ISSN 1066-5307, Mathematical Methods of Statistics, 2009, Vol. 18, No. 4, pp. 310–340. © Allerton Press, Inc., 2009.

Estimation and Detection of Functions from Weighted Tensor Product Spaces

Yu. I. Ingster[1*] **and N. Stepanova**[2**]

[1] *Dept. of Math. II, St.Petersburg State Electrotechnical Univ., Russia*

[2] *School of Math. and Statist., Carleton Univ., Canada*

Received April 5, 2009; in final form, November 4, 2009

Abstract—The problems of estimation and detection of an infinitely-variate signal f observed in the continuous white noise model are studied. It is assumed that f belongs to a certain weighted tensor product space. Several examples of such a space are considered. Special attention is given to the tensor product space of analytic functions with exponential weights. In connection with estimating and detecting unknown signal, the problems of rate and sharp optimality are investigated. In particular, it is shown that the use of a weighted tensor product space makes it possible to avoid the "curse of dimensionality" phenomenon.

Key words: nonparametric estimation, nonparametric goodness-of-fit testing, multivariate functions, white noise model, weighted tensor product space.

2000 Mathematics Subject Classification: primary 62G10; secondary 62G20.

DOI: 10.3103/S1066530709040024

1. INTRODUCTION

Suppose that an unknown signal f defined on $[0,1]^d$ is observed in the continuous white noise model

$$X_\varepsilon = f + \varepsilon N, \tag{1}$$

where $f \in L_2([0,1]^d) = L_2^d$, N is a d-dimensional white Gaussian noise, and $\varepsilon > 0$ is a small parameter (noise intensity). In this model, the "observation" is the function $X_\varepsilon \colon L_2^d \to \mathcal{G}$ taking its values in the set $\mathcal{G}$ of normal random variables such that if $\xi = X_\varepsilon(\phi)$, $\eta = X_\varepsilon(\psi)$, where $\phi, \psi \in L_2^d$, then $\mathbf{E}(\xi) = (f,\phi)$, $\mathbf{E}(\eta) = (f,\psi)$, and $\mathrm{Cov}(\xi,\eta) = \varepsilon^2(\phi,\psi)$. For any $f \in L_2^d$, the observation X_ε determines the Gaussian measure $\mathbf{P}_{\varepsilon,f}$ (see [4], [15] for references), the corresponding expectation is denoted $\mathbf{E}_{\varepsilon,f}$. In this paper we study the case $d = \infty$. Some results are also valid in case of growing dimension $d = d_\varepsilon \to \infty$.

Model (1) can be equivalently represented by the Gaussian sequence space model

$$X_{\boldsymbol\ell} = \theta_{\boldsymbol\ell} + \varepsilon\xi_{\boldsymbol\ell}, \qquad \xi_{\boldsymbol\ell} \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \quad \boldsymbol\ell \in \mathcal{L},$$

where $\mathcal{L}$ is a countable set, $\theta_{\boldsymbol\ell} = (f, \phi_{\boldsymbol\ell})$ are the Fourier coefficients of f with respect to an orthonormal basis $\{\phi_{\boldsymbol\ell}\}_{\boldsymbol\ell\in\mathcal{L}}$ in L_2^d, and the $X_{\boldsymbol\ell} = X_\varepsilon(\phi_{\boldsymbol\ell})$ are the empirical Fourier coefficients.

One problem of interest is to estimate an unknown signal f using quadratic loss. Another problem of interest is to detect f, that is, to test the hypothesis $H_0\colon f = 0$ versus a family of nonparametric alternatives of the form $H_{1\varepsilon}\colon \|f\|_2 \ge r_\varepsilon$, where $\|\cdot\|_2$ is the L_2-norm and $r_\varepsilon \to 0$ is a positive family. These two problems are closely related to each other, see relations (12)–(15) below. Within the framework of the minimax approach, the problems are known to have no good solutions unless some "regularity" constraints on f are imposed. A popular constraint has the form $f \in \mathcal{F}$, where the set

* E-mail: yurii_ingster@mail.ru
** E-mail: nstep@math.carleton.ca

$\mathcal{F} \subset L_2^d$ consists of "regular enough" functions. It is often the case that the regularity constraints are determined in terms of ellipsoids with coefficients $c_{\boldsymbol{\ell}} \geq 0$, $\boldsymbol{\ell} \in \mathcal{L}$:

$$\mathcal{F} = \Big\{ f(t) = \sum_{\boldsymbol{\ell} \in \mathcal{L}} \theta_{\boldsymbol{\ell}} \phi_{\boldsymbol{\ell}}(t) \colon \sum_{\boldsymbol{\ell} \in \mathcal{L}} c_{\boldsymbol{\ell}}^2 \theta_{\boldsymbol{\ell}}^2 \leq 1 \Big\}. \tag{2}$$

Dealing with estimation and detection of multivariate signals, we are often faced with the "curse of dimensionality" when the quality of minimax estimation and testing deteriorates as d gets large.

In particular, for Sobolev balls of σ-smooth functions defined on the cube $[0,1]^d$ the rates $R_\varepsilon(\mathcal{F})$ of minimax estimation (using quadratic loss) and the separation rates $r_\varepsilon^*(\mathcal{F})$ of minimax testing (see Section 2.3 for precise definitions) satisfy

$$R_\varepsilon(\mathcal{F}) \asymp \varepsilon^{2\sigma/(2\sigma+d)}, \qquad r_\varepsilon^*(\mathcal{F}) \asymp \varepsilon^{4\sigma/(4\sigma+d)}. \tag{3}$$

See [4], [13], [17] for estimation and [5] for testing.

For balls in Sobolev spaces of tensor product structure with order σ and domain $[0,1]^d$ (see [12] for definition) one has

$$R_\varepsilon(\mathcal{F}) \asymp \left(\varepsilon \log^{(d-1)/2}(\varepsilon^{-1})\right)^{2\sigma/(2\sigma+1)}, \qquad r_\varepsilon^*(\mathcal{F}) \asymp \left(\varepsilon \log^{(d-1)/4}(\varepsilon^{-1})\right)^{4\sigma/(4\sigma+1)}. \tag{4}$$

See [4], [12] for estimation and [9] for testing.

For classes of analytic functions defined on $[0,1]^d$ (this space has a natural tensor product structure, see [9]) one has

$$R_\varepsilon(\mathcal{F}) \asymp \varepsilon \log^{d/2}(\varepsilon^{-1}), \qquad r_\varepsilon^*(\mathcal{F}) \asymp \varepsilon \log^{d/4}(\varepsilon^{-1}). \tag{5}$$

See [5], [11] for estimation and [9] for testing.

Now, taking formally $d \asymp \log(\varepsilon^{-1})$ in (3) we get $R_\varepsilon(\mathcal{F}) \asymp r_\varepsilon^*(\mathcal{F}) \asymp 1$. Similarly, we may take $d \asymp \log(\varepsilon^{-1})/\log\log(\varepsilon^{-1})$ in (4) and (5) to have the families $R_\varepsilon(\mathcal{F})$ and $r_\varepsilon^*(\mathcal{F})$ bounded away from zero. That is, asymptotically no minimax consistent estimators exist, and the detection of signals in the minimax sense is impossible.

Certainly, these substitutions are not correct since relations (3)–(5) are not uniform in the dimension d. For the functional sets under consideration, the case $d = d_\varepsilon \to \infty$ as $\varepsilon \to 0$ was studied in [7] and [9]. It was shown that for the Sobolev classes and for the balls in tensor product spaces, the asymptotics of quadratic risks and separation rates are of essentially different types for the cases $d \ll \log(\varepsilon^{-1})$ and $d \gg \log(\varepsilon^{-1})$. In particular, if $\log d \asymp \log(\varepsilon^{-1})$, then minimax consistent estimation and testing are impossible.

In a specific real-life problem of estimating and detecting a d-variate signal f, the dimension d is fixed and the noise intensity ε is set at a prescribed level. Under the asymptotic approach, the problem is "embedded" into a family of problems with $\varepsilon \to 0$ and d being fixed or tending to infinity. When estimating and detecting a d-variate analytic function in case of fixed d, the dimensionality only affects the optimal rates of convergence and separation rates on the $\log(\varepsilon^{-1})$ scale. However, in case of growing dimension $d = d_\varepsilon$ satisfying $d \gg \log(\varepsilon^{-1})$ the situation is different: the curse of dimensionality reveals itself by dramatic changes in the optimal rates of convergence and separation rates (see, for example, Theorem 6.3 of [9]).

One way to avoid the curse of dimensionality is to assume that f belongs to a certain weighted tensor product space. The idea behind such a space is to reduce the "working dimension" of the problem. Informally, it is assumed that: (i) f can be approximated by a sum of a small number of functions of a small number of variables, and (ii) the variables are ordered according to their importance, see [10], [12], [16], [18]. The "weighted approach" to the problem of estimating and detecting multivariate functions was recently developed for classes of functions of finite smoothness by Ingster and Suslina [8]. In this paper we treat in a similar manner the case of analytic functions. We also consider several examples of weighted tensor product spaces related to classes of functions of finite smoothness.

The description of a general weighted tensor product space and some important examples of such a space are given in Section 2.1. In Section 2.2, the detailed construction of the tensor product space of analytic functions with exponential weights is presented. Section 2.3 contains the precise statement of the problems and relevant notation.

When studying the rate optimality problem, we are looking for the rate-optimal estimators of f (estimation problem), and the conditions under which the alternative is distiguishable from the null hypothesis (detection problem). The scheme of the study is presented in Sections 4.1—4.2 in general settings for an arbitrary tensor product space determined by two weight sequences. One sequence controls the smoothness properties of the estimated and detected signal, whereas the other characterizes the importance of variables. When applied to the examples of interest, the scheme produces logarithmic asymptotics of quadratic risk of asymptotically minimax estimators and gives conditions under which minimax consistent families of tests exist. The results thus obtained are then improved for tensor product spaces of analytic functions with exponential weights, see Sections 4.3 and 4.4. The main results of Section 4, Theorems 1 and 2, illustrate the advantage of using tensor product spaces in multivariate settings: for such spaces the curse of dimensionality effect can be lifted.

In connection with estimating a signal f, the sharp optimality problem consists of finding sharp rates of asymptotically minimax estimators, i.e., optimal rates including constants. With regard to detection problem, the sharp optimality problem involves finding asymptotics of the total probability error and constructing asymptotically minimax families of tests. In Section 5, using some refined reasoning, we obtain, in an implicit form, the sharp asymptotics for both problems at hand, see Theorems 3 and 4. Finding explicit forms for these asymptotics requires solving nonlinear equations, which can hardly be done analytically.

Appendices I and II serve as references. They contain some well-known results related to estimating and detecting signals in minimax settings for ellipsoids. The proposed estimators and tests described in Appendices I and II have a standard structure. The estimators are of projection type and are based on the "empirical Fourier coefficients" $X_\varepsilon(\phi_{\boldsymbol{\ell}})$ that correspond to "small" coefficients $c_{\boldsymbol{\ell}} \le T$ of the ellipsoid. The tests are based on the normalized χ^2-statistics that correspond to similar collections of $X_\varepsilon(\phi_{\boldsymbol{\ell}})$. The family $T = T_\varepsilon$ is determined by the "balance equations" (12) (or (69)) for the estimation problem and (13) (or (79)) for the detection problem. The key point is the study of asymptotic behavior of the "counting function" $N(t) = \mathrm{card}\{\boldsymbol{\ell} \in \mathcal{L}\colon c_{\boldsymbol{\ell}} \le t\}$, as $t \to \infty$, in the balance equations.

Auxiliary results are collected in Appendix III. The proofs of lemmas are presented in Appendix IV, while the verification of relations (48)–(51) is given in Appendix V.

2. STATEMENT OF THE PROBLEMS

2.1. Weighted Tensor Product Space. Examples

In this section we define a tensor product space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ of the following structure. Let $\boldsymbol{\lambda} = \{\lambda_k\}_{k\in\mathbf{Z}}$ be a sequence such that $\lambda_0 = 1$, $\lambda_k = \lambda_{-k} > 1$, λ_k increases in $k > 0$ and $\lambda_k \to \infty$. Let $\boldsymbol{\gamma} = \{\gamma_j\}_{j\in\mathbf{N}}$ be an increasing sequence such that $\gamma_j \ge 1$, $\gamma_j \to \infty$. Define $\mathbf{Z}_0^d = \{\boldsymbol{\ell} = (l_1, \ldots, l_d, 0, \ldots, 0, \ldots) \in \mathbf{Z}^\infty\}$ and put

$$\mathbf{Z}_0^\infty = \bigcup_{d=1}^{\infty} \mathbf{Z}_0^d.$$

Suppose that $\{\phi_{\boldsymbol{\ell}}(\mathbf{t})\}_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty}$, $\mathbf{t} \in [0,1]^\infty$, is a tensor product basis in L_2^∞, where $\{\phi_k(t)\}_{k\in\mathbf{Z}}$ is an orthonormal basis in $L_2[0,1]$. For a function $f(\mathbf{t}) = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}\phi_{\boldsymbol{\ell}}(\mathbf{t})$ we set

$$\|f\|^2_{\boldsymbol{\lambda},\boldsymbol{\gamma}} = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} c^2_{\boldsymbol{\ell}}\theta^2_{\boldsymbol{\ell}},$$

where

$$c_{\boldsymbol{\ell}} = \prod_{j\colon l_j\ne 0} \gamma_j \lambda_{l_j}, \tag{6}$$

and define the weighted tensor product space as follows:

$$\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}} = \{f \in L_2^\infty\colon \|f\|^2_{\boldsymbol{\lambda},\boldsymbol{\gamma}} < \infty\}.$$

Clearly

$$\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}} = \bigotimes_{j=1}^{\infty} L_{\boldsymbol{\lambda},\gamma_j},$$

where the space $L_{\boldsymbol{\lambda},\gamma_j}$ consists of the functions $f(t) = \sum_{k\in\mathbf{Z}} \theta_k\phi_k(t)$, $t \in [0,1]$, such that

$$\|f\|^2_{\boldsymbol{\lambda},\gamma_j} = \sum_{k\in\mathbf{Z}} b^2_{kj}\theta^2_k < \infty, \qquad b_{0j} = 1, \quad b_{kj} = \gamma_j\lambda_k, \quad k \neq 0. \tag{7}$$

The sequence $\boldsymbol{\lambda} = \{\lambda_k\}$ characterizes smoothness properties of univariate functions in the spaces $L_{\boldsymbol{\lambda},\gamma_j}$, while the sequence $\boldsymbol{\gamma} = \{\gamma_j\}$ characterizes the "importance" of variables. Consider several important examples of the space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$.

Example 1. Let $\lambda_k = e^{\lambda|k|}$, $\lambda > 0$, $k \in \mathbf{Z}$, and $\gamma_j = e^{\mu j}$, $\mu > 0$, $j \in \mathbf{N}$. This is the main example of our interest that corresponds to 1-periodic functions $f(t)$ that can be analytically continued to the complex strip $\{z \in \mathbf{C}\colon |\operatorname{Im} z| \leq \lambda/(2\pi)\}$. The weight sequence $\boldsymbol{\gamma} = \{\gamma_j\}$ is exponential. The space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ with such a choice of $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ corresponds to the space $\mathcal{A}_{\lambda,\mu}$ whose construction is presented in detail in the next section.

Example 2. Let $\lambda_k = e^{\lambda|k|^\alpha}$, $\lambda > 0$, $\alpha > 0$, $k \in \mathbf{Z}$, and $\gamma_j = e^{\mu j^\beta}$, $\mu > 0$, $\beta > 0$, $j \in \mathbf{N}$. This is a generalization of Example 1. The main results obtained for the weighted tensor product space $\mathcal{A}_{\lambda,\mu}$ can be carried over to the present case with arbitrary $\alpha > 0$ and $\beta > 1/2$.

Example 3. Let $\lambda_k = A|k|^\sigma$, $\sigma > 0$, $A > 1$, $k \in \mathbf{Z}$, and $\gamma_j = j^s$, $s > 0$, $j \in \mathbf{N}$. If $A = (2\pi)^\sigma$, then

$$\|f\|^2_{\boldsymbol{\lambda},\gamma_j} = (f,1)^2 + \gamma_j^2\|f^{(\sigma)}\|^2,$$

where $f^{(\sigma)}$ is the σ-derivative of the function f. This norm corresponds to the Sobolev norm of σ-smooth functions. The weight sequence $\boldsymbol{\gamma} = \{\gamma_j\}$ is polynomial. The space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ with such a choice of $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ is a Sloan–Woźniakowski (weighted Sobolev) space, see [10], [16] for references. The rate and sharp asymptotics in the problems of estimating and detecting signals from the Sloan–Woźniakovski space were obtained in [8].

