THEORY PROBAB. APPL.
Vol. 46, No. 4

Translated from Russian Journal

ON THE MINIMAX ESTIMATION PROBLEM OF A FRACTIONAL DERIVATIVE*

G. K. GOLUBEV† AND F. N. ENIKEEVA‡

(*Translated by F. N. Enikeeva*)

Abstract. We consider the problem of minimax estimating the fractional derivative of the order $-\frac{1}{2}$ of an unknown function in the Gaussian white noise model. This problem is closely related to the well-known Wicksell problem. In this paper the second-order minimax approach is developed.

Key words. fractional derivative, second-order minimax risk, Wicksell problem

PII. S0040585X97979251

1. Introduction. In this paper the fractional derivative $f^{(\alpha)}$ of order $\alpha = -\frac{1}{2}$ of an unknown function $f(t)$ is estimated from observations in the Gaussian white noise. The observations are defined as follows:

$$dx(t) = f(t)\,dt + \varepsilon\,dw(t), \qquad t \in [0,1], \quad x(0) = 0, \tag{1.1}$$

where $w(t)$ is the standard Wiener process and ε is a small parameter. The problem is to estimate the fractional derivative $f^{(-1/2)}(t)$, assuming that $f(t)$ belongs to a known class of smooth functions. In fact, we will be concerned with two problems: estimating $f^{(-1/2)}(t)$ at a fixed point t_0, and recovering the derivative $f^{(-1/2)}(t)$ on the unit interval $[0,1]$.

In order to simplify the technical details we suppose that $f(t)$ is a periodic zero-mean function. According to [12] we can define the fractional derivative of the order α as

$$f^{(\alpha)}(t) = \sum_{k=-\infty}^{\infty} \langle f, \varphi_k\rangle\, \varphi_k(t)\,(2\pi i k)^{\alpha},$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbf{L}_2(0,1)$ and $\varphi_k(t) = \exp(2\pi i k t)$ is a trigonometric basis. Let $\theta_k = \langle f, \varphi_k\rangle$ be Fourier coefficients of the function f. Then the problem of estimating the derivative of order $-\frac{1}{2}$ at the point t_0 becomes similar to estimating the linear functional

$$L(\theta) = \sum_{k=-\infty}^{\infty} \frac{\exp(2\pi i k t_0)}{\sqrt{2\pi i k}}\,\theta_k.$$

Likewise, the problem of recovering the fractional derivative of the same order on the unit interval can be reduced to the problem of estimating the vector $(\theta_1/\sqrt{1}, \theta_2/\sqrt{2}, \dots)^T$.

*Received by the editors September 27, 1999.
http://www.siam.org/journals/tvp/46-4/97925.html
†Institute for Information Transmission Problems, Bolshoy Karetnyi, 19, 101447 Moscow, Russia (glbv@iitp.ru).
‡Faculty of Mechanics and Mathematics, Department of Probability Theory, MGU, Vorobjevy Gory, 119899 Moscow, Russia (farida@shade.msu.ru).

Sometimes the estimation problem can be simplified by conversion from the observations in the time domain to the observations in the Fourier coefficients space. Since $\varphi_k(t)$ is a complete orthonormal system in $\mathbf{L}_2(0,1)$, it is easy to see that observations (1.1) are equal to

$$X_k = \theta_k + \varepsilon \xi_k, \qquad k = 0, \pm 1, \pm 2, \dots, \tag{1.2}$$

where ξ_k are independent identically distributed (i.i.d.) complex-valued Gaussian random variables with the parameters $(0,1)$:

$$\xi_k = \int_0^1 \varphi_k(t)\,dw(t), \quad \theta_k = \int_0^1 \varphi_k(t) f(t)\,dt, \quad X_k = \int_0^1 \varphi_k(t)\,dx(t).$$

Prior information about unknown parameters is very important for any statistical problem. In this paper we assume that the unknown function $f(t)$ belongs to an ellipsoid in $\mathbf{L}_2(0,1)$:

$$\theta \in \Theta = \left\{ \theta\colon \sum_{k=1}^{\infty} a_k^2 |\theta_k|^2 \leqq 1 \right\}. \tag{1.3}$$

In particular, if the underlying function belongs to the Sobolev class

$$W_2^\beta = \left\{ f\colon \int_0^1 \left[f^{(\beta)}(t) \right]^2 dt \leqq P \right\},$$

then the axes of the ellipsoid are defined by $a_k^2 = (2\pi k)^{2\beta}/P$.

At first glance the problem of estimating the fractional derivative of order $-\frac{1}{2}$ seems to be of a rather special interest. Actually, this problem is closely related to the well-known Wicksell problem [10], which can be formulated as follows. Suppose that a number of spheres are embedded in an opaque medium. Let the sphere radii be i.i.d. with an unknown distribution function $F(x)$. The item of interest is $F(x)$. Since the medium is opaque, we cannot observe a sample of sphere radii directly. We can only intersect the medium and observe a cross-section showing the circular section of some spheres. Define the radii of the circles in the cross-section by $Y_1, \dots, Y_n$. The problem is to estimate the distribution function $F(x)$ from these observations. It can be easily seen that the variables Y_i are i.i.d.; denote their distribution function by $G(y)$. The relations between F and G are known:

$$\begin{aligned} 1 - G(y) &= \int_y^\infty \sqrt{x-y}\,dF(x) \left(\int_0^\infty \sqrt{x}\,dF(x) \right)^{-1}, \\ 1 - F(x) &= \int_y^\infty \frac{dG(y)}{\sqrt{x-y}} \left(\int_0^\infty \frac{dG(y)}{\sqrt{y}} \right)^{-1}. \end{aligned} \tag{1.4}$$

For an elementary inference of these formulas we refer the reader to [3]. In fact, these formulas express the unknown distribution function $F(x)$ in terms of the derivative of order $\frac{1}{2}$ of the distribution function $G(y)$. Here the integrals present another form of definition of fractional derivatives. Thus the problem is reduced to estimating the functions $G^{(1/2)}(0)$ and $G^{(1/2)}(y)$ from the observations $Y_1, \dots, Y_n$ with unknown density $g(y)$. Obviously, these functions are the derivatives of order $-\frac{1}{2}$ of the distribution density $g(y) = G'(y)$. Undoubtedly, the Wicksell problem does not coincide with the problem of estimation in Gaussian white noise. However, they are closely related. It is well known that the corresponding statistical experiments are asymptotically equivalent in the Le Cam sense (see [7]). We intentionally avoid unimportant details and consider the primitive statistical problem in order to clarify how to construct asymptotically minimax estimates of the second order.

Note that the results concerning asymptotically minimax estimates (as $n \to \infty$) of the first order in the Wicksell problem were achieved rather recently [3] even though the optimal rates of convergence are well known (see [6], [4], [2]).

The aim of this paper is to construct asymptotically minimax estimates of the second order in the model of Gaussian white noise. Transference of the results to the Wicksell problem is not trivial but rather a question of technique. It is natural to apply the second-order minimax theory to this problem. The point is that there are many asymptotically minimax estimates of the first order, and it is impossible to select the best estimator under the first-order theory framework. On the other hand, an asymptotically minimax estimate of the second order is to some extent unique.

2. Statement of the problem and main results. Next we will consider a more general setting of the problem than that in the model (1.2), (1.3). Suppose we observe real random variables

$$X_k = \theta_k + \varepsilon\xi_k, \qquad k = 0, 1, \dots , \tag{2.1}$$

where ξ_k are Gaussian independent random variables with parameters $(0, 1)$. It is also assumed that an unknown vector $\theta = (\theta_1, \theta_2, \dots)^T$ belongs to the ellipsoid

$$\Theta = \left\{\theta\colon \sum_{k=1}^{\infty} a_k^2\theta_k^2 \leqq 1\right\}, \tag{2.2}$$

where the parameters a_k^2 are known. There are two problems related to this statistical model.

