varepsilon,i} \xi(y_i/\varepsilon, z_{\varepsilon,i}), \qquad \xi(t, z) = e^{z^2/2}\cosh(tz) - 1 \tag{4.21}$$

and tests

$$\psi_{\varepsilon,H}^G = \mathbb{1}_{\{l_\varepsilon(y) > H\} \cap \mathcal{Y}_\varepsilon}, \qquad \psi_\varepsilon^D = \mathbb{1}_{\mathcal{Y}_\varepsilon}, \qquad \psi_{\varepsilon,\alpha}^D = \alpha + (1-\alpha)\mathbb{1}_{\mathcal{Y}_\varepsilon}. \tag{4.22}$$

THEOREM 4.7. *Let $q \in (0, 2)$, $a_k = k^\alpha$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$, and set $\lambda = (\alpha+\beta)/2 - \beta/q$.*

(a) *If $\lambda > 0$, then the sharp asymptotics are of the Gaussian type (2.8) with u_ε from (4.15). The tests $\psi_{\varepsilon,H}^G$ of the form (4.22) with $H = H^{(\alpha)}$ and $H = u_\varepsilon/2$ are asymptotically minimax, that is,*

$$\alpha_\varepsilon(\psi_{\varepsilon,H^{(\alpha)}}^G) \leq \alpha + o(1), \qquad \beta_\varepsilon(\Theta(r_\varepsilon), \psi_{\varepsilon,H^{(\alpha)}}^G) = \beta_\varepsilon(r_\varepsilon, \alpha) + o(1),$$

$$\gamma_\varepsilon(\Theta(r_\varepsilon), \psi_{\varepsilon,u_\varepsilon/2}^G) = \gamma_\varepsilon(r_\varepsilon) + o(1).$$

(b) *If* $\lambda \le 0$, *then the sharp asymptotics are of the following degenerate type*: $\beta_\varepsilon(r_\varepsilon, \alpha) = (1-\alpha)\Phi(-D_\varepsilon) + o(1)$, $\gamma_\varepsilon(r_\varepsilon) = \Phi(-D_\varepsilon) + o(1)$, *where* $D_\varepsilon = n_\varepsilon^{-\beta} r_\varepsilon/\varepsilon - \sqrt{2\log(n_\varepsilon)}$, $n_\varepsilon = r_\varepsilon^{-1/\alpha}$. *The tests* ψ_ε^D *and* (*the randomized*) *tests* $\psi_{\varepsilon,\alpha}^D$ *of the form* (4.22) *are asymptotically minimax*, *that is*,

$$\alpha_\varepsilon(\psi_{\varepsilon,\alpha}^D) = \alpha + o(1), \qquad \beta_\varepsilon(\Theta(r_\varepsilon), \psi_{\varepsilon,\alpha}^D) = \beta_\varepsilon(r_\varepsilon, \alpha) + o(1),$$

$$\gamma_\varepsilon(\Theta(r_\varepsilon), \psi_\varepsilon^D) = \gamma_\varepsilon(r_\varepsilon) + o(1).$$

(c) *If* $\lambda > 0$, *then the separation rates are of the form*

$$r_\varepsilon^* = \varepsilon^{(2\alpha+1/q-1/2))/(2(\alpha+\beta)+1/q)}.$$

(d) *If* $\lambda \le 0$, *then the sharp separation rates are of the form*

$$r_\varepsilon^* = \Lambda \varepsilon^{\alpha/(\alpha+\beta)} (\log(\varepsilon^{-1}))^{\alpha/2(\alpha+\beta)}, \qquad \Lambda = (2/(\alpha+\beta))^{\alpha/2(\alpha+\beta)}.$$

Theorem 4.7 is obtained by taking into account the minimax hypothesis testing framework considered in Section 2 and Theorems 4.5 and 6.1 in [11], noting (from their proofs) that the events (thresholding rule) (4.155) in [11] can be replaced by the events (thresholding rule) (4.20). Its proof is omitted. The key ideas of the study are discussed in the supplementary material [10], Section 10.