Example 4. Let $\lambda_k = e^{\lambda|k|}$, $\lambda > 0$, $k \in \mathbf{Z}$, and $\gamma_j = j^s$, $s > 0$, $j \in \mathbf{N}$. This case corresponds to 1-periodic functions $f(t)$ that can be analytically continued to the complex strip $\{z \in \mathbf{C}\colon |\operatorname{Im} z| \leq \lambda/(2\pi)\}$. The weight sequence that controls the importance of variables is polynomial.

Example 5. Let $\lambda_k = A|k|^\sigma$, $\sigma > 0$, $A > 1$, $k \in \mathbf{Z}$, and $\gamma_j = e^{\mu j}$, $\mu > 0$, $j \in \mathbf{N}$. This case corresponds to σ-smooth functions. The weight sequence $\boldsymbol{\gamma} = \{\gamma_j\}$ is exponential.

2.2. Tensor Product Space of Analytic Functions with Exponential Weights

Now we present in detail the construction of the space $\mathcal{A}_{\lambda,\mu}$. We start with the class $\mathcal{F}_\lambda(M)$ of functions $f(\mathbf{t})$, $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbf{R}^d$ such that:

(a) f is 1-periodic in each of its arguments;

(b) f can be analytically continued from $\mathbf{R}^d$ to the strip $S_\lambda = \{\mathbf{z} \in \mathbf{C}^d\colon |\operatorname{Im}\mathbf{z}| < \lambda/(2\pi)\}$, and $|f(\mathbf{z})| \leq M$ for all $\mathbf{z} \in S_\lambda$ and some constant $M > 0$.

Put $\phi_{\boldsymbol{\ell}}(\mathbf{t}) = \prod_{k=1}^d \phi_{l_k}(t_k)$, $\boldsymbol{\ell} = (l_1, \ldots, l_d) \in \mathbf{Z}^d$, where $\{\phi_k(t)\}_{k\in\mathbf{Z}}$ is a standard Fourier basis in $L_2[0,1]$, and denote $\theta_{\boldsymbol{\ell}} = (f, \phi_{\boldsymbol{\ell}})$. Then $f \in \mathcal{F}_\lambda(M)$ can be decomposed into the Fourier series

$$f(\mathbf{t}) = \sum_{\boldsymbol{\ell}\in\mathbf{Z}^d} \theta_{\boldsymbol{\ell}}\phi_{\boldsymbol{\ell}}(\mathbf{t}),$$

with Fourier coefficients $\theta_{\boldsymbol{\ell}}$ decreasing at an exponential rate, cf. [11, Lemma 1]:

$$|\theta_{\boldsymbol{\ell}}| \le M e^{-\lambda \sum_{j=1}^{d} |l_j|}. \tag{8}$$

Using this fact, we now introduce the weighted tensor product space of analytic functions. Our definition is given in terms of the Fourier expansion.

Let $u = (j_1, \dots, j_m)$, $1 \le j_1 < \dots < j_m \le d$, and let $\mathbf{t}_u$ denote the $|u|$-dimensional vector containing those components of $\mathbf{t}$ whose indices belong to the set u. Assume that a function $f \in L_2^d$ can be written in the form

$$f(\mathbf{t}) = \sum_{u \subset \{1,\dots,d\}} f_u(\mathbf{t}_u), \qquad \int_0^1 f_u(\mathbf{t}_u)\, dt_j = 0 \quad \text{for all} \quad j \in u, \tag{9}$$

where the summation extends over all 2^d subsets of the set $\{1, \dots, d\}$, $f_u = \text{const}$ if $u = \emptyset$, and

$$f_u(\mathbf{t}_u) = \sum_{\boldsymbol{\ell} \in \mathbf{Z}^d:\, l_j \ne 0 \text{ iff } j \in u} \theta_{\boldsymbol{\ell}} \phi_{\boldsymbol{\ell}}(\mathbf{t}_u).$$

Each function f_u depends only on variables in $\mathbf{t}_u$ and describes the "interaction" between these variables. Decomposition (9) is known in the statistical literature as the *functional ANOVA decomposition*.

Now let μ be a positive parameter. For the function f in (9) we define its weighted norm $\|f\|_{\lambda,\mu}$ as follows:

$$\|f\|_{\lambda,\mu}^2 = \sum_{u \subset \{1,\dots,d\}} \prod_{j \in u} e^{2\mu j}\, \|f_u\|_\lambda^2, \tag{10}$$

where, in view of (8), we set

$$\|f_u\|_\lambda^2 = \sum_{\boldsymbol{\ell} \in \mathbf{Z}^d:\, l_j \ne 0 \text{ iff } j \in u} \theta_{\boldsymbol{\ell}}^2 a_{\boldsymbol{\ell}}^2, \qquad a_{\boldsymbol{\ell}} = a_{(l_1,\dots,l_d)} = e^{\lambda \sum_{j=1}^{d} |l_j|}.$$

Thus, from (10)

$$\|f\|_{\lambda,\mu}^2 = \sum_{u \subset \{1,\dots,d\}} \prod_{j \in u} e^{2\mu j} \sum_{\boldsymbol{\ell} \in \mathbf{Z}^d:\, l_j \ne 0 \text{ iff } j \in u} \theta_{\boldsymbol{\ell}}^2 a_{\boldsymbol{\ell}}^2 = \sum_{\boldsymbol{\ell} \in \mathbf{Z}^d} \theta_{\boldsymbol{\ell}}^2 \prod_{j \in \{1,\dots,d\}:\, l_j \ne 0} e^{2\mu j + 2\lambda |l_j|} = \sum_{\boldsymbol{\ell} \in \mathbf{Z}^d} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2,$$
$$c_{\boldsymbol{\ell}} = \prod_{j \in \{1,\dots,d\}:\, l_j \ne 0} e^{\mu j + \lambda |l_j|}.$$

The tensor product space of d-variate analytic functions with exponential weights, $\mathcal{A}_{\lambda,\mu}^d$, consists of functions $f \in L_2^d$ such that $\|f\|_{\lambda,\mu} < \infty$. This is a Hilbert space of a tensor product structure:

$$\mathcal{A}_{\lambda,\mu}^d = \bigotimes_{j=1}^{d} L_{\lambda,\mu}^j,$$

where the Hilbert space $L_{\lambda,\mu}^j$ consists of functions $f(t) = \sum_{k \in \mathbf{Z}} \theta_k \phi_k(t)$, $t \in [0, 1]$, with the norm $\|f\|_{j,\lambda,\mu}$ defined by

$$\|f\|_{j,\lambda,\mu}^2 = \sum_{k \in \mathbf{Z}} \theta_k^2 c_{j,k}^2, \qquad c_{j,0} = 1, \quad c_{j,k} = e^{\mu j + \lambda |k|}, \quad k \ne 0. \tag{11}$$

The definition of $\mathcal{A}_{\lambda,\mu}^d$ can be extended to the case of infinitely-variate functions. For $\boldsymbol{\ell} \in \mathbf{Z}_0^\infty$, define the Fourier basis of L_2^∞ by

$$\phi_{\boldsymbol{\ell}}(\mathbf{t}) = \prod_{i=1}^{\infty} \phi_{l_k}(t_k), \qquad \mathbf{t} = (t_1, t_2, \dots) \in \mathbf{R}^\infty,$$

and let $\theta_{\boldsymbol{\ell}} = (f, \phi_{\boldsymbol{\ell}})$ be the Fourier coefficients. We say that the function f belongs to the class $\mathcal{A}_{\lambda,\mu}$ if: (a) $f \in L_2^{\infty}$ and (b) $\|f\|_{\infty,\lambda,\mu} < \infty$, where

$$\|f\|^2_{\infty,\lambda,\mu} = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^{\infty}} \theta^2_{\boldsymbol{\ell}} c^2_{\boldsymbol{\ell}}, \qquad c_{\boldsymbol{\ell}} = \prod_{j\in\mathbf{N}:\, l_j\neq 0} e^{\mu j+\lambda|l_j|}.$$

The space defined in this way is the Hilbert space with an infinite tensor product structure:

$$\mathcal{A}_{\lambda,\mu} = \bigotimes_{j=1}^{\infty} L^j_{\lambda,\mu}.$$

2.3. Estimation and Detection for Ellipsoids

Now we can state the problems more precisely. For $r_\varepsilon > 0$, put

$$\mathcal{F} = \{f \in \mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \colon \|f\|_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \le 1\}, \qquad \mathcal{F}(r_\varepsilon) = \{f \in \mathcal{F} \colon \|f\|_2 \ge r_\varepsilon\},$$

and define the minimax (integrated) quadratic risk by

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

where the infimum is taken over all possible estimators $\tilde f_\varepsilon$ of f based on the observation X_ε. When dealing with the estimation problem, we wish to find *asymptotically minimax* estimator $\hat f_\varepsilon$ of f for which

$$\sup_{f\in\mathcal{F}} \mathbf{E}_{\varepsilon,f}\|f - \hat f_\varepsilon\|_2^2 \sim R^2_\varepsilon(\mathcal{F}), \qquad \varepsilon \to 0,$$

and establish asymptotics for the risk $R^2_\varepsilon(\mathcal{F})$. In the problem of detecting f, we test the hypotheses

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

For a test ψ, define the error probabilities

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

The maximum probability of type II error is then given by

$$\beta_\varepsilon(r_\varepsilon, \psi) = \sup_{f\in\mathcal{F}(r_\varepsilon)} \beta_\varepsilon(\psi, f).$$

The quantity

$$\gamma_\varepsilon(r_\varepsilon) = \inf_{\psi} \gamma_\varepsilon(r_\varepsilon, \psi),$$

where

$$\gamma_\varepsilon(r_\varepsilon, \psi) = \alpha_\varepsilon(\psi) + \beta_\varepsilon(r_\varepsilon, \psi),$$

and the infimum is taken over all tests ψ, is called the *minimax total error probability*. We say that a family of tests ψ^*_ε is *asymptotically minimax* if

$$\gamma_\varepsilon(r_\varepsilon, \psi^*_\varepsilon) = \gamma_\varepsilon(r_\varepsilon) + o(1), \qquad \varepsilon \to 0.$$

We are interested in finding asymptotics of $\gamma_\varepsilon(r_\varepsilon)$ and determining the structure of asymptotically minimax tests. We are also interested in finding asymptotics for the separation rate $r^*_\varepsilon(\mathcal{F})$. We say that a family $r^*_\varepsilon(\mathcal{F})$ is a *separation rate* in the problem of testing H_0 versus $H_{1\varepsilon}$ if $\gamma_\varepsilon(r_\varepsilon) \to 0$ as $r_\varepsilon/r^*_\varepsilon(\mathcal{F}) \to \infty$ and $\gamma_\varepsilon(r_\varepsilon) \to 1$ as $r_\varepsilon/r^*_\varepsilon(\mathcal{F}) \to 0$. In other words, for small ε, it is impossible to distinguish between the null hypothesis and the alternative if the ratio $r_\varepsilon/r^*_\varepsilon(\mathcal{F})$ is small, whereas the alternative is distinguishable from the null hypothesis when the ratio $r_\varepsilon/r^*_\varepsilon(\mathcal{F})$ is large.

The ball $\mathcal{F} = \{f \in \mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \colon \|f\|_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \le 1\}$ in the space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ corresponds to the ellipsoid

$$\Theta = \Big\{\theta = (\theta_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \colon \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2 \le 1\Big\}$$

in the space of Fourier coefficients. Therefore the set $\mathcal{F}(r_\varepsilon)$ that specifies the alternative hypothesis can be equivalently written in the form

$$\Theta_\varepsilon = \Big\{\theta = (\theta_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \colon \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2 \le 1 \quad \text{and} \quad \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 \ge r_\varepsilon^2\Big\}.$$

3. BASIC ELEMENTS OF THE STUDY

3.1. Counting Function

In this section we define the so-called *counting function* $N(t)$ that determines the "distribution" of the coefficients $c_{\boldsymbol{\ell}}$. This function is known to determine the rate asymptotics of quadratic risk in the estimation problem and of separation rate in the detection problem. In addition, it controls the sharp asymptotics of $R_\varepsilon^2(\mathcal{F})$ and $\gamma_\varepsilon(r_\varepsilon)$.

For $t > 0$, consider the set

$$\mathcal{N}(t) = \{\boldsymbol{\ell} \in \mathbf{Z}_0^\infty \colon c_{\boldsymbol{\ell}} \le t\}$$

and put

$$N(t) = \operatorname{card}\{\mathcal{N}(t)\}.$$

Note that in Examples 1–5

$$N(t) < \infty \text{ for any } t > 0, \qquad N(t) \to \infty \text{ as } t \to \infty.$$

The function $N(t)$ is called the *counting function*. It is known (see, for example, [9, Sect. 2]) that for the estimation problem:

$$R_\varepsilon(\mathcal{F}) \asymp T^{-1}, \qquad \text{where} \quad \varepsilon^2 T^2 N(T) \asymp 1, \tag{12}$$

and for the detection problem:

$$r_\varepsilon^*(\mathcal{F}) \asymp T^{-1}, \qquad \text{where} \quad \varepsilon^4 T^4 N(T) \asymp 1. \tag{13}$$

Moreover, if $N(t)$ is a *slowly varying* function, that is, for any $c > 0$,

$$\lim_{t\to\infty} N(ct)/N(t) = 1,$$

then the sharp asymptotics hold true:

$$R_\varepsilon^2(\mathcal{F}) \sim \varepsilon^2 N(T), \qquad \text{where} \quad \varepsilon^2 T^2 N(T) \asymp 1, \tag{14}$$

and

$$\gamma_\varepsilon(r_\varepsilon) = 2\Phi(-u_\varepsilon/2) + o(1), \qquad u_\varepsilon^2 \sim r_\varepsilon^4/2\varepsilon^4 N(T), \tag{15}$$

where

$$\varepsilon^4 T^4 N(T) \asymp 1.$$

The details on the derivation of relations (12)–(15) are given in Appendices I and II. Here the limits are taken as $\varepsilon \to 0$. The relation $A_\varepsilon \sim B_\varepsilon$ means $\lim A_\varepsilon/B_\varepsilon = 1$, and the relation $A_\varepsilon \asymp B_\varepsilon$ means that there exist constants $0 < c < C$ and a number $\varepsilon_0 > 0$ such that $c < A_\varepsilon/B_\varepsilon < C$ for $\varepsilon \in (0, \varepsilon_0)$.

In this paper we present, step-by-step, a general scheme for the study of asymptotics of $N(t)$, $t \to \infty$, in the problems at hand. The results of this study yield asymptotics of $R_\varepsilon(\mathcal{F})$, $r_\varepsilon^*(\mathcal{F})$, and $\gamma_\varepsilon(r_\varepsilon)$. For illustration, the scheme is applied to the space $\mathcal{A}_{\lambda,\mu}$ of analytic functions. Most of the results obtained for $\mathcal{A}_{\lambda,\mu}$ can be carried over to a general weighted tensor product space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ that includes $\mathcal{A}_{\lambda,\mu}$ as a particular case.

Technically, the problem of finding rate and sharp asymptotics is rather difficult. Often, the quantities that determine these asymptotics turn out to be solutions of nonlinear equations that cannot be solved analytically. For this reason, the main results of the paper are given in terms of logarithmic asymptotics.

3.2. Probability Measures

The rate and sharp asymptotics in the problems under study are determined by the function $N(t)$ (see (12)–(15)). In the next section we examine the asymptotic behavior of $N(t)$ as $t \to \infty$. A general method consists of defining a family of prior distributions $\mathbf{P}_h$ on the set of indices $\mathbf{Z}_0^\infty$ and investigating the behavior of the function $N(t) = \operatorname{card}\{\boldsymbol{\ell} \in \mathbf{Z}_0^\infty \colon c_{\boldsymbol{\ell}} \le t\}$ based on representation (20) using probabilistic and analytic tools.

