Problems of Information Transmission, Vol. 37, No. 1, 2001, pp. 24–45. Translated from Problemy Peredachi Informatsii, No. 1, 2001, pp. 28–51.
Original Russian Text Copyright © 2001 by Golubev, Enikeeva.

METHODS OF SIGNAL PROCESSING

Asymptotically Efficient Smoothing in the Wicksell Problem under Squared Losses

G. K. Golubev[1] and F. N. Enikeeva

Received May 17, 2000; in final form, September 28, 2000

Abstract—In the Wicksell problem, it is required to reconstruct a distribution function of radii of balls located in an opaque medium from measurements of radii of circles obtained by intersecting the medium with a certain plane. This problem is intimately bound up with estimating a fractional derivative of order 1/2 for a distribution function concentrated on the positive semi-axis. In this paper, the locally asymptotically minimax risk in the Wicksell problem is evaluated up to a constant. Estimators on which this risk is attained are also constructed.

1. INTRODUCTION

The Wicksell problem [1] is one of the known inverse (statistical) problems. In this problem, it is required to reconstruct a distribution function of radii of balls located in an opaque medium. These radii are independent identically distributed random variables. The obstacle consists in impossibility to directly measure the ball radii since the medium is opaque. However, it is possible to measure radii of circles in the intersection of the medium with a certain hyperplane P in $\mathbb{R}^3$. For definiteness, we will assume that centers of random balls $B_i = B_i(\boldsymbol{v}_i, R_i),\ i = 1, \ldots, n$, are located at random points $\boldsymbol{v}_i$ which are a realization of a stationary Poisson process in $\mathbb{R}^3$ and squares of ball radii are independent random variables $Y_i = R_i^2$ with an unknown distribution function $F(y),\ y \in \mathbb{R}^+$. The problem is to reconstruct the function $F(y)$ for all $y > 0$ from observations of squares of circle radii $X_i = r_i^2$ in the intersection. In [2–4], some applications of the Wicksell problem in biology, stereology, etc., are discussed.

By symmetry of the problem, we choose the coordinate system in $\mathbb{R}^3$ in such a way that $\mathrm{P} = \{\boldsymbol{v} :\ v_3 = 0\}$ and denote by $S_i = S(\boldsymbol{u}_i, r_i),\ \boldsymbol{u}_i = (v_{i1}, v_{i2})$, the circles in the intersection $B_i \cap \mathrm{P}$. Let $G(x)$ be the distribution function of squares of circle radii $X_i = r_i^2$. Denote by λ the rate of a Poisson process in $\mathbb{R}^3$. Then it easily seen that we have the following equalities:

$$\begin{aligned}
&\mathbf{P}\{X_1 > x \mid S_1 \text{ is observed in } v_{11} \in (w_1, w_1 + dw_1), v_{12} \in (w_2, w_2 + dw_2)\}\\
&\qquad = \frac{\mathbf{P}\{v_{11} \in (w_1, w_1 + dw_1), v_{12} \in (w_2, w_2 + dw_2), |v_{13}| < \sqrt{Y_1 - x}\}}{\mathbf{P}\{v_{11} \in (w_1, w_1 + dw_1), v_{12} \in (w_2, w_2 + dw_2), |v_{13}| < \sqrt{Y_1}\}}\\
&\qquad = \frac{2\lambda\, dw_1 dw_2 \int_x^\infty \sqrt{y - x}\, dF(y)}{2\lambda\, dw_1 dw_2 \int_0^\infty \sqrt{y}\, dF(y)} = \frac{\int_x^\infty \sqrt{y - x}\, dF(y)}{\int_0^\infty \sqrt{y}\, dF(y)}.
\end{aligned}$$

Therefore, the distribution functions F and G are connected by the integral equation

$$1 - G(x) = \frac{\int_x^\infty \sqrt{y - x}\, dF(y)}{\int_0^\infty \sqrt{y}\, dF(y)}. \tag{1}$$

[1] Supported in part by the Russian Foundation for Basic Research, project no. 00-01-00266.

Let the distribution function F be Lipschitzian with smoothness greater than $1/2$, then equation (1) can be solved directly. Its solution can be represented as

$$1 - F(y) = \frac{DG(y)}{DG(0)}, \tag{2}$$

where

$$DG(y) = \frac{1}{\sqrt{\pi}} \int_y^\infty \frac{dG(x)}{\sqrt{x-y}} \tag{3}$$

is the derivative of order $1/2$ for the function $G(y)$ at the point y (see e.g., [5]). Formula (2) can easily be proved. Taking into account that F is Lipschitzian with smoothness greater than $1/2$, it is not difficult to show that G is differentiable and its derivative is equal to

$$g(x) = G'(x) = \frac{1}{2\mathbf{E}\sqrt{Y_1}} \int_x^\infty \frac{dF(y)}{\sqrt{y-x}}. \tag{4}$$

Using this formula, we obtain

$$\begin{aligned} \int_y^\infty \frac{dG(x)}{\sqrt{x-y}} &= \frac{1}{2\mathbf{E}\sqrt{Y_1}} \int_y^\infty \frac{1}{\sqrt{x-y}} \int_x^\infty \frac{dF(z)}{\sqrt{z-x}} dx \\ &= \frac{1}{2\mathbf{E}\sqrt{Y_1}} \int_y^\infty \int_y^z \frac{dx}{\sqrt{(z-x)(x-y)}} dF(z) = \frac{\pi[1-F(y)]}{2\mathbf{E}\sqrt{Y_1}}, \end{aligned} \tag{5}$$

which proves (2).

Actually, our problem can now be formulated in very simple way: it is required to estimate the derivative of order $1/2$ of a distribution function from observations of random variables $\boldsymbol{X}^n = X_1, \ldots, X_n$. Below, we consider the situation where the number of observations is large, i.e., $n \to \infty$. In spite of sustained efforts, an asymptotically precise solution of the Wicksell problem at a fixed point was only recently obtained [6] though the orders of convergence rates were found rather long ago [7].

In this paper, we consider the problem of estimating in $\mathbf{L}_2(0,\infty)$ since qualitative behavior of a distribution function is usually more interesting than the value of this function at a fixed point. The risk of an estimator $\widetilde{F}(y, \boldsymbol{X}^n)$ for a distribution function F will be measured by its mean-square error

$$R^n(\widetilde{F}, F) = \mathbf{E} \int_0^\infty [\widetilde{F}(y, \boldsymbol{X}^n) - F(y)]^2 dy.$$

Below, we use a locally-minimax setting of the problem. Denote by $\|\cdot\|$ the ordinary norm in $\mathbf{L}_2(0,\infty)$. Define also the norm $\|\cdot\|_\beta$ for distribution functions by the formula

$$\|F\|_\beta = \|1 - F\| + \|F^{(\beta)}\|,$$

where $F^{(\beta)}(x)$ is the derivative of order β of the function $F(x)$.

Let us fix a distribution function F_0 such that $\|F_0\|_\beta < \infty$ and define its neighborhood $B_\varepsilon(F_0)$ as

$$B_\varepsilon(F_0) = \Big\{F : \ \|F - F_0\|_\beta \le \varepsilon\Big\}.$$

Let $\mathcal{F}_\varepsilon(C)$ be the set of distribution functions such that

$$\int_0^\infty x^{3/2+\varepsilon} dF(x) \le C. \tag{6}$$

Here and below, C denotes a constant whose value is not significant. Our goal is to evaluate the locally minimax risk

$$r(F_0) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \inf_{\widetilde{F}} \sup_{F\in B_\varepsilon(F_0)\cap\mathcal{F}_\varepsilon(C)} \frac{n}{\log n} R^n(\widetilde{F}, F),$$

where the infimum is taken over all estimates, and to construct estimators on which $r(F_0)$ is attained.

Formulas (2) and (3) suggest a simple method for constructing estimates but this method is too naive to succeed. Indeed, instead of the distribution function for squares of circle radii, which we, naturally, do not know, it is possible to substitute its estimate, i.e., the empirical distribution function

$$G_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{X_i < x\}.$$

Thus we arrive at the following estimator for the derivative of order 1/2:

$$DG_n(y) = \frac{1}{n\sqrt{\pi}} \sum_{i=1}^n \frac{1}{\sqrt{X_i - y}} \mathbf{1}\{X_i > y\}.$$

It is not difficult to check that this estimator is unbiased but, unfortunately, its variance is unbounded. Moreover, the function obtained by substitution of this estimator into (2) is undoubtedly not a distribution function since $DG_n(y)$ infinitely grows as y draws near to any of X_i. A natural way to overcome this nuisance is to try to smooth out $DG_n(y)$ by means of a certain filter. Actually, we use two different filters. The first one is used to estimate the derivative $DG(0)$ at the origin and the second one is used to reconstruct $DG(y)$ on the semiaxes. We use the following estimator for the derivative of order 1/2:

$$\widehat{DG}(y, \mu) = \arg\min_m \left\{ \|m - DG_n\|^2 + \mu \|m^{(\beta)}\|^2 \right\}, \tag{7}$$

where μ is a smoothing parameter which will be chosen later. Actually, there are two principal reasons why we choose precisely this estimator. On one hand, the minimization problem of (7) can be solved numerically rather fast since its solution can be reduced to computations using band matrices [8]. The second reason is statistical. It is connected with a boundary effect. The function $DG(x)$ need not be smooth in a neighborhood of the origin. Therefore, classical kernel estimates with symmetrical kernels cannot be used directly.