The first problem is to find the minimax estimate of the infinite-dimensional vector $v(\theta) = (\theta_1 s_1, \theta_2 s_2, \dots)^T$, where the sequence s_k satisfies the condition

$$\lim_{k\to\infty} s_k^2 k = 1. \tag{2.3}$$

From now on we assume that condition (2.3) holds. By $\widehat{v}(X) = (\widehat{v}_1, \widehat{v}_2, \dots)^T$ denote an estimate of the vector $v(\theta)$.

The mean square risk of the estimate $\widehat{v}$ is defined as usual:

$$R^\varepsilon(\widehat{v}, \Theta) = \sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \left\|v(\theta) - \widehat{v}(X)\right\|^2 = \sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \sum_{k=1}^{\infty} \left|v_k(\theta) - \widehat{v}_k\right|^2,$$

where $\mathbf{E}_\theta^\varepsilon$ is the expectation with respect to the measure generated by observations (2.1). The minimax risk is defined by $r^\varepsilon(\Theta) = \inf_{\widehat{v}} R^\varepsilon(\widehat{v}, \Theta)$ over all the estimates of the vector $v(\theta)$. We will show that under some conditions the linear estimates are asymptotically minimax of the second order. More precisely,

$$r^\varepsilon(\Theta) = \inf_{\widehat{v}\in\mathcal{L}} R^\varepsilon(\widehat{v}, \Theta) + o(\varepsilon^2); \tag{2.4}$$

here $\mathcal{L}$ is a class of all linear estimates. We can formulate the following result.

THEOREM 1. *Let the sequence* $|a_k||s_k|$ *be nondecreasing and*

$$\lim_{\varepsilon\to 0} \log^3 \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} a_k^2 \big(|s_k| - \mu|a_k|\big)_+^2 \left[\sum_{k=1}^{\infty} |a_k| \big(|s_k| - \mu|a_k|\big)_+\right]^{-2} = 0, \tag{2.5}$$

$$\lim_{\varepsilon\to 0} \frac{\max_m |a_m|(|s_m| - \mu|a_m|)_+ \sum_{k=1}^{\infty} |a_k|\,(|s_k| - \mu|a_k|)_+}{\sum_{k=1}^{\infty} a_k^2(|s_k| - \mu|a_k|)_+^2} < \infty, \tag{2.6}$$

where μ *is a root of*

$$\varepsilon^2 \sum_{k=1}^{\infty} a_k^2 \left(\frac{|s_k|}{\mu|a_k|} - 1\right)_+ = 1. \tag{2.7}$$

Then the linear estimate

$$\widetilde{v}_k = \left(1 - \frac{\mu|a_k|}{|s_k|}\right)_+ s_k X_k$$

is asymptotically minimax of the second order with the minimax risk

$$r^\varepsilon(\Theta) = R^\varepsilon(\widetilde{v}, \Theta) + o(\varepsilon^2) = \varepsilon^2 \sum_{k=1}^{\infty} |s_k|\big(|s_k| - \mu|a_k|\big)_+ + o(\varepsilon^2).$$

In particular, if $a_k = (\pi k)^\beta/\sqrt{P}$, $\beta > \frac{1}{2}$, and $s_k = k^{-1/2}$, then the asymptotic expansion of the minimax risk ($\varepsilon \to 0$) is

$$r^\varepsilon(\Theta) = \frac{\varepsilon^2}{2\beta+1} \log \frac{(2\beta+1)P}{\pi^{2\beta}\varepsilon^2} + \varepsilon^2 \left(\gamma - \frac{2}{2\beta+1}\right) + o(\varepsilon^2);$$

here and throughout, γ is the Euler constant.

The second problem regards the estimation of the linear functional

$$L(\theta) = \sum_{k=1}^{\infty} \theta_k s_k.$$

Let $\widehat{L}(X)$ be an estimate of the functional L. Its mean square risk is defined by

$$R_0^\varepsilon(\widehat{L}, \Theta) = \sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \big[L(\theta) - \widehat{L}(X)\big]^2. \tag{2.8}$$

The minimax risk, respectively, is

$$r_0^\varepsilon(\Theta) = \inf_{\widehat{L}} R_0^\varepsilon(\widehat{L}, \Theta), \tag{2.9}$$

where the infimum is over all the estimates of functional $L(\theta)$.

The following theorem yields information about the upper and lower bounds of minimax risk.

THEOREM 2. *The following inequalities for the minimax risk hold:*

$$\varepsilon^2 \sum_{k=1}^{\infty} s_k^2 (1 + \pi^2 \varepsilon^2 a_k^2)^{-1} \leqq r_0^\varepsilon(\Theta) \leqq \varepsilon^2 \sum_{k=1}^{\infty} s_k^2 (1 + \varepsilon^2 a_k^2)^{-1}.$$

We see that there is a gap between the upper and lower bounds. It is easy to verify that its size equals $(\varepsilon^2 \log \pi)/\beta$ over the Sobolev ball. Existence of the gap is caused by the fact that the linear estimates are not minimax of the second order. Unfortunately, it is rather difficult to find the explicit minimax estimates in this problem. That is why we reduce our statistical problem to the simpler one; it is the recovering functional problem

$$L_\delta(f) = \int_\delta^\infty \frac{f(t)}{\sqrt{t}}\,dt \tag{2.10}$$

from the observations in Gaussian white noise

$$dX(t) = f(t)\,dt + dw(t), \qquad t \in [0,\infty). \tag{2.11}$$

A priori information about $f(\cdot)$ is

$$f \in \mathcal{F} = \left\{ f \in \mathbf{L}_2(0,\infty)\colon\ \int_0^\infty t^{2\beta} f^2(t)\,dt \leqq 1 \right\}. \tag{2.12}$$

Asymptotic behavior of the minimax risk in the initial problem can be described with accuracy $o(\varepsilon^2)$ in terms of the problem (2.11), (2.12) of estimating the functional $L_\delta(f)$. By

$$\rho = \lim_{\delta\to 0} \left\{ \inf_{\widehat{L}_\delta} \sup_{f\in\mathcal{F}} \mathbf{E}_\theta \left[L_\delta(f) - \widehat{L}_\delta \right]^2 + \log \delta \right\}$$

denote the limit minimax risk in the problem of estimating $L_\delta(f)$. We have the following result.

THEOREM 3. *Let $s_k = k^{-1/2}$ and $a_k^2 = (\pi k)^{2\beta} P^{-1}\,(1+o(1))$ as $k\to\infty$. Then*

$$r_0^\varepsilon(\Theta) = \frac{\varepsilon^2}{2\beta}\,\log \frac{P}{\varepsilon^2} + \varepsilon^2(\gamma + \rho - \log \pi) + o(\varepsilon^2), \tag{2.13}$$

as $\varepsilon \to 0$.

3. Estimation of the derivative on an interval.

3.1. An upper bound. First, to prove Theorem 1 we obtain a trivial upper bound of the minimax risk

$$r^\varepsilon(\Theta) \leqq \inf_{\widehat{v}\in\mathcal{L}} R^\varepsilon(\widehat{v},\Theta).$$

Recall that here $\mathcal{L}$ is the class of all linear estimates. To calculate the minimax risk over the class of linear estimates we will use the well-known saddle point theorem [14].