First, we define a family of probability measures $\mathbf{P}_h$ depending on parameter h on the set $\mathbf{Z}^\infty$:

$$\mathbf{P}_h(\boldsymbol{\ell}) = \prod_{j=1}^{\infty} \mathbf{P}_{h,j}(l_j), \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty.$$

It is convenient to rewrite the coefficients $c_{\boldsymbol{\ell}}$ in the form $c_{\boldsymbol{\ell}} = \exp(S(\boldsymbol{\ell}))$, where, cf. (7),

$$S(\boldsymbol{\ell}) = \sum_{j=1}^{\infty} Y_j(l_j), \quad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty; \qquad Y_j(k) = \begin{cases} 0, & k = 0, \\ \log(\lambda_k) + \log(\gamma_j), & k \ne 0, \end{cases}, \quad k \in \mathbf{Z}, \tag{16}$$

and define the measures $\mathbf{P}_{h,j}$ by

$$\mathbf{P}_{h,j}(k) = \exp\big(- hY_j(k) - Z_j(h)\big), \tag{17}$$

where

$$Z_j(h) = \log\left(\sum_{k \in \mathbf{Z}} \exp(-hY_j(k)) \right) = \log\big(1 + G(h)\gamma_j^{-h}\big), \qquad G(h) = 2\sum_{k=1}^{\infty} \lambda_k^{-h}.$$

This leads to the formula

$$\mathbf{P}_h(\boldsymbol{\ell}) = \prod_{j=1}^{\infty} \mathbf{P}_{h,j}(l_j) = \exp\big(- hS(\boldsymbol{\ell}) - Z(h)\big), \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty, \tag{18}$$

where $Z(h) = \sum_{j=1}^{\infty} Z_j(h)$.

The measures $\mathbf{P}_h$ are well defined for $h > h^* \ge 0$, where

$$h^* = \inf\{h > 0 \colon Z(h) < \infty\} = \max(h^*_{\boldsymbol{\lambda}},\ h^*_{\boldsymbol{\gamma}})$$

with

$$h^*_{\boldsymbol{\lambda}} = \inf\big\{h \colon G(h) < \infty\big\}, \qquad h^*_{\boldsymbol{\gamma}} = \inf\left\{h \colon \sum_{j=1}^{\infty} \gamma_j^{-h} < \infty\right\}.$$

Let us specify the value of h^* in Examples 1–5.

Example 1. The series $\sum_{k=1}^{\infty} \lambda_k^{-h} = \sum_{k=1}^{\infty} e^{-h\lambda k}$ and $\sum_{j=1}^{\infty} \gamma_j^{-h} = \sum_{j=1}^{\infty} e^{-h\mu j}$ are convergent for all $h > 0$, so that $h^* = 0$.

Example 2. The series $\sum_{k=1}^{\infty} \lambda_k^{-h} = \sum_{k=1}^{\infty} e^{-\lambda h k^\alpha}$ and $\sum_{k=j}^{\infty} \gamma_j^{-h} = \sum_{j=1}^{\infty} e^{-\mu h j^\beta}$ are convergent for all $h > 0$. Indeed, setting $m = (\lambda h)^{-1/\alpha}$

$$\begin{aligned} \sum_{k=1}^{\infty} e^{-\lambda h k^\alpha} &= 2m \sum_{k=1}^{\infty} e^{-(k/m)^\alpha} m^{-1} = 2m \int_0^\infty e^{-x^\alpha}\, dx + O(1) \\ &= \frac{2m}{\alpha} \int_0^\infty e^{-y} y^{1/\alpha - 1}\, dy + O(1) = \frac{2m\Gamma(1/\alpha)}{\alpha} + O(1) \\ &= Bh^{-1/\alpha} + O(1), \qquad B = \frac{2\Gamma(1/\alpha)}{\alpha\lambda^{1/\alpha}}. \end{aligned} \tag{19}$$

The same arguments apply to the second series. Therefore, as in Example 1, $h^* = 0$.

Example 3. Here $\sum_{k=1}^{\infty} \lambda_k^{-h} = \sum_{k=1}^{\infty} k^{-\sigma h}$ and $\sum_{k=j}^{\infty} \gamma_j^{-h} = \sum_{j=1}^{\infty} j^{-sh}$. The series are convergent when $\sigma h > 1$ and $sh > 1$, respectively. Therefore $h^* = \max(1/\sigma, 1/s)$.

Example 4. In this case $\sum_{k=1}^{\infty} \lambda_k^{-h} = \sum_{k=1}^{\infty} e^{-\lambda hk}$ and $\sum_{k=j}^{\infty} \gamma_j^{-h} = \sum_{j=1}^{\infty} j^{-sh}$. Applying the above arguments, we get $h^* = \max(0, 1/s) = 1/s$.

Example 5. Similarly, $\sum_{k=1}^{\infty} \lambda_k^{-h} = A\sum_{k=1}^{\infty} k^{-\sigma h}$ and $\sum_{j=1}^{\infty} \gamma_j^{-h} = \sum_{j=1}^{\infty} e^{-h\mu j}$. Therefore $h^* = \max(1/\sigma, 0) = 1/\sigma$.

4. RATE OPTIMALITY

4.1. Rough Log-Asymptotics. Lifting the Curse of Dimensionality

Consider the counting function $N(t) = \mathrm{card}\{\boldsymbol{\ell} \in \mathbf{Z}_0^{\infty} \colon c_{\boldsymbol{\ell}} \le t\}$. Setting

$$H = \log t$$

leads to the representation, for any $h > h^*$,

$$\begin{aligned} N(t) &= \mathrm{card}\{\boldsymbol{\ell} \in \mathbf{Z}_0^{\infty} \colon c_{\boldsymbol{\ell}} \le t\} = \mathrm{card}\{\boldsymbol{\ell} \in \mathbf{Z}_0^{\infty} \colon S(\boldsymbol{\ell}) \le H\} \\ &= e^{Z(h)+hH} \sum_{\boldsymbol{\ell} \in \mathbf{Z}_0^{\infty} \colon S(\boldsymbol{\ell}) \le H} e^{h(S(\boldsymbol{\ell})-H)} \mathbf{P}_h(\boldsymbol{\ell}) = e^{Z(h)+hH} I_h, \end{aligned} \tag{20}$$

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

$$\log N(t) \le Z(h) + hH \qquad \text{for all} \quad h > h^*.$$

This immediately leads to the relation, as $t \to \infty$,

$$\log N(t) \le H(h^* + o(1)) = \log t (h^* + o(1)). \tag{21}$$

At the same time, in view of (12) and (13),

$$\log R_\varepsilon(\mathcal{F}) = -\log T + O(1), \quad \text{where} \quad \log N(T) + 2\log T = 2\log(\varepsilon^{-1}) + O(1), \tag{22}$$

$$\log r_\varepsilon^*(\mathcal{F}) = -\log T + O(1), \quad \text{where} \quad \log N(T) + 4\log T = 4\log(\varepsilon^{-1}) + O(1). \tag{23}$$

Combining these relations with (21) we arrive at the upper bounds

$$\log R_\varepsilon(\mathcal{F}) \le \frac{(2 + o(1))\log \varepsilon}{2 + h^*}, \qquad \log r_\varepsilon^*(\mathcal{F}) \le \frac{(4 + o(1))\log \varepsilon}{4 + h^*}. \tag{24}$$

In particular, for the weighted tensor product spaces in Examples 1 and 2, for which $h^* = 0$, since $\log N(T) \ge 0$ for $T \ge 1$,

$$\log R_\varepsilon(\mathcal{F}) \sim \log r_\varepsilon^*(\mathcal{F}) \sim \log \varepsilon, \qquad \varepsilon \to 0. \tag{25}$$

In order to obtain rough asymptotics in Examples 3–5, we need to establish appropriate lower bounds on $\log N(T)$. First, consider Example 3, where

$$c_{\boldsymbol{\ell}} = \prod_{j \in \mathbf{N} \colon l_j \ne 0} A j^s |l_j|^\sigma, \quad \boldsymbol{\ell} \in \mathbf{Z}_0^{\infty}, \qquad \text{and} \qquad h^* = \min(1/\sigma, 1/s).$$

It can be shown that

$$\log N(T) \ge h^* \log T + O(1). \tag{26}$$

From (26) using (22) and (23) we get the lower bounds

$$\log R_\varepsilon(\mathcal{F}) \ge \frac{2\log \varepsilon}{2 + h^*} + O(1), \qquad \log r_\varepsilon^*(\mathcal{F}) \ge \frac{4\log \varepsilon}{4 + h^*} + O(1),$$

which together with (24) yields

$$\log R_\varepsilon(\mathcal{F}) \sim \frac{2\log \varepsilon}{2 + h^*}, \qquad \log r_\varepsilon^*(\mathcal{F}) \sim \frac{4\log \varepsilon}{4 + h^*}, \quad \varepsilon \to 0. \tag{27}$$

Inequality (26) is based on the following simple arguments. Let $\sigma \le s$. Then the set $\mathcal{N}(T)$ contains all the points $\boldsymbol{\ell}^{(k)} = (k, 0, \dots, 0, \dots)$, $k \in \mathbf{Z}$, $k \neq 0$, $A|k|^\sigma \le T$, and the number of such points is $(T/A)^{1/\sigma}(1 + o(1))$. Hence $\log N(T) \ge (1/\sigma)\log T + O(1)$. Let $s < \sigma$. Then $\mathcal{N}(T)$ includes all the points $\boldsymbol{\ell}^{(j)} = (0, \dots, 0, 1, 0, \dots,)$, with 1 on the jth place, $j \in \mathbf{N}$, $Aj^s \le T$, and the number of such points is $(T/A)^{1/s}(1 + o(1))$. Therefore $\log N(T) \ge (1/s)\log T + O(1)$.

Asymptotics (27) can be also obtained as consequences of Theorems 1 and 2 in [8]. In case of Example 3, these theorems provide the sharp asymptotics for the problems at hand.

The same logarithmic asymptotics (27) hold true in Examples 4 and 5, where

$$c_{\boldsymbol{\ell}} = \prod_{j \in \mathbf{N}:\, l_j \neq 0} j^s e^{\lambda |l_j|}, \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty, \quad h^* = 1/s,$$

$$c_{\boldsymbol{\ell}} = \prod_{j \in \mathbf{N}:\, l_j \neq 0} A|l_j|^\sigma e^{\mu j}, \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty, \quad h^* = 1/\sigma,$$

respectively. In these examples, the above arguments lead to the lower bounds

$$\log N(T) \ge (1/s)\log T + O(1) = h^* \log T + O(1) \quad \text{in Example 4},$$
$$\log N(T) \ge (1/\sigma)\log T + O(1) = h^* \log T + O(1) \quad \text{in Example 5}.$$

Rough asymptotics as in (25) are similar to what we have for univariate analytic functions, whereas (27) are similar to the case of σ^*-smooth univariate functions, where $\sigma^* = 1/h^*$. Thus, the use of weighted tensor product space in estimating and detecting multivariate signals makes it possible to avoid the curse of dimensionality phenomenon.

4.2. Refined Log-Asymptotics: General Case

In order to improve (25), let us rewrite $I_h = \mathbf{E}_h\big(e^{h(S-H)}\mathbb{I}_{\{S \le H\}}\big)$ in the form

$$I_h = \mathbf{E}_h\big(e^{hs_h\tau_h}\mathbb{I}_{\{\tau_h \le 0\}}\big),$$

where the random variables τ_h, $h > 0$, are given by

$$\tau_h = (S - H)/s_h, \qquad s_h > 0,$$

and the parameter $h = h(H)$ is chosen to satisfy the normalization conditions

$$\mathbf{E}_h S = H, \qquad \mathrm{Var}_h\, S = s_h^2.$$

These conditions can be equivalently written in the form

$$Z'(h) = -H, \qquad Z''(h) = s_h^2. \tag{28}$$

Indeed, it can be seen, cf. relations (5.15) in [9], that

$$\mathbf{E}_{h,j}(Y_j) = -Z_j'(h), \qquad \mathrm{Var}_{h,j}(Y_j) = Z_j''(h), \qquad s_{h,j}^4(Y_j) = Z_j^{(4)}(h), \tag{29}$$

where $s^4(Y)$ is the 4th cumulant of Y, and, as shown in the proof of Lemma 1, the kth derivative of $Z(h)$, $k = 1, 2, 3, 4$, satisfies

$$Z^{(k)}(h) = \sum_{j=1}^\infty Z_j^{(k)}(h). \tag{30}$$

With this choice of the parameter $h > 0$, the random variables τ_h are suitably standardized:

$$\mathbf{E}_h\tau_h = 0, \qquad \mathrm{Var}_h\,\tau_h = 1.$$

Denote by $L_h^{(k)}$ the Lyapunov ratio of order $k > 2$:

$$L_h^{(k)} = \frac{\sum_{j=1}^\infty \mathbf{E}_h|Y_j - \mathbf{E}_h Y_j|^k}{\big(\sum_{j=1}^\infty \mathbf{E}_h|Y_j - \mathbf{E}_h Y_j|^2\big)^{k/2}}. \tag{31}$$

Recall (see [14, eq. (1.13)]) that the cumulant $s^k(Y)$ of arbitrary order k is expressed in terms of the moments $\alpha_1 = \mathbf{E}(Y), \dots, \alpha_k = \mathbf{E}(Y^k)$ by the formula

$$s^k(Y) = k! \sum (-1)^{r-1}(r-1)! \prod_{l=1}^{k} \frac{1}{m_l!}\left(\frac{\alpha_l}{l!}\right)^{m_l},$$

where the summation is extended over all non-negative integer solutions of the equation $m_1 + 2m_2 + \ldots + km_k = k$, and $r = m_1 + \ldots + m_k$. This formula implies

$$\mathbf{E}(Y - \mathbf{E}(Y))^4 = s^4(Y) + 3\big(\operatorname{Var}(Y)\big)^2. \tag{32}$$

Taking into account (29), (30), and (32), we can express the Lyapunov ratio of order four as follows:

$$L_h^{(4)} = \frac{Z^{(4)}(h) + 3\sum_{j=1}^{\infty}(Z_j''(h))^2}{(Z''(h))^2}.$$

By the Lyapunov theorem, the relation

$$L_h^{(4)} \to 0, \qquad h \to 0, \tag{33}$$

implies

$$\tau_h \to \xi \sim \mathcal{N}(0,1), \quad h \to 0. \tag{34}$$

This, in its turn, gives

$$\log(I_h) \geq -\delta_h h s_h + \log \delta_h, \quad \delta_h \to 0, \qquad \text{i.e.} \quad \log(I_h) = o(hs_h),$$

and hence by (20)

$$\log N(t) = Z(h) + hH + o(hs_h), \qquad H = \log(t) \to \infty, \tag{35}$$

with h and s_h satisfying (28).

Verification of (33) requires the study of the function $Z(h)$, as $h \to 0$, and its derivatives up to the fourth order. The behavior of $Z(h)$ and its derivatives depends heavily on the weight sequences $\boldsymbol{\lambda} = \{\lambda_k\}$ and $\boldsymbol{\gamma} = \{\gamma_j\}$ that determine $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$. Below we present the corresponding results obtained for the spaces in Examples 1 and 2. The computational technique developed in the next section for the space $\mathcal{A}_{\lambda,\mu}$ (Example 1) will be extended to Example 2 in Section 4.4.

4.3. Refined Log-Asymptotics: Example 1

Consider the space $\mathcal{A}_{\lambda,\mu}$ for which

$$Y_j(k) = \begin{cases} 0, & k = 0, \\ \lambda|k| + \mu j, & k \neq 0, \end{cases} \quad k \in \mathbf{Z}; \qquad S(\boldsymbol{\ell}) = \sum_{j=1}^{\infty} Y_j(l_j), \quad \boldsymbol{\ell} \in \mathbf{Z}_0^{\infty}, \tag{36}$$

and

$$\mathbf{P}_h(\boldsymbol{\ell}) = \prod_{j=1}^{\infty} \mathbf{P}_{h,j}(l_j) = \exp(-hS(\boldsymbol{\ell}) - Z(h)), \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^{\infty}, \quad h > 0,$$

where

$$Z(h) = \sum_{j=1}^{\infty} Z_j(h) = \sum_{j=1}^{\infty} \log\left(1 + G(h)e^{-h\mu j}\right), \qquad G(h) = 2\sum_{k=1}^{\infty} e^{-h\lambda k} = \frac{2}{e^{\lambda h} - 1}.$$

Lemma 1. *Let* $F(h) = h^{-1}\int_0^\infty \log\big(1+G(h)e^{-\mu x}\big)\,dx$, $h>0$. *Then*

$$Z(h) = F(h) + O\big(\log(h^{-1})\big) = \frac{\log^2(h^{-1})}{2\mu h} + O\bigg(\frac{\log(h^{-1})}{h}\bigg), \qquad h \to 0, \tag{37}$$

and for $k = 1,2,3,4$,

$$\begin{aligned} Z^{(k)}(h) &= F^{(k)}(h) + O\bigg(\frac{\log^k(h^{-1})}{h^k}\bigg) \\ &= \frac{(-1)^k k!\log^2(h^{-1})}{2\mu h^{k+1}} + O\bigg(\frac{\log(h^{-1})}{h^{k+1}}\bigg), \qquad h \to 0. \end{aligned}$$

The proof of Lemma 1 is given in Appendix IV.