Having the estimators $\widehat{DG}(y, \mu_1)$ and $\widehat{DG}(0, \mu_2)$ in our disposition, we can now apply formula (2) to estimate the distribution function. Moreover, it should be noted that, in principle, the estimator $\widehat{DG}(0, \mu_2)$ may be very close to zero but the probability of this event is small. Therefore, to avoid dividing by zero, let us define the estimator of the distribution function F by the equality

$$\widehat{F}_{\mu_1,\mu_2}(y, \mathbf{X}^n) = \left[1 - \frac{n\widehat{DG}(y, \mu_1)}{n|\widehat{DG}(0, \mu_2)| + 1}\right]_0^1, \tag{8}$$

where $[x]_0^1 = \min[1, \max(0, x)]$ is a projection of x onto the segment $[0, 1]$.

2. MAIN RESULTS

Let $g_0(x)$ be the probability density of circles corresponding to the distribution function F_0 (see [4]), i.e.,

$$g_0(x) = \frac{1}{2}\int_x^\infty (y-x)^{-1/2} dF_0(y) \Big/ \int_0^\infty \sqrt{y}\, dF_0(y).$$

Note that $g_0(0) > 0$. Below, we will not specially mention this. Set

$$\sigma^2(F_0) = \frac{4}{\pi^2}\left[\int_0^\infty \sqrt{y}\, dF_0(y)\right]^2 \left(\frac{g_0(0)\|1-F_0\|^2}{2\beta-1} + \frac{1}{2\beta}\right).$$

The risk of the estimator $\widehat{F}_{\mu_1,\mu_2}(y, \mathbf{X}^n)$ can be evaluated as follows.

Theorem 1. *Assume that* $\|F_0\|_\beta < \infty$, $\beta \in \mathbb{N}$, $\beta \geq 1$. *Then for estimator* (8) *we have*

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup_{F\in B_\varepsilon(F_0)\cap \mathcal{F}_\varepsilon(C)} \frac{n}{\log n} R^n(\widehat{F}_{\mu_1,\mu_2}, F) \leq \sigma^2(F_0) \tag{9}$$

if $\mu_1 \asymp n^{-1}$ *and* $\mu_2 \asymp n^{-2\beta/(2\beta-1)}$.

Finally, the following theorem states that $\widehat{F}_{\mu_1,\mu_2}(y, \mathbf{X}^n)$ is a locally asymptotically minimax estimator.

Theorem 2. *Assume that* $\|F_0\|_\beta < \infty$, $\beta \in \mathbb{N}$, $\beta \geq 2$. *Then* $r(F_0) = \sigma^2(F_0)$.

Remark 1. Theorem 2 remains also valid for $\beta = 1$ but we do not present here the proof of this fact in order to avoid further complication of technically tedious (even without this assumption) calculations used in the proof of this theorem.

Comparing these results with [6, 7], we observe that the order of convergence rate $\sqrt{\log n/n}$ in the $\mathbf{L}_2$ norm for the Wicksell problem coincides with the convergence rate at a point; naturally, the corresponding constants are different.

3. PROOF OF THEOREM 1

It is well known that the Fourier transform is a very strong tool for the analysis of stationary linear filters due to the fact that the functions $\exp(i\lambda t)$ are eigenfunctions of these filters. In our problem, the Fourier method cannot efficiently be used since the filter from (7) is not time-invariant since the minimization problem is solved in $\mathbf{L}_2(0,\infty)$ but not in $\mathbf{L}_2(-\infty,\infty)$. In this section, we show how one can rather easily overcome this nuisance.

It is not difficult to write out Lagrange's equations and check that the estimator from (7) is a solution of the following boundary-value problem:

$$\begin{gathered} \mu(-1)^\beta \widehat{DG}_x^{(2\beta)}(x,\mu) + \widehat{DG}(x,\mu) = DG_n(x), \quad x \geq 0, \\ \widehat{DG}_x^{(l)}(0,\mu) = 0, \quad l = \beta, \ldots, 2\beta-1. \end{gathered} \tag{10}$$

Note that the solution of this problem can be written as follows:

$$\widehat{DG}(x,\mu) = \int_0^\infty K(x,y,\mu) DG_n(y)dy, \tag{11}$$

where $K(x,y,\mu)$ is the Green function of this boundary-value problem, i.e., the solution of the equation

$$\mu(-1)^\beta K_x^{(2\beta)}(x,y,\mu) + K(x,y,\mu) = \delta(x-y), \quad x \geq 0, \tag{12}$$

with boundary conditions

$$K_x^{(l)}(0,y,\mu)=0,\quad l=\beta,\dots,2\beta-1;$$

here, $\delta(\cdot)$ is the Dirac δ-function.

The parameter μ can be eliminated by renorming. In fact, it can easily be checked that

$$K(x,y,\mu)=\frac{1}{\mu^{1/(2\beta)}}K\left(\frac{x}{\mu^{1/(2\beta)}},\frac{y}{\mu^{1/(2\beta)}}\right),$$

where $K(x,y)$ is a solution of the boundary-value problem

$$\begin{gathered}(-1)^\beta K_x^{(2\beta)}(x,y)+K(x,y)=\delta(x-y),\quad x\ge 0,\\ K_x^{(l)}(0,y)=0,\quad l=\beta,\dots,2\beta-1.\end{gathered}\tag{13}$$

In principle, this problem can easily be solved. In particular, its solution for $\beta=1$ is of the form

$$K(x,y)=\begin{cases}\exp(-x)\cosh(y), & x>y,\\ \exp(-y)\cosh(x), & 0\le x\le y.\end{cases}$$

Note that $K(x,y)$ also satisfies the equation

$$\begin{gathered}(-1)^\beta K_y^{(2\beta)}(x,y)+K(x,y)=\delta(x-y),\quad y\ge 0,\\ K_y^{(l)}(x,0)=0,\quad l=\beta,\dots,2\beta-1.\end{gathered}\tag{14}$$

This can easily be verified applying integration by parts in the equality (see (10))

$$f(x)=\int\limits_0^\infty K(x,y)\left[(-1)^\beta f^{(2\beta)}(y)+f(y)\right]dy,$$

which is valid for any sufficiently smooth function $f(x)$, $x\ge 0$, with $f^{(l)}(0)=0$, $l=\beta,\dots,2\beta-1$.

Multiplying equation (14) by $(x-y)^l$ and using integration by parts, we observe that

$$\int\limits_0^\infty K(x,y)dy=1,\qquad \int\limits_0^\infty (x-y)^l K(x,y)dy=0,\quad l=1,\dots,\beta-1.\tag{15}$$

For $x\ne y$, equation (14) is a homogeneous differential equation with constant coefficients whose general solution can easily be found (see, e.g., [9, p. 44]). Taking into account the boundary conditions, it is not difficult to verify that the function $K(x,y)$ satisfies the inequality

$$|K(x,y)|\le C\mathrm{e}^{-C_\beta|x-y|},\tag{16}$$

where $C_\beta>0$ is a constant.

Consider the following estimator of the function $DG(x)$ at the origin:

$$\begin{aligned}\widehat{DG}(0,\mu)&=\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^\infty K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)DG_n(y)dy\\ &=\frac{1}{\sqrt{\pi}\mu^{1/(2\beta)}n}\sum_{i=1}^n\int\limits_0^\infty K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{\mathbf{1}\{X_i>y\}}{\sqrt{X_i-y}}dy\\ &=\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^\infty K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)DG(y)dy\\ &\qquad+\frac{1}{n}\sum_{i=1}^n\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^\infty K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\left[\frac{\mathbf{1}\{X_i>y\}}{\sqrt{\pi}\sqrt{X_i-y}}-DG(y)\right]dy.\end{aligned}\tag{17}$$

Let us estimate its bias. Equation (14) and the Cauchy–Bunyakovskii inequality yield

$$\begin{aligned}\left|\mathbf{E}\widehat{DG}(0,\mu)-DG(0)\right| &= \mu^{-1/(2\beta)}\left|\int_0^\infty K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)[DG(y)-DG(0)]dy\right| \\ &= \mu^{1-1/(2\beta)}\left|\int_0^\infty K^{(2\beta)}\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)[DG(y)-DG(0)]dy\right| \\ &= \mu^{1/2-1/(2\beta)}\left|\int_0^\infty K^{(\beta)}\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)DG^{(\beta)}(y)dy\right| \\ &\le \mu^{1/2-1/(2\beta)}\left\{\int_0^\infty\left[K^{(\beta)}\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\right]^2dy\right\}^{1/2}\left\{\int_0^\infty\left[DG^{(\beta)}(y)\right]^2dy\right\}^{1/2} \\ &= \mu^{1/2-1/(4\beta)}\left\{\int_0^\infty[K^{(\beta)}(0,y)]^2dy\right\}^{1/2}\left\{\int_0^\infty\left[DG^{(\beta)}(y)\right]^2dy\right\}^{1/2}. \end{aligned} \tag{18}$$

At the same time, the variance is estimated as (see (17))

$$\operatorname{Var}\widehat{DG}(0,\mu)\le\frac{1}{\pi n\mu^{1/\beta}}\int_0^\infty g(x)\left[\int_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^2dx. \tag{19}$$

To calculate this quantity, we use the following statement.