LEMMA 1. *Let μ be a root of* (2.7). *Then the estimate*

$$v_k^* = \left(1 - \mu\,\frac{|a_k|}{|s_k|}\right)_+ s_k X_k$$

is minimax in the class of linear estimates with the minimax risk

$$\inf_{\widehat{v}\in\mathcal{L}} R^\varepsilon(\widehat{v},\Theta) = \varepsilon^2 \sum_{k=1}^\infty s_k^2 \left(1 - \mu\,\frac{|a_k|}{|s_k|}\right)_+ . \tag{3.1}$$

Proof. A mean square error of the linear estimate $\widehat{v}_k = h_k s_k X_k$,

$$\mathbf{E}_\theta \sum_{k=1}^{\infty} |v_k - \widehat{v}_k|^2 = \sum_{k=1}^{\infty} s_k^2 (1-h_k)^2 \theta_k^2 + \varepsilon^2 \sum_{k=1}^{\infty} h_k^2 s_k^2 = F^\varepsilon(h,\theta),$$

is convex with respect to h and linear with respect to θ_k^2. Hence it has a saddle point on the set $l_2(1,\infty) \times \Theta$. We omit simple arithmetic (see, e.g., [14]) showing that the components of the saddle point are as follows:

$$h_k^* = \left(1 - \mu \frac{|a_k|}{|s_k|}\right)_+, \qquad \theta_k^{*2} = \frac{\varepsilon^2 h_k^*}{1-h_k^*} = \varepsilon^2 \left(\frac{|s_k|}{\mu |a_k|} - 1\right)_+. \tag{3.2}$$

Here μ is the root of $\sum_{k=1}^{\infty} a_k^2 \theta_k^{*2} = 1$. Finally, noting that

$$\inf_{\widehat{v} \in \mathcal{L}} R^\varepsilon(\widehat{v}, \Theta) = \inf_h \sup_{\theta \in \Theta} F^\varepsilon(h,\theta) = F^\varepsilon(h^*, \theta^*) = \varepsilon^2 \sum_{k=1}^{\infty} s_k^2 h_k^*$$

we complete the proof.

3.2. A lower bound. We now establish the lower bound of the minimax risk. Our construction is adapted from [14]. Choose an a priori distribution of the parameters θ_k such that the variance of θ_k is close to the saddle point (3.2) and the vector θ lies near the surface of ellipsoid (2.2). More precisely, suppose that θ_k are normally distributed with parameters $(0, \sigma_k^2)$, where

$$\sigma_k^2 = (1-\delta)\, \varepsilon^2 \left(\frac{|s_k|}{\mu |a_k|} - 1\right)_+, \qquad 0 < \delta < 1, \tag{3.3}$$

and μ is the solution of (2.7). Notice that whenever $\delta = 0$ the variance of θ_k is equal to the saddle point θ_k^{*2} from (3.2), which determines the minimax linear estimate. First, we shall show that for small $\delta > 0$ the vector θ does not belong to the ellipsoid with probability tending to zero, as $\varepsilon \to 0$.

LEMMA 2. *Let the sequence* $|a_k s_k|$ *be nondecreasing and condition* (2.6) *hold. Then for any* $\delta \in (0, \delta_0)$

$$\mathbf{P}\left\{\theta \notin \Theta\right\} \leqq \exp\left(-\frac{\delta^2}{4 w_\varepsilon}\right), \qquad \textit{where} \quad w_\varepsilon = \sum_{k=1}^{\infty} a_k^4 \sigma_k^4. \tag{3.4}$$

Proof. Note that

$$\mathbf{P}\{\theta \notin \Theta\} = \mathbf{P}\left\{\sum_{k=1}^{\infty} a_k^2 \theta_k^2 > 1\right\} = \mathbf{P}\left\{\sum_{k=1}^{\infty} a_k^2 (\theta_k^2 - \sigma_k^2) > \delta\right\}.$$

According to Markov's inequality we have

$$\begin{aligned}\mathbf{P}\{\theta \notin \Theta\} &\leqq e^{-\lambda\delta}\, \mathbf{E} \exp\left[\lambda \sum_{k=1}^{\infty} a_k^2 (\theta_k^2 - \sigma_k^2)\right] \\ &= e^{-\lambda\delta} \exp\left[-\lambda \sum_{k=1}^{\infty} a_k^2 \sigma_k^2 - \frac{1}{2} \sum_{k=1}^{\infty} \log(1 - 2\lambda a_k^2 \sigma_k^2)\right] \end{aligned} \tag{3.5}$$

for all λ such that $2\lambda \sup_k a_k^2 \sigma_k^2 < 1$.

Choose $\lambda = \delta/(2\sum_{k=1}^{\infty} a_k^4\sigma_k^4)$ and check the inequality $1 - 2\lambda a_k^2\sigma_k^2 > 0$ for sufficiently small δ. To do this we have to show that

$$\sup_k a_k^2\sigma_k^2 \left(\sum_{k=1}^{\infty} a_k^4\sigma_k^4\right)^{-1} < \infty. \tag{3.6}$$

Combining (3.3) and (2.7) we obtain

$$\frac{\max_k a_k^2\sigma_k^2}{\sum_{k=1}^{\infty} a_k^4\sigma_k^4} = \frac{\max_m |a_m|\,(|s_m| - \mu|a_m|)_+ \cdot \sum_{k=1}^{\infty} |a_k|(|s_k| - \mu|a_k|)_+}{(1-\delta)\sum_{k=1}^{\infty} |a_k|^2(|s_k| - \mu|a_k|)_+^2}.$$

By (2.6), the right-hand side in this equality is bounded, and consequently (3.6) holds. Applying (3.5) and Taylor's formula we conclude that

$$\mathbf{P}\{\theta \notin \Theta\} \leqq e^{-\lambda\delta}\exp\left(\lambda^2\sum_{k=1}^{\infty} a_k^4\sigma_k^4\right) = \exp\left\{-\frac{\delta^2}{4\sum_{k=1}^{\infty} a_k^4\sigma_k^4}\right\},$$

and this is precisely the assertion of the lemma.

LEMMA 3. *Let conditions* (2.5) *and* (2.6) *hold. Then*

$$r^\varepsilon(\Theta) \geqq \varepsilon^2 \sum_{k=1}^{\infty} s_k^2\left(1 - \mu\frac{|a_k|}{|s_k|}\right)_+ + o(\varepsilon^2) \qquad as \quad \varepsilon \to 0, \tag{3.7}$$

where μ is the root of (2.7).

Proof. Let $\hat\theta_k$ be an estimate of the parameter θ_k. By the triangle inequality, we can obtain the following lower bound of the minimax risk:

$$r^\varepsilon(\Theta) = \inf_{\hat\theta}\sup_{\theta\in\Theta}\mathbf{E}_\theta\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2 \geqq \inf_{\hat\theta\in\Theta}\sup_{\theta\in\Theta}\mathbf{E}_\theta\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2. \tag{3.8}$$

Since

$$\sup_{\theta\in\Theta}\mathbf{E}_\theta\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2 \geqq \mathbf{E}\,\mathbf{E}_\theta\mathbf{1}\{\theta\in\Theta\}\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2,$$

we can continue (3.8) in the following way:

$$\begin{aligned} r^\varepsilon(\Theta) &\geqq \inf_{\hat\theta\in\Theta}\mathbf{E}\,\mathbf{E}_\theta\mathbf{1}\{\theta\in\Theta\}\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2 \\ &\geqq \inf_{\hat\theta\in\Theta}\mathbf{E}\,\mathbf{E}_\theta\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2 - \sup_{\hat\theta\in\Theta}\mathbf{E}\,\mathbf{E}_\theta\mathbf{1}\{\theta\notin\Theta\}\sum_{k=1}^{\infty} s_k^2(\theta_k - \hat\theta_k)^2. \end{aligned} \tag{3.9}$$