Remark 1. It can be seen from the proof of Lemma 1 that the major contribution to the asymptotic representation of $Z^{(k)}(h)$, as $h \to 0$, is made by the integral

$$(-1)^k h^{-k}\int_0^J \log\big(1+G(h)e^{-\mu x}\big)\,dx, \qquad J = \mu^{-1}\log G(h).$$

This fact allows us to carry over the main results of the paper from the infinite-dimensional case under consideration to the d-dimensional case of growing dimension d (see Remarks 2 and 3 for details).

Lemma 1 implies that

$$Z'(h) = -\frac{\log^2(h^{-1})}{2\mu h^2} + O\bigg(\frac{\log(h^{-1})}{h^2}\bigg), \tag{38}$$

$$Z''(h) = \frac{\log^2(h^{-1})}{\mu h^3} + O\bigg(\frac{\log(h^{-1})}{h^3}\bigg), \tag{39}$$

$$Z'''(h) = -\frac{3\log^2(h^{-1})}{\mu h^4} + O\bigg(\frac{\log(h^{-1})}{h^4}\bigg), \tag{40}$$

$$Z^{(4)}(h) = \frac{12\log^2(h^{-1})}{\mu h^5} + O\bigg(\frac{\log(h^{-1})}{h^5}\bigg). \tag{41}$$

Using (37)–(41) we can prove the following result (see Appendix IV).

Lemma 2. *As* $h \to 0$, *the Lyapunov ratio* $L_h^{(4)}$ *satisfies*

$$L_h^{(4)} \to 0.$$

Let us return to the log-asymptotics of $N(t)$ given by (35):

$$\log N(t) = Z(h) + hH + o(hs_h), \qquad H = \log t \to \infty.$$

It follows from (28) and (38) that

$$H = -Z'(h) = \frac{\log^2(h^{-1})}{2\mu h^2} + O\bigg(\frac{\log(h^{-1})}{h^2}\bigg), \qquad h \to 0. \tag{42}$$

Solving this equation for h yields the approximate relation

$$h^2 \sim \frac{\log^2(h^{-1})}{2\mu H} \sim \frac{\log^2 H}{8\mu H} \qquad \text{or} \qquad h \sim \frac{\log H}{2\sqrt{2\mu H}}.$$

Next, using (28)

$$hs_h = h\big(Z''(h)\big)^{1/2} \asymp h^{-1/2}\log(h^{-1}) \asymp H^{1/4}\log^{1/2} H. \tag{43}$$

Therefore taking into account (37)

$$\begin{aligned}\log N(t) &= \frac{\log^2(h^{-1})}{\mu h} + O\Big(\frac{\log(h^{-1})}{h}\Big) + o\Big(\frac{\log(h^{-1})}{\sqrt{h}}\Big)\\ &= 2hH\big(1+O(1/\log(h^{-1}))\big) = \frac{\sqrt{H}\log H}{\sqrt{2\mu}}\big(1+O(1/\log H)\big)\\ &= \frac{\sqrt{H}\log H}{\sqrt{2\mu}} + O\left(\sqrt{H}\right), \qquad H=\log t\to\infty. \end{aligned} \tag{44}$$

Now we can state and prove the main result of this section, Theorem 1, which provides the refined log-asymptotics of quadratic risk and separation rate. Compared to the problem of estimating and detecting a d-variate signal $f \in \mathcal{A}^d_{\lambda,\mu}$ observed in the white Gaussian noise (with d being fixed), in which case, cf. [11, Th. 2], $\log R_\varepsilon(\mathcal{F}) = \log\varepsilon + (d/2)\log\log(\varepsilon^{-1}) + O(1)$ and $\log r^*_\varepsilon(\mathcal{F}) = \log\varepsilon + (d/4)\log\log(\varepsilon^{-1}) + O(1)$, the effect of dimensionality in Theorem 1 is completely lifted.

Theorem 1. *Consider the weighted tensor product space $\mathcal{A}_{\lambda,\mu}$. Then, as $\varepsilon\to 0$,*

$$\log R_\varepsilon(\mathcal{F}) = \log\varepsilon + \frac{\sqrt{\log(\varepsilon^{-1})}\log\log(\varepsilon^{-1})}{2\sqrt{2\mu}} + O\left(\sqrt{\log(\varepsilon^{-1})}\right),$$

$$\log r^*_\varepsilon(\mathcal{F}) = \log\varepsilon + \frac{\sqrt{\log(\varepsilon^{-1})}\log\log(\varepsilon^{-1})}{4\sqrt{2\mu}} + O\left(\sqrt{\log(\varepsilon^{-1})}\right).$$

Proof. It follows from (12) that

$$\log R_\varepsilon(\mathcal{F}) \sim \frac{\log N(T)}{2} - \log(\varepsilon^{-1}), \tag{45}$$

where $\log N(T) + 2H = 2\log(\varepsilon^{-1}) + O(1)$. Similarly, (13) implies

$$\log r^*_\varepsilon(\mathcal{F}) \sim \frac{\log N(T)}{4} - \log(\varepsilon^{-1}), \tag{46}$$

where $\log N(T) + 4H = 4\log(\varepsilon^{-1}) + O(1)$. In view of (42) and (44) the above restrictions on $N(t)$ can be written in the form

$$\frac{\sqrt{H}\log H}{\sqrt{2\mu}} + 2H + O\left(\sqrt{H}\right) = 2\log(\varepsilon^{-1}),$$

$$\frac{\sqrt{H}\log H}{\sqrt{2\mu}} + 4H + O\left(\sqrt{H}\right) = 4\log(\varepsilon^{-1}).$$

This yields

$$H = \log(\varepsilon^{-1}) - \frac{\sqrt{\log(\varepsilon^{-1})}\log\log(\varepsilon^{-1})}{2\sqrt{2\mu}} + O\left(\sqrt{\log(\varepsilon^{-1})}\right),$$

and hence

$$\log N(T) = \frac{\sqrt{\log(\varepsilon^{-1})}\log\log(\varepsilon^{-1})}{\sqrt{2\mu}} + O\left(\sqrt{\log(\varepsilon^{-1})}\right).$$

Substituting this relation into (45) and (46) completes the proof. □

Remark 2. In view of Remark 1, the logarithmic asymptotics in Theorem 1 continue to hold for d-variable function f when the dimension $d = d_\varepsilon$ grows to infinity at least as fast as $\log G(h) = O(\log(h^{-1}))$. Since $h \sim \log H/(2\sqrt{2\mu H})$, where $H \sim \log(\varepsilon^{-1})$, it follows that

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

Therefore, Theorem 1 remains valid not only when $d=\infty$, but also when $d = d_\varepsilon$ is at least as large as $O(\log\log(\varepsilon^{-1}))$.

4.4. Refined Log-Asymptotics: Example 2

The results of the previous section can be easily extended to the space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ of analytic functions for which $\lambda_k = e^{\lambda|k|^\alpha}$, $\lambda > 0$, $\alpha > 0$, $k \in \mathbf{Z}$, and $\gamma_j = e^{\mu j^\beta}$, $\mu > 0$, $\beta > 0$, $j \in \mathbf{N}$ (Example 2). In this case, the method described in Section 4.2, and applied in Section 4.3 to the space $\mathcal{A}_{\lambda,\mu}$, continues to work provided $\beta > 1/2$.

According to (16)–(18), the family of probability measures $\mathbf{P}_h$, $h > 0$, is now defined as follows:

$$\mathbf{P}_h(\boldsymbol{\ell}) = \exp\big(-hS(\boldsymbol{\ell}) - Z(h)\big), \qquad \boldsymbol{\ell} \in \mathbf{Z}_0^\infty,$$

where

$$Z(h) = \sum_{j=1}^\infty Z_j(h) = \sum_{j=1}^\infty \log(1 + G(h)e^{-\mu h j^\beta}), \qquad G(h) = 2\sum_{k=1}^\infty e^{-h\lambda k^\alpha} \sim \frac{2\Gamma(1/\alpha)}{\alpha\lambda^{1/\alpha}h^{1/\alpha}}.$$

Introduce a new parameter

$$\delta = \delta(h) = (\mu h)^{1/\beta}.$$

Then, in terms of δ, $Z(h)$ can be conveniently written in the form

$$Z(h) = \tilde{Z}(\delta) = \sum_{j=1}^\infty \log(1 + \tilde{G}(\delta)e^{-(\delta j)^\beta}),$$

where

$$\tilde{G}(\delta) = G(h) = \frac{2\Gamma(1/\alpha)\mu^{1/\alpha}}{\alpha\lambda^{1/\alpha}\delta^{\beta/\alpha}} + O(1).$$

That is,

$$\tilde{Z}(\delta) = \sum_{j=1}^\infty \tilde{f}(j\delta, \delta), \qquad \tilde{f}(x,\delta) = \log(1 + \tilde{G}(\delta)e^{-x^\beta}).$$

By analogy with Example 1, cf. (85),

$$\tilde{Z}(\delta) = \tilde{F}(\delta) + O(\tilde{f}(0,\delta)),$$

where $\tilde{f}(0,\delta) = O(\log(\delta^{-1}))$ and (see relation (82) in Appendix III)

$$\tilde{F}(\delta) = \frac{1}{\delta}\int_0^\infty \tilde{f}(x,\delta)\,dx = \frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(\beta+1)\alpha^{1+1/\beta}\delta} + O\bigg(\frac{\log^{1/\beta}(\delta^{-1})}{\delta}\bigg).$$

Hence, as $h \to 0$,

$$Z(h) = \frac{\beta^{2+1/\beta}\log^{1+1/\beta}(h^{-1})}{(\beta+1)\alpha^{1+1/\beta}\mu^{1/\beta}h^{1/\beta}} + O\bigg(\frac{\log^{1/\beta}(h^{-1})}{h^{1/\beta}}\bigg). \tag{47}$$

Furthermore, using the same reasoning as in the proof of Lemma 1, we can show (see Appendix V) that for $k = 1,2,3,4$, the derivative $\tilde{Z}^{(k)}(\delta)$ can be obtained by differentiating the series $\tilde{Z}(\delta) = \sum_{j=1}^\infty \tilde{f}(j\delta,\delta)$ term-by-term k times, and as $\delta \to 0$,

$$\tilde{Z}^{(k)}(\delta) = \frac{(-1)^k k!\beta^{2+1/\beta}}{(1+\beta)\alpha^{1+1/\beta}}\frac{\log^{1+1/\beta}(\delta^{-1})}{\delta^{k+1}} + O\bigg(\frac{\log^{1/\beta}(\delta^{-1})}{\delta^{k+1}}\bigg). \tag{48}$$

From this, noting that $Z^{(k)}(h) = \dfrac{d^k}{dh^k}\tilde{Z}(\delta(h))$, where $\delta = (\mu h)^{1/\beta}$, we get (see Appendix V)

$$Z'(h) = -\frac{\beta^{1+1/\beta}\log^{1+1/\beta}(h^{-1})}{(\beta+1)\alpha^{1+1/\beta}\mu^{1/\beta}h^{1+1/\beta}} + O\bigg(\frac{\log^{1/\beta}(h^{-1})}{h^{1+1/\beta}}\bigg), \tag{49}$$

$$Z''(h) = \frac{\beta^{1/\beta} \log^{1+1/\beta}(h^{-1})}{\alpha^{1+1/\beta}\mu^{1/\beta}h^{2+1/\beta}} + O\biggl(\frac{\log^{1/\beta}(h^{-1})}{h^{2+1/\beta}}\biggr), \tag{50}$$

$$Z'''(h) \asymp \frac{\log^{1+1/\beta}(h^{-1})}{h^{3+1/\beta}}, \qquad Z^{(4)}(h) \asymp \frac{\log^{1+1/\beta}(h^{-1})}{h^{4+1/\beta}}. \tag{51}$$

Also, using (84) and noting that

$$\frac{\partial^2}{\partial x^2}\tilde f(x,\delta) = \frac{\beta \tilde G(\delta) x^{\beta-2} e^{-x^\beta}\bigl(\beta x^\beta - (\beta-1)(1+e^{-x^\beta})\bigr)}{(1+\tilde G(\delta)e^{-x^\beta})^2},$$

we get as $h \to 0$

$$\begin{aligned}
\sum_{j=1}^\infty \bigl(Z_j''(h)\bigr)^2 &= \sum_{j=1}^\infty \biggl(\frac{\partial^2}{\partial h^2}\tilde Z_j(\delta(h))\biggr)^2 \\
&= (\delta'(h))^2 \sum_{j=1}^\infty \bigl(\tilde Z_j''(\delta)\bigr)^2 + 2\delta'(h)\delta''(h)\sum_{j=1}^\infty \tilde Z_j'(\delta)\tilde Z_j''(\delta) + (\delta''(h))^2 \sum_{j=1}^\infty \bigl(\tilde Z_j'(\delta)\bigr)^2 \\
&\asymp (\delta'(h))^2 \sum_{j=1}^\infty \bigl(\tilde Z_j''(\delta)\bigr)^2 \asymp \frac{(\delta'(h))^2}{\delta^4}\sum_{j=1}^\infty (j\delta)^4 \biggl(\frac{\partial^2}{\partial x^2}\tilde f(j\delta,\delta)\biggr)^2 \\
&\asymp \frac{(\delta'(h))^2}{\delta^5}\int_0^\infty x^4 \biggl(\frac{\partial^2}{\partial x^2}\tilde f(x,\delta)\biggr)^2 dx \asymp \frac{(\delta'(h))^2}{\delta^5}\log^{3+1/\beta}(\delta^{-1}) \\
&\asymp \frac{\log^{3+1/\beta}(h^{-1})}{h^{2-2/\beta}h^{5/\beta}} = \frac{\log^{3+1/\beta}(h^{-1})}{h^{2+3/\beta}}.
\end{aligned}$$

Therefore, applying (50) and (51), for $\beta > 1/2$ the Lyapunov ratio satisfies as $h \to 0$,

$$\begin{aligned}
L_h^{(4)} &= \frac{Z^{(4)} + 3\sum_{j=1}^\infty (Z_j''(h))^2}{(Z''(h))^2} \\
&= O\bigl(h^{1/\beta}\log^{-(1+1/\beta)}(h^{-1})\bigr) + O\bigl(h^{2-1/\beta}\log^{1-1/\beta}(h^{-1})\bigr) = o(1),
\end{aligned}$$

and thus (35) holds true:

$$\log N(t) = Z(h) + hH + o(hs_h), \qquad H = \log t \to \infty.$$

Taking into account (28) and (49),

$$H = \frac{\beta^{1+1/\beta}\log^{1+1/\beta}(h^{-1})}{(\beta+1)\alpha^{1+1/\beta}\mu^{1/\beta}h^{1+1/\beta}} + O\biggl(\frac{\log^{1/\beta}(h^{-1})}{h^{1+1/\beta}}\biggr), \qquad h \to 0. \tag{52}$$

Solving this equation for h, we get the approximate relation

$$h \sim \frac{\beta^2 \log H}{(\beta+1)^{(2\beta+1)/(1+\beta)}\alpha\mu^{1/(\beta+1)}H^{\beta/(1+\beta)}}.$$

Next, thanks to (28) and (50)

$$hs_h = h(Z''(h))^{1/2} \asymp \frac{\log^{(1+1/\beta)/2}(h^{-1})}{h^{1/(2\beta)}} \asymp H^{1/(2\beta+2)}\log^{1/2}(H).$$

Therefore

$$\begin{aligned}
\log N(t) &= \frac{\beta^{1+1/\beta}}{\alpha^{1+1/\beta}\mu^{1/\beta}}\frac{\log^{1+1/\beta}(h^{-1})}{h^{1/\beta}} + O\biggl(\frac{\log^{1/\beta}(h^{-1})}{h^{1/\beta}}\biggr) \\
&= \frac{\beta^2 H^{1/(1+\beta)}\log H}{\alpha(\beta+1)^{\beta/(1+\beta)}\mu^{1/(\beta+1)}}\bigl(1 + O(\log^{-1} H)\bigr)
\end{aligned}$$

$$= \frac{\beta^2 H^{1/(1+\beta)} \log H}{\alpha(\beta+1)^{\beta/(1+\beta)} \mu^{1/(1+\beta)}} + O\big(H^{1/(1+\beta)}\big). \tag{53}$$

Based on the above results, we can state and prove the following extension of Theorem 1.