Lemma 1. *For $h\to 0$, we have*

$$\int_0^x\frac{K(z,y,h)}{\sqrt{x-y}}dy=\begin{cases}O(h^{-1/2})\exp(-C_\beta|z-x|/h), & x<z+h,\\ (x-z)^{-1/2}+O[(x-z)^{-1/2-\beta}]h^\beta, & x\ge z+h.\end{cases} \tag{20}$$

The proof of this lemma can almost directly be derived from (15), (16), and the Taylor formula. In fact, it follows from (16) that for $x<z+h$ and $h\to 0$, we have

$$\begin{aligned}\int_0^x\frac{K(z,y,h)}{\sqrt{x-y}}dy &\le O(h^{-1/2})\int_0^{x/h}\frac{\exp(-C_\beta|(z-x)/h+y|)}{\sqrt{y}}\,dy \\ &\le O(h^{-1/2})\exp(-C_\beta|z-x|/h)\int_0^\infty\frac{e^{-C_\beta y}}{\sqrt{y}}\,dy \\ &= O(h^{-1/2})\exp(-C_\beta|z-x|/h),\end{aligned}$$

which proves the first estimate from (20).

To derive the second estimate, let us substitute the Taylor-series expansion of the function $(x-y)^{-1/2}$ for $x\ge z+h$ into the estimated integral. Using (15) and (16), it is not difficult to show that for $h\to 0$ we have

$$\begin{aligned}\int_0^x\frac{K(z,y,h)}{\sqrt{x-y}}dy &= (x-z)^{-1/2}+O\big[(x-z)^{-\beta-1/2}\big]h^\beta\int_0^{x/h}K\left(\frac{z}{h},t\right)\left(t-\frac{z}{h}\right)^\beta dt \\ &\le (x-z)^{-1/2}+O\big[(x-z)^{-\beta-1/2}\big]h^\beta\int_0^{x/h}\exp(-C_\beta|z/h-t|)\left(t-\frac{z}{h}\right)^\beta dt \\ &= (x-z)^{-1/2}+O\big[(x-z)^{-\beta-1/2}\big]h^\beta,\end{aligned}$$

which implies the second estimate. △

Using (20), the variance of the estimator is calculated in a very simple way:

$$\begin{aligned}\operatorname{Var}\widehat{DG}(0,\mu) &\le \frac{1}{\pi n\mu^{1/\beta}}\left\{\int\limits_0^{\mu^{1/(2\beta)}} g(x)\left[\int\limits_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^2 dx\right.\\ &\qquad\left.+\int\limits_{\mu^{1/(2\beta)}}^{\infty} g(x)\left[\int\limits_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^2 dx\right\}\\ &\le O(n^{-1})+\frac{1}{\pi n}\int\limits_{\mu^{1/(2\beta)}}^{\infty} g(x)x^{-1}dx+O\left[\frac{\mu^{1/2}}{\pi n}\int\limits_{\mu^{1/(2\beta)}}^{\infty} g(x)x^{-1-2\beta}dx\right]\\ &= -[1+o(1)]g(0)n^{-1}(2\pi\beta)^{-1}\log\mu. \end{aligned}\tag{21}$$

Proceeding in just the same way, we obtain that for any integer $p > 1$

$$\begin{aligned}&\int\limits_0^{\infty} g(x)\left[\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^{2p} dx\\ &= \int\limits_0^{\mu^{1/(2\beta)}} g(x)\left[\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^{2p} dx\\ &\quad+\int\limits_{\mu^{1/(2\beta)}}^{\infty} g(x)\left[\frac{1}{\mu^{1/(2\beta)}}\int\limits_0^x K\left(0,\frac{y}{\mu^{1/(2\beta)}}\right)\frac{1}{\sqrt{x-y}}dy\right]^{2p} dx \le O\Big[\mu^{(1-p)/(2\beta)}\Big]. \end{aligned}\tag{22}$$

Now, we can easily prove the following statement.

Lemma 2. *Assume that* $\mu \asymp n^{-2\beta/(2\beta-1)}$. *Then for any integer* $p \ge 1$, *the inequality*

$$\mathbf{E}\Big[\widehat{DG}(0,\mu)-DG(0)\Big]^{2p} \le C_p\left[\frac{g(0)\log n}{n\pi(2\beta-1)}\right]^p \tag{23}$$

holds uniformly in $F \in B_\varepsilon(F_0)$.

Proof. Note that, given μ, the square of the estimator bias has order $O(n^{-1})$ according to (18), which is significantly less than the variance, the latter being of order $O(n^{-1}\log n)$ by (21). Therefore, the assertion follows from (18), (21), (22), and the inequality for the mathematical expectation of sums of independent random variables raised to the power $2p$ (see[10, p. 79]). △

The mean-square risk of the estimator $\widehat{DG}(y,\mu)$ on $[0,\infty)$ is calculated in the following way.

Lemma 3. *Assume that* $\mu \asymp n^{-1}$. *Then, uniformly in* $F \in B_\varepsilon(F_0)$, *we have*

$$\mathbf{E}\left\|\widehat{DG}(\,\cdot\,,\mu)-DG(\cdot)\right\|^2 \le [1+o(1)]\frac{\log n}{2n\pi\beta}.$$

Proof. Let us calculate the bias of the estimator $\widehat{DG}(y,\mu)$. Integrating by parts yields (see (18))

$$\mathbf{E}\widehat{DG}(y,\mu)-DG(y) = \mu^{1/2-1/(2\beta)}\int\limits_0^{\infty} K_x^{(\beta)}\left(\frac{y}{\mu^{1/(2\beta)}},\frac{x}{\mu^{1/(2\beta)}}\right)DG^{(\beta)}(x)dx.$$

From this, using the obvious change of variables, we get

$$\left\|\mathbf{E}\widehat{DG}(\,\cdot\,,\mu)-DG(\cdot)\right\|^2 \le \mu\left\|DG^{(\beta)}\right\|^2 \sup_{\|u\|\le 1}\int\limits_0^\infty\left[\int\limits_0^\infty K_x^{(\beta)}(y,x)u(x)\,dx\right]^2 dy. \tag{24}$$

Then we obtain

$$\sup_{\|u\|\le 1}\int\limits_0^\infty\left[\int\limits_0^\infty K_x^{(\beta)}(y,x)u(x)\,dx\right]^2 dy \le \sup_{\|u\|\le 1}\inf_g\left\{\|g-u\|^2+\|g^{(\beta)}\|^2\right\}\le 1.$$

Therefore, from this and (24), we derive the inequality

$$\left\|\mathbf{E}\widehat{DG}(\,\cdot\,,\mu)-DG(\cdot)\right\|^2 \le \mu\left\|DG^{(\beta)}\right\|^2. \tag{25}$$

Now, applying Lemma 1, calculate the variance of the estimator. Setting for brevity $h=\mu^{1/(2\beta)}$, we have

$$\begin{aligned}
\mathbf{E}\Big[\mathbf{E}\widehat{DG}(y,\mu)-\widehat{DG}(y,\mu)\Big]^2 &\le \frac{1}{\pi n}\int\limits_0^\infty g(x)\left[\int\limits_0^x \frac{K(y,z,h)}{\sqrt{x-z}}dz\right]^2 dx\\
&= \frac{1}{\pi n}\left(\int\limits_0^{y+h}+\int\limits_{y+h}^\infty\right) g(x)\left[\int\limits_0^x \frac{K(y,z,h)}{\sqrt{x-z}}dz\right]^2 dx\\
&\le \frac{C}{hn}\int\limits_0^{y+h} g(x)e^{-2C_\beta|x-y|/h}dx+\frac{1}{\pi n}\int\limits_{y+h}^\infty g(x)\Big[(x-y)^{-1/2}+h^\beta(x-y)^{-1/2-\beta}\Big]^2dx\\
&\le \frac{1}{\pi n}\int\limits_{y+h}^\infty \frac{g(x)}{x-y}dx+\frac{Ch^\beta}{\pi n}\int\limits_{y+h}^\infty \frac{g(x)}{(x-y)^{1+\beta}}dx+\frac{C}{hn}\int\limits_0^{y+h} g(x)e^{-2C_\beta|x-y|/h}dx.
\end{aligned}$$

Integrating this inequality with respect to y and using inequality (25), we complete the proof of Lemma 3. △

We still need a simple but important fact proved similarly to previous lemmas.