REMARK 4.16. Similar to Remark 4.3, we get the sharp asymptotics (4.19) for the sequences $a_k \sim k^\alpha$ and $\sigma_k \sim k^\beta$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$, and similar rate asymptotics for the sequences $a_k \asymp k^\alpha$, $\sigma \asymp k^\beta$, $k \in \mathbb{N}$, $\alpha > 0, \beta > 0$. In both cases, the separation rates are still of the form given in Theorem 4.7.

5. Minimax signal detection in ill-posed inverse problems: Adaptivity and rate optimality. The families of tests described in Section 4.1 (except those described in the supplementary material [10], Theorem 7.1) depend on a parameter $\kappa \in \Sigma \subset \mathbb{R}_+^n \times (0,2]$, $n \ge 2$, associated with the sequences $\{a_k\}_{k\in\mathbb{N}}$ and $\{\sigma_k\}_{k\in\mathbb{N}}$, and $q \in (0,2]$, that are involved in the ill-posed inverse problems under consideration, that are usually unknown in practice. For example, if $a_k = \exp(\alpha k^\tau)$ and $\sigma_k = \exp(\beta k^\gamma)$, $k \in \mathbb{N}$, $\alpha > 0$, $\tau \ge 1$, $\beta > 0$, $\gamma > 1$, and if $q \in (0,2]$, then $\kappa \in \Sigma = \{(\alpha,\tau,\beta,\gamma,q)\} = (0,\infty) \cup [1,\infty) \cup (0,\infty) \cup (1,\infty) \cup (0,2] \subset \mathbb{R}_+^4 \times (0,2]$.

It is of paramount importance to construct families of tests that do not depend on the unknown parameter κ and, at the same time, provide the best possible asymptotical minimax efficiency. These families of tests are called *adaptive* (to the parameter κ), and the formal setting is as follows.

5.1. *Adaptive distinguisability and adaptive separation rates*. Let a set $\Sigma = \{\kappa\}$ and a family $r_\varepsilon(\kappa), \kappa \in \Sigma$, be given, where $\varepsilon > 0$ is small. Let the set $\Theta_\varepsilon(\kappa, r_\varepsilon(\kappa))$ be determined by the constraints (2.4) with $a_k = a_k(\kappa)$, $\sigma_k = \sigma_k(\kappa)$,

$k \in \mathbb{N}$, $q = q(\kappa)$, and $r_\varepsilon = r_\varepsilon(\kappa)$, and set $\Theta_\varepsilon(\Sigma) = \bigcup_{\kappa\in\Sigma} \Theta_\varepsilon(\kappa, r_\varepsilon(\kappa))$. We are interested in the following hypothesis testing problem:

$$H_0 : \eta = 0 \quad \text{versus} \quad H_1 : \eta \in \Theta_\varepsilon(\Sigma).$$

We are aiming to find conditions for either $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 1$ or $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 0$, and to constructing asymptotically minimax adaptive consistent families of tests $\psi_\varepsilon^{\mathrm{ad}}$ such that $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma), \psi_\varepsilon^{\mathrm{ad}}) \to 0$ as $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 0$.

Let $u_\varepsilon(\kappa) = u_\varepsilon(\kappa, r_\varepsilon(\kappa))$ be the value of the extreme problem (3.1) for the set $\Theta_\varepsilon = \Theta_\varepsilon(\kappa, r_\varepsilon(\kappa))$. Set $u_\varepsilon(\Sigma) = \inf_{\kappa\in\Sigma} u_\varepsilon(\kappa)$. We are interested in how large $u_\varepsilon(\Sigma)$ should be in order to provide the relation $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 0$. We say that the family $u_\varepsilon^{\mathrm{ad}} = u_\varepsilon^{\mathrm{ad}}(\Sigma) \to \infty$ characterizes *adaptive distinguishability* if there exist constants $0 < d = d(\Sigma) \le D = D(\Sigma) < \infty$ such that

$$\gamma_\varepsilon\big(\Theta_\varepsilon(\Sigma)\big) \to 1 \qquad \text{as } \lim \sup_{\kappa\in\Sigma} u_\varepsilon(\kappa)/u_\varepsilon^{\mathrm{ad}} < d,$$

$$\gamma_\varepsilon\big(\Theta_\varepsilon(\Sigma)\big) \to 0 \qquad \text{as } \lim \inf_{\kappa\in\Sigma} u_\varepsilon(\kappa)/u_\varepsilon^{\mathrm{ad}} > D.$$