Theorem 2. *Consider the weighted tensor product space $\mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ for which $\lambda_k = e^{\lambda|k|^\alpha}$, $\lambda > 0$, $\alpha > 0$, and $\gamma_j = e^{\mu j^\beta}$, $\mu > 0$, $\beta > 1/2$. Then, as $\varepsilon \to 0$,*

$$\log R_\varepsilon(\mathcal{F}) = \log\varepsilon + \frac{\beta^2 \log^{1/(\beta+1)}(\varepsilon^{-1}) \log\log(\varepsilon^{-1})}{2\alpha(1+\beta)^{\beta/(1+\beta)} \mu^{1/(1+\beta)}} + O\big(\log^{1/(1+\beta)}(\varepsilon^{-1})\big),$$

$$\log r^*_\varepsilon(\mathcal{F}) = \log\varepsilon + \frac{\beta^2 \log^{1/(\beta+1)}(\varepsilon^{-1}) \log\log(\varepsilon^{-1})}{4\alpha(1+\beta)^{\beta/(1+\beta)} \mu^{1/(1+\beta)}} + O\big(\log^{1/(1+\beta)}(\varepsilon^{-1})\big).$$

The proof of Theorem 2 repeats that of Theorem 1, with (52) and (53) in place of (42) and (44), and therefore is omitted.

Remark 3. Similarly to Theorem 1, cf. Remark 2, the logarithmic asymptotics in Theorem 2 continue to hold for a d-variable signal f when $d = d_\varepsilon$ grows to infinity at least as fast as $O\big(\log\log(\varepsilon^{-1})\big)$.

5. SHARP OPTIMALITY

5.1. Sharp Asymptotics for Analytic Functions

In this section we deal with the class $\mathcal{A}_{\lambda,\mu}$ of analytic functions in Example 1. For deriving sharp asymptotics in the estimation and detection problems at hand relation (44) is not sufficient. We need a more accurate result on asymptotics of $N(t)$ as $t \to \infty$. In order to get such a result, we strengthen limiting relation (34) by proving the following version of the Local Limit Theorem.

Lemma 3. *Let $hs_h \to \infty$. Then there exists $\rho = \rho_h > 0$, $\rho_h = o((hs_h)^{-1})$ such that, for any $a = o(1)$,*

$$\mathbf{P}_h\big(\tau_h \in (a, a+\rho)\big) \sim \mathbf{P}(\tau \in (a, a+\rho)), \qquad \tau \sim \mathcal{N}(0,1). \tag{54}$$

The proof of Lemma 3 is given in Appendix IV. Note that

$$\mathbf{P}(\tau \in (a, a+\rho)) \sim \frac{\rho}{\sqrt{2\pi}}, \qquad a \to 0, \quad \rho \to 0.$$

Therefore, as shown in [8, Lemma 5.2], Lemma 3 implies

$$I_h \sim (hs_h\sqrt{2\pi})^{-1}. \tag{55}$$

By (20) this yields, compared with (35),

$$N(t) \sim \frac{\exp(Z(h) + hH)}{hs_h\sqrt{2\pi}}, \qquad H = \log t \to \infty, \tag{56}$$

with h and s_h defined in (28).

According to Proposition 6.4 of [9], under the validity of Lemma 3, the counting function $N(t)$ is slowly varying if $h^* = 0$, $hs_h \to \infty$, and there exists a decreasing positive function $\phi(h)$ such that $Z''(h) \sim \phi^2(h)$ as $h \to 0$. In case of $\mathcal{A}_{\lambda,\mu}$ (Example 1), these assumptions are trivially satisfied in view of (38) and (39). Therefore $N(t)$ is a slowly varying function and hence by (14) and (15)

$$R^2_\varepsilon(\mathcal{F}) \sim \varepsilon^2 N(T), \qquad \text{where} \quad \log N(T) + 2\log T = 2\log\varepsilon^{-1} + O(1),$$
$$u^2_\varepsilon \sim r^4_\varepsilon / 2\varepsilon^4 N(T), \qquad \text{where} \quad \log N(T) + 4\log T = 4\log\varepsilon^{-1} + O(1).$$

At the same time, due to (56)

$$\log N(T) = Z(h) + hH - \log(hs_h) + O(1).$$

Thus we arrive at the following result.

Theorem 3. *For the weighted tensor product space $\mathcal{A}_{\lambda,\mu}$ of analytic functions, we have for the estimation problem:*

$$R^2_\varepsilon(\mathcal{F}) \sim \frac{\varepsilon^2 \exp(Z(h) + hH)}{hs_h\sqrt{2\pi}}, \qquad \varepsilon \to 0, \tag{57}$$

where $H = -Z'(h)$, $s_h = \sqrt{Z''(h)}$, and h satisfies the equation

$$Z(h) - (2 + h)Z'(h) - \log(hs_h) = 2\log(\varepsilon^{-1}) + O(1), \tag{58}$$

and for the detection problem:

$$\gamma_\varepsilon(r_\varepsilon) = 2\Phi(-u_\varepsilon/2) + o(1), \quad u^2_\varepsilon \sim \frac{\sqrt{2\pi}hs_h r^4_\varepsilon}{2\varepsilon^4 \exp(Z(h) + hH)}, \qquad \varepsilon \to 0, \tag{59}$$

where $H = -Z'(h)$, $s_h = \sqrt{Z''(h)}$, and h solves the equation

$$Z(h) - (4 + h)Z'(h) - \log(hs_h) = 4\log(\varepsilon^{-1}) + O(1). \tag{60}$$

Implicit forms of sharp asymptotics (57) and (59) are not very informative. In the next section we use the Euler–Maclaurin expansions of $Z(h)$ and $Z'(h)$ to give a more precise form of Theorem 3.

5.2. Sharp Asymptotics for Analytic Functions: Further Improvements

The Euler–Maclaurin formula that provides a powerful connection between integrals and sums can be written as follows (see [2], formula (11.5))

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,dx + \frac{1}{2}f(n) + C_m + R_{2m}, \tag{61}$$

where $f \in C^{2m}(0,\infty)$, C_m is a constant defined by

$$C_m = \frac{1}{2}f(0) - \frac{B_2}{2!}f'(0) - \ldots - \frac{B_{2m}}{(2m)!}f^{(2m-1)}(0),$$

and R_{2m} is a remainder term of the form

$$R_{2m} = -\int_0^n f^{(2m)}(x)\frac{B_{2m}(x - [x])}{(2m)!}\,dx.$$

Here B_n are the Bernoulli numbers and $B_n(x) = \sum_{k=0}^n \binom{n}{k} B_k x^{n-k}$ are the Bernoulli polynomials. It is well known (see [3, n^o 9.71]) that $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$. Therefore for $0 \le x \le 1$

$$B_2(x) = x^2 - x + 1/6 \qquad \text{and} \qquad \max_{0\le x\le 1}|B_2(x)| = 1/6.$$

Recalling that $Z(h) = \sum_{j=1}^\infty f(jh,h)$, where $f(x,h) = \log(1 + G(h)e^{-\mu x})$, and applying (61) with $m = 1$ we get, as $n \to \infty$,

$$Z(h) = F(h) - \frac{f(0,h)}{2} - \frac{1}{12}\frac{\partial}{\partial x}f(0,h) - \frac{h}{2}\int_0^\infty \frac{\partial^2}{\partial x^2}f(x,h)B_2(x - [x])\,dx, \tag{62}$$

where $F(h) = h^{-1}\int_0^\infty \log\left(1 + G(h)e^{-\mu x}\right)dx \asymp h^{-1}\log^2(h^{-1})$, $f(0,h) \asymp \log(h^{-1})$, $\frac{\partial}{\partial x}f(0,h) \asymp 1$, and since $|B_2(x)| \le 1/6$,

$$\left|\frac{h}{2}\int_0^\infty \frac{\partial^2}{\partial x^2}f(x,h)B_2(x - [x])\,dx\right| \le \frac{h}{12}\left|\frac{\partial}{\partial x}f(0,h)\right| \asymp h.$$

Therefore, as $h \to 0$,

$$Z(h) = F(h) - \frac{\log(1+G(h))}{2} + O(1). \tag{63}$$

Next, upon differentiation of both sides of (62), we obtain as $h \to 0$

$$\begin{aligned}
Z'(h) &= F'(h) - \frac{1}{2}\frac{\partial}{\partial h} f(0,h) - \frac{1}{12}\frac{\partial^2}{\partial x \partial h} f(0,h) \\
&\quad - \frac{d}{dh}\left(\frac{h}{2}\int_0^\infty \frac{\partial^2}{\partial x^2} f(x,h) B_2(x-[x])\,dx\right) \\
&= -\frac{F(h)}{h} + \frac{G'(h)\log(1+G(h))}{\mu h G(h)} - \frac{G'(h)}{2(1+G(h))} + \frac{\mu G'(h)}{12(1+G(h))^2} \\
&\quad - \frac{1}{2}\int_0^\infty \frac{\partial^2}{\partial x^2} f(x,h) B_2(x-[x])\,dx - \frac{h}{2}\int_0^\infty \frac{\partial^3}{\partial x^2 \partial h} f(x,h) B_2(x-[x])\,dx \\
&= -\frac{F(h)}{h} + \frac{G'(h)\log(1+G(h))}{\mu h G(h)} - \frac{G'(h)}{2(1+G(h))} + O(1),
\end{aligned} \tag{64}$$

where $h^{-1}F(h) = O\big(h^{-2}\log^2(h^{-1})\big)$ is the main term, and

$$\frac{G'(h)\log(1+G(h))}{\mu h G(h)} = O\big(h^{-2}\log(h^{-1})\big), \qquad \frac{G'(h)}{2(1+G(h))} = O(h^{-1}).$$

Note also, cf. (39) and (83), that

$$Z''(h) = \frac{2F(h)}{h^2} + O\left(\frac{\log(h^{-1})}{h^3}\right),$$

and hence

$$hs_h = h\sqrt{Z''(h)} = \sqrt{2F(h)(1 + O(1/\log(h^{-1})))} \sim \sqrt{2F(h)}. \tag{65}$$

Let us now return to relations (57)–(60). At this point, it is convenient to represent the integral $F(h)$ as a function of $G = G(h)$:

$$I(G) = \frac{1}{\mu h}\int_0^G \frac{\log(1+t)}{t}\,dt. \tag{66}$$

As seen from (58) and (60), the precision required for solving the equation $H = -Z'(h)$ is controlled by the term $\log(hs_h)$. Using (65) and (83), we have

$$\log(hs_h) = \log\left(h\sqrt{Z''(h)}\right) \sim \frac{1}{2}\log(h^{-1}) + \log\log(h^{-1}) + O(1).$$

Then, due to (63) and (64), with the required precision, $H = H(h)$ that appears in (57) and (59) solves the equation

$$-\frac{I(G)}{\mu h^2} + \frac{G'\log(1+G)}{\mu h G} - \frac{G'}{2(1+G)} = H,$$

and the parameter h is chosen to satisfy, cf. (58),

$$I(G) - \frac{\log(1+G)}{2} + (2+h)H - \frac{\log(h^{-1})}{2} - \log\log(h^{-1}) = 2\log(\varepsilon^{-1}) + O(1)$$

for the estimation problem, and, cf. (60),

$$I(G) - \frac{\log(1+G)}{2} + (4+h)H - \frac{\log(h^{-1})}{2} - \log\log(h^{-1}) = 4\log(\varepsilon^{-1}) + O(1)$$

for the detection problem. Summarizing the above discussion we arrive at the following improvement of Theorem 3.

Theorem 4. *Consider estimating and detecting signal $f \in \mathcal{A}_{\lambda,\mu}$ in the Gaussian white noise model. Then, as $\varepsilon \to 0$,*

$$R_\varepsilon^2(\mathcal{F}) \sim \frac{\varepsilon^2 \exp(T(G))}{2\sqrt{\pi I(G)}},$$

where $G = G(h) = 2/(e^{\lambda h} - 1)$, $I(G)$ is defined by (66),

$$T(G) = I(G) - \frac{\log(1+G)}{2} - \frac{I(G)}{\mu G} + \frac{G' \log(1+G)}{\mu G},$$

and $h = h(\varepsilon)$ solves the equation

$$I(G) - \frac{\log(1+G)}{2} - \frac{(2+h)\,I(G)}{\mu h^2} + \frac{(2+h)\,G'\log(1+G)}{\mu h G} - \frac{G'}{1+G} \\ - \frac{1}{2}\log(h^{-1}) - \log\log(h^{-1}) = 2\log(\varepsilon^{-1}) + O(1), \tag{67}$$

and the total error probability is $\gamma_\varepsilon(r_\varepsilon) \sim 2\Phi(-u_\varepsilon/2) + o(1)$, where

$$u_\varepsilon^2 \sim \frac{r_\varepsilon^4 \sqrt{\pi I(G)}}{\varepsilon^4 \exp(T(G))},$$

with G, $I(G)$, and $T(G)$ as above, and $h = h(\varepsilon)$ satisfying

$$I(G) - \frac{\log(1+G)}{2} - \frac{(4+h)\,I(G)}{\mu h^2} + \frac{(4+h)\,G'\log(1+G)}{\mu h G} - \frac{G'}{1+G} \\ - \frac{1}{2}\log(h^{-1}) - \log\log(h^{-1}) = 4\log(\varepsilon^{-1}) + O(1). \tag{68}$$

Again, the sharp asymptotics of $R_\varepsilon(\mathcal{F})$ and $\gamma_\varepsilon(r_\varepsilon)$ are given in an implicit form. An attempt to obtain explicit expressions encounters an obstacle: one has to solve nonlinear equations (67) and (68). Analytically, this can hardly be done. Unlike the case of analytic functions under consideration, the Sloan–Woźniakowski space in Example 3 admits explicit forms of sharp asymptotics of $R_\varepsilon(\mathcal{F})$ and $\gamma_\varepsilon(r_\varepsilon)$ (see [8, Sec. 3] for details).

APPENDIX I: ESTIMATION PROBLEM

Consider estimating the signal $f \in \mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}}$ in the model (1) for which $d = \infty$. With the tensor product basis $\phi_{\boldsymbol{\ell}}(\mathbf{t}) = \prod_{k=1}^{\infty} \phi_{l_k}(t_k)$, $\mathbf{t} = (t_1, t_2, \dots) \in \mathbf{R}^\infty$, in L_2^∞ and the Fourier coefficients $\theta_{\boldsymbol{\ell}} = (f, \phi_{\boldsymbol{\ell}})$, the norm of the function $f(\mathbf{t}) = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}\phi_{\boldsymbol{\ell}}(\mathbf{t})$ is determined by

$$\|f\|^2_{\boldsymbol{\lambda},\boldsymbol{\gamma}} = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} c_{\boldsymbol{\ell}}^2 \theta_{\boldsymbol{\ell}}^2, \qquad c_{\boldsymbol{\ell}} = \prod_{j\,:\, l_j \neq 0} \gamma_j \lambda_{l_j}.$$

We estimate f by means of the projection-type estimator

$$\hat{f}_{\varepsilon,t}(\mathbf{t}) = \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} X_{\boldsymbol{\ell}}\phi_{\boldsymbol{\ell}}(\mathbf{t}),$$

where $\mathcal{N}(t) = \{\boldsymbol{\ell} \in \mathbf{Z}_0^\infty \colon c_{\boldsymbol{\ell}} \le t\}$ and the $X_{\boldsymbol{\ell}} = X_\varepsilon(\phi_{\boldsymbol{\ell}})$ are the empirical Fourier coefficients. Under the validity of model (1), the $X_{\boldsymbol{\ell}}$ are i.i.d. normal $N(\theta_{\boldsymbol{\ell}}, \varepsilon^2)$ random variables, and the usual variance-bias decomposition leads to

$$\mathbf{E}_{\varepsilon,f}\|\hat{f}_{\varepsilon,t} - f\|_2^2 = \mathbf{E}_{\varepsilon,f}\Big\| \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} (X_{\boldsymbol{\ell}} - \theta_{\boldsymbol{\ell}})\phi_{\boldsymbol{\ell}}\Big\|_2^2 + \Big\| \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty \colon c_{\boldsymbol{\ell}} > t} \theta_{\boldsymbol{\ell}}\phi_{\boldsymbol{\ell}}\Big\|_2^2 = \varepsilon^2 N(t) + \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty \colon c_{\boldsymbol{\ell}} > t} \theta_{\boldsymbol{\ell}}^2,$$

where $N(t) = \operatorname{card}\{\mathcal{N}(t)\}$ is the counting function that satisfies $N(t) \uparrow \infty$ as $t \to \infty$.