Lemma 4. *Assume that $\mu_1\asymp n^{-1}$ and $\mu_2\asymp n^{-2\beta/(2\beta-1)}$. Then, uniformly in $F\in B_\varepsilon(F_0)$, we have*

$$\mathbf{E}\int\limits_0^\infty DG(y)\Big[\widehat{DG}(y,\mu_1)-DG(y)\Big]\Big[\widehat{DG}(0,\mu_2)-DG(0)\Big]dy=O(n^{-1}).$$

Proof. From (18), (25), and the Cauchy–Bunyakovskii inequality, one directly obtains

$$\int\limits_0^\infty DG(y)\Big[\mathbf{E}\,\widehat{DG}(y,\mu_1)-DG(y)\Big]\Big[\mathbf{E}\,\widehat{DG}(0,\mu_2)-DG(0)\Big]dy=O(n^{-1}). \tag{26}$$

On the other hand, denoting for brevity $h_1 = \mu_1^{1/(2\beta)}$ and $h_2 = \mu_2^{1/(2\beta)}$, we have

$$\int_0^\infty DG(y)\Big[\mathbf{E}\,\widehat{DG}(y,\mu_1) - \widehat{DG}(y)\Big]\Big[\mathbf{E}\,\widehat{DG}(0,\mu_2) - \widehat{DG}(0)\Big]dy$$

$$= \frac{1}{n}\int_0^\infty DG(y)\int_0^\infty g(x)\int_0^x\int_0^x \frac{K(y,z,h_1)K(0,t,h_2)}{\sqrt{x-z}\sqrt{x-t}}dz\,dt\,dx\,dy$$

$$+ \frac{1}{n}\int_0^\infty DG(y)\int_0^x\int_0^x K(y,z,h_1)K(0,t,h_2)DG(t)DG(z)\,dz\,dt\,dy. \tag{27}$$

Then, applying Lemma 1, we obtain

$$\int_0^\infty g(x)\frac{K(y,z,h_1)K(0,t,h_2)}{\sqrt{x-z}\sqrt{x-t}}dz\,dt\,dx$$

$$= \left[\int_0^{h_2} + \int_{h_2}^{y+h_1} + \int_{y+h_1}^\infty\right] g(x)\int_0^x\int_0^x \frac{K(y,z,h_1)K(0,t,h_2)}{\sqrt{x-z}\sqrt{x-t}}dz\,dt\,dx$$

$$\le \int_{y+h_1}^\infty g(x)\left(\frac{1}{\sqrt{x}} + \frac{h_2^\beta}{x^{1/2+\beta}}\right)\left(\frac{1}{\sqrt{x-y}} + \frac{h_1^\beta}{(x-y)^{1/2+\beta}}\right)dx$$

$$+ \frac{C}{\sqrt{h_1}}\int_{h_2}^\infty g(x)e^{-C_\beta|y-x|/h_1}\left(\frac{1}{\sqrt{x}} + \frac{h_2^\beta}{x^{1/2+\beta}}\right)dx + Cg(0)h_2^{1/2}h_1^{-1/2}e^{-C_\beta y/h_1}.$$

Multiplying this inequality by $DG(y)$ and integrating with respect to y over $[0,\infty)$, we obtain that the first term on the right-hand side of (27) has order $O(n^{-1})$. The second term is obviously bounded by the same quantity. $\triangle$

Proof of Theorem 1. Denote by A the following subset in $\mathbb{R}^n$:

$$A = \Big\{\mathbf{X}^n : \ \Big|\widehat{DG}(0,\mu_2) - DG(0)\Big| \le DG(0)/\log n\Big\}.$$

By Lemma 2, the probability of the event that the observations belong to A is close to unity. More precisely, applying the Chebyshev inequality and using (23), for any $p > 1$ we have

$$\mathbf{P}\Big\{\mathbf{X}^n \notin A\Big\} \le C(p)\left(\frac{\log^2 n}{n}\right)^p. \tag{28}$$

Represent the risk of the estimator $\widehat{F}_{\mu_1,\mu_2}(y,\mathbf{X}^n)$ as follows:

$$R(\widehat{F}_{\mu_1,\mu_2}, F) = \mathbf{E}\left\|\widehat{F}_{\mu_1,\mu_2} - F\right\|^2 \mathbf{1}\{A\} + \mathbf{E}\left\|\widehat{F}_{\mu_1,\mu_2} - F\right\|^2 \mathbf{1}\{A^c\}, \tag{29}$$

where A^c is the complement of A. Since F is a distribution function, we have

$$\mathbf{E}\left\|\widehat{F}_{\mu_1,\mu_2} - F\right\|^2 \mathbf{1}\{A\} \le \mathbf{E}\left\|F - 1 + \frac{n\widehat{DG}(\,\cdot\,,\mu_1)}{n|\widehat{DG}(0,\mu_2)| + 1}\right\|^2 \mathbf{1}\{A\}. \tag{30}$$

To continue this inequality, let us apply the Taylor formula under the condition that $\mathbf{X}^n \in A$. Then

$$1 - \frac{n\widehat{DG}(\,\cdot\,,\mu_1)}{n|\widehat{DG}(0,\mu_2)|+1} = 1 - \frac{DG(y)}{DG(0)} - \frac{\widehat{DG}(y,\mu_1) - DG(y)}{DG(0)} + (1+o(1))\frac{DG(y)}{DG(0)}\frac{\widehat{DG}(0,\mu_2) - DG(0) + n^{-1}}{DG(0)}.$$

From here, taking into account (30) and Lemmas 2, 3, and 4, we obtain

$$\mathbf{E}\left\|\widehat{F}_{\mu_1,\mu_2} - F\right\|^2 \mathbf{1}\{A\} \le \frac{(1+o(1))\log n}{\pi DG^2(0)n}\left(\frac{1}{2\beta} + \frac{g(0)\|1-F\|^2}{2\beta-1}\right). \tag{31}$$

Since the probability $\mathbf{P}\Big\{\mathbf{X}^n \notin A\Big\}$ is very small, one can calculate the second term in (29) using rough inequalities. Note that, under the condition $y > \max\limits_{i=1,\ldots,n} X_i$, for the estimator $\widehat{DG}(y,\mu_1)$ we have the following relation (see (20)):

$$\widehat{DG}(y,\mu_1) \le C\mu_1^{-1/(2\beta)} \exp\left[-C_\beta \mu_1^{-1/(2\beta)}\left(y - \max_{i=1,\ldots,n} X_i\right)\right].$$

Hence,

$$\widehat{F}_{\mu_1,\mu_2}(y,\mathbf{X}^n) \ge 1 - Cn\mu_1^{-1/(2\beta)} \exp\left[-C_\beta \mu_1^{-1/(2\beta)}\left(y - \max_{i=1,\ldots,n} X_i\right)\right]$$

and, therefore, for any distribution function F, the inequality

$$\begin{aligned}\left\|F - \widehat{F}_{\mu_1,\mu_2}\right\|^2 &\le 2\|1-F\|^2 + 2\left\|1 - \widehat{F}_{\mu_1,\mu_2}\right\|^2 \\ &\le 2\|1-F\|^2 + 2\max_{i=1,\ldots,n} X_i + 2\int\limits_{\max\limits_{i=1,\ldots,n} X_i}^{\infty}\left[1 - \widehat{F}_{\mu_1,\mu_2}(y,\mathbf{X}^n)\right]^2 dy \\ &\le 2\max_{i=1,\ldots,n} X_i + 2\|1-F\|^2 + Cn^2\mu_1^{-1/(2\beta)}\end{aligned} \tag{32}$$

is valid.

Note further that according to (6) and (4), we have

$$\mathbf{E}X_i^{1+\varepsilon} \le C\int\limits_0^\infty\int\limits_0^y \frac{x^{1+\varepsilon}}{\sqrt{y-x}}dx\,dF(y) \le \int\limits_0^\infty y^{3/2+\varepsilon}dF(y) \le C.$$

Thus, $\mathbf{E}\max\limits_{i=1,\ldots,n} X_i^{1+\varepsilon} < Cn$. Therefore, choosing $p = p(\varepsilon)$ sufficiently large and using the Hölder inequality, we obtain

$$\mathbf{E}\left\|\widehat{F}_{\mu_1,\mu_2} - F\right\|^2 \mathbf{1}\{A^c\} \le O(n^{-1}). \tag{33}$$

Finally, taking into account (5), we have

$$DG(0) = \frac{\sqrt{\pi}}{2}\left[\int\limits_0^\infty \sqrt{y}\,dF(y)\right]^{-1},$$

and, therefore, relations (29), (31), and (33) imply the required inequality, (9). △

4. LOWER BOUND

To prove Theorem 2, it remains to check the inequality

$$r(F_0) \geq \sigma^2(F_0). \tag{34}$$

We prove this inequality by applying the technique proposed in [6, 11] and based on the van Trees inequality [12]. As was already mentioned, the problem of reconstruction of the distribution function can conditionally be divided into two subproblems: estimation of the derivative of order 1/2 at the origin and estimation of the same derivative on the whole interval $[0,\infty)$. Therefore, we will prove the lower bound (34) according to this division. First of all, consider the contribution to the minimax risk made by the estimation of $DG(0)$. To do this, we apply the standard technique [13] by constructing a univariate parametric family.

Let $Q(x) \geq 0$ be a function with support on $[-1/2, 1/2]$ such that

$$\int_{-\infty}^{\infty} Q(x)dx = 1, \qquad \sup_{x\in(-\infty,\infty)} |Q^{(\beta-1)}(x)| < \infty.$$

As such a function, we can choose, for example, the function

$$Q(x) = \begin{cases} C(\beta)(1/2-x)^{\beta-1}(x+1/2)^{\beta-1}, & |x| \leq 1/2 \\ 0, & |x| > 1/2, \end{cases}$$

where $C(\beta)$ is the normalizing factor.

Define the function

$$\psi_h(x) = \frac{1}{h}\int_h^{\delta} \frac{1}{y} Q\left(\frac{y-x}{h}\right) dy,$$

where $h < \delta$. Further we assume that δ is rather small and does not depend on the sample size n. Properties of this function used below are collected in the following lemma.

Lemma 5. *For $h \to 0$, we have*

$$\sup_x \psi_h(x) \leq O(h^{-1}), \tag{35}$$

$$\int_0^y \psi_h(x)dx = -(1+o(1))\log h, \quad y > \delta + h, \tag{36}$$

$$\int_0^{\infty} \sqrt{x}\psi_h(x)dx = O(1), \tag{37}$$

$$\|\psi_h^{(\beta-1)}(x)\| \leq O(h^{-\beta+1/2}). \tag{38}$$

The proof is based on elementary calculations and the following formula for the lth derivative of $\psi_h(x)$:

$$\psi_h^{(l)}(x) = \begin{cases} O(h^{-l-1}), & x \leq h, \\ l!\,x^{-l-1} + O(h)x^{-l-2}, & x > h, \end{cases}$$

and therefore is omitted.