We call a family $r_\varepsilon^{\mathrm{ad}}(\kappa)$, $\kappa \in \Sigma$, such that $u_\varepsilon^{\mathrm{ad}} \asymp u_\varepsilon(\kappa, r_\varepsilon^{\mathrm{ad}}(\kappa))$, the family of *adaptive separation rates*.

The relation $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 0$ is possible if $u_\varepsilon(\Sigma) \to \infty$. It was shown in the supplementary material [10], Theorem 7.1, that this relation suffices for the construction of minimax adaptive consistent families of tests for mildly ill-posed inverse problems with the class of analytic functions. This implication, however, does not hold in the remaining ill-posed inverse problems under consideration. In these cases, adaptive distinguishability conditions and adaptive separation rates are sought, and they are the goal of the subsequent sections. In contrast to the above mentioned theorem, there is price to pay for the adaptation. We show that $u_\varepsilon^{\mathrm{ad}} = \sqrt{\log\log \varepsilon^{-1}}$ for the mildly ill-posed inverse problems with the Sobolev class of functions and $u_\varepsilon^{\mathrm{ad}} = \log\log\varepsilon^{-1}$ for other problems under consideration (except the case mildly ill-posed inverse problems with the class of analytic functions). These yield a loss in the separation rates in terms of an extra $\sqrt[4]{\log\log\varepsilon^{-1}}$ factor for the mildly ill-posed inverse problems with the Sobolev class, and in terms of an extra $\sqrt{\log\log\varepsilon^{-1}}$ factor for severely problems with analytic classes of functions. (A similar loss in the separation rates for a well-posed signal detection problem was first observed in [13].)

As we shall show below, the derived families of tests are of simple structure. In particular, for the mildly ill-posed inverse problems with the Sobolev class of functions, these are of the form

$$\psi_\varepsilon^{\mathrm{ad}} = \mathbb{1}_{\{\sup_k t_{\varepsilon,m_k} > H_k\}}, \qquad m_k = 2^k, \qquad H_k = \sqrt{C\log(k)}, \qquad k \ge L,$$

where $C > 2$, for an integer-valued family $L = L_\varepsilon$, $L_\varepsilon \to \infty$, and

$$t_{\varepsilon,m} = \frac{1}{\sqrt{2m}} \sum_{k=1}^{m} \big((y_k/\varepsilon)^2 - 1\big) \tag{5.1}$$

are centered and normalized version of χ^2-statistics that correspond to the first m observations.

For the severely ill-posed inverse problems with the Sobolev class of functions or the class of analytic functions, the derived families of tests are of the form

$$\psi_\varepsilon^{\mathrm{ad}} = \mathbb{1}_{\{\sup_k |y_k| > \varepsilon H_k\}}, \qquad H_k = \sqrt{2\log(k)}, \qquad k < L,$$
$$H_k = \sqrt{C\log(k)}, \qquad k \ge L,$$

where $C > 2$, for an integer-valued family $L = L_\varepsilon$, $L_\varepsilon \to \infty$.

Finally, for the severely ill-posed inverse problems with the generalized analytic class of functions, the derived tests are of the form

$$\psi_\varepsilon^{\mathrm{ad}} = \mathbb{1}_{\{\sup_k |y_k| > \varepsilon T_{\varepsilon,k}\}}, \qquad T_{\varepsilon,k} = \max\bigl(T_\varepsilon, \sqrt{2\bigl(\log(k) + \log\log(k)\bigr)}\bigr)$$

for a family $T_\varepsilon \to \infty$.