Recall that variables θ_k are independent normally distributed with parameters $\mathcal{N}(0,\sigma_k^2)$, where σ_k^2 are determined in (3.3). Thus $\theta_k = \sqrt{1-\delta}\,\theta_k^*\xi_k$, where ξ_k are $\mathcal{N}(0,1)$-i.i.d. and θ_k^* is the second component of the saddle point (3.2). Combining these with Lemma 1 we see that

$$\inf_{\hat\theta\in\Theta} \mathbf{E}\,\mathbf{E}_\theta \sum_{k=1}^\infty s_k^2(\theta_k-\hat\theta_k)^2 \geqq \inf_{\hat\theta} \mathbf{E}\,\mathbf{E}_\theta \sum_{k=1}^\infty s_k^2\Big(\sqrt{1-\delta}\,\theta_k^*\xi_k-\hat\theta_k\Big)^2$$

$$\geqq (1-\delta)\inf_{\hat\theta} \mathbf{E}\,\mathbf{E}_\theta \sum_{k=1}^\infty s_k^2(\theta_k^*\xi_k-\hat\theta_k)^2$$

$$= (1-\delta)\inf_{h_k} \mathbf{E}\,\mathbf{E}_\theta \sum_{k=1}^\infty s_k^2(\theta_k^*\xi_k-h_kX_k)^2$$

$$= (1-\delta)\,\varepsilon^2 \sum_{k=1}^\infty s_k^2\left(1-\mu\,\frac{|a_k|}{|s_k|}\right)_+ . \tag{3.10}$$

Now we obtain the lower bound for the last term in the right-hand side of (3.9). Since θ_k are Gaussian random variables, it follows that

$$\mathbf{E}\left(\sum_{k=1}^\infty s_k^2\theta_k^2\right)^2 = 3\sum_{k=1}^\infty s_k^4\sigma_k^4 + 2\sum_{k\neq l} s_k^2 s_l^2\sigma_k^2\sigma_l^2 \leqq 3\left(\sum_{k=1}^\infty s_k^2\sigma_k^2\right)^2 \leqq C.$$

From this and the Cauchy–Schwarz inequality we have

$$\sup_{\hat\theta\in\Theta} \mathbf{E}\,\mathbf{E}_\theta \mathbf{1}\{\theta\notin\Theta\}\sum_{k=1}^\infty s_k^2(\theta_k-\hat\theta_k)^2$$

$$\leqq \big[\mathbf{P}\{\theta\notin\Theta\}\big]^{1/2}\sup_{\hat\theta\in\Theta}\left\{\mathbf{E}\mathbf{E}_\theta\left[\sum_{k=1}^\infty s_k^2(\theta_k-\hat\theta_k)^2\right]^2\right\}^{1/2} \leqq C\,\big[\mathbf{P}\{\theta\notin\Theta\}\big]^{1/2}.$$

Combining these, (3.9), (3.10), and Lemma 2 we conclude that for a constant C

$$r^\varepsilon(\Theta) \geqq (1-\delta)\,\varepsilon^2\sum_{k=1}^\infty s_k^2\Big(1-\mu\sqrt{k}\,|a_k|\Big)_+ - C\exp\left(-\frac{\delta^2}{Cw_\varepsilon}\right). \tag{3.11}$$

Let us “improve” the lower bound, maximizing with respect to δ the right-hand side of this inequality. For abbreviation, we denote

$$\rho^\varepsilon = \varepsilon^2\sum_{k=1}^\infty s_k^2\Big(1-\mu\sqrt{k}\,|a_k|\Big)_+ .$$

Choose $\delta=\sqrt{-C_0 w_\varepsilon\log(\rho^\varepsilon w_\varepsilon)}$, where C_0 is sufficiently large. It is easy to see that

$$\min_\delta\left\{\delta\rho^\varepsilon + C\exp\left(-\frac{\delta^2}{Cw_\varepsilon}\right)\right\} \leqq \rho^\varepsilon\sqrt{w_\varepsilon}\,\log^{1/2}\frac{1}{\rho^\varepsilon w_\varepsilon}$$

$$\leqq C\rho^\varepsilon\sqrt{w_\varepsilon}\,\log^{1/2}\frac{1}{\rho^\varepsilon\sqrt{w_\varepsilon}}. \tag{3.12}$$

In addition, by (2.7)

$$w_\varepsilon = \sum_{k=1}^\infty a_k^4\sigma_k^4 = \frac{\varepsilon^2}{\mu^2}\sum_{k=1}^\infty a_k^2 s_k^2\left(1-\mu\,\frac{|a_k|}{|s_k|}\right)_+^2$$

$$= \sum_{k=1}^\infty a_k^2 s_k^2\left(1-\mu\,\frac{|a_k|}{|s_k|}\right)_+^2\left[\sum_{k=1}^\infty |a_k||s_k|\,\left(1-\mu\,\frac{|a_k|}{|s_k|}\right)_+\right]^{-2}. \tag{3.13}$$

To continue (3.12), let us define an integer

$$N = \min\left\{k\colon\ 1 - \mu\,\frac{|a_k|}{|s_k|} < 0\right\}. \tag{3.14}$$

Obviously, $\rho^\varepsilon \leqq \varepsilon^2 \log N$. Then (2.7) gives $N < \varepsilon^{-2}$. Hence $\rho^\varepsilon \leqq \varepsilon^2 \log \varepsilon^{-2}$. This, (3.11), (3.12), and (2.5) give inequality (3.7).

The proof of Theorem 1 immediately follows from Lemmas 1 and 3.

3.3. The asymptotic behavior of the minimax risk over the Sobolev class. In this section we will look more closely at the asymptotic behavior of the minimax risk over the Sobolev class with coefficients $a_k = (\pi k)^\beta/\sqrt{P}$, where $\beta > \frac{1}{2}$ and $s_k = k^{-1/2}$.

Let N be defined in (3.12). Then we have the following simple relation for N and μ: $\mu = (1 + o(1))|s_N|/|a_N|$ as $\varepsilon \to 0$. Thus we can rewrite (2.7) for N:

$$\varepsilon^2 \sum_{k=1}^{N} a_k^2 \left(\frac{|s_k a_N|}{|a_k s_N|} - 1\right) = 1.$$

It yields the equation for N,

$$\sum_{k=1}^{N} k^{2\beta}\left[\left(\frac{N}{k}\right)^{\beta+1/2} - 1\right] = \frac{P}{\pi^{2\beta}\varepsilon^2}.$$

An easy computation shows that

$$N = \big(1 + o(1)\big)\left[\frac{P(2\beta+1)}{\pi^{2\beta}\varepsilon^2}\right]^{1/(2\beta+1)} \qquad \text{as} \quad \varepsilon \to 0. \tag{3.15}$$

To check (2.5) and (2.6), note that

$$\begin{aligned}
\sum_{k=1}^{\infty} |a_k s_k| \left(1 - \mu\,\frac{|a_k|}{|s_k|}\right)_+ &\asymp \sum_{k=1}^{N} k^{\beta-1/2}\left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right] \asymp N^{\beta+6/2},\\
\sum_{k=1}^{\infty} a_k^2 s_k^2 \left(1 - \mu\,\frac{|a_k|}{|s_k|}\right)_+^2 &\asymp \sum_{k=1}^{N} k^{2\beta-1}\left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right]^2 \asymp N^{2\beta},\\
\max_k |a_k s_k| \left(1 - \mu\,\frac{|a_k|}{|s_k|}\right)_+ &\asymp \max_k k^{\beta-1/2}\left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right] \asymp N^{\beta-1/2}.
\end{aligned}$$

Let us examine the asymptotic behavior of the minimax risk as $\varepsilon \to 0$. For any $\delta \in (0,1)$ we have

$$\begin{aligned}
r^\varepsilon(\Theta) &= \varepsilon^2 \sum_{k=1}^{N} s_k^2 \left(1 - \mu\,\frac{\pi^\beta k^\beta}{|s_k|\sqrt{P}}\right) = \varepsilon^2 \sum_{k=1}^{N} s_k^2 \left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right] + o(\varepsilon^2)\\
&= \varepsilon^2 \sum_{k=1}^{\delta N} s_k^2 \left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right] + \varepsilon^2 \sum_{k=\delta N+1}^{N} s_k^2 \left[1 - \left(\frac{k}{N}\right)^{\beta+1/2}\right] + o(\varepsilon^2)\\
&\equiv \varepsilon^2\big[r_1(\varepsilon,\delta) + r_2(\varepsilon,\delta)\big] + o(\varepsilon^2).
\end{aligned}$$