Define the family of distribution functions depending on an unknown parameter $\theta \in [0,1]$ in the following way:

$$F_\theta^n(x) = \left(F_0(x) + \theta n^{-1/2} \int\limits_0^x \psi_h(y)dy\right) \Bigg/ \left(1 + \theta n^{-1/2} \int\limits_0^\infty \psi_h(y)dy\right),$$

where $F_0(x)$ is a distribution function. Lemma 5 directly implies the statement below.

Lemma 6. *Assume that* $\|F_0\|_\beta < \infty$ *and*

$$h = \left(\frac{\log n}{n}\right)^{1/(2\beta-1)}. \tag{39}$$

Then $\|F_\theta^n - F_0\|_\beta < \varepsilon/2$ *for all sufficiently large* n.

Let us observe another simple fact easily obtained by means of (36) and the Taylor formula: for $x > \delta + h$, we have the equality

$$F_\theta^n(x) = F_0(x) + [1 + O(n^{-1/2}\log h)][1 - F_0(x)]\frac{\theta}{\sqrt{n}} \int\limits_0^\infty \psi_h(y)dy. \tag{40}$$

In particular, this fact implies the following lower bound for estimation of the distribution function $F_\theta^n(x)$:

$$n \inf_{\widehat{F}} \sup_\theta \mathbf{E}_\theta \int\limits_{h+\delta}^\infty \left|\widehat{F}(x) - F_\theta^n(x)\right|^2 dx$$
$$\geq (1 + o(1)) \log^2 h \inf_{\widehat{\theta}} \sup_\theta \mathbf{E}_\theta (\widehat{\theta} - \theta)^2 \int\limits_{h+\delta}^\infty [1 - F_0(x)]^2 dx; \tag{41}$$

here, $\mathbf{E}_\theta$ is averaging over the measure generated by n independent random variables with the density (see (4))

$$g_\theta(x) = \frac{1}{2} \int\limits_x^\infty \frac{dF_\theta^n(y)}{\sqrt{y-x}} \Bigg/ \int\limits_0^\infty \sqrt{y}\, dF_\theta^n(y). \tag{42}$$

Therefore, to continue inequality (41), it remains to compute the Fisher information

$$I^n(\theta) = \int\limits_0^\infty \frac{[g_\theta'(x)]^2}{g_\theta(x)} dx.$$

In this formula and below, $g_\theta'(x)$ denotes the derivative with respect to θ. Actually, the only nontrivial point in the computation of this quantity is calculation of the derivative of order $-1/2$ for the function $\psi_h(x)$. This derivative is computed in the following lemma.

Lemma 7. *For* $h \to 0$, *we have the following relations:*

$$\int\limits_x^\infty \frac{\psi_h(y)}{\sqrt{y-x}} dy = \begin{cases} 0, & x > \delta, \\ \pi x^{-1/2} + O\left(\delta^{-1/2}\right), & h \leq x \leq \delta, \\ O\left(h^{-1/2}\right), & 0 \leq x \leq h. \end{cases}$$

The proof of this result becomes obvious if one uses the formula

$$\int_u^\infty \frac{1}{t\sqrt{t-u}}\,dt = \pi \mathbf{1}\{u \ge 0\} u^{-1/2}. \quad \triangle$$

It is not difficult to estimate the Fisher information $I^n(\theta)$ applying Lemmas 5 and 6.

Lemma 8. *Assume that* $\|F_0\|_\beta < \infty$ *and* h *is defined in* (39). *Then, for* $n \to \infty$, *we have*

$$I^n(\theta) \le (1+\rho(\delta))\pi^2 g_0^{-1}(0)\left(2\int_0^\infty \sqrt{y}\,dF_0(y)\right)^{-2} n^{-1}\log h^{-1},$$

where $\rho(\delta) \to 0$ *as* $\delta \to 0$.

Proof. Using (42), the Taylor formula, and Lemmas 5 and 7, it is not difficult to verify that for $n \to \infty$ we have

$$g_\theta(x) = \frac{1+o(1)}{2}\int_x^\infty \frac{dF_0(y)}{\sqrt{y-x}} \Big/ \int_0^\infty \sqrt{y}\,dF_0(y). \tag{43}$$

Similarly, we get

$$\sqrt{n}g_\theta'(x) = \frac{1+o(1)}{2}\int_x^\infty \frac{\psi_h(y)}{\sqrt{y-x}}dy \Big/ \int_0^\infty \sqrt{y}\,dF_0(y)$$
$$- \frac{1+o(1)}{2}\int_x^\infty \frac{dF_0(y)}{\sqrt{y-x}}\left(\int_0^\infty \sqrt{y}\,dF_0(y)\right)^{-2}\int_0^\infty \sqrt{y}\psi_h(y)dy. \tag{44}$$

The contribution of the second term on the right-hand side of (44) to the Fisher information is of order $O(n^{-1})$ since $DF_0 \in \mathbf{L}_2(0,\infty)$ and the latter integral is bounded (see (37)).

To calculate the main component of the Fisher information, it suffices to note that $D\psi_h(x)$ does not vanish in a δ-neighborhood of the origin only and $DF_0(x)$ is continuous at the origin. More precisely, $|DF_0(x) - DF_0(0)| < C\sqrt{x}\max_y F_0'(y)$. Therefore, noting that, by Lemma 7,

$$\int_0^\infty \left(\int_x^\infty \frac{\psi_h(y)}{\sqrt{y-x}}dy\right)^2 dx = (1+o(1))\pi^2 \log h^{-1},$$

and taking into account (43) and (44), we obtain the required inequality. $\triangle$

Now, we can continue lower bound (41) using the van Trees inequality (see [12, p. 83]). Assume that the parameter θ is a random variable with probability density $\pi(x)$, $x \in [0,1]$. Moreover, we also assume that the Fisher information

$$I = \int_0^1 \frac{\pi'^2(x)}{\pi(x)}dx$$

is finite. Then

$$\inf_{\widehat\theta}\sup_{\theta\in[0,1]} \mathbf{E}_\theta\,(\widehat\theta - \theta)^2 \ge \inf_{\widehat\theta} \mathbf{E}\,\mathbf{E}_\theta\,(\widehat\theta-\theta)^2 \ge (nI^n(\theta)+I)^{-1}.$$

Then, since δ is arbitrary, we obtain by Lemma 8 and (41) that

$$n \inf_{\widehat{F}} \sup_{\theta} \mathbf{E}_\theta \int_0^\infty \left|\widehat{F}(x) - F_\theta^n(x)\right|^2 dx$$
$$\geq (1+o(1)) \frac{4}{\pi^2(2\beta-1)} \left[\int_0^\infty \sqrt{y}\, dF_0(y)\right]^2 \|1 - F_0\|^2 \log n. \qquad (45)$$

A significant contribution to the lower bound is also made by estimation of the function $DG(y)$ for $y \in [0,\infty)$. We take this fact into account by means of introducing a new parametric family. Put

$$\varphi_{2k}^T(x) = \sqrt{\frac{2}{T}} \cos\left(\frac{2\pi k x}{T}\right), \qquad \varphi_{2k+1}^T(x) = \sqrt{\frac{2}{T}} \sin\left(\frac{2\pi k x}{T}\right)$$

and define the normalized trigonometric polynomial $P_\nu(x,W)$ of order W in the following way:

$$P_\nu(x,W) = \frac{1}{\sqrt{n}} \sum_{k=\log W}^{W} a_k \nu_k \varphi_k^T(x),$$

where $a_k = (2\pi\lfloor k/2\rfloor/T)^{1/2}$; here, $\lfloor y\rfloor$ denotes the integral part of y and the quantity $T \geq 1$ does not depend on n. Along with asymptotics as $n \to \infty$, we are interested in the case where T is large. Denote for brevity

$$\chi_T(x) = \frac{1}{\delta} \int_{-\infty}^{T+5/2\delta} Q\left(\frac{x-y}{\delta}\right) dy.$$

This function is a smooth approximation of $\mathbf{1}\{t \leq T + 5/2\delta\}$.