5.1.1. *Mildly ill-posed inverse problems with the Sobolev class of functions.* Consider first the "standard" case $q = 2$. Let $a_k = k^\alpha$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$. Set $\kappa = (\alpha, \beta)$, and let Σ be a compact subset of $\mathbb{R}_+^2$. We show that, under a weak assumption on the set Σ, $u_\varepsilon^{\mathrm{ad}} = \sqrt{\log\log(\varepsilon^{-1})}$. This corresponds to the adaptive separation rates

$$r_\varepsilon^{\mathrm{ad}}(\kappa) = \bigl(\varepsilon\sqrt[4]{\log\log(\varepsilon^{-1})}\bigr)^{4\alpha/(4\alpha+4\beta+1)}. \tag{5.2}$$

The rate optimal adaptive family of tests is of the following structure. Take a collection $m_k = 2^k$, $k \in \mathbb{N}$, $k \ge L = L_\varepsilon$, for an integer-valued family $L_\varepsilon \to \infty$, $L_\varepsilon = o(\log(\varepsilon^{-1}))$, and a family of test statistics t_{ε,m_k} of the form (5.1). Consider the thresholds and tests

$$H_k = \sqrt{C\log(k)}, \qquad \mathcal{Y}_\varepsilon = \{y : t_{\varepsilon,m_k} > H_k, \forall k \ge L_\varepsilon\}, \qquad \psi_\varepsilon = \mathbb{1}_{\mathcal{Y}_\varepsilon}, \tag{5.3}$$

where $C > 2$. Denote also

$$\phi(\kappa) = \frac{4}{4\alpha + 4\beta + 1}, \qquad \phi(\Sigma) = \bigl\{\phi(\kappa) : \kappa \in \Sigma\bigr\} \subset (0, \infty). \tag{5.4}$$

THEOREM 5.1. *Let $q = 2$, $a_k = k^\alpha$ and $\sigma_k = k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.*

(a) (*Lower bounds.*) *Let the set $\phi(\Sigma)$ given by (5.4) contain an interval $[a, b]$, $0 < a < b < \infty$. Then, there exists constant $d > 0$ such that if $\limsup_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\sqrt{\log\log(\varepsilon^{-1})} \le d$, then $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 1$.*

(b) (*Upper bounds.*) *For the family of tests ψ_ε given by (5.3), $\alpha(\psi_\varepsilon) = o(1)$, and there exists constant $D = D(\Sigma) > 0$ such that if $\liminf_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\sqrt{\log\log(\varepsilon^{-1})} > D$, then $\beta_\varepsilon(\psi_\varepsilon, \Theta_\varepsilon(\Sigma)) = o(1)$.*

(c) (*Adaptive separation rates.*) *The adaptive separation rates* $r_\varepsilon^{\mathrm{ad}}(\kappa)$, $\kappa \in \Sigma$, *are given by* (5.2).

The proof is given in the supplementary material [10], Section 11.11.

REMARK 5.1. In view of Remark 4.3, similar rate optimality results and the same adaptive separation rates hold for the sequences $a_k \sim k^\alpha$ and $\sigma \sim k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$, and the sequences $a_k \asymp k^\alpha$ and $\sigma \asymp k^\beta$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.

As in the case of rate and sharp asymptotics, the "sparse" case $q \in (0, 2)$ is not directly linked will Theorem 4.1. Rate-optimal adaptive tests in this case, however, can also be constructed, based on the family of tests considered in Section 7.4.1 of [11], Chapter 7. Their construction is omitted.