Let us estimate $r_1(\varepsilon,\delta)$. Since $k \leqq \delta N$, we have $1-(k/N)^{\beta+1/2} = 1+O(\delta^{\beta+1/2})$ as $\delta \to 0$. Consequently,

$$r_1(\varepsilon,\delta) = \log N + \gamma + \log\delta + O\Big(\delta^{\beta+1/2}\log N\Big) + o(1). \tag{3.16}$$

Let us turn to $r_2(\varepsilon,\delta)$. It is easy to see that

$$\begin{aligned} r_2(\varepsilon,\delta) &= \int_\delta^1 x^{-1}\Big(1-x^{\beta+1/2}\Big)\,dx + O(N^{-1}\delta^{-1}) \\ &= -\log\delta - \frac{2}{2\beta+1}\Big(1-\delta^{\beta+1/2}\Big) + O\big(N^{-1}\delta^{-1}\big). \end{aligned} \tag{3.17}$$

Choose $\delta = (\log N)^{-1-q}$, where $q > 0$. Therefore, combining (3.16) with (3.17) yields

$$r_1(\varepsilon,\delta) + r_2(\varepsilon,\delta) = \log N + \gamma - \frac{2}{2\beta+1} + o(1).$$

Hence by (3.15) we have the following expansion for the minimax risk as $\varepsilon \to 0$:

$$r^\varepsilon(\Theta) = \frac{\varepsilon^2}{2\beta+1}\,\log\frac{P(2\beta+1)}{\varepsilon^2\pi^{2\beta}} + \varepsilon^2\left(\gamma - \frac{2}{2\beta+1}\right) + o(\varepsilon^2).$$

4. Estimation of the derivative at a fixed point.

4.1. An upper bound. In this section we establish the upper bound of the minimax risk $r_0^\varepsilon(\Theta)$ in the problem of estimating the linear functional $L(\theta) = \sum_{k=1}^\infty s_k\theta_k$. We shall look for it in the class of linear estimates $\mathcal{L}$.

LEMMA 4. *The minimax risk of estimating the functional $L(\theta)$ in the class of linear estimates is*

$$\inf_{\widehat{L}\in\mathcal{L}} R_0^\varepsilon(\widehat{L},\Theta) = \varepsilon^2\sum_{k=1}^\infty s_k^2(1+\varepsilon^2a_k^2)^{-1}. \tag{4.1}$$

The estimate

$$\widehat{L}_h(X) = \sum_{k=1}^\infty (1+\varepsilon^2a_k^2)^{-1}s_kX_k$$

is a minimax linear estimate.

Proof. Consider $\widehat{L}_h(X) = \sum_{k=1}^\infty h_ks_kX_k$. It is easily seen that

$$R_0^\varepsilon(\widehat{L}_h,\Theta) = \sup_{\theta\in\Theta}\left[\sum_{k=1}^\infty \theta_k(1-h_k)\,s_k\right]^2 + \varepsilon^2\sum_{k=1}^\infty h_k^2s_k^2.$$

Applying the Cauchy–Schwarz inequality, we have

$$\begin{aligned} R_0^\varepsilon(\widehat{L}_h,\Theta) &= \sum_{k=1}^\infty a_k^2\theta_k^2\sum_{k=1}^\infty a_k^{-2}(1-h_k)^2s_k^2 + \varepsilon^2\sum_{k=1}^\infty h_k^2s_k^2 \\ &= \sum_{k=1}^\infty a_k^{-2}s_k^2(1-h_k)^2 + \varepsilon^2\sum_{k=1}^\infty h_k^2s_k^2. \end{aligned}$$

It is a simple matter to check that the minimum with respect to h_k of the right-hand side is attained by $\widehat{h}_k = (1+\varepsilon^2a_k^2)^{-1}$. This completes (4.1).

4.2. A lower bound. We now find the lower bound of the minimax risk $r_0^\varepsilon(\Theta)$. To do this we will use the standard arguments of [13]. Assume that $\theta_k = \theta b_k$, where θ is a random variable and b_k is a fixed sequence. Then we have to estimate the parameter

$$L(\theta) = \theta \sum_{k=1}^{\infty} b_k s_k \tag{4.2}$$

from the observations $X_k = \theta b_k + \varepsilon \xi_k$. Note that $\sum_{k=1}^{\infty} b_k X_k$ is a sufficient statistic. Thus we only need the observation

$$Y = \sum_{k=1}^{\infty} b_k X_k = \theta \sum_{k=1}^{\infty} b_k^2 + \varepsilon^2 \sum_{k=1}^{\infty} b_k \xi_k$$

for estimating the parameter θ. Hence we have an equivalent problem of estimating the parameter $L(\theta)$ (see (4.2)) from the observation

$$Y' = \theta + \varepsilon^2 \xi \, \|b\|^{-1}, \tag{4.3}$$

where $\xi \sim \mathcal{N}(0,1)$ and $\|\cdot\|$ is a norm in $l_2(1,\infty)$. At the same time, condition (2.2) yields the following restriction on θ:

$$\theta^2 \leqq \left(\sum_{k=1}^{\infty} a_k^2 b_k^2 \right)^{-1}. \tag{4.4}$$

LEMMA 5. *The following lower bound for the risk $r_0^\varepsilon(\Theta)$ holds:*

$$r_0^\varepsilon(\Theta) \geqq \varepsilon^2 \sum_{k=1}^{\infty} s_k^2 \Big(1 + \varepsilon^2 \pi^2 a_k^2 \Big)^{-1}. \tag{4.5}$$

Proof. Set $A = (\sum_{k=1}^{\infty} a_k^2 b_k^2)^{-1/2}$. Note that

$$\begin{aligned} r_0^\varepsilon(\Theta) &\geqq \inf_{\widehat{L}} \sup_{|\theta| \leqq A} \mathbf{E}_\theta \left[\theta \sum_{k=1}^{\infty} b_k s_k - \widehat{L}(Y') \right]^2 \\ &= \left(\sum_{k=1}^{\infty} b_k s_k \right)^2 \inf_{\hat\theta} \sup_{|\theta| \leqq A} \mathbf{E}_\theta (\theta - \hat\theta)^2. \end{aligned} \tag{4.6}$$

Let ν be the a priori probability density of the parameter θ, which is supported in the interval $[-A, A]$. Then we have

$$\sup_{|\theta| \leqq A} \mathbf{E}_\theta (\theta - \hat\theta)^2 \geqq \int \mathbf{E}_\theta (\theta - \hat\theta)^2 \nu(\theta) \, d\theta. \tag{4.7}$$

Then from the Van Trees inequality [11] we obtain

$$\int \mathbf{E}_\theta (\theta - \hat\theta)^2 \nu(\theta) \, d\theta \geqq \big[\mathbf{E} I(p_\theta) + I(\nu) \big]^{-1}. \tag{4.8}$$

The Fisher information in the right-hand side of this inequality consists of, respectively,

$$I(p_\theta) = \int \frac{{p'_\theta}^2(x)}{p_\theta(x)}\,dx, \qquad I(\nu) = \int_{-A}^{A} \frac{\nu'^2(x)}{\nu(x)}\,dx;$$

here $p_\theta(\cdot)$ is the probability density of observations (4.3). Minimizing the Fisher information $I(\nu)$ with respect to the prior density ν, we can easily assert that $\inf_\nu I(\nu) = I(\nu^*) = \pi^2 A^2$, where $\nu^*(x) = A^{-1}\cos^2[\pi x/(2A)]$. Since $\mathbf{E}I(p_\theta) = \|b\|^2/\varepsilon^2$, we see from (4.6)–(4.8) that

$$r_0^\varepsilon(\Theta) \geqq \left(\sum_{k=1}^{\infty} b_k s_k\right)^2 \left(\|b\|^2\varepsilon^{-2} + \pi^2 \sum_{k=1}^{\infty} a_k^2 b_k^2\right)^{-1}.$$

Let us "improve" this lower bound, maximizing with respect to b_k the right-hand side of the inequality. It is easy to see that the maximum is attained by $b_k^* = s_k(\varepsilon^{-2} + \pi^2 a_k^2)^{-1}$. This yields (4.5).