The arguments that lead to relations (12) and (14) involve the van Trees inequality and can be found, for example, in [9, Sec. 2.1]. For the reader's convenience, these arguments are presented below.

First, consider the rate optimality problem. Let $B = B_\varepsilon \asymp 1$ and let $T = T(B_\varepsilon)$ satisfy the *balance equation*

$$T = \sup\{t \colon \varepsilon^2 t^2 N(t) \le B\}. \tag{69}$$

Choose $t_1 < T$ and $t_2 > T$. Then the balance equation implies the following upper bounds:

$$R_\varepsilon(\mathcal{F}) \le t_1^{-1}\sqrt{1+B}, \qquad R_\varepsilon(\mathcal{F}) \le \varepsilon\sqrt{N(t_2)(1+1/B)}. \tag{70}$$

The appropriate lower bounds are obtained by using the van Trees inequality, a Bayesian analog of the Cramér–Rao inequality. Recall that the ball $\mathcal{F} = \{f \in \mathcal{L}_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \colon \|f\|_{\boldsymbol{\lambda},\boldsymbol{\gamma}} \le 1\}$ corresponds to the ellipsoid

$$\Theta = \Big\{\theta = (\theta_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \colon \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2 \le 1\Big\}$$

in the space of Fourier coefficients. The prior distribution on Θ is defined as follows. First, note that Θ contains the $N(t)$-dimensional ball $B^{N(t)}(0, 1/t) = \{\theta \colon \sum_{\boldsymbol{\ell}\colon c_{\boldsymbol{\ell}} \le t} \theta_{\boldsymbol{\ell}}^2 \le 1/t^2\}$ of radius t^{-1}, and hence it also contains the $N(t)$-dimensional cube with side $2L$, $L = (t\sqrt{N(t)})^{-1}$. Using this fact, the prior distribution Π is taken to be

$$\Pi(d\theta) = \prod_{j=1}^{N(t)} L^{-1}h(L^{-1}\theta_j)\,\lambda(d\theta), \qquad h(x) = \cos^2\Big(\frac{\pi x}{2}\Big)\mathbb{I}(|x| \le 1),$$

where λ is the Lebesgue measure on $\mathbf{R}^{N(t)}$. It is easy to see that $\Pi(\Theta) = 1$. Let $R(\Pi)$ denote the Bayes risk with respect to prior Π, that is,

$$R^2(\Pi) = \inf_{\hat{\theta}_\varepsilon} \int_\Theta \mathbf{E}_{\varepsilon,\theta}\|\theta - \hat{\theta}_\varepsilon\|^2 \Pi(d\theta) = \inf_{\hat{\theta}_\varepsilon} \int_\Theta \mathbf{E}_{\varepsilon,\theta}\Big(\sum_{j=1}^{N(t)} |\theta_j - \hat{\theta}_{\varepsilon,j}|^2\Big)\Pi(d\theta).$$

Applying the van Trees inequality (see [1, Prop. 1]) and recalling that $L^2 = (t^2N(t))^{-1}$, we get

$$R_\varepsilon^2(\mathcal{F}) = R_\varepsilon^2(\Theta) = \inf_{\hat{\theta}_\varepsilon}\sup_{\theta\in\mathcal{F}} \mathbf{E}_{\varepsilon,\theta}\|\theta - \hat{\theta}_\varepsilon\|^2 \ge R^2(\Pi) \ge \frac{\varepsilon^2 N(t)}{1 + \pi^2\varepsilon^2 t^2 N(t)}. \tag{71}$$

Now take $t_1 > T$ and $t_2 < T$, where T is defined in (69). Then (71) implies the following lower bounds:

$$R_\varepsilon(\mathcal{F}) \ge t_1{}^{-1}\sqrt{B/(1+\pi^2 B)}, \qquad R_\varepsilon(\mathcal{F}) \ge \varepsilon\sqrt{N(t_2)/(1+\pi^2 B)}. \tag{72}$$

Comparing (70) and (72) we arrive at relation (12):

$$R_\varepsilon(\mathcal{F}) \asymp T^{-1}, \qquad \text{where} \quad \varepsilon^2 T^2 N(T) \asymp 1.$$

Let us now turn to the sharp optimality problem and show that (14) holds true provided the function $N(t)$ is slowly varying. Note that the asymptotics of $N(T)$, where $T = T(B_\varepsilon)$ is determined by the balance equation (69), does not depend on the family $B_\varepsilon \asymp 1$. Indeed, if $B_1 = B_{1,\varepsilon} \asymp 1$ is another family that appears in (69), then $T(B_\varepsilon) \sim T(B_{1,\varepsilon})$ and $N(T(B_\varepsilon)) \sim N(T(B_{1,\varepsilon}))$. Using again the fact that $N(t)$ is a slowly varying function, we can choose positive families $B_\varepsilon \to \infty$ and $b_\varepsilon \to 0$ such that

$N(T(B_\varepsilon)) \sim N(T(b_\varepsilon)) \sim N(T(1))$. Finally, setting $t_2 = T(B_\varepsilon)$ in (70) and $t_1 = T(b_\varepsilon)$ in (72) leads to (14):

$$R_\varepsilon(\mathcal{F}) \sim \varepsilon N^{1/2}(t), \qquad \text{where} \quad \varepsilon^2 T^2 N(T) \asymp 1.$$

APPENDIX II: DETECTION PROBLEM

This Appendix is here for the reader's convenience. It presents some well-known results of the theory of nonparametric goodness-of-fit testing under Gaussian models scattered among various publications and chapters of the book [6].

Consider testing

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

Below we provide simple arguments that lead to (13) and (15). The proof of these asymptotic relations is performed in two steps (lower bound and upper bound).

Lower bound. The lower bound on the minimax total error probability $\gamma_\varepsilon(r_\varepsilon)$ is obtained by reducing the problem of testing H_0 versus $H_{1\varepsilon}$ to the problem of testing H_0 versus a suitable family of *finite-dimensional* alternatives. For this, consider the family of $(N(t)-1)$-dimensional spheres that correspond to the first $N(t)$ coordinates of $\theta = (\theta_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty}$ in l^2:

$$S^{N(t)-1}(r_\varepsilon) = \Big\{\theta \in \Theta_t\colon \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2 = r_\varepsilon^2\Big\},$$

where $\Theta_t = \big\{\theta = \{\theta_{\boldsymbol{\ell}}\}\colon \theta_{\boldsymbol{\ell}} = 0 \text{ if } \boldsymbol{\ell} \notin \mathcal{N}(t)\big\}$, and t is chosen to have $tr_\varepsilon \le 1$. Then for any $\theta \in S^{N(t)-1}(r_\varepsilon)$,

$$\sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2 \le t^2 \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2 = t^2 r_\varepsilon^2 \le 1,$$

and hence $S^{N(t)-1}(r_\varepsilon) \subset \Theta_\varepsilon$, where

$$\Theta_\varepsilon = \Big\{\theta = (\theta_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty}\colon \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 c_{\boldsymbol{\ell}}^2 \le 1 \text{ and } \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 \ge r_\varepsilon^2\Big\}.$$

The lower bound now follows from [6, Prop. 2.15]. Namely, if $t \to \infty$ and $\varepsilon \to 0$ in such a way that $tr_\varepsilon \le 1$, then

$$\gamma_\varepsilon(r_\varepsilon) \ge \gamma_\varepsilon\big(S^{N(t)-1}(r_\varepsilon)\big) = \Phi(-u_\varepsilon(t)/2) + o(1), \qquad u_\varepsilon(t) = \frac{r_\varepsilon^2}{\varepsilon^2\sqrt{2N(t)}}, \tag{73}$$

where $\gamma_\varepsilon(S^{N(t)-1}(r_\varepsilon))$ is the minimax total error probability that corresponds to testing $H_0\colon \theta = 0$ versus $H'_{1\varepsilon}\colon \theta \in S^{N(t)-1}(r_\varepsilon)$.

Upper bound. The required upper bound is achieved for the χ^2-type test $\psi_{\varepsilon,t} = \mathbb{I}\{T_{\varepsilon,t} > h_\varepsilon(t)/2\}$ based on the statistic

$$T_{\varepsilon,t} = \frac{1}{\sqrt{2N(t)}} \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \big((X_{\boldsymbol{\ell}}/\varepsilon)^2 - 1\big), \tag{74}$$

where $h_\varepsilon(t)$ is defined in (77) and the $X_{\boldsymbol{\ell}} = X_\varepsilon(\phi_{\boldsymbol{\ell}})$ are the empirical Fourier coefficients. Under H_0, the statistic $T_{\varepsilon,t}$ is asymptotically normal with mean 0 and variance 1. Under the alternative, the $(X_{\boldsymbol{\ell}}/\varepsilon)$ are i.i.d. normal $\mathcal{N}(\theta_{\boldsymbol{\ell}}/\varepsilon, 1)$ random variables. Therefore

$$h_\varepsilon(\theta, t) = \mathbf{E}_\theta T_{\varepsilon,t} = \frac{1}{\varepsilon^2\sqrt{2N(t)}} \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2. \tag{75}$$

When studying the total probability error, we can assume that $\mathbf{E}_\theta T_{\varepsilon,t} = O(1)$, see [6, Corollary 3.1]. For $\theta \in \Theta_\varepsilon$,

$$\sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2 = \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 - \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty:\, c_{\boldsymbol{\ell}}>t} \theta_{\boldsymbol{\ell}}^2 \geq \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty} \theta_{\boldsymbol{\ell}}^2 - t^{-2} \sum_{\boldsymbol{\ell}\in\mathbf{Z}_0^\infty:\, c_{\boldsymbol{\ell}}>t} c_{\boldsymbol{\ell}}^2\theta_{\boldsymbol{\ell}}^2 \geq r_\varepsilon^2 - t^{-2}$$
$$= r_\varepsilon^2\big(1-(tr_\varepsilon)^{-2}\big) \geq r_\varepsilon^2\big(1-C^{-2}\big),$$

where the constant C is such that $tr_\varepsilon > C > 1$, and hence

$$h_\varepsilon(\theta,t) \geq \frac{r_\varepsilon^2}{\varepsilon^2\sqrt{2N(t)}}(1-C^{-2}). \tag{76}$$

Next, using (75)

$$\operatorname{Var}_\theta T_{\varepsilon,t} = \frac{1}{2N(t)} \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \operatorname{Var}_\theta(X_{\boldsymbol{\ell}}/\varepsilon)^2 = \frac{1}{2N(t)} \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \big(2+4\mathbf{E}_\theta^2(X_{\boldsymbol{\ell}}/\varepsilon)\big)$$
$$= 1 + \frac{2}{\varepsilon^2 N(t)} \sum_{\boldsymbol{\ell}\in\mathcal{N}(t)} \theta_{\boldsymbol{\ell}}^2 = 1 + \frac{4}{\sqrt{2N(t)}}\,\mathbf{E}_\theta T_{\varepsilon,t}.$$

From this, recalling that $\mathbf{E}_\theta T_{\varepsilon,t} = O(1)$ we get as $t\to\infty$,

$$\operatorname{Var}_\theta T_{\varepsilon,t} = 1 + o(1).$$

Set

$$h_\varepsilon(t) = \inf_{\theta\in\Theta_\varepsilon} h_\varepsilon(\theta,t) = \inf_{\theta\in\Theta_\varepsilon} \mathbf{E}_\theta T_{\varepsilon,t}. \tag{77}$$

It now follows from [6, Corollary 3.2] that the test

$$\psi_{\varepsilon,t} = \mathbb{I}\{T_{\varepsilon,t} > h_\varepsilon(t)/2\}$$

satisfies

$$\gamma_\varepsilon(r_\varepsilon, \psi_{\varepsilon,t}) \leq 2\Phi(-h_\varepsilon(t)/2) + o(1), \qquad \varepsilon\to 0.$$

Taking into account (76), we get

$$\gamma_\varepsilon(r_\varepsilon,\psi_{\varepsilon,t}) \leq 2\Phi\Big(-\frac{u_\varepsilon(t)}{2}(1-C^{-2})\Big) + o(1), \qquad u_\varepsilon(t) = \frac{r_\varepsilon^2}{\varepsilon^2\sqrt{2N(t)}}. \tag{78}$$

The comparison of the lower and upper bounds implies that the separation rate $r_\varepsilon^*(\mathcal{F})$ must satisfy the relation $r_\varepsilon^*(\mathcal{F}) \asymp T^{-1}$, where $T = T_\varepsilon$ is determined by the balance equation

$$T = \sup\{t>0\colon \varepsilon^4 t^4 N(t) \leq B\}, \qquad B = B_\varepsilon \asymp 1. \tag{79}$$

Asymptotically minimax tests. The above results will now be applied to construct a family of asymptotically minimax tests. Recall that the family of tests $\{\psi_\varepsilon\}$ is called *asymptotically minimax* if

$$\gamma_\varepsilon(r_\varepsilon,\psi_\varepsilon) \leq \gamma_\varepsilon(r_\varepsilon) + o(1), \qquad \varepsilon\to 0,$$

where $\gamma_\varepsilon(r_\varepsilon) = \inf_\psi \gamma_\varepsilon(r_\varepsilon,\psi)$. If the counting function $N(t)$ is slowly varying then the χ^2-type test

$$\psi_{\varepsilon,t} = \mathbb{I}\{T_{\varepsilon,t} > u_\varepsilon(t)/2\}, \qquad u_\varepsilon(t) = \frac{r_\varepsilon^2}{\varepsilon^2\sqrt{2N(t)}}, \tag{80}$$

is asymptotically minimax. Indeed, since $N(t)$ is a slowly varying function, it follows that the asymptotics for $u_\varepsilon(T_\varepsilon)$ do not depend on the family B_ε that appears in the balance equation. Therefore it is sufficient to consider the case $r_\varepsilon \asymp r_\varepsilon^*(\mathcal{F})$, or equivalently, the case $r_\varepsilon T_\varepsilon \asymp 1$. Let the family t_ε satisfy $t_\varepsilon \sim (2r_\varepsilon)^{-1} \asymp$

T_ε. Then $r_\varepsilon t_\varepsilon \sim 1/2$, and hence the lower bound (73) holds with $t = t_\varepsilon$. For the upper bound, choose the family C_ε in such a way that $\lim_{\varepsilon\to 0} N(C_\varepsilon T_\varepsilon)/N(T_\varepsilon) = 1$. Then for $t = t_\varepsilon = C_\varepsilon T_\varepsilon$,

$$\inf_{\theta\in\Theta_\varepsilon} h_\varepsilon(\theta, t) \ge u_\varepsilon(t)(1 + o(1)),$$

and hence

$$\gamma_\varepsilon(r_\varepsilon, \psi_{\varepsilon,t}) \le 2\Phi\left(-u_\varepsilon(t)/2\right) + o(1).$$

Asymptotic equality of the lower and upper bounds implies minimaxity of test (80).