5.1.2. *Severely ill-posed inverse problems with the class of analytic functions.* Let $a_k = \exp(\alpha k)$ and $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$. Set $\kappa = (\alpha, \beta, q)$, and let Σ be a compact subset of $\mathbb{R}_+^2 \times (0, 2]$. We show that, under a weak assumption on the set Σ, $u_\varepsilon^{\mathrm{ad}} = \log\log(\varepsilon^{-1})$. This corresponds to the adaptive separation rates

$$r_\varepsilon^{\mathrm{ad}}(\kappa) = \left(\varepsilon\sqrt{\log\log(\varepsilon^{-1})}\right)^{\alpha/(\alpha+\beta)}. \tag{5.5}$$

The rate optimal adaptive family of tests is of the following structure. Take an integer-valued family $L = L_\varepsilon$, $L_\varepsilon \to \infty$, $L_\varepsilon = o(\log\log(\varepsilon^{-1}))$. Consider the families of thresholds and tests

$$H_k = \begin{cases} \sqrt{2\log(L)}, & k < L, \\ \sqrt{C\log(k)}, & k \ge L, \end{cases} \qquad \mathcal{Y}_\varepsilon = \{y : |y_k| > \varepsilon H_k\}, \qquad \psi_\varepsilon = \mathbb{1}_{\mathcal{Y}_\varepsilon}, \tag{5.6}$$

where $C > 2$. Denote also

$$\phi(\kappa) = \frac{1}{2(\alpha+\beta)}, \qquad \phi(\Sigma) = \{\phi(\kappa) : \kappa \in \Sigma\} \subset (0, \infty). \tag{5.7}$$

THEOREM 5.2. *Let* $a_k = \exp(\alpha k)$, $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.

(a) (*Lower bounds.*) *Let the set* $\phi(\Sigma)$ *given by* (5.7) *contains an interval* $[a, b], 0 < a < b < \infty$. *Then, there exists constant* $d > 0$ *such that if* $\limsup_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\log\log(\varepsilon^{-1}) \le d$, *then* $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 1$.

(b) (*Upper bounds.*) *For the family of tests* ψ_ε *given by* (5.6), $\alpha(\psi_\varepsilon) = o(1)$ *and there exists constant* $D = D(\Sigma) > 0$ *such that if* $\liminf_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\log\log(\varepsilon^{-1}) > D$, *then* $\beta_\varepsilon(\Theta_\varepsilon(\Sigma), \psi_\varepsilon) = o(1)$.

(c) (*Adaptive separation rates.*) *The adaptive separation rates* $r_\varepsilon^{\mathrm{ad}}(\kappa)$, $\kappa \in \Sigma$, *are given by* (5.5).

The proof is given in the supplementary material [10], Section 11.12.

REMARK 5.2. In view of Remark 4.5, similar rate optimality results and the adaptive separation rates (5.5) hold for the sequences $a_k \asymp \exp(\alpha k)$ and $\sigma \asymp \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.

5.1.3. *Severely ill-posed inverse problems with the Sobolev class of functions.* Let $a_k = k^\alpha$ and $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$. Set $\kappa = (\alpha, \beta, q)$, and let Σ be a compact subset of $\mathbb{R}_+^2 \times (0, 2]$. We show that, under a weak assumption on the set Σ, $u_\varepsilon^{\mathrm{ad}} = \log\log(\varepsilon^{-1})$. This corresponds to the adaptive separation rates

$$
\begin{aligned}
r_\varepsilon^{\mathrm{ad}}(\kappa) &= \left(\frac{2\beta}{2\log(\varepsilon^{-1}) - 2\alpha\log\log(\varepsilon^{-1}) - \log\log\log(\varepsilon^{-1})}\right)^\alpha \\
&\sim \left(\frac{\beta}{\log(\varepsilon^{-1})}\right)^\alpha.
\end{aligned}
\tag{5.8}
$$

The rate optimal adaptive family of tests is of the following structure. Take an integer-valued family $L = L_\varepsilon$, $L_\varepsilon \to \infty$, $L_\varepsilon = o(\log\log(\varepsilon^{-1}))$, and consider the families of thresholds and tests given by (5.6).

THEOREM 5.3. *Let $a_k = k^\alpha$ and $\sigma_k = \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.*

(a) (*Lower bounds.*) *Let the set Σ contains an interval of $(\alpha, \beta) : \beta \in [1/2b, 1/2a]$, $0 < a < b < \infty$, and a fixed $\alpha > 0$. Then there exists constant $d > 0$ such that if* $\limsup_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\log\log(\varepsilon^{-1}) \le d$, *then* $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 1$.