Below, we consider the following parametric family of distribution functions depending on unknown parameters θ and ν_k, $k \in [\log W, W]$:

$$F_{\theta,\nu}^n(x) = \frac{F_\theta^n(x) + \int_0^x \chi_T(u) P_\nu(u,W)\,du + \delta^{2\beta} \int_0^x \chi_{T+\delta}(u)\,du}{1 + \int_0^\infty \chi_T(u) P_\nu(u,W)\,du + \delta^{2\beta} \int_0^\infty \chi_{T+\delta}(u)\,du}.$$

Lemma 9. *Assume that $\nu_k^2 \leq \log n$ and $W = (n/\log^2 n)^{1/(2\beta)}$. Then $\|F_{\theta,\nu}^n - F_0\|_\beta \leq \varepsilon$ for all sufficiently large n and sufficiently small δ.*

Proof. First of all, let us check that $F_{\theta,\nu}^n(x)$ is a distribution function. To this end, it suffices to show that

$$\inf_{x\geq 0}\Big[\chi_T(x) P_\nu(x,W) + \delta^{2\beta} \chi_{T+\delta}(x)\Big] \geq 0.$$

Since $\chi_T(x) = 0$ for $x \geq T + 3\delta$ and $\chi_{T+\delta}(x) = 1$ for $x \leq T + 3\delta$, it suffices to verify that $\sup_{x\geq 0} |P_\nu(x,W)|\delta^{-2\beta} < 1$. This inequality is fulfilled for sufficiently large n, since for $\beta \geq 2$ we have

$$\sup_{x\geq 0} |P_\nu(x,W)| \leq \left(\frac{\log n}{n}\right)^{1/2} W^{3/2} \leq n^{-1/8} \log^{1/2} n.$$

Therefore, it remains to check that the functions $F_{\theta,\nu}^n(x)$ are smooth. Indeed, for any $m \leq \beta - 1$, we have

$$\int_0^{2T} [P_\nu^{(m)}(x,W)]^2 dx = \frac{2}{n} \sum_{k=\log W}^{W} \nu_k^2 a_k^{2m+1} \leq O\left(\frac{W^{2m+2}\log n}{n}\right).$$

Then, observing that $\|\chi_{T+\delta}^{(2m)}\| = \|\chi_T^{(2m)}\| \le C\delta^{1-2m}$ for all $0 < m \le \beta - 1$, we obtain

$$\int_0^\infty \left[\frac{d^{\beta-1}}{dx^{\beta-1}} P_\nu(x, W)\chi_T(x)\right]^2 dx \le C\delta^{2\beta-2}W^2 n^{-1}\log n + CTW^{2\beta}n^{-1}\log n \le O(\log^{-1} n),$$

which obviously completes the proof of the lemma. △

For subsequent calculations, we shall use the Taylor formula for the parametric family $F_{\theta,\nu}^n(x)$. Denote for brevity

$$F_0^{\delta,T}(x) = \left(F_0(x) + \delta^{2\beta}\int_0^x \chi_{T+1}(u)du\right) \Big/ \left(1 + \delta^{2\beta}\int_0^\infty \chi_{T+1}(u)du\right).$$

Omitting elementary computations, for $x > \delta + h$ we arrive at the following formula:

$$F_{\theta,\nu}^n(x) = F_0^{\delta,T}(x) + (1 + O(\delta))(1 - F_0(x))\frac{\theta}{\sqrt{n}}\int_0^\infty \psi_h(u)du$$
$$+ (1 + O(\delta)\left[\int_0^x \chi_T(u)P_\nu(u, W)du - F_0^{\delta,T}(x)\int_0^\infty \chi_T(u)P_\nu(u, W)du\right]. \qquad (46)$$

Lemma 10. *Under the conditions of Lemma* 9, *we have*

$$n\inf_{\widehat{F}}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}\int_0^\infty \left|\widehat{F}(x) - F_{\theta,\nu}^n(x)\right|^2 dx$$
$$\ge (1 + O(\delta))\|1 - F_0\|^2\log^2 h\inf_{\widehat{\theta}}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}|\widehat{\theta} - \theta|^2$$
$$+ (1 + O(\delta))\inf_{\widehat{\nu}_k}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}\sum_{k=\log W}^{W}(\widehat{\nu}_k - \nu_k)^2 a_k^{-2} + O(1). \qquad (47)$$

Proof. Taking into account that $\chi_T(x) = 1$ for $x \in [2\delta, T + 2\delta]$, let us use the trivial inequality

$$\int_0^\infty \left|\widehat{F}(x) - F_{\theta,\nu}^n(x)\right|^2 dx \ge \int_{2\delta}^{T+2\delta} \left|\widehat{F}(x) - F_{\theta,\nu}^n(x)\right|^2 dx.$$

Then, substituting (46) into the inequality above and estimating the interference terms by means of integrating by parts, we arrive at the required inequality. △

Proof of Theorem 2. A lower bound for the first term on the right-side of (47) is already obtained in (45). To estimate the second term, put $W = (n/\log^2 n)^{1/(2\beta)}$ (see Lemma 9), divide the index set $\{\log W, \dots, W\}$ into blocks B_s of length $\log\log W$, and lower bound the risk over the block B_s

$$\rho_s = \inf_{\widehat{\nu}_k}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}\sum_{k\in B_s}(\widehat{\nu}_k - \nu_k)^2.$$

Moreover, the parameter θ is assumed to be known. Then, assuming that the quantities ν_k are independent random variables whose Fisher information has order $\log^{-1} n$, apply the van Trees inequality [12]. Note that, by (4), for $n \to \infty$ we have

$$\begin{aligned}\sqrt{n}\frac{\partial g_{\theta,\nu}(x)}{\partial \nu_k} &= \frac{1}{2}\frac{\partial}{\partial \nu_k}\int\limits_x^\infty \frac{dF^n_{\theta,\nu}(y)}{\sqrt{x-y}}dy \Big/ \int\limits_0^\infty \sqrt{y}\,dF^n_{\theta,\nu}(y) \\ &= (1+o(1))\frac{a_k}{2}\int\limits_x^\infty \frac{\chi_T(y)\varphi_k^T(y)}{\sqrt{y-x}}dy \Big/ \int\limits_0^\infty \sqrt{y}\,dF_0^{\delta,T}(y) \\ &- (1+o(1))\frac{a_k}{2}\int\limits_x^\infty \frac{dF_0^{\delta,T}(y)}{\sqrt{x-y}}dy\left[\int\limits_0^\infty \sqrt{y}\,dF_0^{\delta,T}(y)\right]^{-2}\int\limits_0^\infty \sqrt{y}\chi_T(y)\varphi_k^T(y)dy. \end{aligned} \tag{48}$$

Thus, we arrive at the inequality

$$\rho_s \geq \operatorname{tr}(I + \log^{-1} n E)^{-1}; \tag{49}$$

here, E is the identity matrix and I is the mean value of the Fisher information matrix

$$I_{kj} = \mathbf{E}\int\limits_0^{T+3\delta} g^{-1}_{\theta,\nu}(x)\frac{\partial g_{\theta,\nu}(x)}{\partial \nu_k}\frac{\partial g_{\theta,\nu}(x)}{\partial \nu_j}dx.$$

Denote for brevity

$$s_k(x) = a_k\int\limits_x^{T+3\delta}\frac{\varphi_k^T(y)\chi_T(y)}{\sqrt{y-x}}dy.$$

To compute these functions, let us use the identity

$$\int\limits_x^\infty \frac{\exp(i\lambda t)}{\sqrt{t-x}}dt = (\pi/\lambda)^{1/2}e^{i(\lambda x - \pi/4)},$$

or, equivalently,

$$\int\limits_x^\infty \frac{\varphi_k^T(x)}{\sqrt{t-x}}dt = \sqrt{\pi}a_k^{-1}\varphi_k^T\left(x - \frac{T}{4k}\right).$$

Therefore, integrating by parts, we obtain, for $x \le T$, that

$$\begin{aligned} s_k(x) &= \sqrt{\pi}\varphi_k^T\left(x-\frac{T}{4k}\right) + a_k\int\limits_x^\infty \frac{\varphi_k^T(y)[1-\chi_T(y)]}{\sqrt{y-x}}dx \\ &= \sqrt{\pi}\varphi_k^T\left(x-\frac{T}{4k}\right) + \varepsilon_k(x), \end{aligned} \tag{50}$$

where

$$\varepsilon_k(x) = \sqrt{\frac{2}{T}}O\left(\frac{1}{\sqrt{T+3\delta-x}}\right)a_k^{-1}.$$

Note also that $|\varepsilon_k(x)| \le T^{-1/2}$ for $x \in [T, T+3\delta]$. From this, we immediately get

$$\lim_{n\to\infty}\sum_{s\in B_s}\int\limits_0^\infty \frac{\varepsilon_k^2(x)}{g_{\theta,\nu}(x)}dx = 0.$$

Then, again integrating by parts, we note that the relation

$$\int_0^\infty \sqrt{y}\chi_T(y)\varphi_k^T(y)dy = O(a_k^{-2})$$

holds. Therefore, taking into account two latter relations, formula (48), and Lemma 12 (see Appendix), we obtain

$$\operatorname{tr}(I + \log^{-1} nE)^{-1} = (1+o(1))4\pi^{-1}\left[\int_0^\infty \sqrt{y}\, dF_0^{\delta,T}(y)\right]^2 \operatorname{tr} J^{-1}, \tag{51}$$

where

$$\begin{aligned} J_{jk} &= \mathbf{E}\int_0^{T+3\delta} g_{\theta,\nu}^{-1}(x)\varphi_k^T\left(x - \frac{T}{4k}\right)\varphi_j^T\left(x - \frac{T}{4j}\right)dx \\ &= \mathbf{E}\, T\int_0^{1+3\delta/T} g_{\theta,\nu}^{-1}(xT)\varphi_k^T\left(xT - \frac{T}{4k}\right)\varphi_j^T\left(xT - \frac{T}{4j}\right)dx. \end{aligned}$$

To lower bound the quantity $\operatorname{tr} J^{-1}$, let us use inequality (60) (see Appendix). We have

$$\begin{aligned} \operatorname{tr} J^{-1} &\geq (1+o(1))\log\log W\int_{3\delta/T}^1 \left[\mathbf{E}\, g_{\theta,\nu}^{-1}(xT)\right]^{-1}dx \\ &\geq (1+o(1))T^{-1}\log\log W\int_{3\delta}^T g_0(x)dx. \end{aligned} \tag{52}$$