APPENDIX III: ASYMPTOTIC REPRESENTATIONS FOR $F(h)$ AND $\tilde{F}(\delta)$

For $\delta = \delta(h) = (\mu h)^{1/\beta} > 0$, consider the function

$$\tilde{F}(\delta) = \frac{1}{\delta}\int_0^\infty \log(1 + \tilde{G}(\delta)e^{-x^\beta})\, dx,$$

where, cf. (19),

$$\tilde{G}(\delta) = 2\sum_{k=1}^\infty e^{-h\lambda k^\alpha} = \frac{2\Gamma(1/\alpha)\mu^{1/\alpha}}{\alpha\lambda^{1/\alpha}\delta^{\beta/\alpha}} + O(1), \qquad \delta \to 0.$$

Let $\tilde{J}$ be such that $\tilde{G}(\delta)e^{-\tilde{J}^\beta} = 1$, that is, $\tilde{J} = \log^{1/\beta}\tilde{G}(\delta)$. Then

$$\tilde{F}(\delta) = \frac{1}{\delta}\int_0^{\tilde{J}} (\log\tilde{G}(\delta) - x^\beta)\, dx + \frac{1}{\delta}\int_0^{\tilde{J}} \log\left(1 + \frac{e^{x^\beta}}{\tilde{G}(\delta)}\right) dx$$

$$+ \frac{1}{\delta}\int_{\tilde{J}}^\infty \log\left(1 + \tilde{G}(\delta)e^{-x^\beta}\right) dx = K_1 + K_2 + K_3.$$

The first integral is equal to

$$K_1 = \frac{\tilde{J}\log\tilde{G}(\delta)}{\delta} - \frac{\tilde{J}^{1+\beta}}{(1+\beta)\delta} = \frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta} + O\left(\frac{\log^{1/\beta}(\delta^{-1})}{\delta}\right).$$

The second integral satisfies

$$K_2 \le \frac{(\log 2)\tilde{J}}{\delta} \asymp \frac{\log^{1/\beta}(\delta^{-1})}{\delta} = o(K_1).$$

Next, using the inequality $\log(1+x) \le x$ and the asymptotic relation

$$\int_t^\infty e^{-x^\beta}\, dx \sim \beta^{-1}t^{1-\beta}e^{-t^\beta}, \qquad t \to \infty, \tag{81}$$

the third integral is estimated as follows:

$$K_3 \le \frac{\tilde{G}(\delta)}{\delta}\int_{\tilde{J}}^\infty e^{-x^\beta}\, dx \asymp \frac{\tilde{J}^{1-\beta}}{\delta} \asymp \frac{\log^{1/\beta-1}(\delta^{-1})}{\delta} = o(K_1).$$

Therefore as $\delta \to 0$,

$$\tilde{F}(\delta) = \frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta} + O\left(\frac{\log^{1/\beta}(\delta^{-1})}{\delta}\right). \tag{82}$$

In particular, by taking $\alpha = \beta = 1$ in formula (82) we get

$$F(h) = \frac{\log^2(h^{-1})}{2\mu h} + O\left(\frac{\log(h^{-1})}{h}\right). \tag{83}$$

APPENDIX IV: PROOFS OF LEMMAS

Proof of Lemma 1. First, we show that the derivatives $Z^{(k)}(h)$, $k = 1, 2, 3, 4$, can be obtained by differentiating the series $Z(h) = \sum_{j=1}^{\infty} f(jh, h)$ term-by-term k times, and that

$$Z^{(k)}(h) = F^{(k)}(h) + O\left(\frac{\log^k(h^{-1})}{h^k}\right),$$

where

$$F^{(k)}(h) = \frac{(-1)^k k! \log^2(h^{-1})}{2\mu h^{k+1}} + O\left(\frac{\log(h^{-1})}{h^{k+1}}\right), \qquad h \to 0.$$

The case $k = 1$ is considered in detail. The cases $k = 2, 3, 4$ are treated similarly.

Recall that

$$Z(h) = \sum_{j=1}^{\infty} \log\left(1 + G(h)e^{-\mu hj}\right) = \sum_{j=1}^{\infty} f(jh, h),$$

where $f(x, h) = \log(1 + G(h)e^{-\mu x})$ and $G(h) = 2/(e^{\lambda h} - 1)$. For the sequel we need the following result.

Proposition 1 (Prop. 1.3 of [2]). *Let $f(x)$ be continuously differentiable for all $x \geq 0$. Then*

$$\left|\sum_{j=0}^{n} f(j) - \int_0^n f(x)\,dx\right| \leq \int_0^n |f'(x)|\,dx + |f(0)|.$$

This proposition implies, for $h > 0$,

$$\left|\sum_{j=0}^{n} f(jh) - \frac{1}{h}\int_0^{nh} f(x)\,dx\right| \leq \int_0^{nh} |f'(x)|\,dx + |f(0)|.$$

If, in addition, we assume that f and f' are integrable on $(0, \infty)$, then as $n \to \infty$

$$\begin{aligned}\left|\sum_{j=0}^{\infty} f(hj) - \frac{1}{h}\int_0^{\infty} f(x)\,dx\right| &\leq \int_0^{\infty} |f'(x)|\,dx + |f(0)| \\ &\leq (2m)\max_{x\geq 0}|f(x)| + 2|f(0)|, \qquad (84)\end{aligned}$$

where m is the number of local extrema of function $f(x)$ on the interval $[0, \infty)$.

Applying (84) to the function $f(x, h) = \log(1 + G(h)e^{-\mu x})$, we get

$$\begin{aligned}Z(h) = \sum_{j=1}^{\infty} f(jh, h) &= \frac{1}{h}\int_0^{\infty} f(x, h)\,dx + O\left(\int_0^{\infty} \frac{\partial}{\partial x} f(x, h)\,dx + f(0, h)\right) \\ &= F(h) + O\big(f(0, h)\big), \qquad (85)\end{aligned}$$

with $F(h)$ satisfying (83).

Now, on the one hand, noting that

$$\frac{\partial}{\partial h} f(x, h) = \frac{G'(h)e^{-\mu x}}{1 + G(h)e^{-\mu x}}, \qquad G(h) \sim \frac{2}{\lambda h}, \quad G'(h) \sim \frac{2}{\lambda h^2},$$

and differentiating both sides of (85) gives

$$Z'(h) = F'(h) + O\left(\max_{x\geq 0}\left|\frac{\partial}{\partial h} f(x, h)\right|\right)$$

$$
\begin{aligned}
&= -\frac{1}{h^2}\int_0^\infty f(x,h)\,dx + \frac{1}{h}\int_0^\infty \frac{\partial}{\partial h} f(x,h)\,dx + O(h^{-1}) \\
&= I_1 + I_2 + O(h^{-1}),
\end{aligned}
$$

where according to (83)

$$
I_1 = -\frac{1}{h^2}\int_0^\infty f(x,h)\,dx = -\frac{F(h)}{h} = -\frac{\log^2(h^{-1})}{2\mu h^2} + O\Big(\frac{\log(h^{-1})}{h^2}\Big).
$$

For estimating I_2, take $J = \mu^{-1}\log G(h) \sim -\mu^{-1}\log(h^{-1})$. Then

$$
\begin{aligned}
I_2 &= \frac{1}{h}\int_0^\infty \frac{\partial}{\partial h} f(x,h)\,dx = \frac{G'(h)}{h}\int_0^\infty \frac{e^{-\mu x}}{1+G(h)e^{-\mu x}}\,dx \\
&= \frac{G'(h)}{h}\Big(\int_0^J \frac{e^{-\mu x}}{1+G(h)e^{-\mu x}}\,dx + \int_J^\infty \frac{e^{-\mu x}}{1+G(h)e^{-\mu x}}\,dx\Big) \\
&\le \frac{G'(h)}{h}\Big(\frac{J}{G(h)} + \frac{1}{G(h)}\Big) = O\Big(\frac{\log(h^{-1})}{h^2}\Big).
\end{aligned}
$$

Therefore

$$
Z'(h) = -\frac{\log^2(h^{-1})}{2\mu h^2} + O\Big(\frac{\log(h^{-1})}{h^2}\Big). \tag{86}
$$

On the other hand, differentiating the series $Z(h) = \sum_{j=1}^\infty f(jh,h)$ term-by-term and applying formula (84) yields

$$
\begin{aligned}
Z'(h) &= \sum_{j=1}^\infty \frac{\partial}{\partial h} f(jh,h) = \frac{1}{h}\sum_{j=1}^\infty (jh)\frac{\partial}{\partial x} f(jh,h) + \sum_{j=1}^\infty \frac{\partial}{\partial h} f(jh,h) \\
&= \frac{1}{h}\Big[\frac{1}{h}\int_0^\infty x\frac{\partial}{\partial x} f(x,h) + O\Big(\max_{x\ge 0}\Big|x\frac{\partial}{\partial x} f(x,h)\Big|\Big)\Big] \\
&\quad + \frac{1}{h}\int_0^\infty \frac{\partial}{\partial h} f(x,h)\,dx + O\Big(\max_{x\ge 0}\Big|\frac{\partial}{\partial h} f(x,h)\Big|\Big).
\end{aligned}
$$

Integrating by parts in the first integral and using the fact that

$$
\lim_{x\to\infty} x^k f(x,h) = 0, \qquad k = 0,1,\dots,
$$

we may continue

$$
\begin{aligned}
Z'(h) &= -\frac{1}{h^2}\int_0^\infty f(x,h) + \frac{1}{h}\int_0^\infty \frac{\partial}{\partial h} f(x,h)\,dx + \Delta_1 + \Delta_2 \\
&= F'(h) + \Delta_1 + \Delta_2,
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_1 &= h^{-1} O\Big(\max_{x\ge 0}\Big|x\frac{\partial}{\partial x} f(x,h)\Big|\Big) = O\big(h^{-1}\log(h^{-1})\big), \\
\Delta_2 &= O\Big(\max_{x\ge 0}\Big|\frac{\partial}{\partial h} f(x,h)\Big|\Big) = O\Big(\frac{\partial}{\partial h} f(0,h)\Big) = O\big(h^{-1}\log(h^{-1})\big).
\end{aligned}
$$

Combining the above calculations, we again arrive at (86):

$$Z'(h) = F'(h) + O\biggl(\frac{\log(h^{-1})}{h}\biggr) = -\frac{\log^2(h^{-1})}{2\mu h^2} + O\biggl(\frac{\log(h^{-1})}{h^2}\biggr).$$

Thus, we have shown that it is allowable to differentiate the series $Z(h) = \sum_{j=1}^{\infty} f(jh, h)$ term-by-term and have found the asymptotic representation for $Z'(h)$. This approach easily extends to the case of derivatives $Z''(h)$, $Z'''(h)$, and $Z^{(4)}$.

The validity of (37) follows immediately from (83) and (85). The proof of Lemma 1 is completed. □

Proof of Lemma 2. Thanks to (84) we have

$$\sum_{j=1}^{\infty} (Z_j''(h))^2 = \sum_{j=1}^{\infty} \biggl(j^2 \frac{\partial^2}{\partial x^2} f(jh,h) + 2j \frac{\partial^2}{\partial x \partial h} f(jh,h) + \frac{\partial^2}{\partial h^2} f(jh,h)\biggr)^2$$
$$\asymp \frac{1}{h^4} \sum_{j=1}^{\infty} (jh)^4 \biggl(\frac{\partial^2}{\partial x^2} f(jh,h)\biggr)^2 \asymp \frac{1}{h^5} \int_0^{\infty} x^4 \biggl(\frac{\partial^2}{\partial x^2} f(x,h)\biggr)^2 dx,$$

where $\frac{\partial^2}{\partial x^2} f(x,h) = \frac{\mu^2 G(h) e^{-\mu x}}{(1 + G(h)e^{-\mu x})^2}$. Then it is easily seen that

$$\int_0^{\infty} x^4 \left(\frac{\partial^2}{\partial x^2} f(x,h)\right)^2 = O(\log^4(h^{-1})),$$

and hence

$$\sum_{j=1}^{\infty} (Z_j''(h))^2 = O\bigl(h^{-5} \log^4(h^{-1})\bigr).$$

From this, taking into account (39) and (41) as $h \to 0$

$$\begin{aligned} L_h^{(4)} &= \frac{Z^{(4)}(h) + 3\sum_{j=1}^{\infty}(Z_j''(h))^2}{(Z''(h))^2} \\ &= \frac{O\bigl(h^{-5}\log^2(h^{-1})\bigr) + O\bigl(h^{-5}\log^4(h^{-1})\bigr)}{h^{-6}\log^4(h^{-1})} = O(h) \to 0. \end{aligned} \tag{87}$$

The proof is completed. □

Proof of Lemma 3. The proof follows the method of [9, Sec. 7]. Let ξ be a normal $\mathcal{N}(0, \zeta^2)$ random variable independent of $\tau_h = (S - H)/s_h$ and such that

$$\zeta = o(\rho/\sqrt{\log(\rho^{-1})}), \qquad \rho = o((hs_h)^{-1}), \quad hs_h \to \infty. \tag{88}$$

Put

$$\tau_{h,1} = \tau_h + \xi.$$

It follows from (88) that for any $b > 0$,

$$\mathbf{P}(|\xi| > b\rho) = o(\rho).$$

Therefore the proof will be complete if we show that (54) holds true with $\tau_{h,1}$ in place of τ_h. Take $c \in (0, 1/2)$ and let $g^-(t)$ be a function in $C^1(\mathbf{R})$ with support on $(-1/2, 1/2)$ taking its values in the interval $[0,1]$ such that $g^-(t) = 1$ for $|t| < c$. Set

$$g^+(t) = g^-(2ct), \qquad g_\rho^\pm(t) = g^\pm\bigl((t - a - \rho/2)/\rho\bigr).$$

It is sufficient to check (see the proof of Lemma 5.2 in [8] for details) that, for any $c \in (0, 1/2)$,

$$\mathbf{E}_h g_\rho^\pm(\tau_{h,1}) = \mathbf{E} g_\rho^\pm(\tau) + o(\rho), \tag{89}$$

where $\tau \sim \mathcal{N}(0,1)$. The proofs are similar for $g^-(t)$ and $g^+(t)$. Consider the case $g = g^-$.

Let ϕ be the Fourier transform of g. Then the Fourier transform of g_ρ is equal to

$$\phi_\rho(u) = \rho e^{i(a+\rho/2)u}\phi(\rho u).$$

Further, denote by $f_h(u)$, $f_{h,1}$, and $f(u)$ the characteristic functions of τ_h, $\tau_{h,1}$, and τ, respectively. Then $f(u) = e^{-u^2/2}$ and $f_{h,1}(u) = f_h(u)e^{-u^2\zeta^2/2}$. Since $\phi_\rho \in L_1(\mathbf{R})$, relation (89) is equivalent to $\int_{\mathbb{R}} \phi_\rho(u)\bar f_{h,1}(u)\,du = \int_{\mathbb{R}} \phi_\rho(u)\bar f(u)\,du + o(\rho)$. The latter equality follows from the relation

$$\int_{\mathbb{R}} |\phi(\rho u)|\,|(f_h(u)e^{-u^2\zeta^2/2} - f(u))|\,du = o(1). \tag{90}$$

Thus, the problem is reduced to proving (90).

Let $L_h^{(k)}$ be the Lyapunov ratio defined by (31). Set

$$N = (4L_h^{(3)})^{-1} \geq (16L_h^{(4)})^{-1/2} = N_1,$$

where by Lemma 2 $N_1 \to \infty$. Due to Lemma 1 of [14, Ch. 5]

$$|f_h(u) - f(u)| \leq 4N^{-1}|u|^3 e^{-u^2/3}, \qquad |u| \leq N. \tag{91}$$

Since $\phi(u)$ is bounded and $\zeta = o(1)$, it follows that

$$\int_{-N_1}^{N_1} |\phi(\rho u)|\,|(f_h(u)e^{-u^2\zeta^2/2} - f(u))|\,du = o(1).$$

Now observe that $\int_{|u|>N_1} f(u)\,du = o(1)$, and for any $M > N_1$,

$$\int_{|u|>N_1} |f_h(u)e^{-u^2\zeta^2/2}|\,du \leq \int_{M>|u|>N_1} |f_h(u)|\,du + \int_{|u|>M} e^{-u^2\zeta^2/2}\,du,$$

where, as $M \geq 2\zeta^{-1}(\log(\zeta^{-1}))^{1/2}$,

$$\int_M^\infty e^{-u^2\zeta^2/2}\,du = \sqrt{2\pi}\zeta^{-1}\Phi(-M\zeta) = O\big(\zeta \log^{1/2}(\zeta^{-1})\big) \to 0, \qquad \zeta \to 0.$$

Taking into account (88) and (91) we see that in order to obtain (90) it suffices to show that there exists $M = M_h > N_1$, $M \gg hs_h \log(hs_h)$ such that

$$\int_{N_1<|u|<M} |f_h(u)|\,du = o(1), \qquad h \to 0. \tag{92}$$

Recall that

$$\tau_h = (S - H)/s_h, \qquad S = S(\boldsymbol{\ell}) = \sum_{j=1}^\infty Y_j(l_j),$$

where the random variables $Y_j(l_j)$ are independent with respect to the product measure $\mathbf{P}_h(\boldsymbol{\ell}) = \prod_{j=1}^\infty \mathbf{P}_{h,j}(l_j)$. Therefore the characteristic function of τ_h satisfies

$$|f_h(u)| = \prod_{j=1}^\infty |f_{h,j}(u/s_h)|,$$

where $f_{h,j}(v)$ is the characteristic function of the random variable $Y_j(l_j)$ with respect to the measure $\mathbf{P}_{h,j}$, that is,

$$f_{h,j}(v) = \sum_{k\in\mathbf{Z}} p_{jk}\exp(ivY_j(k)), \qquad p_{jk} = \mathbf{P}_{h,j}(k).$$

Due to (16) and (17)

$$p_{jk} = \begin{cases} e^{j\mu h}/(G(h)+e^{j\mu h}), & k=0,\\ e^{-|k|\lambda h}/(G(h)+e^{j\mu h}), & k\neq 0.\end{cases}$$

Using the inequality $\log(1+x)\le x$ and setting $v = u/s_h$, where $N_1 < |u| < M$,

$$\begin{aligned}\log|f_h(u)| &= \frac{1}{2}\sum_{j=1}^{\infty}\log|f_{h,j}(v)|^2 \le -\frac{1}{2}\sum_{j=1}^{\infty}(1-|f_{h,j}(v)|^2)\\ &= -\sum_{j=1}^{\infty}\sum_{(k,l)\in\mathbf{Z}^2} p_{jk}p_{jl}\sin^2\big(v(Y_j(l)-Y_j(k))/2\big) \le -R_h(v)S_h, \end{aligned} \tag{93}$$

where, cf. (36),

$$R_h(v) = \sum_{(k,l)\in\mathcal{N}} e^{-h\lambda(k+l)}\sin^2(v\lambda(l-k)/2), \qquad \mathcal{N} = \{(k,l)\in\mathbf{N}^2,\ k<l\},$$

$$S_h = \sum_{j=1}^{\infty}(G(h)+e^{j\mu h})^{-2}.$$

Let us bound the terms $R_h(v)$ and S_h from below. Recalling (43) and (87), we take

$$M = hs_h\log^2(hs_h) = O\big(h^{-1/2}\log^3(h^{-1})\big) > N_1 = O(h^{-1/2}).$$

Then the constraint $|u|\le M$ corresponds to the inequality

$$|v| = u/s_h \le M/s_h = h\log^2(hs_h) = V \asymp h\log^2(h^{-1}) = o(1),$$

and the constraint $|u| > N_1$ corresponds to $|v| > (bh)/\log(h^{-1})$ for some $b>0$.