(b) (*Upper bounds.*) *For the family of tests ψ_ε given by* (5.6), $\alpha(\psi_\varepsilon) = o(1)$ *and there exists constant $D = D(\Sigma) > 0$ such that if* $\liminf_{\kappa\in\Sigma} u_\varepsilon(\kappa)/\log\log(\varepsilon^{-1}) > D$, *then* $\beta_\varepsilon(\psi_\varepsilon, \Theta_\varepsilon(\Sigma)) = o(1)$.

(c) (*Adaptive separation rates.*) *The adaptive separation rates $r_\varepsilon^{\mathrm{ad}}(\kappa)$, $\kappa \in \Sigma$, are given by* (5.8).

The proof is given in the supplementary material [10], Section 11.13.

REMARK 5.3. It is worth mentioning that a stronger result is possible in this case. In view of (11.21) in [10], relation (5.8) determines *sharp adaptive separation rates* $r_\varepsilon^{\mathrm{ad}}(\kappa)$, $\kappa \in \Sigma$, in the following sense:

(a) if $\liminf(r_\varepsilon(\kappa)/r_\varepsilon^{\mathrm{ad}}(\kappa)) > 1$, then $u_\varepsilon \to \infty$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 0$;

(b) if $\limsup(r_\varepsilon(\kappa)/r_\varepsilon^{\mathrm{ad}}(\kappa)) < 1$, then $u_\varepsilon \to 0$, that is, $\gamma_\varepsilon(r_\varepsilon) \to 1$, and the minimax testing is impossible.

REMARK 5.4. In view of Remark 4.8, similar rate optimality results and the adaptive separation rates (5.8) (as well as the sharp adaptive separation rates mention in Remark 5.3) hold for the sequences $a_k \sim k^\alpha$ and $\sigma \asymp \exp(\beta k)$, $k \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$.

5.1.4. *Extremely ill-posed inverse problems with the class of generalized analytic functions.* We consider the case $q = 2$ only. By the results of Section 4.6, in order to obtain distinguishability conditions, we can replace $u_\varepsilon(\kappa)$, $\kappa \in \Sigma$, by $u_\varepsilon^{\mathrm{lin}}(\kappa) = u_\varepsilon^{\mathrm{lin}}(\kappa, r_\varepsilon(\kappa))$, determined by (4.12), by $a_k = a_k(\kappa)$ and $\sigma_k = \sigma_k(\kappa)$, $k \in \mathbb{N}$. Set $u_\varepsilon^{\mathrm{lin}}(\Sigma) = \inf_{\kappa\in\Sigma} u_\varepsilon^{\mathrm{lin}}(\kappa)$. We are interested in how large $u_\varepsilon^{\mathrm{lin}}(\Sigma)$ should be in order to $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 0$.

Assume below the uniform version of (4.10): let $\sigma_k(\kappa)$ and $a_k(\kappa)$, $k \in \mathbb{N}$, be increasing sequences such that, for all $\kappa \in \Sigma$ and some constants $0 < b < B$,

$$b \le a_1(\kappa) \le B, \qquad b \le \sigma_1(\kappa) \le B \tag{5.9}$$

and, for some increasing sequence $\tau_k > 1$, $\tau_k \to \infty$ and some $c_0 > 1$ for all $\kappa \in \Sigma$ and $k \in \mathbb{N}$,

$$\sigma_{k+1}(\kappa)/\sigma_k(\kappa) \ge \tau_k, \qquad a_{k+1}(\kappa)/a_k(\kappa) \ge c_0. \tag{5.10}$$

Similar to $u_\varepsilon^{\mathrm{ad}}$, one can consider a family $u_{\varepsilon,\mathrm{ad}}^{\mathrm{lin}}$ which characterizes adaptive distinguishability. We show that, under some assumption on the set Σ, one has $u_{\varepsilon,\mathrm{ad}}^{\mathrm{lin}} = \log\log(\varepsilon^{-1})$.