The proof of Theorem 2 immediately follows from Lemmas 4 and 5.

4.3. The asymptotic behavior of the minimax risk over the Sobolev class. Here we investigate the asymptotic behavior of the bounds for the minimax risk as $\varepsilon \to 0$. Set $a_k = (\pi k)^\beta/\sqrt{P}$ and $s_k = k^{-1/2}$. Let $N = \max\{k\colon\ |a_k| < 1/\varepsilon\}$. Then we have

$$\begin{aligned} \sum_{k=1}^{\infty} s_k^2(1+\varepsilon^2 a_k^2)^{-1} &= \varepsilon^2 \sum_{k=1}^{\delta N} k^{-1}(1+\varepsilon^2 a_k^2)^{-1} + \varepsilon^2 \sum_{k=\delta N+1}^{\infty} k^{-1}(1+\varepsilon^2 a_k^2)^{-1} \\ &\equiv \varepsilon^2\big[S_1(\varepsilon,\delta) + S_2(\varepsilon,\delta)\big], \end{aligned} \tag{4.9}$$

where $\delta \in (0,1)$ is a number depending on ε, which will be chosen later. Note that as $\varepsilon \to 0$

$$S_1(\varepsilon,\delta) = \sum_{k=1}^{\delta N} k^{-1} + O\big(\delta^{2\beta}\big) = \log N + \log\delta + \gamma + O\big(\delta^{2\beta}\big). \tag{4.10}$$

Taking into account the fact that $N = (1+o(1))\,\pi^{-1}(P/\varepsilon^2)^{1/(2\beta)}$, we obtain

$$\begin{aligned} S_2(\varepsilon,\delta) &= \sum_{k=\delta N+1}^{\infty} k^{-1}\left[1 + \frac{\varepsilon^2(\pi k)^{2\beta}}{P}\right]^{-1} \\ &= \int_\delta^\infty x^{-1}(1+x^{2\beta})^{-1}\,dx + O\big(N^{-1}\delta^{-1}\big). \end{aligned} \tag{4.11}$$

On the other hand, it is easy to show that

$$\int_\delta^\infty x^{-1}(1+x^{2\beta})^{-1}\,dx = -\log\delta + (2\beta)^{-1}\log(1+\delta^{2\beta}).$$

Then choosing $\delta = \log^{-1}\varepsilon^{-2}$ and applying (4.9)–(4.11), we have the asymptotic expansion for the upper bound:

$$r_0^\varepsilon(\Theta) \leqq \sum_{k=1}^{\infty} s_k^2(1+\varepsilon^2 a_k^2)^{-1} = \frac{\varepsilon^2}{2\beta}\log\frac{P}{\varepsilon^2} + \varepsilon^2\,(\gamma - \log\pi) + o(\varepsilon^2). \tag{4.12}$$

Obviously, the asymptotic expansion of the lower bound (4.5) is similar, with the only difference being in replacing P in (4.12) by P/π^2. Therefore the asymptotic behavior for the lower bound is

$$r_0^\varepsilon(\Theta) \geqq \sum_{k=1}^{\infty} s_k^2(1+\varepsilon^2\pi^2 a_k^2)^{-1} = \frac{\varepsilon^2}{2\beta}\log\frac{P}{\pi^2\varepsilon^2} + \varepsilon^2(\gamma - \log\pi) + o(\varepsilon^2).$$

Thus there is a gap between the upper and lower bounds. It equals $\varepsilon^2\beta^{-1}\log\pi$ and decreases with respect to increasing the smoothness β.

4.4. Nonlinear estimation. In this section we express the minimax risk $r_0^\varepsilon(\Theta)$ in terms of the problem (2.10)–(2.12). It is assumed that $a_k^2 = (\pi k)^{2\beta}/P$ and $s_k = k^{-1/2}$.

Set $N = \max\{k\colon\ |a_k| < 1/\varepsilon\}$ as before. Let

$$R_\delta = \inf_{\widehat{L}_\delta}\sup_{f\in\mathcal{F}} \mathbf{E}_\theta\left[L_\delta(f) - \widehat{L}_\delta\right]^2 + \log\delta$$

be the normalized minimax risk.

Lemma 6. *The following inequality holds*:

$$r_\varepsilon^2(\Theta) \geqq \frac{\varepsilon^2}{2\beta}\log\frac{P}{\varepsilon^2} + \varepsilon^2\gamma + \varepsilon^2\lim_{\delta\to 0} R_\delta + o(\varepsilon^2) \qquad \textit{as} \quad \varepsilon\to 0.$$

Proof. Let us divide the set of indices $\{1,2,\dots\}$ into two subsets $K_1 = \{1,\dots,\delta N\}$ and $K_2 = \{\delta N+1,\dots\}$, where δ is sufficiently small. Then the functional to be estimated can be rewritten as $L(\theta) = L_1(\theta) + L_2(\theta)$, where

$$L_1(\theta) = \sum_{k\in K_1} s_k\theta_k, \qquad L_2(\theta) = \sum_{k\in K_2} s_k\theta_k.$$

Further, for a number $\alpha\in(0,1)$ fix two sets

$$\begin{aligned} \Theta_1 &= \left\{\theta_k,\ k\in K_1\colon\ \sum_{k\in K_1} a_k^2\theta_k^2 \leqq \alpha\right\}, \\ \Theta_2 &= \left\{\theta_k,\ k\in K_2\colon\ \sum_{k\in K_2} a_k^2\theta_k^2 \leqq 1-\alpha\right\}. \end{aligned} \tag{4.13}$$

For abbreviation, we write $\mathbf{X}_1 = (X_1,\dots,X_{\delta N})^T$ and $\mathbf{X}_2 = (X_{\delta N+1},\dots)^T$. Let $\pi_1(\theta)$, $\theta\in\Theta_1$ and $\pi_2(\theta)$, $\theta\in\Theta_2$ be the prior distribution densities on the sets Θ_1 and Θ_2; let $p_1(\mathbf{X}_1|\theta)$, $\theta\in\Theta_1$ and $p_2(\mathbf{X}_2|\theta)$, $\theta\in\Theta_2$ be, respectively, the distribution densities of the vectors $\mathbf{X}_1$, $\mathbf{X}_2$.

Therefore

$$\inf_{\widehat{L}}\sup_{\theta\in\Theta}\mathbf{E}_\theta\left[L_1(\theta)+L_2(\theta)-\widehat{L}\right]^2 \geqq \inf_{\widehat{L}}\mathbf{E}\,\mathbf{E}_{\theta_1,\theta_2}\left[L_1(\theta_1)+L_2(\theta_2)-\widehat{L}\right]^2; \tag{4.14}$$

here $\mathbf{E}$ is the expectation with respect to the measure with density $\pi_1(\theta_1)\,\pi_2(\theta_2)$. Note that the infimum of the right-hand side is attained, whence $\widehat{L}$ is the Bayes estimator:

$$\widehat{L}(X) = \frac{\int_{\Theta_1}\int_{\Theta_2}[L_1(\theta_1)+L_2(\theta_2)]\,p_1(\mathbf{X}_1\mid\theta_1)\,p_2(\mathbf{X}_2\mid\theta_2)\,\pi_1(\theta_1)\pi_2(\theta_2)\,d\theta_1\,d\theta_2}{\int_{\Theta_1}\int_{\Theta_2}p_1(\mathbf{X}_1\mid\theta_1)\,p_2(\mathbf{X}_2\mid\theta_2)\,\pi_1(\theta_1)\,\pi_2(\theta_2)\,d\theta_1\,d\theta_2}.$$