Assume that $|v|\in[bh/\log(h^{-1}), V]$ and for $k\in\mathbf{N}$, set

$$\mathcal{N}_k = \big\{l\in\mathbf{N}\colon 0<\lambda|v||l-k|/2\le 1\big\} = \big\{l\colon k+o(1) < l \le k+2/(\lambda|v|)+o(1)\big\},$$

and $\mathcal{N}_k\neq\emptyset$ for $|v|<1/\lambda$. Recalling that $|v|\le V=o(1)$ we have, with $r=l-k$,

$$\begin{aligned} R_h(v) &\ge bv^2\sum_{k=1}^{\infty}\exp(-h\lambda k)\sum_{l\in\mathcal{N}_k}(l-k+o(1))^2\exp(-h\lambda l)\\ &= bv^2\sum_{k=1}^{\infty}\exp(-2h\lambda k)\sum_{1\le r\le 2/(\lambda|v|)+o(1)}(r^2+o(r))\exp(-h\lambda r)\\ &\asymp v^2D(v)\sum_{k=1}^{\infty}\exp(-2h\lambda k) \asymp v^2D(v)h^{-1}, \end{aligned}$$

where, for $K=(h\lambda)^{-1}\to\infty$,

$$VK \asymp \log^2(h^{-1}) = o(h^{-1}), \qquad VK \ge |v|K \ge b/(\lambda\log(h^{-1})),$$

and

$$D(v) = \sum_{1\le r\le 2/(\lambda|v|)+o(1)} r^2\exp(-h\lambda r) = K^2\sum_{1\le r\le 2/(\lambda|v|)+o(1)}(r/K)^2\exp(-r/K)$$

$$\asymp K^3 \int_{h\lambda}^{c/|v|K+o(h)} e^{-x}x^2\,dx \asymp \begin{cases} K^3 & \text{as } v \in [bh/\log(h^{-1}), h], \\ |v|^{-3} & \text{as } v \in [h, V], \end{cases} \qquad c = 2/\lambda.$$

Thus, for some $b > 0$, $b_1 > 0$, $b_2 > 0$,

$$R_h(v) \geq b_1 \begin{cases} K^3 h^{-1} v^2 \geq b_2 (h\log(h^{-1}))^{-2} & \text{as } |v| \in [bh/\log(h^{-1}), h], \\ (hv)^{-1} \geq b_1 (hV)^{-1} \geq b_2 (h\log(h^{-1}))^{-2} & \text{as } |v| \in [h, V]. \end{cases}$$

Let us now bound the term S_h. Denote $J = (\mu h)^{-1}\log G(h)$ and let J_0 be the integer part of J. Since $G(h) \geq e^{j\mu h}$ for $j \leq J$ and $G(h) \asymp h^{-1}$, it follows that

$$S_h = \sum_{j=1}^{\infty} \frac{1}{(G(h) + e^{j\mu h})^2} \geq \sum_{j=1}^{J_0} \frac{1}{(G(h) + e^{j\mu h})^2} \geq \frac{J_0}{4G^2(h)} \asymp h\log(h^{-1}).$$

Therefore, due to (93) for some $b_3 > 0$,

$$|f_h(u)| \leq \exp(-b_3/(h\log(h^{-1})))$$

and by the choice of M

$$\int_{N_1<|u|<M} |f_h(u)|\,du \leq M\exp(-b_3/(h\log(h^{-1}))) \to 0, \qquad h \to 0.$$

Thus (92) is proved, and Lemma 3 follows. □

APPENDIX V: VERIFICATION OF RELATIONS (48)–(51)

The proof of (48)–(51) is largely based on the arguments that lead to Lemma 1. These arguments are presented in Appendix IV. We start with the equality, cf. (85),

$$\tilde{Z}(\delta) = \tilde{F}(\delta) + O(\tilde{f}(0,\delta)), \qquad \delta \to 0,$$

where $\tilde{f}(x,\delta) = \log(1 + \tilde{G}(\delta)e^{-x^\beta})$, so that $\tilde{f}(0,\delta) \asymp \log(\delta^{-1})$, and in accordance with (82)

$$\tilde{F}(\delta) = \frac{1}{\delta}\int_0^{\infty} \tilde{f}(x,\delta)\,dx = \frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta} + O\bigg(\frac{\log^{1/\beta}(\delta^{-1})}{\delta}\bigg).$$

Next, as in the case of function $Z(h) = \sum_{j=1}^{\infty} f(jh,h)$, the function $\tilde{Z}(\delta) = \sum_{j=1}^{\infty} \tilde{f}(j\delta,\delta)$ may be differentiated term-by-term, and

$$\begin{aligned} \tilde{Z}'(\delta) &= \tilde{F}'(\delta) + O\bigg(\max_{x\geq 0}\Big|\frac{\partial}{\partial\delta}\tilde{f}(x,\delta)\Big|\bigg) \\ &= -\frac{\tilde{F}(\delta)}{\delta} + \frac{1}{\delta}\int_0^{\infty}\frac{\partial}{\partial\delta}\tilde{f}(x,\delta)\,dx + O\bigg(\max_{x\geq 0}\Big|\frac{\partial}{\partial\delta}\tilde{f}(x,\delta)\Big|\bigg) \\ &= T_1 + T_2 + T_3. \end{aligned}$$

Thanks to (82)

$$T_1 = -\frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta^2} + O\bigg(\frac{\log^{1/\beta}(\delta^{-1})}{\delta^2}\bigg).$$

Let $\tilde{J}$ be such that $\tilde{G}(\delta)e^{-\tilde{J}^\beta} = 1$, that is, $\tilde{J} = \log^{1/\beta}\tilde{G}(\delta) \asymp \log^{1/\beta}(\delta^{-1})$. Noting that

$$\frac{\partial}{\partial\delta}\tilde{f}(x,\delta) = \frac{\tilde{G}'(\delta)e^{-x^\beta}}{1 + \tilde{G}(\delta)e^{-x^\beta}}, \qquad \tilde{G}(\delta) \asymp \frac{1}{\delta^{\beta/\alpha}}, \quad \tilde{G}'(\delta) \asymp \frac{1}{\delta^{1+\beta/\alpha}}$$

and using relation (81), the second integral is estimated as follows:

$$T_2 = \frac{\tilde G'(\delta)}{\delta}\int_0^{\tilde J}\frac{e^{-x^\beta}}{1+\tilde G(\delta)e^{-x^\beta}}\,dx + \frac{\tilde G'(\delta)}{\delta}\int_{\tilde J}^{\infty}\frac{e^{-x^\beta}}{1+\tilde G(\delta)e^{-x^\beta}}\,dx$$
$$\le \frac{\tilde J\tilde G'(\delta)}{\delta\tilde G(\delta)} + \frac{\tilde G'(\delta)}{\delta}\int_{\tilde J}^{\infty} e^{-x^\beta}\,dx \sim \frac{\tilde J\tilde G'(\delta)}{\delta\tilde G(\delta)}\left(1+\frac{1}{\beta\tilde J^\beta}\right) \asymp \frac{\log^{1/\beta}(\delta^{-1})}{\delta^2} = o(T_1).$$

Consider the term T_3. We have

$$\max_{x\ge 0}\left|\frac{\partial}{\partial\delta}\tilde f(x,\delta)\right| = \frac{\partial}{\partial\delta}\tilde f(0,\delta) \asymp \tilde G'(\delta)e^{-\tilde J^\beta} = \frac{\tilde G'(\delta)}{\tilde G(\delta)} \asymp \frac{1}{\delta} = o(T_1).$$

Therefore for $k=1$, relation (48) holds true:

$$\tilde Z'(\delta) = -\frac{\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta^2} + O\left(\frac{\log^{1/\beta}(\delta^{-1})}{\delta^2}\right).$$

Now recalling that $\delta=(\mu h)^{1/\beta}$ and substituting $\delta'(h)=\beta^{-1}\mu^{1/\beta}h^{1/\beta-1}$, we get

$$Z'(h) = \tilde Z'(\delta(h))\delta'(h) = -\frac{\beta^{2+1/\beta}\log^{1+1/\beta}(h^{-1})}{(1+\beta)\alpha^{1+1/\beta}(\mu h)^{2/\beta}}\,\frac{\mu^{1/\beta}h^{1/\beta-1}}{\beta} + O\left(\frac{\log^{1/\beta}(h^{-1})}{h^{2/\beta}h^{1-1/\beta}}\right)$$
$$= -\frac{\beta^{1+1/\beta}\log^{1+1/\beta}(h^{-1})}{(1+\beta)\alpha^{1+1/\beta}\mu^{1/\beta}h^{1+1/\beta}} + O\left(\frac{\log^{1/\beta}(h^{-1})}{h^{1+1/\beta}}\right).$$

Thus (49) is also verified.

Let us turn to the second derivative $\tilde Z''(\delta)$. It is allowed to differentiate the series $\tilde Z(\delta) = \sum_{j=1}^\infty \tilde f(j\delta,\delta)$ term-by-term twice, and by (84)

$$\tilde Z''(\delta) = \frac{1}{\delta^2}\sum_{j=1}^\infty (j\delta)^2\frac{\partial^2}{\partial x^2}\tilde f(j\delta,\delta) + \frac{2}{\delta}\sum_{j=1}^\infty (j\delta)\frac{\partial^2}{\partial x\partial\delta}\tilde f(j\delta,\delta) + \sum_{j=1}^\infty \frac{\partial^2}{\partial h^2}\tilde f(j\delta,\delta)$$
$$= \frac{1}{\delta^3}\int_0^\infty x^2\frac{\partial^2}{\partial x^2}\tilde f(x,\delta)\,dx + \frac{2}{\delta^2}\int_0^\infty x\frac{\partial^2}{\partial x\partial\delta}\tilde f(x,\delta)\,dx$$
$$+ \frac{1}{\delta}\int_0^\infty \frac{\partial^2}{\partial\delta^2}\tilde f(x,\delta)\,dx + O(\delta^{-2}).$$

Integrating by parts yields

$$\tilde Z''(\delta) = \frac{2}{\delta^3}\int_0^\infty \tilde f(x,\delta)\,dx - \frac{2}{\delta^2}\int_0^\infty \frac{\partial}{\partial\delta}\tilde f(x,\delta)\,dx + \frac{1}{\delta}\int_0^\infty\frac{\partial^2}{\partial\delta^2}\tilde f(x,\delta)\,dx + O(\delta^{-2})$$
$$= \frac{2\tilde F(\delta)}{\delta^2} + O\left(\frac{\log^{1/\beta}(\delta^{-1})}{\delta^3}\right).$$

Recalling (82) we arrive at (48) with $k=2$:

$$\tilde Z''(\delta) = \frac{2\beta^{2+1/\beta}\log^{1+1/\beta}(\delta^{-1})}{(1+\beta)\alpha^{1+1/\beta}\delta^3} + O\left(\frac{\log^{1/\beta}(\delta^{-1})}{\delta^3}\right).$$

Returning to the parameter h,

$$Z''(h) = \tilde Z''(\delta(h))(\delta'(h))^2 + \tilde Z'(\delta(h))\delta''(h),$$

where $\delta'(h) = \beta^{-1}\mu^{1/\beta}h^{1/\beta-1}$ and $\delta''(h) = \beta^{-1}(\beta^{-1}-1)\mu^{1/\beta}h^{1/\beta-2}$. Therefore (50) also holds true:

$$Z''(h) = \frac{\beta^{1/\beta}\log^{1+1/\beta}(h^{-1})}{\alpha^{1+1/\beta}\mu^{1/\beta}h^{2+1/\beta}} + O\left(\frac{\log^{1/\beta}(h^{-1})}{h^{2+1/\beta}}\right).$$

The proof of relations in (51) requires elementary calculations similar to those for $Z'(h)$ and $Z''(h)$. The proof is purely technical, and we omit it.

ACKNOWLEDGMENTS

The research of the first author has been partially supported by the Russian Foundation for Basic Research grant 08-01-00692 and by the grant NSh–638.2008.1. The research of the second author has been supported by an NSERC grant.

REFERENCES

1. E. N. Belitser and B. Y. Levit, "On Minimax Filtering over Ellipsoids", Math. Methods Statist. **4**, 259–273 (1995).
2. M. V. Fedoryuk, *Asymptotics: Integrals and Series* (Nauka, Moscow, 1987) [in Russian].
3. I. S. Gradshtein and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. (Academic Press, New York, 2000).
4. I. A. Ibragimov and R. Z. Khasminskii, "Some Estimation Problems on Infinite Dimensional Gaussian White Noise", in *Festschrift for Lucien Le Cam. Research papers in Probability and Statistics* (Springer, New York, 1997), pp. 275–296.
5. Yu. I. Ingster, "Asymptotically Minimax Testing of Nonparametric Hypotheses", in *Proc. of the Fourth Vilnius Conf., Vilnius, 1987*, Vol. 1, Probab. Theory and Math. Statist., 553–573 (1987).
6. Yu. I. Ingster and I. A. Suslina, *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*, in *Lecture Notes in Statistics* (Springer, New York, 2002), Vol. 169.
7. Yu. I. Ingster and I. A. Suslina, "On Estimation and Detection of Smooth Functions of Many Variables", Math. Methods Statist. **14**, 299–331 (2005).
8. Yu. I. Ingster and I. A. Suslina, "Estimation and Detection of High-Variable Functions from Sloan–Woźniakowski Space", Math. Methods Statist. **16**, 318–353 (2007).
9. Yu. I. Ingster and I. A. Suslina, "On Estimation and Detection of a Function from Tensor Product Spaces", Zapiski Nauchn. Sem. POMI **351**, 180–218 (2007) [J. Math. Sci. **152**, 897–920 (2008)].
10. F. Y. Kuo and J. H. Sloan, "Lifting the Curse of Dimensionality", Notices of AMS **52**, 1320–1329 (2005).
11. B. Levit and N. Stepanova, "Efficient Estimation of Multivatiate Analytic Functions in Cube-Like Domains", Math. Methods Statist. **13**, 253–281 (2004).
12. Y. Lin, "Tensor product space ANOVA models", Ann. Statist. **28**, 734–755 (2000).
13. M. Nussbaum, "Optimal Filtering of a Function of Many Variables in Gaussian White Noise", Problems Inform. Transmission **19**, 23–29 (1983).
14. V. V. Petrov, *Limit Theorems of Probability Theory* (Oxford Univ. Press, New York, 1995).
15. A. V. Skorohod, *Integration in Hilbert Spaces* (Springer, Berlin–New York, 1974).
16. I. H. Sloan and H. Woźniakowski, "When are Quazi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?", J. Complexity **14**, 1–33 (1998).
17. C. Stone, "Additive Regression and Other Nonparametric Models", Ann. Statist. **13**, 689–705 (1985).
18. H. Woźniakowski, "Tractability of Multivariate Problems for Weighted Spaces of Functions", in *Appoximation and Probability, Banach Center Publications* (Institute of Mathematics, Polish Academy of Science, Warszawa, 2006), Vol. 72.