Hence, from (49), (51), and (52), we derive

$$\begin{aligned} &\inf_{\widehat{\nu}_k}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}\sum_{k=\log W}^{W}(\widehat{\nu}_k - \nu_k)^2 a_k^{-2} \\ &\geq \sum_{s=1}^{W/\log W} a_{\log W + s\log\log W}^{-2}\inf_{\widehat{\nu}_k}\sup_{\theta,\nu}\mathbf{E}_{\theta,\nu}\sum_{k\in B_s}(\widehat{\nu}_k - \nu_k)^2 \\ &\geq (1+o(1))\frac{4}{\pi^2}\int_{3\delta}^T g_0(x)dx\left[\int_0^\infty \sqrt{y}\, dF_0^{\delta,T}(y)\right]^2\sum_{s=1}^{W/\log W}\frac{\log\log W}{\log W + s\log\log W} \\ &\geq (1+o(1))\frac{4}{\pi^2}\int_{3\delta}^T g_0(x)dx\left[\int_0^\infty \sqrt{y}\, dF_0^{\delta,T}(y)\right]^2\log W. \end{aligned}$$

Noting that, for $\delta \to 0$ and $T \to \infty$, we have

$$\int_{3\delta}^T g_0(x)dx \to 1 \qquad \text{and} \qquad \int_0^\infty \sqrt{y}\, dF_0^{\delta,T}(y) \to \int_0^\infty \sqrt{y}\, dF_0(y),$$

we complete the proof of the theorem. △

APPENDIX

Here, some simple results about traces of the Fisher information inverse matrices of a special form are given. In principle, these results could be obtained by the standard moment method [14] but, using a statistical interpretation of the problem, we derive them by means of rather simple reasoning.

Denote for brevity

$$\varphi_{2k}(t) = \sqrt{2}\cos[2\pi(k+M)t + \alpha_k], \qquad \varphi_{2k+1}(t) = \sqrt{2}\sin[2\pi(k+M)t + \alpha_k].$$

Let A be an $m \times m$ matrix with entries

$$A_{ij} = \int_0^1 \frac{\varphi_i(t)\varphi_j(t)}{g^2(t)}dt;$$

here, $g(t)$ is a function such that $\int_0^1 g^{-2}(t)dt < \infty$ and M is an integer. We are interested in the behavior of $\operatorname{tr} A^{-1}$ for large m.

Lemma 11. *We have*

$$\frac{1}{m}\operatorname{tr} A^{-1} \le \int_0^1 g^2(t)dt \tag{53}$$

and, if $p(t)$ is a trigonometric polynomial of order Q such that $p^2(t) \le g^2(t)$, then

$$\frac{1}{m}\operatorname{tr} A^{-1} \ge \left(1 - \frac{8Q+4}{m}\right)\int_0^1 p^2(t)dt. \tag{54}$$

Proof. Let us use a statistical interpretation of the matrix A^{-1}. Assume that it is required to estimate a vector $(\theta_1, \ldots, \theta_m)^T$ from the observations

$$y(t) = \sum_{k=1}^m \theta_k\varphi_k(t) + g(t)n(t), \quad t \in [0,1],$$

where $n(t)$ is a standard white Gaussian noise. It is well known that, in the problem considered, the matrix A is a Fisher information matrix and therefore (see e.g., [13, pp. 112–114]) we have

$$\inf_{\widehat{\theta}_k}\sup_{\theta_k} \mathbf{E}\sum_{k=1}^m (\widehat{\theta}_k - \theta_k)^2 = \operatorname{tr} A^{-1}, \tag{55}$$

where $\mathbf{E}$ is averaging with respect to the measure generated by the observations $y(t)$ and the infimum is taken over all estimators.

To prove inequality (53), consider the estimator

$$\widehat{\theta}_k = \int_0^1 \varphi_k(t)y(t)dt = \theta_k + \int_0^1 \varphi_k(t)g(t)n(t)dt.$$

Clearly,

$$\mathbf{E}(\widehat{\theta}_k - \theta_k)^2 = \int_0^1 g^2(t)\varphi_k^2(t)dt.$$

Therefore, using the equality $\varphi_{2s+1}^2(t)+\varphi_{2s}^2(t)=2$ and formula (55), we obtain (53).

The proof of inequality (54) is more complicated. Consider new observations

$$\widetilde{y}(t)=\sum_{k=1}^{m}\theta_k\varphi_k(t)+p(t)\widetilde{n}(t),\quad t\in[0,1], \tag{56}$$

where $p(t)$ is a trigonometric polynomial of order Q such that $p^2(t)\le g^2(t)$ and $\widetilde{n}(t)$ is a white Gaussian noise independent of $n(t)$. Then the observations $y(t)$ can be represented as $y(t)=\widetilde{y}(t)+\sqrt{g^2(t)-p^2(t)}n_0(t)$, where $n_0(t)$ is a white Gaussian noise independent of $\widetilde{n}(t)$. It is obvious that adding an additional noise to observations can only enlarge the minimax risk; in other words, it is true that

$$\inf_{\widehat{\theta}_k}\sup_{\theta_k}\mathbf{E}\sum_{k=1}^{m}(\widehat{\theta}_k-\theta_k)^2\ge\inf_{\widehat{\theta}_k}\sup_{\theta_k}\widetilde{\mathbf{E}}\sum_{k=1}^{m}(\widehat{\theta}_k-\theta_k)^2, \tag{57}$$

where $\widetilde{\mathbf{E}}$ is averaging with respect to the measure generated by the observations $\widetilde{y}(t)$. To continue this inequality, multiply observations (56) by the functions φ_k and then integrate over the interval $[0,1]$. Then we arrive at the following model of observations:

$$\widetilde{y}_k=\widetilde{\theta}_k+\xi_k,\quad \widetilde{\theta}_k=\theta_{2k}+i\theta_{2k+1},\quad \xi_k=\sqrt{2}\int_0^1\exp[i(2\pi ikt+\alpha_k)]p^2(x)\widetilde{n}(t)dt,$$

where index k takes all values beginning with $-M$ and the quantities $\widetilde{\theta}_k$ do not vanish for $k\in[1,m/2]$ only. Then, note that $p^2(t)$ is a trigonometric polynomial of order $2Q$, which means independence of the Gaussian random variables ξ_k and ξ_s for $|k-s|>2Q$.

Since there are no any restrictions on $\widetilde{\theta}_k$, we can put

$$\widetilde{\theta}_k=A\sqrt{2}\int_0^1\exp[i(2\pi kt+\alpha_k)]p^2(t)n_1(t)dt,$$

where A is an arbitrary number and $n_1(t)$ is a white Gaussian noise independent of $\widetilde{n}(t)$. Since the quantities $\widetilde{\theta}_k$ are now Gaussian random variables, the optimal estimator is linear, $\widehat{\widetilde{\theta}}_k=\sum_s h_s\widetilde{y}_s$, where h_s is a solution of the equation

$$\mathbf{E}\left(\widetilde{\theta}_k-\sum_s h_s\widetilde{y}_s\right)\widetilde{y}_l=0.$$

From this, it is not difficult to obtain that, for $k\in[2Q+1,m-2Q-1]$, the estimator $\widehat{\widetilde{\theta}}_k$ is of the form $\widehat{\widetilde{\theta}}_k=A/(A+1)\widetilde{y}_k$ and its risk is equal to $\widetilde{\mathbf{E}}|\widehat{\widetilde{\theta}}_k-\widetilde{\theta}_k|^2=2A/(A+1)$. Therefore,

$$\inf_{\widehat{\theta}_k}\sup_{\theta_k}\widetilde{\mathbf{E}}\sum_{k=1}^{m}(\widehat{\theta}_k-\theta_k)^2\ge\sum_{k=2Q+1}^{m/2-2Q-1}\widetilde{\mathbf{E}}\left|\widehat{\widetilde{\theta}}_k-\widetilde{\theta}_k\right|^2\ge(m-8Q-4)A/(A+1)$$

since the minimax risk is lower bounded by the Bayesian risk. Since the quantity A is arbitrary, the above inequality and (57) imply (54). △

From Lemma 11, it is not difficult to obtain some useful facts about the behavior of $\operatorname{tr}A^{-1}/m$ as $m\to\infty$. For example, if $g^2(t)$ is strictly bounded away from zero and is continuous everywhere except for a finite number of points, then we have

$$\lim_{m\to\infty}\frac{1}{m}\operatorname{tr}A^{-1}=\int_0^1 g^2(t)dt. \tag{58}$$

This fact easily follows from the second Weierstrass theorem [15, p. 26] and the fact that the space of continuous functions on $[0,1]$ is dense in $\mathbf{L}_2[0,1]$.