It is easy to see that $\widehat{L}(X) = \widehat{L}_1(\mathbf{X}_1) + \widehat{L}_2(\mathbf{X}_2)$, where $\widehat{L}_1(\cdot)$ and $\widehat{L}_2(\cdot)$ are Bayes estimators of the vectors θ_1 and θ_2, respectively. Since these estimators are nonbiased, we have

$$\inf_{\widehat{L}} \mathbf{E}\,\mathbf{E}_{\theta_1,\theta_2}\big[L_1(\mathbf{X}_1) - L_2(\mathbf{X}_2) - \widehat{L}\big]^2 = \mathbf{E}\,\mathbf{E}_{\theta_1}\big[L_1(\theta_1) - \widehat{L}_1(\mathbf{X}_1)\big]^2 + \mathbf{E}\,\mathbf{E}_{\theta_2}\big[L_2(\theta_2) - \widehat{L}_2(\mathbf{X}_2)\big]^2. \tag{4.15}$$

Let us recall that the densities π_1 and π_2 were chosen arbitrarily. Consequently,

$$\sup_{\pi_i} \mathbf{E}\,\mathbf{E}_{\theta_i}\big[L_i(\theta_i) - \widehat{L}_i(\mathbf{X}_i)\big]^2 = \sup_{\theta_i\in\Theta_i} \mathbf{E}_{\theta_i}\big[L_i(\theta_i) - \widehat{L}_i(\mathbf{X})\big]^2, \qquad i = 1,2.$$

Thus combining (4.14) and (4.15) we conclude that

$$r_0^\varepsilon(\Theta) \geqq r_0^\varepsilon(\Theta_1) + r_0^\varepsilon(\Theta_2). \tag{4.16}$$

Hence our problem can be divided into two independent ones. These two problems deal with estimating the functionals $L_1(\theta)$, $\theta \in \Theta_1$ and $L_2(\theta)$, $\theta\in\Theta_2$, respectively.

The lower bound for $r_0^\varepsilon(\Theta_1)$ follows by the same method as in Lemma 5. Set $\theta_k = b_k\zeta$ with $k \in K_1$, where ζ is some random variable. The condition $\theta\in\Theta_1$ yields the restriction on ζ: $\zeta^2\sum_{k=1}^{\delta N} a_k^2 b_k^2 \leqq \alpha$. Therefore it can be easily seen (see the proof of Lemma 5) that

$$r_0^\varepsilon(\Theta_1) \geqq \varepsilon^2 \sum_{k=1}^{\delta N} k^{-1}\left(1 + \frac{\varepsilon^2\pi^2 a_k^2}{\alpha}\right)^{-1}. \tag{4.17}$$

Note also that

$$\begin{aligned}
\sum_{k=1}^{\delta N} k^{-1}\left(1 + \frac{\varepsilon^2\pi^2 a_k^2}{\alpha}\right)^{-1} &\geqq \sum_{k=1}^{\delta N} k^{-1}\left(1 - \frac{\pi^2 a_k^2\varepsilon^2}{\alpha}\right)\\
&= \log\delta N + \gamma + o(1) - \frac{\varepsilon^2\pi^2}{\alpha}\sum_{k=1}^{\delta N}\frac{a_k^2}{k}\\
&= \log\delta N + \gamma + o(1) - \alpha^{-1}O\big(\delta^{2\beta}\big).
\end{aligned}$$

Choosing $\delta^{2\beta}/\alpha = o(1)$ as $\varepsilon\to 0$, we infer from these and (4.17) the inequality

$$r_0^\varepsilon(\Theta_1) \geqq \varepsilon^2\big[\log N + \log\delta + \gamma + o(1)\big]. \tag{4.18}$$

Consider now the second term in the right-hand side of (4.16). Notice that we can estimate the functional $L_\delta^A(\theta) = \sum_{k=\delta N}^{AN}\theta_k/\sqrt{k}$ instead of $L_2(\theta)$ since

$$r_0^\varepsilon(\Theta_2) \geqq \inf_{\widehat{L}} \sup_{\theta\in\Theta_2} \mathbf{E}_\theta^\varepsilon\big[L_\delta^A(\theta) - \widehat{L}\big]^2. \tag{4.19}$$

Let us consider a new loss function

$$w_B(x) = \begin{cases} x^2, & |x| < B,\\ B^2, & |x| \geqq B,\end{cases}$$

where B is a positive number. Since $x^2 \geqq w_B(x)$, (4.19) gives

$$r_0^\varepsilon(\Theta_2) \geqq \varepsilon^2 \inf_{\widehat{L}} \sup_{\theta\in\Theta_2} \mathbf{E}_\theta^\varepsilon\, w_B\big[\varepsilon^{-1}(L_\delta^A(\theta) - \widehat{L})\big]. \tag{4.20}$$

Let $\mathcal{F}_\alpha(A,Q)$ be the set of all functions f such that $f(x) = 0$ for $x \notin [0, A]$, $\sup_x f'(x) \leqq Q$, and

$$\int_0^A x^{2\beta} f(x)\,dx \leqq 1 - 2\alpha. \tag{4.21}$$

Set

$$\theta_k = \frac{P^{1/2}}{N^{\beta+1/2}\,\pi^\beta}\, f\left(\frac{k}{N}\right). \tag{4.22}$$

Since the derivative $f'(t)$ is bounded on the interval $[0, A]$, we see that

$$\sum_{\delta N}^{AN} a_k^2\theta_k^2 \leqq \frac{1}{N}\sum_{k=1}^{AN}\left(\frac{k}{N}\right)^{2\beta} f^2\left(\frac{k}{N}\right) = \int_0^A t^{2\beta} f^2(t)\,dt + O\left(\frac{Q}{N}\right).$$

This gives $\theta \in \Theta_2$. In the same manner we can see that the functional being estimated can be approximated in the following way:

$$L_\delta^A(\theta) = \frac{\varepsilon}{N}\sum_{k=\delta N}^{AN} \frac{f(k/N)}{\sqrt{k/N}} = \varepsilon \int_\delta^A \frac{f(t)}{\sqrt{t}}\,dt + o(\varepsilon). \tag{4.23}$$

Further note that the observations rewritten in terms of $f(\cdot)$ are

$$X_k = \frac{P^{1/2}}{N^{\beta+1/2}\pi^\beta}\, f\left(\frac{k}{N}\right) + \varepsilon\xi_k.$$

Noting that $\sqrt{P}/(N^\beta\pi^\beta) = (1+o(1))\,\varepsilon$, we deduce the equivalent observations

$$Y_k = f\left(\frac{k}{N}\right) + \big(1+o(1)\big)\sqrt{N}\,\xi_k. \tag{4.24}$$

Let us pass from these observations to the equivalent observations with continuous time. Denote by $\bar f(t)$ the step function such that $\bar f(t) = f(k/N)$ as $|t - k/N| \leqq 1/(2N)$. Therefore the observations (4.24) are equivalent to

$$d\overline{Y}(t) = \bar f(t)\,dt + dw(t), \qquad t \in [0, A], \tag{4.25}$$

where $w(t)$ is the standard Wiener process. Note that $\|\bar f - f\| \to 0$ as $\varepsilon \to 0$. This means that the problem of estimating the functional

$$L_\delta^A(f) = \int_\delta^A \frac{f(t)}{\sqrt{t}}\,dt$$

from observations (4.25) is asymptotically equivalent (see, e.g., [1]) to the problem of estimating this functional from the observations

$$dY(t) = f(t)\,dt + dw(t), \qquad t \in [0, A]. \tag{4.26}$$

This means that

$$\inf_{\widehat L}\ \sup_{f\in\mathcal{F}_\alpha(A,Q)} \mathbf{E}_f w_B\big\{L_\delta^A(f) - \widehat L(\overline{Y})\big\} \geqq \inf_{\widehat L}\ \sup_{f\in\mathcal{F}_\alpha(A,Q)} \mathbf{E}_f w_B\big\{L_\delta^A(f) - \widehat L(Y)\big\} + o(1).$$

This, (4.23), and (4.20) yield

$$\lim_{\varepsilon\to 0} \varepsilon^{-2} r_0^\varepsilon(\Theta_2) \geqq \inf_{\widehat{L}} \sup_{f\in\mathcal{F}_\alpha(A,Q)} \mathbf{E}_f w_B\big[L_\delta^A(f) - \widehat{L}\big].$$

A passage to the limit as $B\to\infty$, $Q\to\infty$, and $\alpha\to 0$ (see also (4.16) and (4.18)) completes the proof of the lemma.