For $\kappa \in \Sigma$ and for $A > 0$ large enough, let an integer $m = m(A, \kappa)$ be defined by the relations

$$a_{m-1}(\kappa)\sigma_{m-1}(\kappa) \le A < a_m(\kappa)\sigma_m(\kappa). \tag{5.11}$$

Under (5.9) and (5.10), one has $a_{m-1}\sigma_{m-1} \ge b^2 c_0^{m-2} \prod_{k=1}^{m-2} \tau_k$, which yields

$$\sup_{\kappa\in\Sigma} m(A, \kappa) = o\big(\log(A)\big) \qquad \text{as } A \to \infty. \tag{5.12}$$

Set $\mathcal{M}(A, \Sigma) = \{m(A, \kappa) \in \mathbb{N} : \kappa \in \Sigma\}$, $M(A, \Sigma) = \#(\mathcal{M}(A, \Sigma))$. Since $M(A, \Sigma) \le \max_{m\in\mathcal{M}(A,\Sigma)} m$, one has, by (5.12), $M(A, \Sigma) = o(\log(A))$ as $A \to \infty$. Let $m = m(A, \kappa)$ be defined by (5.11) and set $L(A, \Sigma) := \sup_{\kappa\in\Sigma} \log(m(A, \kappa))$. By (5.12) we have, as $A \to \infty$,

$$\limsup L(A, \Sigma)/\log\log(A) \le 1. \tag{5.13}$$

For the lower bounds we suppose one can find quantities $b > 0$, $C \ge 1$ such that

$$\begin{gathered} \liminf_{A\to\infty} \log\big(M(A, \Sigma)\big)/\log\log(A) = b, \\ \sup_{\kappa\in\Sigma} u_\varepsilon^{\mathrm{lin}}\big(\kappa, r_\varepsilon(\kappa)\big) \le C u_\varepsilon^{\mathrm{lin}}(\Sigma). \end{gathered} \tag{5.14}$$

[The first relation in (5.14) is fulfilled for the example mentioned in Remark 4.15, at least if the set $\Sigma = \{(\alpha, \tau, \beta, \gamma)\}$ contains an interior point.]

The rate optimal adaptive family of tests is of the following structure. Take a family $T_\varepsilon \to \infty$ such that $T_\varepsilon = o(\sqrt{\log\log(\varepsilon^{-1})})$, and take a family of sequences $T_{\varepsilon,k}$ of the form $T_{\varepsilon,k} = \max(T_\varepsilon, \sqrt{2(\log(k) + \log\log(k))})$.

Consider the families of events and of tests

$$\mathcal{Y}_\varepsilon = \big\{y : |y_k| > \varepsilon T_{\varepsilon,k}, \forall k \in \mathbb{N}\big\}, \qquad \psi_\varepsilon = \mathbb{1}_{\mathcal{Y}_\varepsilon}. \tag{5.15}$$

THEOREM 5.4. *Consider the GSM* (2.2) *and the hypothesis testing problem* (2.3) *and* (2.4) *for* $q = 2$. *Let* $\{a_k\}_{k\in\mathbb{N}}$ *and* $\{\sigma_k\}_{k\in\mathbb{N}}$ *be increasing sequences satisfying* (5.9) *and* (5.10).

(a) (*Lower bounds*.) *Assume* (5.14). *Then there exists a constant* $d > 0$ *such that if* $\limsup u_\varepsilon^{\mathrm{lin}}(\Sigma)/\log\log(\varepsilon^{-1}) \le d$, *then* $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma)) \to 1$.

(b) (*Upper bounds*.) *Assume* (5.13). *For the family of tests* ψ_ε *given by* (5.15), *there exists a constant* $D > 0$ *such that if* $u_\varepsilon^{\mathrm{lin}}(\Sigma) > D\log\log(\varepsilon^{-1})$, *then* $\gamma_\varepsilon(\Theta_\varepsilon(\Sigma), \psi_\varepsilon) = o(1)$.