Consider another simple example. Put

$$B_{ij} = \int_0^2 \frac{\varphi_i(t)\varphi_j(t)}{g^2(t)}dt;$$

i.e., B is the Fisher information matrix in the problem of estimating the parameters $\theta_1,\dots,\theta_m$ from the observations

$$y(t) = \sum_{k=1}^m \theta_k\varphi_k(t) + g(t)n(t), \quad t\in[0,2].$$

Since $\sum\limits_{k=1}^m \theta_k\varphi_k(t)$ is a periodic function of period 1, the observations $y(t)$ are equivalent to the following ones:

$$\bar y(t) = \sum_{k=1}^m \theta_k\varphi_k(t) + \bar g(t)n(t), \quad t\in[0,1], \quad \text{where} \quad \bar g(x) = \frac{g(t)g(t+1)}{\sqrt{g^2(t)+g^2(t+1)}}.$$

Therefore, traces of the Fisher information matrices in the problem of estimation from the observations $y(t)$, $t\in[0,2]$, and $\bar y(t)$, $t\in[0,1]$, coincide. If the function $g^2(t)$ is bounded away from zero on the segment $[0,2]$ and is continuous everywhere except for a finite number of points, then Lemma 11 implies

$$\lim_{m\to\infty}\frac{1}{m}\operatorname{tr}B^{-1} = \int_0^1 \bar g^2(t)dt. \tag{59}$$

Observe one more simple inequality. Define a matrix C by the equalities

$$C_{ij} = \int_0^{1+\rho} \frac{\varphi_i(t)\varphi_j(t)}{g^2(t)}dt,$$

where $\rho\in(0,1)$. If the function $g^2(t)$ is strictly bounded away from zero on the segment $[0,1+\rho]$ and is continuous everywhere except for a finite number of points, then we have

$$\lim_{m\to\infty}\frac{1}{m}\operatorname{tr}C^{-1} \ge \int_\rho^1 g^2(t)dt. \tag{60}$$

In certain statistical problems, the Fisher information matrix can slightly differ from the ideal matrix A. To compute the trace of the inverse matrix in this case, consider the "perturbed" $m\times m$ matrix

$$A_{ij}^\varepsilon = \int_0^1 \frac{[\varphi_i(t)+\varepsilon_i(t)][\varphi_j(t)+\varepsilon_j(t)]}{g^2(t)}dt,$$

where the quantities $\varepsilon_i(t)$ are some functions from a Hilbert space $\mathbf{L}_2^g(0,1)$ with the norm

$$\|f\|_g^2 = \int_0^1 \frac{f^2(t)}{g^2(t)}dt.$$

The following lemma estimates the trace of the matrix $[A^\varepsilon]^{-1}$ by that of A^{-1}.

Lemma 12. *We have*

$$\left|\operatorname{tr}[A^{\varepsilon}]^{-1} - \operatorname{tr} A^{-1}\right| \le 4m\left(\sum_{k=1}^{m}\|\varepsilon_k\|_g^2\right)^{1/2}.$$

Proof. Denote by P_{φ}, $\mathrm{P}_{\varphi+\varepsilon}$, and $\mathrm{P}_{\varphi,\varepsilon}$ the subspaces spanned by the functions $\{\varphi_1,\dots,\varphi_m\}$, $\{\varphi_1+\varepsilon_1,\dots,\varphi_m+\varepsilon_m\}$, and $\{\varphi_1,\varepsilon_1,\dots,\varphi_m,\varepsilon_m\}$ respectively. Denote also by Π_{φ}, $\Pi_{\varphi+\varepsilon}$, and $\Pi_{\varphi,\varepsilon}$ the projection operators onto these spaces respectively.

Let $n^g(t) = n(t)g(t)$ be a white Gaussian noise in $\mathbf{L}_2^g(0,1)$. Then, as is well known,

$$\begin{aligned} \operatorname{tr} A^{-1} &= \mathbf{E}\|\Pi_{\varphi} n^g\|_g^2 = \mathbf{E}\|\Pi_{\varphi}\Pi_{\varphi,\varepsilon} n^g\|_g^2, \\ \operatorname{tr}[A^{\varepsilon}]^{-1} &= \mathbf{E}\|\Pi_{\varphi+\varepsilon} n^g\|_g^2 = \mathbf{E}\|\Pi_{\varphi+\varepsilon}\Pi_{\varphi,\varepsilon} n^g\|_g^2. \end{aligned} \tag{61}$$

Note that for any elements $x, y, z \in \mathbf{L}_2^g(0,1)$ we have

$$\|x-y\|_g - \|x-z\|_g \le \|y-z\|_g.$$

Therefore, putting

$$d = \inf_{\substack{\|y\|_g\le 1 \\ y\in \mathrm{P}_{\varphi}}} \sup_{\substack{\|z\|_g\le 1 \\ z\in \mathrm{P}_{\varphi+\varepsilon}}} \|y-z\|_g,$$

we obtain that for any x such that $\|x\|_g = 1$, we have the following inequality:

$$\inf_{\substack{\|y\|_g\le 1 \\ y\in \mathrm{P}_{\varphi}}} \sup_{\substack{\|z\|_g\le 1 \\ z\in \mathrm{P}_{\varphi+\varepsilon}}} \left\{\|x-y\|_g - \|x-z\|_g\right\} = \left(1-\|\Pi_{\varphi}x\|_g^2\right)^{1/2} - \left(1-\|\Pi_{\varphi+\varepsilon}x\|_g^2\right)^{1/2} \le d.$$

Hence, for any $x \in \mathbf{L}_2^g(0,1)$, we have

$$\|\Pi_{\varphi+\varepsilon}x\|_g^2 - \|\Pi_{\varphi}x\|_g^2 \le 2d\|x\|^2 \tag{62}$$

and, by symmetry,

$$\|\Pi_{\varphi}x\|_g^2 - \|\Pi_{\varphi+\varepsilon}x\|_g^2 \le 2d\|x\|^2. \tag{63}$$

Note that the dimension of the subspace $\mathrm{P}_{\varphi,\varepsilon}$ is less than $2m$. Therefore $\mathbf{E}\|\Pi_{\varphi,\varepsilon}n^g\|_g^2 \le 2m$ and hence, from (61)–(63) we derive

$$\left|\operatorname{tr}[A^{\varepsilon}]^{-1} - \operatorname{tr} A^{-1}\right| \le 4md.$$

To complete the proof of the lemma, it suffices to note that the distance d can be estimated as follows:

$$d = \sup_{\substack{\|z\|_g\le 1 \\ z\in \mathrm{P}_{\varphi+\varepsilon}}} \inf_{\substack{\|y\|_g\le 1 \\ y\in \mathrm{P}_{\varphi}}} \|y-z\|_g \le \sup_{\sum v_k^2\le 1} \left\|\sum_{k=1}^{m} v_k\varepsilon_k\right\|_g \le \left(\sum_{k=1}^{m}\|\varepsilon_k\|_g^2\right)^{1/2}. \quad \triangle$$

In conclusion, we would like to express our sincere gratitude to the reviewer for valuable remarks, which helped us to improve this paper.

REFERENCES

1. Wicksell, S.D., The Corpuscle Problem: A Mathematical Study of a Biometric Problem, *Biometrica*, 1925, vol. 17, no. 1/2, pp. 84–99.

2. Stoyan, D., Kendall, W.S., and Mecke, J., *Stochastic Geometry and Its Applications*, Chichester: Wiley, 1987.

3. Hall, P. and Smith, R.L., The Kernel Method for Unfolding Sphere Size Distributions, *J. Comp. Physics*, 1998, vol. 74, pp. 409–421.

4. Hoogendoorn, A.W., Estimating the Weight Undersize Distribution for the Wicksell Problem, *Statist. Neerl.*, 1992, vol. 46, no. 4, pp. 259–282.

5. Zygmund, A., *Trigonometric Series*, Cambridge: Cambridge Univ. Press, 1959, vol. 2, 2nd ed. Translated under the title *Trigonometricheskie ryady*, Moscow: Mir, 1965.

6. Golubev, G. and Levit, B., Asymptotically Efficient Estimation in the Wicksell Problem, *Ann. Statist.*, 1998, vol. 26, no. 6, pp. 2407–2419.

7. Groeneboom, P. and Jongbloed, G., Isotonic Estimation and Rates of Convergence in Wicksell's Problem, *Ann. Statist.*, 1995, vol. 23, no. 5, pp. 1518–1542.

8. Green, P.J. and Silverman, B.W., *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, London: Chapman & Hall, 1994.

9. Pontriagin, L.S., *Obyknovennye differentsial'nye uravneniya*, Moscow: Nauka, 1970, 3rd ed. First edition translated under the title *Ordinary Differential Equations*, Reading: Addison-Wesley, 1962.

10. Petrov, V.V., *Summy nezavisimykh sluchainykh velichin*, Moscow: Nauka, 1972. Translated under the title *Sums of Independent Random Variables*, Berlin: Springer, 1975.

11. Golubev, G. and Levit, B., On the Second Order Minimax Estimation of Distribution Function, *Math. Methods Statist.*, 1996, vol. 5, pp. 1–31.

12. van Trees, H.L., *Detection, Estimation, and Modulation Theory*, New York: Wiley, 1968, vol. 1. Translated under the title *Teoriya obnaruzheniya, otsenok i modulyatsii*, Moscow: Sov. Radio, 1975.

13. Ibragimov, I.A. and Khasminskii, R.Z., *Asimptoticheskaya teoriya otsenivaniya*, Moscow: Nauka, 1979. Translated under the title *Statistical Estimation—Asymptotic Theory*, New York: Springer, 1981.

14. Grenander, U. and Szégö, G., *Toeplitz Forms and Their Applications*, Berkeley: Univ. of California Press, 1958. Translated under the title *Teplitsevy matritsy i ikh primeneniya*, Moscow: Inostrannaya Literatura, 1961.

15. Natanson, I.P., *Konstruktivnaya teoriya funktsii*, Moscow: GITTL, 1949. Translated under the title *Constructive Function Theory*, New York: Ungar, 1964.