Let us show now that the bound obtained cannot be improved.

LEMMA 7. *The following inequality holds as* $\varepsilon\to 0$:

$$r_0^\varepsilon(\Theta) \leqq \varepsilon^2 \log N + \varepsilon^2\gamma + \varepsilon^2 \lim_{\delta\to 0} R_\delta + o(\varepsilon^2).$$

Proof. We can divide the observations into two parts in the same manner as in the proof of the lower bound. The underlying functional $L(\theta)$ can be rewritten as the sum of two functionals $L_1(\theta)$ and $L_2(\theta)$. Take the projection estimator $\widehat{L}_1(X) = \sum_{k=1}^{\delta N} X_k/\sqrt{k}$ as an estimate of the functional $L_1(\theta) = \sum_{k=1}^{\delta N}\theta_k/\sqrt{k}$. Then it is easy to estimate the risk

$$\begin{aligned} r_0^\varepsilon(\Theta) &\leqq \inf_{\widehat{L}_2}\sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \left[L_1(\theta) + L_2(\theta) - \widehat{L}_1(X) - \widehat{L}_2(X)\right]^2 \\ &= \varepsilon^2 \sum_{k=1}^{\delta N} k^{-1} + \inf_{\widehat{L}_2}\sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \left[L_2(\theta) - \widehat{L}_2(X)\right]^2. \end{aligned} \tag{4.27}$$

It is clear that we have to obtain an upper bound for the last term. To do this, let us reduce our problem with continuous time (2.10)–(2.12) to the initial discrete problem. To be precise, we shall prove that

$$\begin{aligned} &\inf_{\widehat{L}_2}\sup_{\theta\in\Theta} \mathbf{E}_\theta^\varepsilon \left[L_2(\theta) - \widehat{L}_2(X)\right]^2 + \varepsilon^2\log\delta \\ &\qquad \leqq \varepsilon^2 \inf_{\widehat{L}_\delta}\sup_{f\in\mathcal{F}} \mathbf{E}_f \left[\widehat{L}_\delta(X) - L_\delta(f)\right]^2 + \varepsilon^2\log\delta + o(\varepsilon^2); \end{aligned} \tag{4.28}$$

here the functional $L_\delta(f)$ is defined in (2.10). Let $\overline{\mathcal{F}}$ be the set of all step functions from $\mathcal{F}$ such that

$$\bar f(x) = \sum_{k=\delta N}^{\infty} f\left(\frac{k}{N}\right) \mathbf{1}\left\{\frac{k}{N} \leqq x < \frac{k+1}{N}\right\}.$$

Obviously, for any estimate $\widehat{L}_\delta(X)$ the following inequality holds:

$$\sup_{f\in\mathcal{F}} \mathbf{E}_f \left[L_\delta(f) - \widehat{L}_\delta(X)\right]^2 \geqq \sup_{f\in\overline{\mathcal{F}}} \mathbf{E}_f \left[L_\delta(f) - \widehat{L}_\delta(X)\right]^2. \tag{4.29}$$

At the same time, observations (2.11) are equivalent to

$$X_k^\varepsilon = f\left(\frac{k}{N}\right) + \sqrt{N}\,\xi_k. \tag{4.30}$$

Let θ_k be defined in (4.22). Then observations (4.30) are equivalent to

$$Z_k^\varepsilon = \theta_k + (1+o(1))\,\varepsilon\xi_k. \tag{4.31}$$

In terms of θ_k, the underlying functional is

$$L_\delta(\bar f) = \int_\delta^\infty \frac{\bar f(t)}{\sqrt{t}}\,dt = \frac{\pi^\beta N^{\beta+1/2}}{\sqrt{P}} \sum_{k=\delta N}^\infty \theta_k \int_{k/N}^{(k+1)/N} \frac{dt}{\sqrt{t}}$$

$$= \frac{1+o(1)}{\varepsilon} \sum_{k=\delta N}^\infty \frac{\theta_k}{\sqrt{k}} - \frac{1+o(1)}{\varepsilon} \sum_{k=\delta N}^\infty \theta_k \frac{1}{\sqrt{k}(\sqrt{k+1}+\sqrt{k})^2}. \tag{4.32}$$

The restrictions on $\bar f$ can be recalculated in the restrictions on θ_k:

$$1 \geqq \int_0^\infty t^{2\beta} \bar f^2(t)\,dt = \frac{\pi^{2\beta}}{P(2\beta+1)} \sum_{k=\delta N}^\infty \theta_k^2 \Big[(k+1)^{2\beta+1} - k^{2\beta+1}\Big] \geqq \frac{\pi^{2\beta}}{P} \sum_{k=\delta N}^\infty \theta_k k^{2\beta},$$

i.e.,

$$\sum_{k=\delta N}^\infty a_k^2 \theta_k^2 \leqq 1. \tag{4.33}$$

Thus applying the Cauchy–Schwarz inequality, we obtain from this and (4.32)

$$\left[\varepsilon L_\delta(\bar f) - \sum_{k=\delta N}^\infty \frac{\theta_k}{\sqrt{k}}\right]^2 \leqq O\left((\delta N)^{-2\beta-2}\right) + o\left((\delta N)^{-2\beta}\right).$$

Consequently, combining this inequality with (4.33), (4.31), and (4.29), we obtain (4.28). The lemma immediately follows from (4.27).

The proof of Theorem 3 is straightforward (see Lemmas 6 and 7).

REFERENCES

[1] L. D. Brown and M. G. Low, *Asymptotic equivalence of nonparametric regression and white noise*, Ann. Statist., 24 (1996), pp. 2384–2398.

[2] B. van Es and A. W. Hoogendoorn, *Kernel estimation in Wicksell's corpuscle problem*, Biometrika, 77 (1990), pp. 139–145.

[3] G. Golubev and B. Levit, *Asymptotically efficient estimation in the Wicksell problem*, Ann. Statist., 26 (1998), pp. 2407–2419.

[4] P. Groeneboom and G. Jongbloed, *Isotonic estimation and rates of convergence in Wicksell's problem*, Ann. Statist., 23 (1995), pp. 1518–1542.

[5] P. Hall and R. L. Smith, *The kernel method for unfolding sphere size distributions*, J. Comput. Phys., 74 (1988), pp. 409–421.

[6] A. W. Hoogendoorn, *Estimating the weight undersize distribution for the Wicksell problem*, Statist. Neerlandica, 46 (1992), pp. 259–282.

[7] M. Nussbaum, *Asymptotic equivalence of density estimation and Gaussian white noise*, Ann. Statist., 24 (1996), pp. 2399–2430.

[8] C. C. Taylor, *A new method for unfolding sphere size distributions*, J. Microscopy, 132 (1983), pp. 57–66.

[9] D. S. Watson, *Estimating functionals of particle size distributions*, Biometrika, 58 (1971), pp. 483–490.

[10] S. D. Wicksell, *The corpuscle problem. A mathematical study of a biometric problem*, Biometrika, 17 (1925), pp. 84–99.

[11] H. L. Van Trees, *Detection, Estimation, and Modulation Theory.* Part I, John Wiley and Sons, New York, 1968.

[12] A. Zygmund, *Trigonometric Series.* Vol. II, Cambridge University Press, London, 1968.

[13] I. A. Ibragimov and R. Z. Hasminskii, *Statistical Estimation. Asymptotic Theory*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1981.

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