(c) (*Adaptive separation rates*.) *The adaptive separation rates* $r_\varepsilon^{\mathrm{ad}}(\kappa), \kappa \in \Sigma$, *are determined by the relation* $u_{\varepsilon,\mathrm{ad}}^{\mathrm{lin}} \asymp u_\varepsilon^{\mathrm{lin}}(\kappa, r_\varepsilon^{\mathrm{ad}}(\kappa))$.

The proof is given in the supplementary material [10], Section 11.14.

Acknowledgments. We would like to thank the Editor, the Associate Editor and two anonymous referees for useful comments and suggestions.

SUPPLEMENTARY MATERIAL

Detailed proofs and other material (DOI: 10.1214/12-AOS1011SUPP; .pdf). In this supplement, we present relevant material and the detailed proofs of the previous sections.

REFERENCES

[1] CAVALIER, L. (2008). Nonparametric statistical inverse problems. *Inverse Problems* **24** 034004, 19. MR2421941

[2] CAVALIER, L., GOLUBEV, Y., LEPSKI, O. and TSYBAKOV, A. (2004). Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Theory Probab. Appl*. **48** 426–446.

[3] CAVALIER, L. and TSYBAKOV, A. (2002). Sharp adaptation for inverse problems with random noise. *Probab. Theory Related Fields* **123** 323–354. MR1918537

[4] ERMAKOV, M. S. (1990). Minimax detection of a signal in Gaussian white noise. *Theory Probab. Appl*. **35** 667–679.

[5] GOLUBEV, G. and KHASMINSKII, R. (2001). A statistical approach to the Cauchy problem for the Laplace equation. In *State of the Art in Probability and Statistics* (*Leiden*, 1999) (M. de Gunst, C. Klaassen and A. van der Vaart, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **36** 419–433. IMS, Beachwood, OH. MR1836573

[6] GOLUBEV, G. K. and KHAS'MINSKIĬ, R. Z. (1999). A statistical approach to some inverse problems for partial differential equations. *Problemy Peredachi Informatsii* **35** 51–66. MR1728907

[7] IBRAGIMOV, I. A. and KHAS'MINSKIĬ, R. Z. (1985). On nonparametric estimation of the value of a linear functional in Gaussian white noise. *Theory Probab. Appl*. **29** 18–32.

[8] INGSTER, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I, II, III. *Math. Methods Statist*. **2** 85–114, 171–189, 249–268.

[9] INGSTER, Y. I., SAPATINAS, T. and SUSLINA, I. A. (2011). Minimax nonparametric testing in a problem related to the radon transform. *Math. Methods Statist*. **20** 347–364. MR2886641

[10] INGSTER, Y. I., SAPATINAS, T. and SUSLINA, I. A. (2012). Supplement to "Minimax signal detection in ill-posed inverse problems." DOI:10.1214/12-AOS1011SUPP.

[11] INGSTER, Y. I. and SUSLINA, I. A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. Springer, New York. MR1991446

[12] LAURENT, B., LOUBES, J. M. and MARTEAU, C. (2012). Nonasymptotic minimax rates of testing in signal detection with heterogeneous variances. *Electron*. *J*. *Stat*. **6** 91–122.

[13] SPOKOINY, V. G. (1996). Adaptive hypothesis testing using wavelets. *Ann*. *Statist*. **24** 2477–2498. MR1425962

[14] WAHBA, G. (1990). *Spline Models for Observational Data*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. SIAM, Philadelphia, PA. MR1045442

Y. I. INGSTER
DEPARTMENT OF MATHEMATICS II
SAINT PETERSBURG ELECTROTECHNICAL UNIVERSITY
PROF. POPOV STR., 5
197376, ST. PETERSBURG
RUSSIA
E-MAIL: yurii_ingster@mail.ru

T. SAPATINAS
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF CYPRUS
P.O. BOX 20537
NICOSIA 1678
CYPRUS
E-MAIL: fanis@ucy.ac.cy

I. A. SUSLINA
DEPARTMENT OF MATHEMATICS
SAINT PETERSBURG UNIVERSITY OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
KRONVERKSKIY PR., 49
197101, ST. PETERSBURG
RUSSIA
E-MAIL: isuslina@mail.ru